8-2022

Discrete Choice Experiments in Agricultural and Food Economics: Two Essays on Information Provision Modalities, Uncertainty Adjustment, and Hypothetical Bias

Monsoïa Arsène Juste Agossadou
University of Arkansas, Fayetteville

Discrete Choice Experiments in Agricultural and Food Economics: Two Essays on Information Provision Modalities, Uncertainty Adjustment, and Hypothetical Bias

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science in Agricultural Economics

by

Monsoïa Arsène Juste Agossadou
University of Abomey-Calavi
Master of Science in Rural Economics, Sociology and Extension, 2014

August 2022
University of Arkansas

This thesis is approved for recommendation to the Graduate Council.

____
Rodolfo M. Nayga, Jr, Ph.D.
Thesis Director

Di Fang, Ph.D.	Alvaro Durand-Morat, Ph.D.
Committee Member	Committee Member

Abstract

Stated discrete choice experiments are extensively used in applied economics to study preferences and valuation for new products, as well as costs and benefits of new policies and programs. Moreover, information provision experiments widely use the method to examine information effects on different outcomes. This thesis explores two methodological issues in discrete choice experiments: (i) information provision modalities and (ii) hypothetical bias. The first study examines the effect of information modality by testing the effect of using combined text script and audio clip (treatment) versus only text script (control) to convey information in discrete choice experiments. Specifically, the study elicits willingness to accept agricultural field jobs of low-skilled nonmigrant workers in the US amid the COVID-19. Using an online discrete choice experiment, subjects were randomly assigned to one of the information modalities. The findings indicate that respondents treated with the combined text script and audio clip were willing to accept more for three out of seven attributes. The treatment effect was detected for two more attributes when estimates were conditional on attribute attendance. Moreover, the treatment lowered the prevalence of attribute non-attendance. The second paper assesses the effectiveness of budget reminder to mitigate the hypothetical bias relative to cheap talk and cheap talk with budget reminder. Moreover, it explores the impact of accounting for uncertainty directly in the choice tasks on respondents' choices. We conducted a laboratory experiment and randomly assigned subjects to a control group and five treatments to elicit their willingness to pay for animal-based and plant-based burgers. The results suggest that the choice task uncertainty adjustment reduced the likelihood of choosing the no-buy option. Furthermore, respondents exhibited hypothetical bias and overstated their willingness to pay by a factor of 1.29 and 1.40 for the animal-based and the plant-based burgers, respectively. Budget reminder reduced the hypothetical bias for the animal-based burger, while cheap talk and cheap talk combined with

budget reminder eliminated it for both products. This thesis revisits the potential of budget reminder and cheap talk to mitigate hypothetical bias and magnifies the importance of information provision modalities in discrete choice experiments.

Acknowledgements

This thesis is the outcome of the tremendous support I have received from different organizations and people. I am extremely grateful to the William J. Fulbright Foreign Student Program for this opportunity to pursue my studies in the United States. The program has not only strengthened my knowledge and skills but has enriched my understanding of the world.

Many people have contributed to my Master's degree program in Agricultural Economics. Words cannot express my gratitude to my advisor, Dr. Rodolfo M. Nayga, Jr. You are an outstanding teacher, advisor, and mentor. You cared not only about my academic performance but also about my overall well-being. Thank you for providing me with opportunities to enhance my knowledge and skills in experimental economics through external training. Your guidance and involvement during the research cycle were pivotal. I feel blessed to have you as advisor and mentor. My sincere gratitude to Dr. Di Fang, for her contribution to the proposal development, survey design, laboratory experiment, data analysis, and the improvement of the final version of the thesis. I would like to thank, Dr. Alvaro Durand-Morat, for being my committee member and providing his comments on the thesis.

Thanks should also go to Wei Yang. Thank you for your assistance during the experimental design, the survey design in Qualtrics, and your permanent availability to help me with my R codes. I am thankful to Eva Caroline, Ada-Miette Thomas, Vera Adabrah Danquah, Seitov Madaminbek, and Alexis Skinner for their assistance during the data collection.

Lastly, I would be remiss in not mentioning my family, especially my wife and daughter. Thank you for your unwavering support and sacrifice.

Dedication

In memory of:

- my father Jacques Agossadou; and
- my mother Pascaline Hessou.

To

- my lovely spouse, Marguerite; and
- my daughter, Faith.

Table of contents

List of Tables

List of Figures

INTRODUCTION

Discrete choice experiments (DCE) have become the most used stated preference approaches over the last decades. They are used in a wide range of applied economics fields, including agricultural, food, environmental, and health economics (de-Magistris, Gracia, and Nayga Jr 2013). DCE are instrumental as they allow to assess preferences and valuation for new products, preferences, costs, and benefits of new policies and programs (Haghani et al. 2021; Broadbent 2014) to inform policy-making and business decisions (Fang et al. 2021). With DCEs, one can estimate marginal values of different attributes and, as a result, their importance on choice behavior. This advantage has increased the popularity of this approach (Broadbent 2014).

The importance and popularity of DCE have led researchers to investigate many methodological aspects to improve its design and welfare estimates. A group of studies focused on the complexity and cognitive burden of DCE. Evidence suggests that the number of alternatives (Chung, Boyer, and Han 2011; Dellaert, Donkers, and van Soest 2012), the number of attributes and levels (Dellaert, Donkers, and van Soest 2012; Kragt and Bennett 2012; Caussade et al. 2005), the presentation formats (DeLong et al. 2021; Shr et al. 2019; Tarfasa et al. 2017), and the number of choice tasks (Chung, Boyer, and Han 2011; Caussade et al. 2005) can affect choices and welfare estimates. The second group of research emphasized improving the statistical efficiency of DCE (Bliemer, Rose, and Chorus 2017; Rose et al. 2008; Scarpa and Rose 2008). A third group assessed strategies to eliminate or mitigate hypothetical bias (HB) (de-Magistris, Gracia, and Nayga Jr 2013; Fang et al. 2021; Broadbent 2014), the main weakness of hypothetical DCE. However, there is a need for more research in this third area.

Uncertainty has also gained attention in the literature in the past decade. It is generally assumed in DCE that respondents know their true preference with certainty while making a choice. However, they can be uncertain about their choices, and failing to account for preferences uncertainty could affect valuation estimates (Olsen et al. 2011). A wealth of studies that accounted for uncertainty used the ex-post approach. This approach has been however criticized, and there is a need to explore new approach to consider uncertainty in DCE.

An unexplored methodological consideration in DCE is the effect of information provision modality on choices and welfare estimates. The number of studies that assessed the impact of different types of information on selected outcomes has increased over the last decade (Haaland, Roth, and Wohlfart 2022). Information treatment effects on choices and welfare estimates are of particular interest in agricultural and food economics. For example, some studies have investigated the effect of different information provision modalities on choices and valuation (e.g., Zossou et al. 2022; Channa et al. 2019). However, these studies mainly used a single modality such as text, audio, or video. It is however possible to use a dual modality (e.g., text script and audio clip) to provide information to respondents. Given the popularity of DCE and information provision experiments, it is crucial to assess whether a dual modality yields a different result than a single modality.

This thesis examines three methodological issues in DCE. First, it answers whether using a dual information provision modality (text script combined with audio clip) affects choices, willingness to accept (WTA) and attribute non-attendance (ANA) in DCE and to what extent. Second, it assesses the effectiveness of budget reminder to mitigate or eliminate HB relative to popular methods such as cheap talk and cheap talk combined with budget reminder in DCE. Moreover, it explores the effect of adding an uncertainty option in the choice tasks on respondents' choices. We investigated these objectives in two separate essays.

Essay 1 : Information Effects in Discrete Choice Experiments: Does Single Versus Dual Modality Delivery Matter for Valuation Estimates and Attribute Non-Attendance?

Introduction

The potential of information to influence decision-making has gained increasing attention in the literature. Consequently, information provision experiments — economic experiments in which information are provided and/or varied to investigate the choice of economic agents (Haaland, Roth, and Wohlfart 2022), have increased over the last decade. For instance, Haaland, Roth, and Wohlfart (2022) showed that the number of information provision experiments published in top economics journals in 2020 is more than five times the number published in the same outlets in 2010, suggesting the growing importance of these experiments.

Information plays several roles in choice making. It helps economic agents to mitigate uncertainty related to the goods of interests (Handel and Schwartzstein 2018) and is critical in their preference formation (Bateman et al. 2009). Information also affect beliefs and perceived constraints, which are essential in choice making (Haaland, Roth, and Wohlfart 2022). These functions of information explain the growing attention to information effects in general and on preferences and welfare estimates, in particular. Over the last decades, there has been an increasing interest in the effects of information on preferences and valuations.

Discrete choice experiments (DCE) have become a widely used approach in information provision experiments. Applied economists have used experimental auctions, DCE, and contingent valuation methods, to investigate information effects in different contexts and for different public and private goods (e.g., Lombardi et al. 2019; Nayga, Aiew, and Nichols 2005; Weir, Uchida, and Vadiveloo 2021; Wuepper, Wree, and Ardali 2019; Channa et al. 2019). In the above examples of studies, the effects of different types of information (e.g., positive, negative, and two-sided information; novel technology, nutritional information) on one or several attributes are typically assessed. Among available methods, DCE have been largely used by applied economists (Fang et al. 2021) to investigate the effect of information on behavior.

In DCE, the most common way to provide respondents with information is to ask them to read carefully a text script either on paper or on a screen. For example, Hoke et al. (2017) assessed the effect of negative taste, positive health, and two-sided information on US consumers' preferences and willingness to pay (WTP) for aronia berries after respondents read information in an online DCE survey. Similarly, Weir, Uchida, and Vadiveloo (2021) used the same approach to investigate the impact of positive and negative information on genetically modified (GM) technology on GM salmon in the US. The effect of balanced information related to GM bread is also assessed by Wuepper, Wree, and Ardali (2019) in Germany using a within-subjects comparison. Regarding Willingness to Accept (WTA), Luckstead, Nayga Jr, and Snell (2022) conducted an information provision experiment investigating US workers' willingness to accept meatpacking jobs amid the COVID-19 pandemic. A common characteristic of these studies is that they provided the information in written form only.

However, it is also possible that information scripts can be given to respondents as an audio clip, an image (visual), video, narrative, or using a dual modality in economic experiments. The dual modality consists of using two different modes (e.g., text and audio clips) of presentation to provide information to subjects (Chang and Millett 2015; Moreno and Mayer 2002). For instance, using experimental auctions, Channa et al. (2019) investigated the effect of providing information through text scripts, audio clips, and videos on farmers' relative WTP for storage bags. Bateman et al. (2009), on the other hand, used virtual reality to investigate differences in preferences and WTP compared with the standard presentation in a DCE. Mokas et al. (2021) conducted a similar study but used video and virtual reality formats of the product profile compared to the standard text format. However, none of Bateman et al. (2009) and Mokas et al. (2021) assessed the effect of information provision. Rather, they exploited virtual reality to make the choice environment more salient. In fact, the rapid development of digital technology (Rogowsky, Calhoun, and Tallal 2016) has induced a substantial increase in digital audio consumption (Baskin and Harris 1995; Daniel and Woody 2010) and has facilitated the use of audio-assisted reading (Chang and Millett 2015). Consequently, listening, reading, or doing both simultaneously using digital devices has become popular (Rogowsky, Calhoun, and Tallal 2016).

Evidence suggests that the modality used to provide information can influence information comprehension and recall (Chang 2009; Chang and Millett 2015; Daniel and Woody 2010; Diao and Sweller 2007; Lund 1991; Moreno and Mayer 2002). For instance, a comparison of single modalities revealed that reading provides better results in terms of information understanding and recall than listening (Daniel and Woody 2010; Lund 1991). Moreover, using a dual modality including a combination of text scripts and audio clips instead of text only scripts significantly increased information comprehension and recall. Therefore, scientists have called for the use of this dual modality to convey information (Chang 2009; Chang and Millett 2015; Daniel and Woody 2010; Diao and Sweller 2007). No other study, however, has directly examined whether the way that information is conveyed to respondents can affect choice behavior and valuation estimates in DCE. This issue is important given the immense popularity of DCE and information testing.

The information provision modality in a DCE could affect outcomes through a three-step mechanism. First, it could directly impact respondents’ level of understanding and the amount of information retained (Chang 2009; Chang and Millett 2015; Daniel and Woody 2010; Diao and

Sweller 2007). This step is critical since participants are usually asked to perform repeated choice tasks. To ensure that respondents recall information as they go through these repeated processes, researchers usually ask participants to carefully read the information and keep it in mind as they make their choices. However, there is no guarantee or proof that they do so.

Second, the amount of information kept in memory could in turn influence how respondents consider attributes, especially those related to the information provided, as they complete the choice tasks. This possibility is crucial given the main DCE assumption that respondents consider all product attributes when making their choice (Kragt 2013). Nevertheless, it is now well established that participants may ignore one or several product attributes during choice tasks in DCE. This phenomenon is known as attribute non-attendance (ANA) (Caputo et al. 2018a; Hensher, Rose, and Greene 2012).

Finally, how respondents attend to attributes could affect their valuation estimates. Consequently, understanding the attributes ignored and their drivers in DCE is crucial for accurate assessment of valuation estimates and welfare predictions (Caputo et al. 2018a). Hence, investigating the effect of the information provision modality is paramount. This investigation is also becoming more crucial with the increasing number of online DCE studies being conducted overall. The use of online DCE could raise concerns regarding the accuracy of the responses. For example, Savage and Waldman (2008) assessed the difference between online and mail survey modes in the use of a stated DCE and show that the variability in the error of participants' utility was 25% higher in the online survey mode than in the mail mode.

In this paper, we examine whether providing information using a dual modality (with a text script and an audio clip presenting the same information) versus a single modality (with a text script only) in a DCE affects choices, valuation estimates, and ANA. We also investigate the difference in the duration of choice tasks completion across groups. While we could also assess the effect of using an audio clip only (listening only), we decided not to investigate this information modality because, as discussed, previous research has shown that information comprehension and recall are lower when listening only than when reading only. We discuss this point in more detail in the next section. To achieve our objective, we randomly assigned respondents to one of the information modalities in an online DCE that investigates US nonmigrant workers' WTA agricultural field jobs amid the COVID-19 pandemic.

Respondents' utility was estimated using a random parameter logit with error component (RPL-EC) model under full attribute attendance and conditional on self-reported serial attendance. Moreover, consistency between self-reported serial ANA and choice behavior was assessed using a validation model (Caputo et al. 2018a). The findings indicate that conveying information using both text script and audio clip affected relative importance of job attributes, WTA, and ANA. Respondents treated with the dual modality were willing to accept more for three out of seven attributes in comparison to those who received information in only text format. Conditional on ANA, WTA of those in the treatment were higher than WTA for those in the control group for five out of seven attributes. Moreover, the dual modality lowered the prevalence of ANA.

To the best of our knowledge, this study is the first to assess the effect of the information provision modality on respondents' choices, WTA, and ANA in a DCE. It differs from other studies in many ways. First, our study differs from works that used presentation formats (e.g., images and virtual reality) of product profiles to improve product evaluation, realism, and respondents' certainty in DCE (e.g., Bateman et al. 2009; Matthews, Scarpa, and Marsh 2017; Mokas et al. 2021). We assess the treatment effect of using a dual modality by providing the same information to respondents and determining the effect of type of modality on choices and valuations. Therefore, the paper does not investigate the effect of a specific information on respondents' choices and valuation. As such, it extends the previous literature on DCE and information effects by providing new insights into how the modality used to provide information to respondents could affect the results of DCE. Furthermore, it can help applied economists improve the design of information and the assessment of how information affects behavior. Similarly, companies and policy decision makers can use our findings to improve their interventions.

The remainder of the paper proceeds as follows. The following section presents the theoretical background related to the information modality and literature review followed by the section describing the methods. Afterward, the results are presented followed by the discussion and the conclusion.

Theoretical Background and Literature Review

We use dual processing theory and cognitive load theory to explain the difference between the effects of dual versus single information provision modalities. Dual processing theory postulates that visual and audio information are processed by two different working memories. The visual mode is processed in visual working memory, while the audio mode is processed in auditory working memory (Moreno and Mayer 2002). These two systems are independent, and they contribute to building a better and coherent representation of information. In fact, people can process information from both types of working memory without being cognitively overloaded. According to this theory, subjects retain more verbally redundant information than information presented in a single mode without visual material provided simultaneously (Moreno and Mayer 2002). Furthermore, the use of the dual modality reduces the perceived difficulty of a passage, compels respondents to pay greater attention to the information provided (Chang 2009; Chang and Millett 2015) and reduces the time spent reading the material (Chang and Millett 2015).

Cognitive load theory, on the other hand, provides the opposite explanation. This theory posits that relative to conveying information using a single medium, using a dual modality deteriorates learning and comprehension (Diao and Sweller 2007). According to this theory, listening interferes with reading and requires more cognitive activity, which hinders comprehension. This detrimental effect is driven mainly by the attention split that is required to process two information modalities (Diao and Sweller 2007; Moreno and Mayer 2002) since two different sensory signals are used (Moreno and Mayer 2002).

We hypothesize that if respondents process information differently when it is provided in text versus text and audio form, then it is possible that this difference would influence the amount and comprehension of information retained, which would consequently affect how they attend to attributes, choices, and valuation estimates. In fact, some of the initial information provided to subjects could be lost, which would hamper their capacity to assess effectively presented choice alternatives. Moreover, there are differences in individuals' ability to process information (Turner and Makhija 2012). Given the popularity of studying information effects using DCE, it is important to investigate which information provision medium maximizes information recall during choice tasks.

Studies related to the effect of information modality can be classified into three groups. The first group includes works that have found no impact of information modality on comprehension and information retention. For instance, Rogowsky, Calhoun, and Tallal (2016) found no treatment effect of digital audio, electronic text, and dual (electronic text and digital audio) modalities on respondents' immediate and two-week retention of material regardless of the subject's gender. The second group of works reported reading as superior to listening in terms of understanding**.** Lund (1991) demonstrated that readers performed better on a comprehension test than listeners and therefore concluded that the former retained more information and details than listeners did. Daniel and Woody (2010) found similar results. They showed that relative to listening, reading increased comprehension even when no statistically significant difference was found in the time spent by each group reading or listening to information. Unlike previous studies, Diao and Sweller (2007) compared reading to reading and listening to the same information at the same time. They concluded that reading was more effective than reading and listening to the same material simultaneously.

Lastly, studies in the third group have proven the superiority of the dual modality of reading and listening simultaneously over reading or listening only. Chang (2009) showed that relative to listening only, the dual modality increased respondents' comprehension test score by 10%. In addition, the findings suggest that the majority of respondents preferred the dual modality over listening. Similarly, Chang and Millett (2015) reported that respondents in the reading and listening group retained significantly more information and performed well on the immediate and three-month delayed comprehension tests. In addition, those exposed to the dual modality read at a relatively faster pace. In the same line, Moreno and Mayer (2002) demonstrated that reading and listening simultaneously increased information recall and generated significantly more conceptual solutions than reading only. It appears that reading provides better results than listening. However, providing information using text and audio clips could provide better results than reading only.

Methods

This section describes the experimental design, the data collection, the conceptual framework, the empirical models, and the hypothesis testing.

Experimental Design

A DCE was implemented to investigate US nonmigrant workers' WTA agricultural field jobs. The study used an unlabeled DCE. Job options were described using hourly wages and a combination of nonpecuniary attributes that included health insurance, transportation, housing, food/clothing allowances, and duration of employment in months (Table 1.1). The hourly wage rate had five levels ranging from $5.15/hour to $23.60/hour, while the duration (in months) of employment had four levels (three, six, nine, and 12). For each of the remaining nonmonetary attributes, two levels were used: offered or not offered. Each choice task included two alternatives and a "neither of these" option. It was clearly explained to respondents that the latter option meant that they prefer their current employment status to an agricultural field job.

Table 1.1 DCE Attributes and Attribute Levels

Attribute	Attributes levels
Health insurance	Offered
	Not offered (baseline)
Transportation	Offered
	Not offered (baseline)
Housing	Offered
	Not offered (baseline)
Food and clothing allowance	Offered
	Not offered (baseline)
Duration of employment	3 months
	6 months
	9 months
	12 months (baseline)
Hourly wage	$5.15/hour
	$9.76/hour
	$14.38/hour
	$18.99/hour
	$23.60/hour

The choice tasks were generated in NGene using a D-efficient Bayesian design. Following Scarpa et al. (2013), a three-step procedure was used. First, a pilot study was conducted with 130 respondents using an optimal orthogonal design. Next, coefficient estimates were generated by estimating a multinomial logit. Finally, these estimates were used as Bayesian prior mean values to generate an efficient Bayesian design. Twenty choice tasks were generated and broken into 2 blocks of 10 choice tasks each to prevent fatigue (Savage and Waldman 2008) with an efficiency D-error of 0.105. The number of choice tasks we used is reasonable given that some studies used more than 10 choice tasks (e.g., DeLong et al. 2021; De Marchi et al. 2016). Choice tasks and alternative orders were randomized to prevent order effects (Hoke et al. 2017; Van Loo et al. 2015). Figure 1.1 presents an example of a choice task.

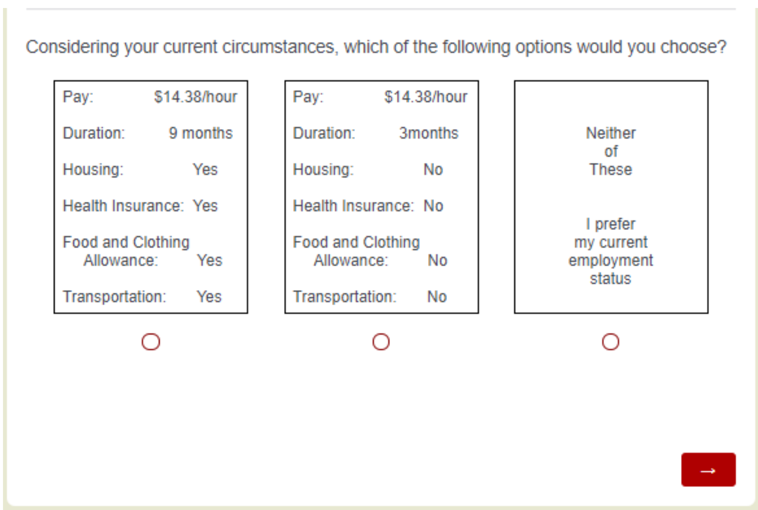

Figure 1.1 Example of a Choice Task in the Agricultural Field Jobs DCE

The DCE tasks were performed following a rigorous procedure. First, respondents were provided with general instructions, and a practice session with one choice task. Respondents were then provided with a set of information. The study used a combination of oath, consequentiality, and a short cheap talk to mitigate potential hypothetical bias. The choice tasks were presented after the hypothetical mitigation methods.

Experimental Treatments

The design used a between-subject approach to estimate the effect of the combined text and audio modality on respondents' choices, valuation estimates, and ANA. Respondents were randomly assigned to the text and audio treatment or the text only modality (control group). In the latter, respondents were provided with the text information through the conventional approach, i.e., they were asked to read the text script. In the former, which is the treatment group, respondents were given the same information in both text format and an audio clip. For simplicity, in the remainder of the paper, we use "ToT" to denote the control and "TAT" to represent the text and audio treatment. In the treatment group, respondents read and listened to the same information simultaneously. Subjects were provided with 2 sets of information of approximately 200 words. The information included a description of the labor shortage in agricultural production, especially the disruption created by COVID-19 on the labor market in the agricultural sector. The second set of information describes the use of H-2A agricultural guest workers and the challenges they face in their jobs (see Appendix A.1). In the TAT group, the "next" button appeared only after each audio clip was fully played. Likewise, in the ToT group, the "next" button appeared only after the lapse of sufficient time to read the text script.

Data Collection

This paper used data collected through an online DCE in April 2020 amid the COVID-19. The Institutional Review Board of the University of Arkansas provided full approval for the survey (protocol #1910222312). The data were collected using Qualtrics, and the recruitment and randomization of respondents were performed using Dynata. The study targeted only respondents aged between 18 and 65 years old, are low-skilled, and capable of competing with H-2A workers. Moreover, targeted respondents are without a college degree, with an annual individual income below $50,000, and with no disability preventing them from lifting at least 10 pounds. In addition to these criteria, the TAT was restricted to those who passed the technical audio pretest. This test was performed to ensure that respondents could clearly hear the audio clip on their devices.

The survey questionnaires include four parts. First, subjects consent and oath were collected, and participants were provided with a description of the agricultural workers duties. Second, respondents were provided with the information in either the dual modality format or the text only format, depending on their group. Third, respondents indicated their choices for each randomly presented choice task. The DCE was followed by a group of debriefing questions, including ANA questions. Stated serial ANA, which consists of asking respondents an ANA question at the end of all choice tasks (Caputo et al. 2018a), was used. The serial ANA question was framed as follows: "During the previous choice task questions, which of the following aspects of the job offers did you ignore when making a decision?" Finally, sociodemographic data on respondents and their households were collected. 679 respondents were surveyed.

Conceptual Framework

We employed random utility theory of McFadden (1973) and the Lancaster theory of consumer behavior of Lancaster (1966) as the theoretical framework. According to random utility theory, the utility that respondent n obtains from alternative i in choice task s is defined as follows:

(1.1) $U_{ins} = \beta' X_{ins} + \varepsilon_{ins}$

where β' is a vector of taste parameters characterizing choices, X_{ins} is a function of observable variables and ε_{ins} is the idiosyncratic unobserved error, which is assumed to be independent and homogenous across participants, alternatives, and choice tasks in the standard multinomial logit. However, based on the unlikelihood of meeting this assumption (Hensher, Rose, and Greene

2005) and evidence of heterogeneity in preferences in past studies (Caputo et al. 2018a; Van Loo et al. 2015), we used a random parameter logit with an error component to estimate respondents' WTA. RPL-EC considers not only heterogeneity across participants but also correlation across utilities. Heterogeneity is considered in the model by allowing the coefficients of alternatives' attributes to vary randomly across respondents and to deviate from the population mean. Correlation across utilities is included by identifying the additional variance in the utility of the two designed alternatives (Scarpa, Willis, and Acutt 2007). Utility in an RPL-EC context is defined as follows (Train 2009).

$$(1.2) \quad U_{ins} = \beta_i' X_{ins} + \mathrm{d}_i(\eta_{in}) + \varepsilon_{ins}$$

where $\mathrm{d}_i(.)$ is a binary function indicator with a value of 1 for the two hypothetical alternatives and 0 for the no-buy alternative. η_{in} represents a zero-mean normal individual-specific idiosyncratic error, the error component. It is related only to the two hypothetical alternatives and is absent in the utility of the no-buy alternative (Scarpa, Willis, and Acutt 2007). ε_{ins} is assumed to be an independently and identically distributed (IID) extreme value error term.

The choice probability of choosing alternative i is expressed as follows (Train 2009).

$$(1.3) \quad P_{ins} = \int \left(\frac{e^{\beta_i' X_{ins} + \mathrm{d}_i(\eta_{in})}}{\sum_{j=1}^{J} e^{\beta_j' X_{jns} + \mathrm{d}_j(\eta_{jn})}} \right) \emptyset(\beta | b, W) d\beta$$

where $\emptyset(\beta | b, W)$ is the normal density with mean b and covariance W to be computed.

Choosing the "neither of these" option means that the respondent prefers the current job. Therefore, values of the attributes in that option were not null for those with a job, and these values were adjusted. For instance, for those with a job at the time of the survey, the hourly wage in the no-buy option is the hourly wage of their job at the time of the survey. Furthermore, if the job provides any nonwage benefits such as insurance, food/clothes, transportation, or housing, the values were adjusted under the no-buy option for the respondent.

Empirical Specifications

Standard RPL-EC Assuming Full Attribute Attendance

Several empirical models were implemented using dummy coding for attributes and attribute levels. We start with a standard (unconditional) RPL-EC under the assumption that respondents fully attended to all attributes. Equation 1.4 presents the utility function of a nonmigrant worker. The buy option represents the alternative specific constant (ASC). To capture the error component, the parameter of the buy option is set to be normally distributed. The model was estimated for both ToT and TAT.

$$(1.4)\quad U_{ins} = \beta_0 Buy_{ins} + \beta_{1n} Wage_{ins} + \beta_{2n} Length3_{ins} + \beta_{3n} Length6_{ins} + \beta_{4n} Length9_{ins} + \beta_{5n} Housing_{ins} + \beta_{6n} Insure_{ins} + \beta_{7n} Food_{ins} + \beta_{8n} Transport_{ins} + \beta_{9n}(Housing_{ins} * Insure_{ins}) + \beta_{10n}(Food_{ins} * Transport_{ins}) + \varepsilon_{ins}$$

Modeling Serial Stated ANA

We first investigated difference in ANA prevalence as well as well ANA for each job attribute across groups using a Pearson's chi-squared and Fisher's exact tests.

To consider serial stated ANA in the empirical modeling, two different models are tested for each group. First, the model referred to as the conventional ANA model by Caputo et al. (2018a) was estimated. The coefficient estimates in these models were conditional on attribute attendance in the serial task (Campbell and Lorimer 2009), and self-reported ignored attributes were set to 0 (Caputo et al. 2018a). The conventional ANA utility function is presented in equations 1.5.

$$\text{(1.5)} \quad U_{ins} = \delta_0 Buy_{ins} + \delta_{1n}(Awag_n * Wage_{ins}) + \delta_{2n}(Al3_n * Length3_{ins}) + \delta_{3n}(Al6_n * Length6_{ins}) + \delta_{4n}(Al9_n * Length9_{ins}) + \delta_{5n}(Ahous_n * Housing_{ins}) + \delta_{6n}(Ainsu_n * Insure_{ins}) + \delta_{7n}(Afood_n * Food_{ins}) + \delta_{8n}(Atrans_n * Transport_{ins}) + \delta_{9n}((Ahous_n * Housing_{ins}) * (Ainsu_n * Insure_{ins})) + \delta_{10n}((Afood_n * Food_{ins}) * (Atrans_n * Transport_{ins})) + \varepsilon_{ins}$$

where $Awag$, $Al3$, $Al6$, and $Al9$ are respective dummy variables that take a value of 1 if the attributes of hourly wage and the proposed job durations of three months, six months, and nine months, respectively, were considered and 0 otherwise. Similarly, $Ahous$, $Ainsu$, $Afood$, and $Atrans$ are dummy variables that indicate whether the attributes of housing, insurance, food/clothing, and transport, respectively, were attended to during the experiment.

The second model was developed following Scarpa et al. (2013) and Caputo et al. (2018a). It aims to test whether the choice behavior of respondents declaring non-attendance matches their responses to the serial ANA questions (Scarpa et al. 2013) in the two groups. This model, referred to as the validation model by Caputo et al. (2018a), reveals whether attribute non-attendants actually ignored the attributes. In this model, two different coefficients are estimated for each attribute. One coefficient is related to the attribute for respondents who consider it, and the second is associated with the attribute for those who declare having ignored it. These outcomes were estimated by decomposing β_n for the attributes in equations 1.4 into two different coefficients, thus capturing the parameters conditional on attribute attendance and ANA as follows:

$$\text{(1.6)} \quad \beta_{nk} = \delta_{nk} + \gamma_{nk}$$

where δ_{nk} represents the parameters for individual n and attribute k as in equations 1.5, which show the utility conditional on attribute attendance. γ_{nk} represents the parameters for individual n and attribute k conditional on attribute k having been declared ignored by individual n. The validation model is presented in equations 1.7.

$$(1.7)\quad U_{ins} = \delta_0 Buy_{ins} + (\delta_{kn}A_{kn} + \gamma_{kn}ANA_{kn}) * [Wage_{ins} + Length3_{ins} + Length6_{ins} + Length9_{ins} + Housing_{ins} + Insure_{ins} + Food_{ins} + Transport_{ins}] + \delta_n^{c1}((Ahous_n * Housing_{ins}) * (Ainsu_n * Insure_{ins})) + \delta_n^{c2}((Afood_n * Food_{ins}) * (Atrans_n * Transport_{ins})) + \gamma_n^{c1}((ANAhous_n * Housing_{ins}) * (ANAInsu_n * Insure_{ins})) + \gamma_n^{c2}((ANAfood_n * Food_{ins}) * (ANAtrans_n * Transport_{ins})) + \varepsilon_{ins}$$

where ANA_{nk} is an indicator variable that takes a value of 1 if attribute k is ignored by individual n , and 0 otherwise. $c1$ denotes the interaction between housing and insurance, and $c2$ the interaction between food and transport.

There is concern that including ANA in the utility function can lead to biased estimates due to potential endogeneity of self-reported ANA (Hole, Kolstad, and Gyrd-Hansen 2013). However, as pointed out by the same authors, this issue is not a concern when heterogeneity in preferences is accounted for in the estimation process, as we did in this study. All models were estimated in preference space with 500 iterations for the simulation using the package “gmnl” version 1.1.3.2 (Sarrias and Daziano 2017) in RStudio 2022.02.2+485. We used preference space since we were interested not only in respondents’ WTA but also preferences for job attributes.

Prior to interpreting results per group, we tested for preferences equality between TAT and ToT by performing a likelihood ratio (LR) test following Zellner (1962) using the standard

RPL-EC model. Rejecting the null hypothesis suggests that it is appropriate to estimate models separately for each group. The LR statistics is as follows (Zellner 1962).

$$(1.8) \quad LRstatistic = -2\left(LL_j - \sum LL_i\right)$$

where LL_j is the log likelihood of the pooled data and LL_i the loglikelihood of the model of each group. The LR statistic is compared to χ^2 critical value with $K(M-1)$ degree of freedom; K indicates the number of parameters in the model, and M is the number of treatments.

Testing the WTA Difference between TAT and ToT

We computed WTA as β_k/β_w where β_k denotes attribute k's parameter estimates, and β_w indicates the hourly wage parameter estimate. Dividing two random parameters could yield infinite moments (Revelt and Train 1999); hence, we fixed the wage and allowed all nonwage attributes to be randomly distributed in all the models.

Given the nonlinearity between coefficients and WTA, a common way of testing the difference between groups may be erroneous (Howard 2017). We therefore tested the difference in WTA across groups using the complete combinatorial test recommended by Poe, Giraud, and Loomis (2005). The null hypothesis of equal WTA is tested for each attribute based on the WTA generated using the parameter estimates from the standard RPL-EC and the conventional ANA model. This process was performed in two steps. First, we computed the average WTA for each attribute using the parametric bootstrapping method of Krinsky and Robb (1986). Following Haab and McConnell (2002) and Carlsson, Mørkbak, and Olsen (2012), we used 10,000 replications to generate the average WTA estimates. The combinatorial test was implemented using these 10,000 values from each treatment to test the difference in valuation estimates between the TAT and ToT. Furthermore, standard t tests were performed. Using a one-sided t test, we also tested whether respondents in the dual modality group spent on average less time completing the choice tasks than those in the single information provision modality.

Results

Descriptive Statistics

Table 1.2 reports the sociodemographic characteristics of respondents and the results of the balance test across TAT and ToT. A total of 331 nonmigrant workers were randomly assigned to the TAT, while 343 were assigned to the ToT, for a total of 674 participants[1]. Fifty-three percent of the respondents were female, and 92% had completed primary or high school as the highest level of education (no certificate). However, there were more respondents with no certificate (96%) in ToT than in TAT. The majority (80%) of respondents reported never having an agricultural field job in both experiment groups. Thirty-eight percent and 32% of the respondents had a household income less than 20,000 and between \$20,000 to \$34,999, respectively. Overall, 52% of subjects did not lost their job or experienced a pay cut due to the COVID-19. The results of the χ^2 tests show that balance was achieved (p value >0.05) for all nine sociodemographic characteristics except for education. Consequently, this latter variable was controlled for in the models.

[1] Initially, the sample was 679, but 5 respondents (1 in the text and audio group and 4 in the text-only group) were dropped. These respondents stated that they would not provide honest answers to the questions.

Table 1.2 Descriptive Statistics of US Nonmigrant Workers and Balance Test

Characteristic	N	Overall, N = 674[a]	Text_Audio, N = 331	Text_only, N = 343	p-value[b]
Gender	674				>0.9
Male		317 (47.0%)	155 (46.8%)	162 (47.2%)	
Female		357 (53.0%)	176 (53.2%)	181 (52.8%)	
Age class	674				0.2
[18,24)		142 (21.1%)	65 (19.6%)	77 (22.4%)	
[24,32)		129 (19.1%)	56 (16.9%)	73 (21.3%)	
[32,42)		144 (21.4%)	72 (21.8%)	72 (21.0%)	
[42,55)		129 (19.1%)	63 (19.0%)	66 (19.2%)	
[55,64]		130 (19.3%)	75 (22.7%)	55 (16.0%)	
Education	674				<0.001
Hold a certificate[c]		57 (8.5%)	42 (12.7%)	15 (4.4%)	
No		617 (91.5%)	289 (87.3%)	328 (95.6%)	
Number of adults	674				0.7
[1,3)		453 (67.2%)	225 (68.0%)	228 (66.5%)	
[3,7]		221 (32.8%)	106 (32.0%)	115 (33.5%)	
Number of children	674				0.6
0		474 (70.3%)	238 (71.9%)	236 (68.8%)	
[1,3)		145 (21.5%)	69 (20.8%)	76 (22.2%)	
[3,6]		55 (8.2%)	24.0 (7.3%)	31 (9.0%)	
Agricultural field job	674				>0.9
Yes		57 (8.5%)	29 (8.8%)	28 (8.2%)	
In the past		77 (11.4%)	37 (11.2%)	40 (11.7%)	
No		540 (80.1%)	265 (80.1%)	275 (80.2%)	
Employment status	674				0.4
Employed full time		232 (34.4%)	111 (33.5%)	121 (35.3%)	
Employed part time		142 (21.1%)	65 (19.6%)	77 (22.4%)	
Unemployed looking for work		145 (21.5%)	70 (21.1%)	75 (21.9%)	
Unemployed not looking for work		155 (23.0%)	85 (25.7%)	70 (20.4%)	
Household income class	674				0.4
<$20,000		253 (37.5%)	120 (36.3%)	133 (38.8%)	
$20,000 to $34,999		219 (32.5%)	111 (33.5%)	108 (31.5%)	
$35,000 to $49,999		171 (25.4%)	89 (26.9%)	82 (23.9%)	
$50,000 to $69,999		18 (2.7%)	6 (1.8%)	12 (3.5%)	
$70,000 to $109,999		6 (0.9%)	1 (0.3%)	5 (1.5%)	
110,000 or more		7 (1.0%)	4 (1.2%)	3 (0.9%)	
Loss of income due to COVID-19	673				0.12
Yes, permanently		45 (6.7%)	24 (7.3%)	21 (6.1%)	
Yes, temporarily		148 (22.0%)	60 (18.1%)	88 (25.7%)	
No but payment cut		127 (18.9%)	66 (19.9%)	61 (17.8%)	
No		353 (52.5%)	181 (54.7%)	172 (50.3%)	

Notes: [a] n (%).
[b] Pearson's chi-squared test; Fisher's exact test.
[c] participants with a certificate are those who attended a career school (technical or vocational) or junior school/community college with degree or certificate awarded

Standard RPL-EC and WTA Comparison across Groups

The details of model selection are presented in appendix B.1. The likelihood ratio test of the standard RPL-EC model suggests the strong rejection of the null hypothesis of equality between the pooled model and models from segmented samples (($\chi 2 = 107.50$; $p < 0.01$). We, therefore, reports the results of the estimation for each group separately.

Table 1.3 presents the parameter estimates for the standard RPL-EC across groups and for the pooled data. This model assumes that respondents completely attended to all attributes. For both groups, the coefficient estimates of the ASC for the buy option are negative and significant, revealing that US nonmigrant workers' utility decreased when they choose any of the two proposed jobs.

As expected, the hourly wage coefficient estimate is positive and significant. This finding is consistent with the a priori expectation that utility increases as the proposed hourly wage increases. The estimates of the standard deviations are significant for all nonmonetary attributes, showing evidence of substantial unobserved heterogeneity in preferences for job attributes across respondents. Additionally, the estimate of the error component is significant, indicating that it is more suitable to allow for a more flexible substitution pattern than that reflected under the usual assumption made in the multinomial logit model (Alemu et al. 2013).

In the TAT, the coefficients estimate of insurance, food/clothing, transportation, and the interaction between housing and insurance are positive and significant at the 5 % significance level. Similar results are found in the control group. Moreover, housing coefficient estimate is barely significant in both groups ($p < 0.1$). In both groups, the coefficient estimates of the durations of three months, six months, and nine months of the hypothetical jobs are negative and significant. In other words, relative to a 12-months job, choosing a job lasting three, six, or nine months decreased participants' utility. This result suggests that both information provision modalities yielded the same result regarding the effect of job duration on respondents' utilities. This is not the case for the interaction between food and transport. The association of food and transport decreased utility in TAT, while there is no significant effect in ToT. This indicate that there is one more determinant of utility in TAT. The findings suggest that holding a certificate increased the marginal utility of wages for treated respondents but decreased the marginal utility of a job duration of three months and insurance, holding other factors in the TAT constant. However, in the ToT, holding a certificate decreased the marginal utility with respect to food ($p < 0.1$). The relative importance of attributes is slightly different across groups. Food was the second most important attribute in TAT, while transport held the same ranking in ToT.

Panel A in Table 1.4 presents the WTA estimates for the standard RPLE-EC model. Both treatments resulted in a positive WTA for the nonmonetary attributes housing, insurance, food/clothing, and transport. The three levels of the duration attribute recorded a negative WTA. However, the ranking of WTAs is different across groups. In the TAT, food/clothing is the most valued attribute, with a WTA of $9.50/hr., followed by transport ($8.02/hr.) and insurance ($4.15/hr.). In the ToT, transport records the highest WTA, with an average WTA of $5.08/hr. followed by food/clothing ($4.58/hr.), insurance ($4.33/hr.), and housing ($2.82/hr.). In both groups, the WTAs for the attributes related to the duration of the hypothetical job are negative for durations of three months, six months, and nine months relative to that for 12 months. This result suggests that respondents valued more the 12-months job compared to three, six, or nine-months job. Comparison of the average WTAs across TAT and ToT resulted in the rejection of the null hypothesis for duration of nine months ($p<0.1$), food/clothing ($p< 0.01$), and transport ($p<0.1$).

Table 1.3 Estimate from the Standard RPL-EC (model 1) Across the TAT and ToT[a]

	Text and audio (TAT)		Text only (ToT)		Pooled data	
	Estimate (se)[b]	Std. (se)	Estimate (se)	Std. (se)	Estimate (se)	Std. (se)
Buy	-2.387*** (0.412)		-2.664*** (0.386)		-1.615*** (0.261)	
Error component		6.677*** (0.466)		6.619*** (0.442)		6.845*** (0.344)
Wage	0.172*** (0.012)	-	0.196*** (0.013)	-	0.181*** (0.009)	
Duration of 3 months	-0.562*** (0.191)	1.908*** (0.213)	-0.563*** (0.175)	1.386*** (0.252)	-0.621*** (0.118)	1.421*** (0.137)
Duration of 6 months	-0.962*** (0.265)	2.100*** (0.282)	-0.689*** (0.227)	0.543** (0.262)	-1.047*** (0.178)	1.625*** (0.192)
Duration of 9 months	-1.176*** (0.268)	1.943*** (0.284)	-0.729*** (0.223)	0.671** (0.291)	-1.115*** (0.177)	1.375*** (0.190)
Housing	0.571* (0.323)	1.889*** (0.163)	0.554* (0.301)	2.080*** (0.180)	0.455** (0.211)	1.833*** (0.114)
Insurance	0.714** (0.335)	1.817*** (0.176)	0.850*** (0.286)	1.920*** (0.169)	0.701*** (0.201)	1.632*** (0.107)
Food	1.636*** (0.253)	0.865*** (0.160)	0.898*** (0.239)	1.294*** (0.177)	1.265*** (0.171)	0.845*** (0.104)
Transport	1.380*** (0.249)	0.835*** (0.143)	0.996*** (0.235)	0.939*** (0.127)	1.219*** (0.167)	0.732*** (0.094)
Housing x Insurance	1.717*** (0.394)	-	1.533*** (0.391)	-	1.519*** (0.272)	-
Food x Transport	-1.149*** (0.362)	-	-0.433 (0.347)	-	-0.907*** (0.253)	-
Certificate*Wage	0.083** (0.033)	-	0.014 (0.032)	-	-0.028 (0.021)	-
Certificate*Duration3	-1.080** (0.465)	-	0.145 (0.772)	-	-0.409 (0.361)	-
Certificate*Duration6	0.327 (0.606)	-	0.149 (1.089)	-	0.346 (0.498)	-
Certificate*Duration9	0.637 (0.621)	-	0.305 (1.069)	-	0.476 (0.502)	-
Certificate*Insurance	-1.485*** (0.458)	-	0.841 (0.772)	-	-0.644* (0.384)	-
Certificate*Housing	-0.170 (0.420)	-	-0.876 (0.717)	-	-0.281 (0.366)	-
Certificate*Food	0.231 (0.306)	-	-1.080* (0.651)	-	-0.010 (0.278)	-
Certificate*Transport	-0.058 (0.288)	-	0.440 (0.532)	-	-0.066 (0.243)	-
N[c]	3310		3430		6740.000	
Log-likelihood	-2177.514		-2091.01		-4322.273	
AIC	4465.029		4292.012		8754.545	
BIC	4800.788		4629.729		9129.415	

Notes: [a] The standard RPL-EC assumes full attendance to attributes in the choice tasks.
[b] ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.
[c] Denotes the number of observations (choices scenarios for all respondents).

Thus, we conclude that the WTAs for duration of nine months, food/clothing, and transport are significantly different between TAT and ToT. Respondents in TAT were willing to accept more for these attributes in comparison to subjects in ToT. They valued more the 12 months jobs relative to a nine-month job. Nonmigrants workers in TAT would require an additional $6.83/hr. to accept a nine-month job while those in ToT would require only $3.72/hr.

Table 1.4 WTA Estimates from the Standard RPL-EC and the Conventional ANA Model Across Groups

Attributes	Text and Audio (TAT) MWTA[a]	Text only (ToT) MWTA	Difference Poe p value[b]	T p value[c]
Panel A. Results based on the assumption of complete attribute attendance (Standard RPL-EC)				
Duration of 3 months	-3.267	-2.871	0.396	0.423
Duration of 6 months	-5.591	-3.515	0.16	0
Duration of 9 months	-6.832	-3.72	0.067	0
Housing	3.316	2.823	0.422	0.393
Insurance	4.148	4.333	0.464	0.455
Food/clothing	9.504	4.579	0.007	0
Transport	8.018	5.078	0.07	0
Panel B. Results conditional on attribute attendance (Conventional ANA model)				
Duration of 3 months	-2.686	-3.233	0.337	0.057
Duration of 6 months	-4.446	-2.714	0.152	0
Duration of 9 months	-5.67	-2.54	0.034	0
Housing	6.261	2.676	0.031	0
Insurance	6.571	3.622	0.052	0
Food/clothing	5.489	2.296	0.013	0
Transport	4.641	2.334	0.051	0

Notes: [a] WTA are computed with Krinsky-Robb's bootstrapping approach with 10,000 replications using the original covariance matrix from the standard RPL-EC and the conventional ANA model results.
[b] 1-tailed test of equal WTAs across the TAT and ToT based on the complete combinatorial test of Poe, Giraud, and Loomis (2005). P values of the Poe test indicate the proportion of estimates in the complete combinatorial where the WTA from the group obtaining the highest mean WTA is higher than WTA from the other group.
[c] 1-tailed test of equal WTAs across the TAT and ToT based on the t test. Alternative hypothesis is that the greater value is higher than the lower value.

Prevalence of Stated Serial ANA

Results of the prevalence of ANA are presented in table 1.5. We present the results by education category to account for difference in education. We focus on respondents with no certificate, as there is no difference in the proportion of attribute nonattendants between the TAT and ToT for respondents holding a certificate. Among respondents with no certificate, 38% and

48% reported having ignored at least one attribute in the TAT and ToT, respectively. That is, the prevalence of ANA is 10 percentage points lower in the treatment group than in the control group. This reduction in attribute non-attendance is significant at the 5% significance level. In both groups, the most frequently ignored attributes were transport, food/clothing, and housing. Comparing the proportion of nonattenders to individual attributes reveals that except in relation to the hourly wage attribute, fewer respondents reported having not attended to attributes in the TAT than in the ToT. However, the difference between treatments is significant at the 5% significance level for three (insurance, transport, and food/clothing) of the five nonmonetary attributes.

Table 1.5 ANA Prevalence Across Treatments and Education Categories

	No certificate				Hold a certificate			
Variable	N	Text and audio (TAT), N = 289[a]	Text only (ToT), N = 3281	p-value[b]	N	Text and audio (TAT), N = 42	Text only (ToT), N = 15	p-value 2
ANA prevalence	617			0.017	57			>0.9
Nonattendants		110 (38.1%)	156 (47.6%)			13 (31.0%)	5 (33.3%)	
Attendants[d]		179 (61.9%)	172 (52.4%)			29 (69.0%)	10 (66.7%)	
ANA for individual attributes								
Wage	617	18 (6.2%)	14 (4.3%)	0.3	57	2 (4.8%)	0 (0.0%)	>0.9
Duration	617	29 (10.0%)	34 (10.4%)	0.9	57	6 (14.3%)	0 (0.0%)	0.3
Insurance	617	21 (7.3%)	40 (12.2%)	0.041	57	4 (9.5%)	1 (6.7%)	>0.9
Housing	617	34 (11.8%)	46 (14.0%)	0.4	57	7 (16.7%)	3 (20.0%)	0.7
Transport	617	42 (14.5%)	74 (22.6%)	0.011	57	8 (19.0%)	2 (13.3%)	>0.9
Food/Clothing	617	38 (13.1%)	64 (19.5%)	0.034	57	7 (16.7%)	3 (20.0%)	0.7
Number of ANA	617			0.2	57			0.5
0		179 (61.9%)	172 (52.4%)			29 (69.0%)	10 (66.7%)	
1		61 (21.1%)	76 (23.2%)			4 (9.5%)	2 (13.3%)	
2		34 (11.8%)	55 (16.8%)			4 (9.5%)	2 (13.3%)	
3		10 (3.5%)	17 (5.2%)			0 (0.0%)	1 (6.7%)	
4		3 (1.0%)	6 (1.8%)			4 (9.5%)	0 (0.0%)	
5		1 (0.3%)	1 (0.3%)					
6		1 (0.3%)	1 (0.3%)			1 (2.4%)	0 (0.0%)	

Notes: [a] n (%).

[b] Pearson's chi-squared test; Fisher's exact test for respondents with no education certificate or degree.

[c] Pearson's chi-squared test; Fisher's exact test for respondents with an education certificate or degree from career or junior/community schools. .[d] Attendants denote respondents who ignore none of the attributes, while nonattendants include respondents who ignore at least one attribute.

Conventional ANA Model and WTA Tests across Treatments

The results of the conventional ANA RPL-EC are shown in Table 1.6. The model controls for education, as the proportion of education levels was unbalanced across groups. Unsurprisingly, results related to the wage coefficient estimate, and the three duration options were similar to the result of the standard RPL-EC. Having housing, transport, food/clothing, insurance, or a combination of housing and insurance among the nonmonetary job benefits significantly increased utility. However, the relative ranking of the effects of attributes with a positive impact was different across groups. In TAT, insurance was the most important attribute followed by the combination of housing and insurance, housing, food, transport, and the wage. In ToT, the combination of housing and insurance was the most preferred attribute followed by insurance, housing, transport, food, and wage. All standard deviation estimates of the random parameters and the error component are significant (p value<0.01). These results suggest that the relative ranking and importance of job attributes were different across groups when attribute attendance was accounted for.

Compared with the WTAs under the complete attendance assumption, WTAs estimated from the model accounting for ANA displayed remarkable discrepancies across groups. Panel B in Table 1.4 shows the marginal WTAs conditional on attribute attendance as well as the Poe test and the t test result. Except for job duration, all nonmonetary attributes had positive WTAs. In both groups, insurance and housing recorded the highest WTA, respectively, \$6.57/hr. and 6.26/hr. in the treatment group and \$3.62/hr. and \$2.68/hr. in the control group. The Poe test revealed a significant difference in terms of WTA between the treatment and the control groups for job durations of nine months (p<0.05), housing (p<0.05), insurance (p<0.10), food/clothing (p<0.05), and transport (p<0.10).

Table 1.6 Estimate from the Conventional ANA RPL-EC Across the TAT and ToT

	Text and audio (TAT)		Text only (ToT)	
	Estimate (se)	Std. (se)	Estimate (se)	Std. (se)
Buy	-2.188***		-1.342***	
	(0.354)		(0.360)	
Error component		6.666***		7.278***
		(0.484)		(0.503)
Wage	0.170***	-	0.229***	-
	(0.011)		(0.013)	
Duration of 3 months	-0.458**	1.938***	-0.741***	1.597***
	(0.200)	(0.219)	(0.167)	(0.239)
Duration of 6 months	-0.757***	1.453***	-0.622***	1.142***
	(0.213)	(0.281)	(0.225)	(0.253)
Duration of 9 months	-0.966***	1.368***	-0.582**	0.941***
	(0.221)	(0.259)	(0.227)	(0.265)
Housing	1.067***	1.623***	0.613**	1.794***
	(0.257)	(0.162)	(0.259)	(0.141)
Insurance	1.119***	1.590***	0.830***	1.716***
	(0.240)	(0.161)	(0.245)	(0.153)
Food	0.935***	0.800***	0.526***	1.110***
	(0.184)	(0.169)	(0.201)	(0.170)
Transport	0.791***	0.781***	0.535***	0.858***
	(0.184)	(0.129)	(0.192)	(0.143)
Housing x Insurance	1.090***	-	1.184***	-
	(0.309)		(0.336)	
Food x Transport	-0.240	-	0.440	-
	(0.269)		(0.306)	
Certificate*Wage	0.077**	-	-0.078**	-
	(0.034)		(0.033)	
Certificate*Duration3	-1.634***	-	0.001	-
	(0.576)		(0.692)	
Certificate*Duration6	-0.408	-	-0.209	-
	(0.559)		(0.939)	
Certificate*Duration9	0.017	-	-0.070	-
	(0.594)		(0.979)	
Certificate*Insurance	-0.828*	-	1.979***	-
	(0.448)		(0.698)	
Certificate*Housing	0.446	-	-0.780	-
	(0.424)		(0.668)	
Certificate*Food	0.631*	-	-0.877	-
	(0.344)		(0.642)	
Certificate*Transport	0.258	-	0.036	-
	(0.327)		(0.537)	
N[a]	3310.000		3430.000	
Log-likelihood	-2204.638		-2088.806	
AIC	4519.276		4287.612	
BIC	4855.034		4625.329	

Notes: The conventional ANA RPL-EC model is conditional on attendance to attributes during choice tasks.
***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.
[a] Denotes the number of observations (choices scenarios for all respondents).

On average respondents in the treatment group valued more these attributes compared to their counterparts in the control group. Moreover, relative to low-skilled workers in the control group, those in the treatment group would require an additional $3.13/hr. more to accept a nine-months job.

Validation ANA Model Across Treatments

Table 1.7 presents the results of the validation model across groups. Two different coefficients were estimated for each attribute in this model. One coefficient reveals marginal utility for those who self-declared that they considered the attributes, while the second represents the choice behavior of attribute nonattenders.

The signs and significance of the coefficient estimates of the considered attributes were consistent with expectations in both groups. Regarding the coefficient estimates of attributes reported to be ignored, significant discrepancies were found. In the control group, of the eight individual attributes and attribute levels in the utility function, only the estimates for hourly wage, duration of three months, and insurance are significantly different from zero, indicating that respondents did not actually ignore these three attributes. In the treatment group, however, seven of the eight estimates for attributes are significant. The estimates for wage, six-month duration, nine-month duration, housing, insurance, transport, and food/clothing are significantly different from zero. Hence, respondents who were provided information using both text scripts and audio clips did not completely ignore almost all attributes that they reported having ignored.

Table 1.7 Estimates from the Validation ANA RPL-EC Across the TAT and ToT[a]

	Text and audio (TAT)		Text-only (ToT)	
	Estimate (se)[b]	Std. (se)	Estimate (se)	Std. (se)
Buy	-2.617*** (0.364)	-	-3.618*** (0.412)	-
Error component	-	7.438*** (0.550)		7.350*** (0.510)
Attributes attended				
Wage	0.198*** (0.014)	-	0.238*** (0.016)	-
Duration of 3 months	-0.795*** (0.229)	2.120*** (0.228)	-0.572*** (0.221)	2.045*** (0.252)
Duration of 6 months	-0.834*** (0.260)	1.556*** (0.297)	-0.720*** (0.255)	1.485*** (0.304)
Duration of 9 months	-1.090*** (0.276)	1.916*** (0.276)	-0.636** (0.252)	0.617** (0.289)
Housing	1.171*** (0.258)	2.169*** (0.194)	1.545*** (0.323)	2.408*** (0.208)
Insurance	1.197*** (0.237)	1.960*** (0.202)	1.653*** (0.304)	2.297*** (0.197)
Food	1.339*** (0.224)	1.032*** (0.163)	0.694*** (0.235)	1.396*** (0.168)
Transport	1.065*** (0.221)	1.162*** (0.145)	0.809*** (0.222)	1.080*** (0.151)
Housing * Insurance	1.335*** (0.325)	-	0.718* (0.379)	-
Food * Transport	-0.548* (0.323)	-	0.345 (0.309)	-
Certification*Wage	0.087** (0.037)	-	-0.127*** (0.029)	-
Certification*Duration3	-1.747*** (0.651)	-	0.074 (0.729)	-
Certification*Duration6	-0.349 (0.663)	-	-0.216 (1.021)	-
Certification*Duration9	0.126 (0.697)	-	-0.167 (1.033)	-
Certification*Insurance	-1.300** (0.510)	-	2.042*** (0.660)	-
Certification*Housing	0.327 (0.483)	-	-1.771* (0.994)	-
Certification*Food	0.531 (0.395)	-	-1.392* (0.759)	-
Certification*Transport	0.203 (0.378)	-	0.409 (0.579)	-

Table 1.7 (Cont.)

	Text and audio (TAT)		Text-only (ToT)	
	Estimate (se)[b]	Std. (se)	Estimate (se)	Std. (se)
Attributes non-attended				
Wage	0.133***	-	0.083**	-
	(0.036)		(0.036)	
Duration of 3 months	-0.674	1.133**	-1.123*	1.610**
	(0.460)	(0.537)	(0.616)	(0.656)
Duration of 6 months	-1.840**	3.454***	-0.214	2.101**
	(0.753)	(0.868)	(0.688)	(0.846)
Duration of 9 months	-1.503**	3.382***	-0.783	1.881**
	(0.759)	(0.708)	(0.710)	(0.872)
Housing	-0.933*	3.071***	0.562	1.589****
	(0.536)	(0.586)	(0.376)	(0.408)
Insurance	1.404***	2.828***	1.523***	1.124***
	(0.482)	(0.611)	(0.369)	(0.330)
Food	0.968***	1.508***	0.055	1.011***
	(0.302)	(0.303)	(0.345)	(0.292)
Transport	0.954***	0.707***	0.415	0.678**
	(0.292)	(0.241)	(0.255)	(0.291)
Housing * Insurance	-0.242	-	0.483	-
	(1.393)		(0.919)	
Food * Transport	-0.415	-	-0.209	-
	(0.571)		(0.555)	
N	3310		3430	
Log-likelihood[c]	-2137.6		-2044.2	
AIC	4573.29		4386.3	
BIC	5482.89		5301.2	

Notes: [a] In the validation ANA model, two coefficients are estimated for each attribute and the interaction of attributes. One coefficient is estimated for respondents who consider the attributes, and a second coefficient is estimated for respondents who ignore the attributes.
[b] ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.
[c] Denotes the number of observations (choices scenarios for all respondents).

Discussion

This study focused on US nonmigrant workers' choice behavior and WTAs for agricultural field jobs. The findings suggest a difference in relative ranking of attributes, WTAs, and ANA between the treated and the control groups. Under the complete attendance assumption, we found one more attribute that significantly affect utility in the TAT than in the ToT. Additionally, WTAs were significantly larger among treated respondents for three attributes.

We hypothesized that providing information using a dual modality (text and audio) versus a single modality (text only) could affect respondents' choices and WTAs through how they attended to attributes. To investigate this conjecture, we examined ANA in terms of the proportion of respondents ignoring attributes and its effects on coefficients estimates and WTAs across treatments. The results related to ANA suggest a significant difference between the TAT and the ToT. Indeed, only 38% of respondents ignored at least one attribute in the treatment group, while 48% did so in the control group. Similarly, a lower proportion of respondents self-reported having ignored individual attributes in the treatment group relative to the control group. A statistically significant difference ($p < 0.05$) was found for three of the six attributes[2], indicating that using text and audio reduces attribute non-attendance. This confirms our hypothesis. In terms of coefficient estimates, change in the order of attributes ranking was more noticeable relative to the situation assuming full attribute attendance. The Poe test on WTAs conditional on ANA revealed significant differences for five attributes, including attributes that already exhibited significant divergence in the model under complete attendance. According to Scarpa et al. (2013), the significance of the coefficients of attributes declared to be ignored

[2] Employment duration is one attribute but has four levels.

highlights discrepancies between respondents' serial self-reported ANA and their behavior. Such inconsistencies were found in our validation model when one observes the significance of the coefficients for the non-attended attributes.

However, the results were significantly different across groups. Respondents in the control group did not completely ignore three out of eight attributes that they stated to have ignored. In contrast, their counterparts in the treatment group did not actually ignore seven out of eight attributes that they claimed to have ignored, indicating four more attributes respondents revealed inconsistent behavior about. Such inconsistent behavior found in TAT was reported in previous studies (e.g., Alemu et al. 2013; Caputo et al. 2018a; Scarpa et al. 2013; Hess and Hensher 2010). The strong discrepancy between serial stated-ANA and behavior found in the treatment group are close to findings reported by Caputo et al. (2018a). These authors showed a significant difference for seven of nine ostensibly ignored attributes. As explained by Hess and Hensher (2010), the difference between self-reported ANA and choices suggests that rather than ignoring attributes they averred to have ignored, respondents simply lower the weight of these attributes in their utility. This is more noticeable in the TAT group. Overall, the dual modality treatment resulted in a low proportion of respondents claiming to have ignored attributes during the choice tasks. Moreover, participants declaring ANA did not actually ignore most of the attributes.

The significant difference between the treatment and the control groups can be explained using dual processing theory. It is possible that after the treatment, respondents in TAT had a better understanding and recall of the message conveyed as showed in previous studies (e.g., Chang 2009; Chang and Millett 2015; Moreno and Mayer 2002). This effect of the dual modality may even be more critical given that two sets of information script of about 200 words were used. Moreover, the difference in the amount of information retained and recalled probably plays a pivotal role as respondents complete repeated choice tasks. In fact, one plausible explanation to the higher WTA observed in TAT is that respondents treated with the dual modality retained more information and understood how urgent and valuable were workers in the agricultural sector amid the COVID-19 pandemic. This perspective is consistent with dual processing theory, which supports that individuals retain more information with redundant verbal information provision than with provision through a single mode (Chang 2009; Chang and Millett 2015).

In contrast to Chang (2009) and Chang and Millett (2015), who showed that respondents in the TAT completed the assigned cognitive task in a shorter time, we found no significant evidence that respondents treated with the dual modality spent on average less time than their counterparts in the single reading modality. Respondents in the TAT spent on average 137.35 seconds, while those in the ToT spent on average 141.59 seconds in completing the choice experiment tasks. This can be explained by the time setting implemented during the experiment. Although no significant difference was found, respondents might have considered the choice tasks slightly more complex in the ToT. Although this statement is speculative, it has been found that a larger amount of time devoted to choice tasks indicates a more complex task (Alós-Ferrer, Fehr, and Netzer 2021). This complexity can be explained by the significant effort required to recall the information presented prior to the tasks.

Conclusion

This study assesses the effect of information provision through a dual modality (text scripts and audio clips) on choice behavior, WTA, and attribute non-attendance (ANA) in discrete choice experiments. We randomly assigned respondents to either the dual modality treatment or the single modality (text script only). The treatment effect was investigated using an online discrete choice experiment on US low-skilled nonmigrant workers' WTA for agricultural field jobs. We used a random parameter logit with error component, the conventional attribute nonattendance model, and the validation model to estimate respondents' behavior related to jobs attributes in agricultural fields.

The results yield three main conclusions. First, there is a treatment effect with respect to the relative ranking of job attributes. Second, the dual modality yields a higher WTA for three attributes while one assumes full attendance to attributes. Third, the dual modality has a noticeable effect on attribute non-attendance, WTA conditional on attribute attendance, and the consistency between model estimates and serial stated ANA. A significantly lower proportion of respondents self-reported having ignored at least one attribute in the treatment group compared to the control group. A similar pattern was found for individual attributes for three of the six job attributes. Furthermore, when ANA is accounted for, there is a significant treatment effect on WTA for five out of seven attributes. The findings also suggest that providing information using the dual modality leads respondents to consider more attributes, even attributes they claim to have ignored them. This result implies that using text scripts combined with audio clips to provide information may generate results closer to meeting the continuity axiom, which suggests that participants consider all product attributes in choice tasks (Kragt 2013). Finally, we identified no evidence that respondents in the treatment group perform the choice tasks in a short period of time compared to those in the control group.

Our findings have several implications for future studies related to the effect of information on valuations. First, researchers might consider using a combination of text and audio clips to provide information in DCE focusing on WTA. Second, given the effect of the dual modality on ANA, the use of the dual modality may have improved respondents' attendance to attributes, thus affecting their valuation estimates.

This study is not without limitations. First, in the control group, we used only one single modality, text scripts. Future research could include other standalone modalities, such as audio clips only or videos only. Channa et al. (2019) investigated differences in WTP across text, audio, or video treatments, and found no difference based on the modality used to convey information to farmers. However, unlike our study, they used an experimental auction for maize storage technology, which is different from the goods and methods we used. Moreover, as discussed in the introduction and the literature review sections, it is unlikely that audio clips alone can yield better results than text scripts only. Therefore, video clips could be a main modality of interest that could be tested in the future. Third, the experiment was conducted online, which could raise some concerns about the treatment delivery and the preciseness of respondents' answers. Respondents might not have devoted the required attention to the experiment, especially the part relating to the information provision. However, because of the random assignment of respondents, we believe that if these effects occurred, they would also be randomly balanced across treatments.

This study is the first to evaluate the effect of how information is provided to respondents on their choice behavior, valuation and ANA in discrete choice experiments. This topic is important since DCE are arguably the most popular preference elicitation method used by economists, and testing of different types of information is prevalent in the DCE literature. Future work could replicate this study using the same types of products (for example, food products, environmental products, or public goods) and assess the effect of the modality on both WTP and WTA. Moreover, an investigation of the effect of the modality type on comprehension and information recall in DCE would be valuable. The effect of the dual modality in a context of various choice tasks number and/or different script length and complexity could be a potential area of investigation. Future research could also assess ANA by including the reasons for not attending to an attribute.

APPENDICES

Appendix A.1: Information Scripts

Information 1: Since the 2008 Great Recession, U.S. farmers continually struggle to fill jobs with U.S. domestic workers, particularly in fruit and vegetable production. Therefore, farmers have been turning to H-2A agricultural guest-worker visas to fill temporary or seasonal jobs with foreign laborers, many from Mexico and Central America.

H-2A visas require approval by the U.S. government. Farm employers must show that
(i) there are not enough domestic workers to fill the positions and
(ii) hiring H-2A workers will not adversely impact prevailing wages and working conditions of domestic workers.

Once hired, employers must provide guest workers housing and transportation. The U.S. government screens any foreign workers that enter the U.S. with H-2A visas. The H-2A program is complex, and many farmers hire outside consultants to manage the cost and web of regulations. While many employers of H-2A workers follow the laws and regulations, violations do occur. H-2A workers do not know labor laws as well as their domestic counter parts. Thus, H-2A workers are more vulnerable to exploitation, which can range from wage theft to recruiter payoffs leading to foreign worker debt.

Other problems facing H-2A workers include: no bargaining power, no political representation, no job security, a lack of career path, and a lack of flexibility.

Information 2: Following the first confirmed case of COVID-19 on January 20, 2020, the new coronavirus rapidly spread throughout the United States. By the start of April 2020, 45 US states have ordered "shelter in place" orders to slow the spread, allowing only essential travel. Infectious disease experts predict that COVID-19 will cause between 100,000 and 240,000 deaths within the United States alone.

The high infection rate and rapid transmission have stressed the food supply chain (agricultural producers, food distribution, and grocery stores) as consumers are stockpiling food. Agricultural workers are also at high risk of contracting coronavirus because remaining at least six feet apart is not possible in many cases. If there is an outbreak among agricultural workers, the US food supply chain could be harmed since an adequate supply of agricultural workers is key to the production and steady supply of labor-intensive agricultural products.

Consequently, the federal government has deemed agricultural workers as essential, exempting them from shelter-in-place orders. Furthermore, while the US Department of State has formally suspended routine visa services, they continue to process H-2 visas (temporary worker visas requested by employers) and have temporarily waived the in-person interview requirement because the "H-2 program is essential to the economy and food security."

Appendix B.1: Model Selection

We hypothesized that a random parameter logit (RPL) with error component (EC) would better fit the data. However, we started by estimating an RPL using the main attributes, and we compared model performance with the RPL-EC model. Results indicate that the latter outperforms the former based on the Log-likelihood, the AIC, and the BIC as shown in table 1.8. Next, different RPL-EC models with interaction between attributes were estimated, and the model used in the body of the paper performs better. We do not present the results of other models with interaction here.

Table 1.8 Main attributes Estimate from the Random Parameter Logit and the Standard RPL-EC Across Groups[a]

	Random parameter Logit				Random parameter Logit with error component			
	Text and Audio (TAT)		Text only (ToT)		Text and Audio (TAT)		Text only (ToT)	
	Estimate (se)[b]	Std. (se)	Estimate (se)	Std. (se)	Estimate se)	Std. (se)	Estimate (se)[b]	Std. (se)
Buy	-1.416***		-1.264***		-2.802***		-3.047***	
	(0.137)		(0.136)		(0.332)		(0.386)	
Error component			-			7.272***		7.115***
						(0.526)		(0.579)
Wage	0.085***	-	0.120***	-	0.173***		0.198***	
	(0.008)		(0.009)		(0.012)		(0.014)	
Duration of 3 months	-0.441**	1.967***	-0.902***	1.263***	-0.480**	1.944***	-0.554***	1.662***
	(0.170)	(0.179)	(0.143)	(0.172)	(0.188)	(0.205)	(0.183)	(0.206)
Duration of 6 months	-0.744***	2.414***	-0.835***	2.931***	-0.945***	1.911***	-0.657***	0.241
	(0.232)	(0.218)	(0.242)	(0.291)	(0.223)	(0.281)	(0.217)	(0.328)
Duration of 9 months	-0.877***	2.598***	-0.645***	3.028***	-1.186***	1.588***	-0.858***	0.838***
	(0.247)	(0.237)	(0.237)	(0.305)	(0.232)	(0.270)	(0.228)	(0.306)
Housing	1.028***	1.913***	0.963***	2.298***	1.732***	1.989***	1.502***	2.215***
	(0.144)	(0.150)	(0.155)	(0.178)	(0.187)	(0.156)	(0.197)	(0.194)
Insurance	1.265***	1.923***	1.242***	2.020***	1.785***	1.956***	1.743***	2.260***
	(0.143)	(0.162)	(0.147)	(0.157)	(0.190)	(0.173)	(0.226)	(0.201)
Food	0.452***	1.205***	0.288**	1.666***	0.947***	0.839***	0.788***	1.149***
	(0.111)	(0.128)	(0.131)	(0.159)	(0.132)	(0.146)	(0.159)	(0.149)
Transport	0.485***	1.079***	0.474***	1.151***	0.704***	0.784***	0.808***	0.915***
	(0.098)	(0.125)	(0.105)	(0.133)	(0.115)	(0.123)	(0.128)	(0.133)

Table 1.8 (Cont.)

	Random parameter Logit				Random parameter Logit with error component			
	Text and Audio (TAT)		Text only (ToT)		Text and Audio (TAT)		Text only (ToT)	
	Estimate (se)[b]	Std. (se)	Estimate (se)[b]	Std. (se)	Estimate (se)[b]	Std. (se)	Estimate (se)[b]	Std. (se)
Certification*Wage	0.040*	-	-0.007	-	0.098***	-	-0.069**	-
	(0.023)		(0.030)		(0.033)		(0.031)	
Certification*Duration3	-1.191***	-	-0.123	-	-1.762***	-	0.794	-
	(0.456)		(0.645)		(0.523)		(0.750)	
Certification*Duration6	-0.190	-	-0.982	-	0.029	-	0.198	-
	(0.595)		(1.005)		(0.606)		(0.957)	
Certification*Duration9	-0.169	-	-0.774	-	0.274	-	0.493	-
	(0.621)		(1.073)		(0.626)		(1.059)	
Certification*Insurance	-1.015**	-	0.645	-	-1.355***	-	1.217*	-
	(0.393)		(0.767)		(0.420)		(0.629)	
Certification*Housing	-0.006	-	-1.458**	-	0.009	-	-0.663	-
	(0.384)		(0.723)		(0.398)		(0.672)	
Certification*Food	0.626**	-	-1.460**	-	0.691**	-	-0.849	-
	(0.309)		(0.618)		(0.328)		(0.646)	
Certification*Transport	-0.082	-	-0.283	-	0.290	-	0.159	-
	(0.268)		(0.495)		(0.301)		(0.528)	
N[c]	3310		3430		3310		3430	
Log-likelihood	-2529.324		-2430.224		-2187.275		-2105.853	
AIC	5148.647		4950.447		4480.551		4317.706	
BIC	5423.359		5226.761		4804.100		4643.142	

Notes:
[a] Models were estimated assuming full attendance to jobs attributes. All models allowed for correlation.
[b] ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Standard errors are in parentheses.
[c] Denotes the number of observations (choices scenarios for all respondents).

Essay 2 : Budget Reminder and Hypothetical Bias in Discrete Choice Experiments: An Application to Fast-Food Products

Introduction

Hypothetical bias (HB) is a well-established concept in stated preference methods, including discrete choice experiments (DCE). Bohm (1972) first observed that hypothetical values for public goods were higher than actual values, and raised concerns about estimates obtained in hypothetical settings. This fact was formally conceptualized as HB by Schulze, d'Arge, and Brookshire (1981), who defined it as "the potential error induced by not confronting the individual with an actual situation, i.e., an organized market with well-defined prices" (Schulze, d'Arge, and Brookshire 1981, 183).

The literature on HB has been flourishing since the 1980s. However, HB is diversely (Haghani et al. 2021) and inconsistently (Carson, Groves, and List 2014) defined in the literature. Haghani et al. (2021) reviewed 33 studies and reported that HB definitions vary across studies and focus on different aspects of the bias. Some studies defined HB as overstating the true value in a hypothetical setting, while others described it as the divergence between statement in a hypothetical setting and a "real" setting. To address such inconsistency, these authors proposed a broad and conceptualizable definition: hypothetical bias is "the deviation in a predefined aggregate or disaggregate measure due to choice data being collected in a hypothetical setting instead of a more realistic (but not necessarily naturalistic) setting." (Haghani et al. 2021, 3). There is also a divergence on findings related to the magnitude and directions of HB in studies (Haghani et al. 2021; Fang et al. 2021). Nevertheless, evidence from meta-analysis suggests that HB is ubiquitous, and decision makers overstate their WTP in a hypothetical setting by a factor up to three on average (Penn and Hu 2018; List and Gallet 2001; Murphy et al. 2005).

Given that HB threatens the legitimacy of stated DCE, there is a growing number of strategies in the literature to mitigate or eliminate it. There are two categories of strategies: the ex-ante approaches applied before the choice tasks while the ex-post calibration methods are implemented after the choice tasks (Haghani et al. 2021). The former includes cheap talk (CT) (Cummings and Taylor 1999; Silva et al. 2011; Broadbent 2014), solemn oath (Jacquemet et al. 2013; de-Magistris and Pascucci 2014), honesty priming (de-Magistris, Gracia, and Nayga Jr 2013; Gschwandtner and Burton 2020), consequentiality (Drichoutis et al. 2017), repeated opt-out reminder (Ladenburg and Olsen 2014; Alemu and Olsen 2018), and defaulting the status quo option (Penn and Hu 2021). It also encompasses the time-to-think method (Cook et al. 2007; Cook et al. 2012), inferred valuation (Carlsson, Daruvala, and Jaldell 2010; Lusk and Norwood 2009a, b), and virtual reality (Fang et al. 2021). The ex-post methods include ex-post calibration (Fox et al. 1998), certainty calibration approach (certainty follow-ups) (Dekker et al. 2016; Lundhede et al. 2009), and consequentiality (Herriges et al. 2010).

However, there is no consensus regarding the independent effects of the mitigation strategies, and studies on the effectiveness of the proposed approaches tend to yield mixed results. Penn and Hu (2018) performed a meta-analysis using 131 studies and found that CT, consequentiality, and certainty follow-up were the most effective in reducing HB. Some studies however reported that these methods were not effective. For instance, Andor, Frondel, and Vance (2017) showed that CT was ineffective, while Lusk (2003) found that this method did not reduce the bias for knowledgeable subjects. Solemn Oath has been found to be ineffective at eliminating HB (Mamkhezri et al. 2020), and it even yielded undesirable effects on WTP (De-Magistris, Akaichi, and Ben Youssef 2016). Concerning consequentiality, Andor, Frondel, and Vance (2017) also showed that it did not affect hypothetical bias. Lundhede et al. (2009) proved no significant effect of choice certainty calibration on WTP, while Dekker et al. (2016) reported that this approach increased WTP as opposed to the expected impact.

Comparative studies also generate inconsistent results. For example, Gschwandtner and Burton (2020) concluded that CT was more effective than honesty priming, while Gschwandtner, Jang, and McManus (2020) found the opposite result. Moreover, there was no difference between the effect of solemn oath and CT (Lawton et al. 2020) and between CT and opt-out reminder (Ladenburg and Olsen 2014). Therefore, there is a need to further investigate HB in many contexts as possible to provide more insight into the phenomenon (Haghani et al. 2021).

Despite the growing literature on HB, there are still uninvestigated approaches. For example, the independent effect of budget reminder (BR) has never been investigated in discrete choice experiments (Haghani et al. 2021). In fact, many studies (e.g., Bello and Abdulai 2016; de-Magistris, Gracia, and Nayga Jr 2013; de-Magistris, Akaichi, and Ben Youssef 2016; de-Magistris and Pascucci 2014) mentioned budget constraint as a cause of hypothetical bias but referred to the script as just cheap talk. The short mention of budget in the previous cheap talk scripts may not play a crucial role in the effectiveness of the method. Furthermore, the framing of the scripts in these studies did not explain how respondents' decision may affect their budget and abilities to purchase other goods. It is therefore pivotal to assess the independent effect of BR. Loomis, Gonzalez-Caban, and Gregory (1994) were the first to study the sole budget reminder effect on HB using a contingent valuation method and reported no effect on WTP. No other study however has examined this topic in DCE, and we aim to fill this gap.

A common assumption while studying choices is that respondents have precise knowledge of their preferences. This strong assumption may be violated for non-market and market goods (Alberini, Boyle, and Welsh 2003). In some situations, respondents are uncertain or indifferent to the choice options presented, which hinder their ability to provide a definitive answer (Fenichel et al. 2009). They therefore may choose the no-buy option while uncertain and in absence of an uncertainty option (Alberini, Boyle, and Welsh 2003). Uncertainty refers to the fact that choice decision makers are not sure about the actual value of the goods (Bobinac 2019) or their preferences among the choice alternatives (Olsen et al. 2011) presented to them in a choice scenario. Uncertainty is more likely to occur in stated preference approaches since respondents complete choice tasks based on the partial description of the good or have little prior experience with the good (Alberini, Boyle, and Welsh 2003). Preference uncertainty has been a critical issue in stated preferences approaches as it could affect welfare estimates (Olsen et al. 2011). This issue has received a great attention in stated preferences methods, especially in DCE in the last decade. There are two main approaches to elicit uncertainty: 1) incorporate some levels of uncertainty in the choices and 2) elicit degree of certainty after making a choice (ex-post approach). The latter is widely used in DCE to capture uncertainty.

However, there are concerns related to this approach. First, it is less reliable since responses to certainty questions can be malleable (Bobinac 2019). Moreover, there is no agreement on the best methods to collect certainty questions (Beck, Rose, and Hensher 2013). Beck, Rose, and Hensher (2013) also pointed out that econometric methods used to include certainty questions significantly influence estimates, and they warned researchers about using these methods. Finally, the approach can yield to certainty that differs from the true certainty (Lundhede et al. 2009). It is possible to account for uncertainty during choice tasks implementation in DCE (Balcombe and Fraser 2011). While “unsure” or “I don’t know” have been already used in contingent valuation to capture uncertainty (Alberini, Boyle, and Welsh 2003), there are scant DCE studies that used the "unsure" or "I don’t know" as a distinct option from the no-buy option. Exception is Fenichel et al. (2009), who provided a general treatment of the "I don't know" alternative in DCE.

Given the importance of HB and the mixed results in the literature, it is important to continue exploring new approaches to mitigate or eliminate this bias in diverse contexts as possible. Therefore, we aim to answer whether BR is effective at mitigating or eliminating HB and investigate its effectiveness relative to CT and CT with BR. Furthermore, we seek an answer to whether the new approach we refer to as "choice task uncertainty adjustment" affects choices. We designed a DCE in laboratory to elicit consumers' willingness to pay for animal-based and plant-based burgers. Given that we were interested in whether BR will eliminate or mitigate the HB (with and without a cheap talk) and whether the uncertainty adjustment approach will affect choices, we randomly assigned participants to six groups, including one non-hypothetical treatment. The five hypothetical groups included one hypothetical group where no mitigation strategy was used. In three treatments, we used BR, CT, CT combined with BR as HB mitigation methods. We used the choice task uncertainty adjustment in the last treatment.

The objectives of the study are hence threefold. First, it investigates the effectiveness of BR to mitigate or eliminate HB in DCE. Second, it compares the effectiveness of BR to the independent effect of CT and the combined effect of CT with BR. Lastly, it explores the potential of the "choice task uncertainty adjustment" to affect choices. Proposed causes of HB include lack of consequentiality (Vossler, Doyon, and Rondeau 2012), lack of incentive compatibility (Harrison 2006), social desirability (Norwood and Lusk 2011), cognitive dissonance and falling to considering budget constraints (Loomis, Gonzalez-Caban, and Gregory 1994; Ding, Grewal, and Liechty 2005), and uncertainty (Dekker et al. 2016; Lundhede et al. 2009). We focused on budget constraints and uncertainty.

The results from our Generalized Multinomial type II model indicated that consumers exhibited hypothetical bias and overstated their willingness to pay by a factor of 1.29 and 1.40, respectively, for the animal-based and the plant-based burgers. Budget reminder reduced the hypothetical bias for the animal-based burger but failed to do so for the plant-based burger. Cheap talk and cheap talk combined with budget reminder, on the other hand, eliminated the bias for both products. This implies that budget reminder in cheap talk script may not be needed. Moreover, the choice task uncertainty adjustment reduced respondents' likelihood to choose the no-buy option. This suggests that under uncertainty, some respondents just choose the no-buy option if an uncertainty option is not provided in the choice tasks. The findings are crucial as they shed light on the potential of BR to reduce HB depending on the product as well the effectiveness of CT without BR. BR is simpler and uses short script compared to cheap talk, and could also reduce the cognitive effort needed from respondent as well as the survey duration.

Our study is crucial as it addresses one aspect of external validity, and the accuracy of welfare estimates in DCE. Because the method is widely used mainly for policy and business decisions, it is pivotal to accurately elicit preferences and willingness to pay even while using hypothetical DCE. The study is unique as it deals with two approaches never studied in the literature on DCE. Moreover, it provides a comparison between BR and other widely used mitigation techniques. By doing so, we contribute to the understanding of the effectiveness of different methods to reduce or eliminate HB. The results of this study can benefit researchers, especially those using DCE, business, and policymakers.

The balance of the chapter is as follows. The following section presents the methods. Afterward, the results, discussion, and conclusion sections are presented.

Methods

This section discusses the product selection, the experimental treatments, the experimental design, and the econometric models.

Product Selection

The study used hamburgers, a widely consumed food product in the United States. We included plant-based burgers for three reasons. First, plant-based meat is considered an alternative to animal-based meat. Animal-based meat production uses extensive water and land and significantly contributes to greenhouse gas emissions (Pimentel and Pimentel 2003). Therefore, plant-based meats are perceived as a strategy for reducing livestock production and its effects on the environment. Second, even though demand for plant-based meat has increased over the last few years, the market is still very much dominated by animal-based meat. Plus, consumers' attitude and preferences toward this type of meat are not well known (Slade 2018). Lastly, since we used a non-hypothetical choice experiment as a treatment, burger products can be sourced directly on campus. Having the source on site reduces the logistic burden of the experiment.

Experimental Treatments

We used a between-subject design to disentangle the treatment effects. Subjects were randomly assigned to one control group and five treatments. The randomization was automated in Qualtrics. All control and treatments groups faced the same choice tasks. Subjects in the control group were presented with hypothetical choices with no HB mitigation strategy. In one treatment, we used a neutral cheap talk without budget reminder as a mitigation strategy. Our CT script is similar to the one used by de-Magistris, Gracia, and Nayga Jr (2013) and Bello and

Abdulai (2016). However, we removed the budget part in the script to avoid a confounded effect of budget reminder.

Budget constraint is a critical determinant of consumer choices; hence in the second treatment, we used only budget reminder as a mitigation strategy to test the sole effect of BR. Following Gschwandtner and Burton (2020), we reminded subjects in this group that if they spend more on one product, they will have less to buy other products. In the third treatment, we presented respondents with both the CT and the budget reminder scripts before the choice tasks.

In the fourth treatment, we used the choice task uncertainty adjustment approach. We included uncertainty as a choice option in each choice task in addition to the two burger alternatives and the "No buy" option. No other mitigation strategy was delivered for this treatment. Specifically, before the choice tasks, we informed respondents to choose the unsure/I do not know option if they would like to buy one of the products, but they were uncertain which one to choose. We provided this information to prevent ambiguity between the no-buy and the uncertainty option given that Alberini, Boyle, and Welsh (2003) reported that some researchers see the "I don't know" option as equivalent to a no-buy option. We used hypothetical choice experiments in all the above-described control and treatment groups.

Finally, we implemented a non-hypothetical DCE. Before the choice tasks, we incentivized respondents to reveal their true preferences by informing them that one of the choice tasks would be randomly selected at the end of the experiment as the binding choice task. That is, they would have to pay the price indicated for the chosen alternative if any of the two burgers products was chosen in that choice task. We indicated to respondents that each choice task has the same probability to be selected as binding. This treatment allows eliciting preferences and

WTP for animal-based and plant-based burgers in a realistic and revealed preference condition. Appendix A.2 presents all scripts used in the experiment.

In the rest of the paper, we will refer to the control group as HCONT, the cheap talk only treatment as HCT, the budget reminder only treatment as HBR, cheap talk and budget reminder treatment as HCTBR, the choice-task uncertainty adjustment treatment as HUA, and the non-hypothetical treatment as NH.

Experimental Design

As discussed above, both hypothetical and non-hypothetical DCEs were used to elicit subjects' preferences and WTP for a 5-Oz burger. Following Papoutsi et al. (2015) and Fang et al. (2021), we used a labeled design with each alternative having two attributes, the label and the price. We generated the choice tasks using these two attributes. We labeled the burgers as "animal-based" and "plant-based" burgers. Price levels were generated based on the actual prices found in fast food restaurants. The price levels include \$3.00, \$4.50, \$6.00, and \$8.50.

The experiment was designed in three-steps using a D-efficient Bayesian design (Scarpa and Rose 2008; Scarpa et al. 2013). This approach minimizes the standard error of the coefficient estimates but requires priors to generate the experimental design. The priors can be determined through a pilot study, the literature, or expert judgment (Walker et al. 2018). In the first stage, we employed an efficient design including eight choice tasks and conducted a pilot study with 20 respondents using a non-hypothetical DCE. An expert provided the priors used in the pilot phase. Choice data from the first step was used in the second stage to estimate a multinomial logit model to get the priors (means values and standard errors). Finally, we employed these priors as Bayesian mean values to generate the final experimental design, which includes one block of eight choice tasks with a D-error of 0.079. We used a visual presentation format (Figure 2.1 and 2.2) and randomized the order of choice tasks and the two alternatives within each choice task.

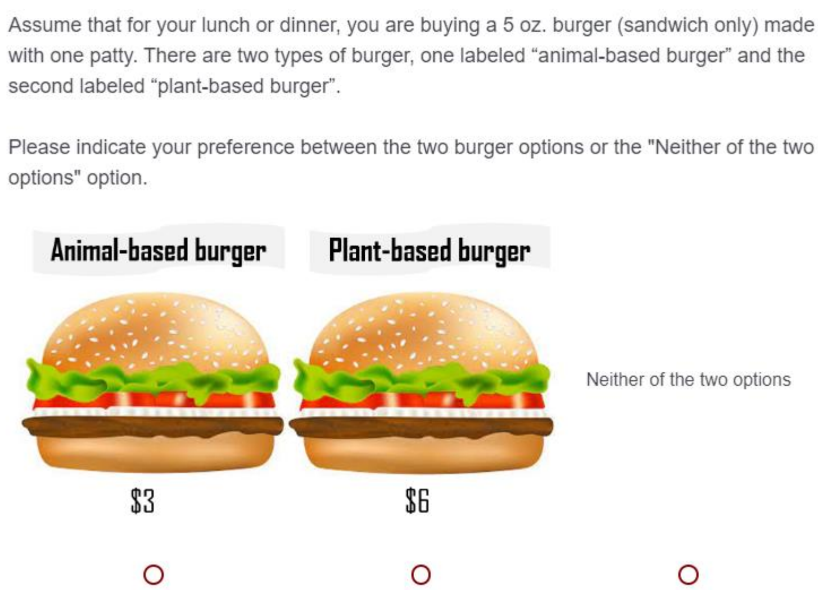

Figure 2.1 Example of a Choice Task in the burger DCE

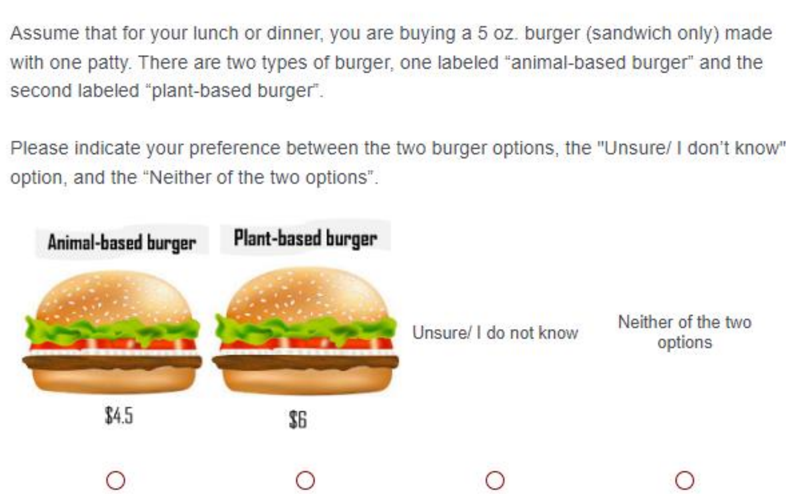

Figure 2.2 Example of a Choice Task with Uncertainty in the burger DCE

Laboratory Experiment and Data Collection

The data collection was approved by the Institutional Review Board of the University of Arkansas (protocol # 2107344220). We conducted a laboratory experiment using students from the University of Arkansas from February 7th to May 20th, 2022. We posted the study announcement in the university news. Interested students could then read a brief description of the study and sign up for a date and time slot. We used surveyors to recruit students near the experiment room close to the main food court.

Surveyors stayed in front of the main entrance and recruited students from those entering the food court. They randomly approached students and informed potential participants about the possibility to engage in a burger survey and being rewarded up to a $10 e-gift card. Interviewers then screened interested students through the lens of three questions. First, they checked whether the student eat animal-based meat. Second, they ensured that the student was at least eighteen years old and had purchased a burger within the last six months. Interested students who passed the screening questions were sent to the experiment room. They were then asked to read a quick instruction on a screen displayed in the room and afterward to start the survey. Once in the survey, each respondent was randomly assigned by Qualtrics to one of the six groups. 317 students participated in the actual experiment.

The survey was developed in Qualtrics, and it included five parts. First, the consent form included a brief description of the study, the reward, and the timeline in which they would receive it, and participants' right. We provided only a generic description of the purpose of the study to avoid experimenter demand effect. However, we emphasized that the final participation fee will depend on the outcome of the experiment. Those in the hypothetical groups received $10. Respondents in the non-hypothetical treatment, on the other hand, received the difference between their participation fee and the price of the product they chose in the binding choice task, if any. If the respondent chose the no-buy option, he/she received the whole participation fee of $10.

In the second part of the experiment, all respondents performed a blind tasting of a sample of sandwich of an animal-based burger and a plant-based burger made of black beans. Taste is an essential determinant of food choices and consumer behavior (Lewis, Grebitus, and Nayga 2016; Van Loo et al. 2010; Hoyer and Stokburger-Sauer 2012). For instance, Van Loo et al. (2010) found that taste was the most crucial meat quality attribute for consumers. Lewis, Grebitus, and Nayga (2016) used an experimental auction for soft drinks and concluded that including a tasting task in the design allows an accurate elicitation of consumers' WTP for highly consumed foods. Two types of taste can be measured, experienced taste and predicted taste (Lewis, Grebitus, and Nayga 2016). We used the experienced taste by asking respondents to taste each of the two types of burgers sample. The order of the burgers was randomized to prevent the order effect.

Specifically, once the respondents reached the tasting section, the instruction asked them to raise their hand, and we checked on their screen the first burger to taste. They then went on the table displaying the burger pieces to pick a sample of that burger for tasting[3]. After tasting, respondents provided the overall rating for the experienced taste of the first burger. They rinsed their mouths, picked the second burger, tasted it, and provided the overall rating for its taste. The burgers were sourced by a restaurant on campus. Each burger was cut into four similar pieces. We provided no information about the brand to avoid brand effect. The sensory test is a unique

[3] Respondents performed the tasting at their place in the experiment room.

feature of our design. It allows us to estimate the experienced taste rather than the perceived taste. Previous studies on plant-based meat did not include a tasting part in their design. We are only aware of Caputo, Sogari, and Van Loo (2022) as the study that included a tasting task.

Moreover, respondents were provided with instructions related to the choice tasks, a practice session of the choice task, and the script of the hypothetical bias mitigation strategy depending on their treatments. They completed the eight choices tasks after all the instructions were clearly provided to them.

In the fourth part of the survey, we collected additional information including burger consumption habits, time and risk preferences, attitude toward animal welfare, attitude toward the environment, and sociodemographic information. We used the New Ecological Paradigm (NEP) scale developed by Dunlap et al. (2000) to elicit participants' attitudes toward the environment. The NEP uses a five Likert scale ranging from “Strongly agree” to “Strongly disagree” to collect respondents’ perceptions on fifteen statements related to the relation between humans and the environment. To elicit the attitude toward animal welfare, we followed Cembalo et al. (2016) and used the seven-point Likert scale to measure participants’ perceptions of eight statements expressing different aspects of the relationship between humans and animals. For risk attitude, we employed the 11-point- Likert scale as in Dohmen et al. (2011) and elicited risk preference in general and in the domain of health and financial matters. Similar to De Marchi et al. (2016), we exploited the 7- point ordinal Likert scale (“not at all like me (1)” to “very much like me (7)”) on fourteen statements to categorize respondents' time preferences.

Lastly, we determined the final reward of each respondent by automating the process in Qualtrics. In the non-hypothetical treatment, the binding choice task was randomly selected, and this determined final amount and the burger type received by the respondent, if any. The screen clearly showed the amount and the burger the respondent would receive. The experimenter then recorded the information on this screen to be able to distribute products according to the experiment results. The brand was removed from products given to respondents.

Regression Model

The random utility theory of McFadden (1973) served as the theoretical framework for studying respondents' choices for the burger products. The theory posits that a decision-maker seeks for the highest utility while making a choice among different alternatives. Specifically, in a finite set of i alternatives ($i = 1, \dots, I$) with associate utility, U_i, the behavioral model indicates that the decision-maker n will choose an alternative i only and only if $U_{ni} > U_{nj}\ \forall j \neq i$. Moreover, the utility that the decision-maker n derives from alternative i in the choice task t can be written as follows.

$$(2.1) \quad U_{nit} = V_{nit} + \varepsilon_{nit}$$

where $V_{nit} = \beta' X_{nit}$ represents the deterministic component of the utility with X_{nit} a vector of observed attributes. β' indicates a vector of mean attribute utility weight or preferences parameters. ε_{nit} is the idiosyncratic error term. Under the basic multinomial logit model, the error term is considered independent and homogenous for all alternatives across respondents (i.i.d.). Moreover, the preference parameters in the deterministic part of the utility are identical for all respondents. These two restrictive assumptions yield the Identical and Irrelevant Alternatives (IIA) property. However, evidence suggests that consumers' preferences are heterogeneous (Caputo et al. 2018a; Fang et al. 2021), and the IIA assumption is unlikely to hold (Hensher, Rose, and Greene 2005). We, therefore, accounted for consumer heterogeneity by allowing individual-specific preference parameters to deviate from the population mean weight attribute as follows.

(2.2) $U_{nit} = (\beta' + \eta_n)\mathrm{X}_{nit} + \varepsilon_{nit}$

Where η_n represents the consumer-specific preference divergence from the mean conditional on X_{nit}. This model, known as Random Parameter Logit (RPL) (or mixed logit), relaxes the IIA property, but the error term is still i.i.d. To overcome this issue, the utility can be scaled by the individual-specific scale parameter μ_n . Using for example equation 2.1, we obtain $V_{nit} = (\beta'\mu_n)X_{nit} + \varepsilon_{nit}$. This model is known as the scale heterogeneity model (S-MNL) and seems to perform better than the RPL. An alternative solution to the RPL and the S-MNL is the "generalized multinomial logit" (G-MNL) which is obtained by nesting both RPL and S-MNL (Fiebig et al. 2010). In this model, the utility becomes (Fiebig et al. 2010):

(2.3) $U_{nit} = \beta_n'\mathrm{X}_{nit} + \varepsilon_{nit}$

with $\beta_n' = \mu_n\beta' + \gamma\eta_n + (1-\gamma)\mu_n\eta_n$. γ in β_n' ranges from 0 to 1 and shows how the variance of the residual taste heterogeneity varies with scale. Plugging $\gamma = 0$ in β_n' yields $\beta_n' = \mu_n(\beta' + \eta_n)$, which is the G-MNL-II model and indicates that the taste heterogeneity is proportional to the scale parameter. When $\gamma = 1$, on the other hand, $\beta_n' = \mu_n\beta' + \eta_n$ and shows that the standard deviation is independent of the scaling of parameter β' (G-MNL-I model).

Since μ_n, the individual-specific scale of the idiosyncratic error should be positive (Fiebig et al. 2010), we assumed it is random and follow a log-normal distribution with mean 1 and standard deviation τ. According to Fiebig et al. (2010), the simulated choice probability in the G-MNL model is:

(2.4) $$P(i|X_{nt}) = \frac{1}{D}\sum_{d=1}^{D}\frac{exp(\mu^d\beta' + \gamma\eta^d + (1-\gamma)\mu^d\eta^d)X_{nit}}{\sum_{k=1}^{I} exp(\mu^d\beta' + \gamma\eta^d + (1-\gamma)\mu^d\eta^d)X_{nit}}$$

with d indicating the number of draws in the simulation.

Moreover, the likelihood functions that respondent n will choose a sequence of choices $\{y_{nit}\}_{t=1}^{T}$ is:

(2.5) $$\hat{P}_n = \frac{1}{D}\sum_{d=1}^{D}\prod_{t}\prod_{i}\left(\frac{exp(\mu^d\beta' + \gamma\eta^d + (1-\gamma)\mu^d\eta^d)X_{nit}}{\sum_{k=1}^{I} exp(\mu^d\beta' + \gamma\eta^d + (1-\gamma)\mu^d\eta^d)X_{nit}}\right)^{y_{nit}}$$

Where y_{nit} is dummy variable that indicates whether the respondent choose the alternative.

Empirical Specification

We first estimated a main effect model with interactions. These interactions were created between the alternative specific constant (ASC) of each burger in the non-hypothetical treatment and other treatments. Equation 2.6 shows the empirical specification. It includes the ASC of the animal-based ($Animal_{based}$) and plant-based burgers ($Plant_{based}$) in NH, and their interactions with the remaining groups. We refer to this model as model 1. We allowed the effect of the burger products and their ASC to vary randomly and follow a normal distribution. The price is assumed to follow a log-normal distribution because respondents will prefer lower price levels.

(2.6) $$U_{nit} = [\mu_n\beta' + \gamma\eta_n + (1-\gamma)\mu_n\eta_n](Price + Animal_{based} + Plant_{based} + Animal_{based} * HCONT + Plant_{based} * HCONT + Animal_{based} * HCT + Plant_{based} * HCT + Animal_{based} * HBR + Plant_{based} * HBR + Animal_{based} * HCTBR + Plant_{based} * HCTBR) + \varepsilon_{nit}$$

Next, we followed Fang et al. (2021) and estimated a second (model 2) in which we used a set of respondents' characteristics to better predict the scale parameter. The set of characteristics includes familiarity with animal-based and plant-based burgers, the importance of the type of meat, time preferences, and the attitude toward animal welfare. Given that respondents tasted burgers products before making their choices, this may explain heterogeneity among participants. We, therefore, included experienced taste variables as predictors of the scale parameter. We exploited a principal component analysis (PCA) to reduce the 14 time-preference questions into two components, CFC_I and CFC_F including seven questions each. CFC_I characterizes high time preference respondents (high orientation toward the present), and CFC_F describes low time preference respondents. The details of the PCA are presented in Appendix B.2. In model 2, we defined $\mu_n = \exp(\mu + \theta_1 Typ + \theta_2 Fam_{ab} + \theta_3 Fam_{pb} + \theta_4 CFC_I + \theta_5 CFC_F + \theta_6 AF + \theta_7 Taste_{ab} + \theta_8 Taste_{pb} + \iota\epsilon_0)$ where *Typ* indicates the score of the importance of the type of meat, *Fam* is the familiarity score for each burger type, CFC_I is the score for high time preference category, CFC_F the score for low time preference category, and AF is the animal welfare score. *Taste* is a binary variable that takes values 1 (taste score 4 and 5) if the respondent likes the taste of the burger and 0 otherwise. ab denotes animal-based and pb plant-based burgers.

It is common in the literature to estimate the utility in preferences space and compute WTP as the ratio of the mean parameter estimate of a specific non-monetary attribute and the mean parameter estimate of the price attribute. However, the present study estimated utility in WTP space. Deriving WTP from models estimated in preference space may yield significantly high estimates with long upper tail distributions (Scarpa, Thiene, and Train 2008). Consequently, recent studies used the WTP space (e.g., Fang et al. 2021; Luckstead et al. 2022) where the coefficients of the non-monetary attributes represent the WTP. We estimated the WTP for each burger product using equation 2.7. γ can be fixed at 0,1 or considered random. Greene and Hensher (2010) recommend fixing γ at 0 to estimate the model in WTP. We, therefore, fixed γ at

0 in the specification. However, we estimated a model with γ set to be random as in Fang et al. (2021) and the G-MNL-I model for robustness check. The price attribute was fixed at -1 and the scale parameter μ_{nt} was used to account for respondent heterogeneity.

$$(2.7) \quad WTP_{nit} = \mu_n\beta' + \gamma\eta_n + (1-\gamma)\mu_n\eta_n$$

All models are estimated with 500 iterations for the simulation using the package "gmnl" version 1.1.3.2 (Sarrias and Daziano 2017) in RStudio 2022.02.2+485.

Hypotheses Testing and Market Share Simulation

The presence of hypothetical bias and the effectiveness of the proposed mitigation strategies were investigated by comparing WTPs between the hypothetical and the non-hypothetical groups. We first explored the presence of hypothetical bias by comparing the WTP in HCONT to WTP in NH. The rejection of the null hypothesis below will demonstrate the existence of HB.

$$H0_1: WTP_{HCONT} - WTP_{NH} = 0, and$$

$$H1_1: WTP_{HCONT} - WTP_{NH} > 0$$

Second, we tested each HB mitigation strategy relative to the control group (HCONT) through hypotheses H_2 to H_4 (Table 2.1). Rejection of the null hypotheses will prove that the proposed method is effective at mitigating HB. Third, the effectiveness of each mitigation approach to eliminate HB was assessed by comparing WTP in the HB mitigation treatments with the non-hypothetical group. Failing to reject the null hypothesis in any of the hypotheses H_5 to H_7 will be evidence that the mitigation strategy used in the treatment eliminated the HB. Finally, we compared BR with CT and CT with BR. If we fail to reject the null hypothesis in hypotheses

H_8 or H_9, we can conclude that budget reminder is effective as CT or CT with BR. We also compared cheap talk only with cheap talk combined with the budget reminder (H_{10}).

Table 2.1 Hypotheses

Hypotheses to test effectiveness of methods to mitigate HB	**Hypotheses to test effectiveness of methods to eliminate HB**	**Comparison of the effectiveness between HB mitigation methods**
$H0_2: WTP_{HCONT} - WTP_{HCT} = 0$,	$H0_5: WTP_{HCT} - WTP_{NH} = 0$,	$H0_8: WTP_{HBR} - WTP_{HCT} = 0$,
$H1_2: WTP_{HCONT} - WTP_{HCT} > 0$	$H1_5: WTP_{HCT} - WTP_{NH} > 0$	$H1_8: WTP_{HBR} - WTP_{HCT} > 0$
$H0_3: WTP_{HCONT} - WTP_{HBR} = 0$,	$H0_6: WTP_{HBR} - WTP_{NH} = 0$,	$H0_9: WTP_{HBR} - WTP_{HCTBR} = 0$,
$H1_3: WTP_{HCONT} - WTP_{HBR} > 0$	$H1_6: WTP_{HBR} - WTP_{NH} > 0$	$H1_9: WTP_{HBR} - WTP_{HCTBR} > 0$
$H0_4: WTP_{HCONT} - WTP_{HCTBR} = 0$,	$H0_7: WTP_{HCTBR} - WTP_{NH} = 0$,	$H0_{10}: WTP_{HCT} - WTP_{HCTBR} = 0$,
$H1_4: WTP_{HCONT} - WTP_{HCTBR} > 0$	$H1_7: WTP_{HCTBR} - WTP_{NH} > 0$	$H1_{10}: WTP_{HCT} - WTP_{HCTBR} > 0$

To test our hypotheses, we first performed a t-test between groups using predicted individual WTP. Second, we performed the complete combinatorial test recommended by Poe, Giraud, and Loomis (2005). We did this in two stages. First, we bootstrapped individual WTP estimates 10,000 times and used the mean values to perform the Poe test in the second stage. Furthermore, a price sensitivity analysis of the market share was performed using equation 2.4. We simulated the market share using 5000 iterations at different prices from \$3 to \$8.5, which is the price range used in the experiment. This yields the demand curve for each product and each treatment.

Results

Descriptive Statistics and Balance Test

Table 2.2 details the summary statistics for the sample, per treatment, and the balance test result across treatments regarding socioeconomic and burger consumption characteristics. Three hundred and seventeen subjects participated in the experiment. However, we removed five participants who showed insufficient attention during the survey[4].

A significant share of participants (69%) were undergraduate students and a similar proportion (68%) earn less than $15,000, annually. Nearly 50% of participants fell in the age range of 18-20, and 57% were male. Ninety-three percent of the respondents often consume animal-based burger against only 6% for plant-based burger, and 38 % were regular burger consumers. Thirty-four percent of the respondents ranked protein as the first or the second most important nutritional attributes. The average score of the importance of taste and type of meat were 4.59/5 and 3.51/5, respectively. Respondents had an average familiarity score, respectively, of 3.79/5 and 2.47/5 for animal-based burger and plant-based burger prior to the experiment. Fifty-six and 37% of respondents, respectively, liked the taste of the sample of animal-based and plant-based burgers during the sensory test[5]. Respondents' characteristics are balanced across treatment ($p>0.1$), suggesting that respondents were similar with respect to their characteristics. These balanced samples rule out these characteristics as potential causes of any divergence in WTP.

[4] They completed the survey in less than 10 minutes, a duration which is unlikely for a respondent who paid attention to the questions during the survey. The average completion time is 18.78 minutes.

[5] A binary variable was defined using the five-point Likert scale. It takes 1 for those above the average mean and 0 otherwise. Respondents who chose 1, 2, and 3 were below the average mean for both burger types.

Table 2.2 Descriptive Statistics and Balance Test

Characteristic	N	Full sample, N = 312[a]	Treatments Received						p-value[b]
			HCONT, N = 53	HCT, N = 52	HBR, N = 53	HCTBR, N = 51	HUA, N = 51	NH, N = 52	
Age category	312								0.7
[18-20]		144 (46.1%)	30 (56.6%)	19 (36.5%)	25 (47.2%)	24 (47.1%)	20 (39.2%)	26 (50.0%)	
[21-23]		76 (24.4%)	10 (18.9%)	13 (25.0%)	12 (22.6%)	12 (23.5%)	17 (33.3%)	12 (23.1%)	
24+		92 (29.5%)	13 (24.5%)	20 (38.5%)	16 (30.2%)	15 (29.4%)	14 (27.5%)	14 (26.9%)	
Gender	312								0.7
Male		177 (56.7%)	29 (54.7%)	25 (48.1%)	30 (56.6%)	31 (60.8%)	32 (62.7%)	30 (57.7%)	
Female		125 (40.1%)	21 (39.6%)	26 (50.0%)	21 (39.6%)	19 (37.3%)	19 (37.3%)	19 (36.5%)	
Other		10 (3.2%)	3 (5.7%)	1 (1.9%)	2 (3.8%)	1 (2.0%)	0 (0.0%)	3 (5.8%)	
Study level	312								>0.9
Undergraduate		216 (69.2%)	39 (73.6%)	32 (61.5%)	39 (73.6%)	34 (66.7%)	37 (72.5%)	35 (67.3%)	
Graduate		86 (27.6%)	13 (24.5%)	18 (34.6%)	13 (24.5%)	15 (29.4%)	13 (25.5%)	14 (26.9%)	
Other		10 (3.2%)	1 (1.9%)	2 (3.9%)	1 (1.9%)	2 (3.9%)	1 (2.0%)	3 (5.8%)	
Personal annual income	312								0.8
Less than $15,000		213 (68.3%)	32 (60.4%)	35 (67.3%)	40 (75.5%)	38 (74.5%)	34 (66.7%)	34 (65.4%)	
$15,000 to $29,999		74 (23.7%)	15 (28.3%)	13 (25.0%)	10 (18.9%)	11 (21.6%)	11 (21.6%)	14 (26.9%)	
30,000+		25 (8.0%)	6 (11.3%)	4 (7.7%)	3 (5.6%)	2 (3.9%)	6 (11.7%)	4 (7.7%)	
Burger type often consumed	312								0.7
Animal-based burger		290 (92.9%)	49 (92.4%)	50 (96.2%)	49 (92.4%)	48 (94.1%)	47 (92.2%)	47 (90.4%)	
Plant-based burger		19 (6.1%)	2 (3.8%)	2 (3.8%)	3 (5.7%)	3 (5.9%)	4 (7.8%)	5 (9.6%)	
Other		3 (1.0%)	2 (3.8%)	0 (0.0%)	1 (1.9%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	
Type of burger consumer[c]	312								0.2
Regular consumer		118 (37.8%)	20 (37.7%)	12 (23.1%)	23 (43.4%)	20 (39.2%)	24 (47.1%)	19 (36.5%)	
Irregular consumer		194 (62.2%)	33 (62.3%)	40 (76.9%)	30 (56.6%)	31 (60.8%)	27 (52.9%)	33 (63.5%)	
Protein rank among nutritional attributes	312								0.9
1st or 2nd		106 (34.0%)	18 (34.0%)	18 (34.6%)	17 (32.1%)	14 (27.5%)	19 (37.3%)	20 (38.5%)	
3rd to 6th		206 (66.0%)	35 (66.0%)	34 (65.4%)	36 (67.9%)	37 (72.5%)	32 (62.7%)	32 (61.5%)	
BMI category	312								0.4
Healthy		169 (54.2%)	28 (52.8%)	32 (61.5%)	34 (64.2%)	26 (51.0%)	23 (45.1%)	26 (50.0%)	
Not healthy		143 (45.8%)	25 (47.2%)	20 (38.5%)	19 (35.8%)	25 (49.0%)	28 (54.9%)	26 (50.0%)	
Overweight or obese[d]	312	136 (43.6%)	22 (41.5%)	20 (38.5%)	17 (32.1%)	24.(47.1%)	28 (54.9%)	25 (48.1%)	0.2

Table 2.2 (Cont.)

Characteristic	N	Full sample, N = 312	Treatments Received						p-value[b]
			HCONT, N = 53	HCT, N = 52	HBR, N = 53	HCTBR, N = 51	HUA, N = 51	NH, N = 52	
Taste importance	310	4.6 (0.6)	4.5 (0.5)	4.6 (0.6)	4.6 (0.6)	4.5 (0.7)	4.7 (0.6)	4.6 (0.5)	0.8
Like animal-based burger taste	312	176 (56.4%)	30 (56.6%)	27 (51.9%)	24 (45.3%)	36 (70.6%)	29 (56.9%)	30 (57.7%)	0.2
Like plant-based burger taste	312	116 (37.2%)	21 (39.6%)	18 (34.6%)	23 (43.4%)	19 (37.3%)	14 (27.5%)	21 (40.4%)	0.6
Familiarity score- animal-based burger	312	3.8 (0.9)	3.7 (1.0)	3.7 (0.9)	4.0 (0.8)	3.7 (0.9)	3.9 (1.0)	3.8 (0.9)	0.3
Familiarity score- plant-based burger	312	2.5 (1.0)	2.4 (1.1)	2.3 (0.9)	2.4 (1.0)	2.4 (1.1)	2.6 (1.0)	2.5 (1.1)	0.7
Type of meat importance	312	3.5 (1.1)	3.3 (1.3)		3.5 (1.1)	3.5 (1.0)	3.7 (1.1)	3.6 (1.1)	0.7
High time preference score	312	20.6 (6.9)	20.0 (7.9)	19.4 (6.7)	21.0 (6.7)	21.1 (6.4)	21.5 (7.5)	20.6 (6.4)	0.7
Low time preferences score	312	33.9 (6.6)	34.0 (5.9)	33.4 (6.5)	34.5 (5.8)	33.0 (7.6)	33.8 (6.8)	35.0 (7.2)	0.7
Animal welfare score[e]	312	6.4 (3.1)	6.7 (3.1)	6.0 (3.1)	6.2 (3.1)	6.8 (2.8)	6.4 (3.2)	6.2 (3.6)	0.8

Notes:
[a] n (%); Mean (SD)
[b] Pearson's Chi-squared test; Fisher's exact test; One-way ANOVA
[c] Regular consumer category include those who consume burger at least once a week
[d] Overweight or obese includes respondents with a BMI that is greater or equal to 25 as defined by Centers for Disease Control and Prevention (CDC).
[e] The two statements that captures respondents' consideration of animal welfare while making food choice are used.

Respondents' Choice Characterization

Respondents preferred the animal-based burger to the plant-based burger in the majority of choice situations. Overall, respondents chose the animal-based burger in 55% of the choice situations (Table 2.3). This proportion varies from 49% in the NH group to 58% in HCONT and HUA treatments. The plant-based burger was chosen in only 31% of choice situations in the entire sample and ranges from 27% in the HCTBR treatment to 34% in the HBR treatment. Compared to the non-hypothetical treatment (78%), respondents chose more frequently (92%) to buy one of the two products in the purely hypothetical group (HCONT). This is an indication of a potential hypothetical bias in our study. Another interesting result in Table 2.3 is related to the HUA treatment. Respondents in this treatment chose the unsure option in 4.4% of the choice scenarios. The choice task uncertainty adjustment significantly reduced the proportion of the No-buy option (8.8%) which was chosen on average in 14% of the cases in the sample and ranges from 8.3% in HCONT to 22% in the NH treatment.

Table 2.3 Choices Characterization per Treatment

Choice characteristic	Full sample[a], N = 2,495[b]	Treatments Received						p-value[c]
		HCONT, N = 424	HCT, N = 416	HBR, N = 424	HCTBR, N = 408	NH, N = 416	HUA, N = 407	
Choices during choice tasks								<0.001[d]
Animal based	1,364 (54.7%)	247 (58.3%)	217 (52.2%)	237 (55.9%)	225 (55.1%)	202 (48.6%)	236 (58.0%)	
Plant based	772 (30.9%)	142 (33.5%)	137 (32.9%)	144 (34.0%)	111 (27.2%)	121 (29.1%)	117 (28.7%)	
No buy	341 (13.7%)	35 (8.3%)	62 (14.9%)	43 (10.1%)	72 (17.6%)	93 (22.4%)	36 (8.8%)	
Uncertain	18 (0.7%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	0 (0.0%)	18 (4.4%)	
Buying decision								<0.001
No buy	341 (13.7%)	35 (8.3%)	62 (14.9%)	43 (10.1%)	72 (17.6%)	93 (22.4%)	36 (8.8%)	
Buy	2,154 (86.3%)	389 (91.7%)	354 (85.1%)	381 (89.9%)	336 (82.4%)	323 (77.6%)	371 (91.2%)	

Notes: [a] n (%)

[b] The number of observations is equal to the number of respondents times the number of choices sets.

c Pearson's Chi-squared test; [d] the test does not include Unsure/I do not know option which is only present in HUA.

WTP for Animal-based and Plant-based Burgers

The results of the G-MNL-II model in the WTP space are shown in Table 2.4. The ASC under each burger type represents the WTP for the burger in the non-hypothetical treatment since it was used as the reference group during the estimation. The interaction of the ASC with each treatment, on the other hand, reveals the variation of WTP in the given treatment relative to NH.

Model 1 indicates that consumers' preferences and WTP were heterogeneous. The standard deviation of the ASC of each burger product is highly significant ($p<0.01$), suggesting that individuals' valuations deviated from the sample mean. Moreover, the intercept of the scale parameter and its standard deviation Tau (τ) are highly significant. This indicates that scaling the utility by an individual-specific scale parameter is appropriate in our study.

Consumer's WTP for the animal-based burger ranges from \$6.43/5 Oz to \$8.28/5 Oz [6]. As one might expect, WTP in the non-hypothetical treatment was the lowest at \$6.43/5oz, while HCONT recorded the highest WTP (\$8.28/5 Oz), suggesting the existence of HB for the animal-based burger. Furthermore, compared to the WTP in NH, respondents were willing to pay more for the animal-based burger in all treatments with an HB mitigation strategy (HCT, HBR, and HCTBR). WTPs in these treatments are generally lower than in the HCONT group. These lower values suggest that the mitigation strategies may have reduced the hypothetical bias but did not eliminate it for the animal-based burger. CT only and CT combined with BR yielded the lowest WTPs.

In the non-hypothetical condition (NH), consumers were willing to pay a smaller amount for the plant-based burger (Table 2.4). The WTP ranges from \$5.12/5 Oz to \$7.19/5 Oz. Like for

[6] The upper bound is obtained by adding the high coefficient (ASC*HCONT) to the ASC in model 1 (Table 4).

the animal-based burger, HCONT recorded the highest WTP, indicating that hypothetical bias was potentially present while consumers chose plant-based burger in a hypothetical scenario. Like for the animal-based burger, HCT and HCTBR recorded the lowest WTP among the hypothetical treatments. One difference is that the coefficient of HCTBR is not significant, suggesting there was no difference between WTP in the HCTBR treatment and WTP in the non-hypothetical treatment. If this is confirmed, one can conclude that HCTBR has eliminated the hypothetical bias for the plant-based burger.

Table 2.4 WTP ($) Estimates Using the G-MNL-II Model

Parameter	Model 1		Model 2[b]	
	Coefficients[a]	Std.err.	Coefficients	Std.err.
Animal-based burger				
ASC	6.431***	0.202	6.404***	0.202
sd. (ASC)	2.243***	0.115	2.205***	0.111
ASC*HCONT	1.850***	0.343	1.929***	0.339
ASC*HCT	0.618**	0.267	0.651**	0.266
ASC*HBR	1.158***	0.280	1.187***	0.273
ASC*HCTBR	0.743***	0.273	0.746***	0.274
Plant-based burger				
ASC	5.120***	0.219	5.103***	0.223
sd. (ASC)	2.532***	0.106	2.548***	0.107
ASC*HCONT	2.071***	0.381	1.989***	0.380
ASC*HCT	0.674**	0.285	0.674**	0.286
ASC*HBR	1.368***	0.309	1.343***	0.306
ASC*HCTBR	0.062	0.294	0.038	0.300
Consumer heterogeneity				
Intercept	0.821***	0.220	0.813	0.581
Tau	1.055***	0.140	1.009***	0.139
Importance type of meat			-0.092	-0.096
Familiarity with plant-based burger			0.051	0.045
Familiarity with animal-based burger			0.107	0.106
Hight time preference			0.008	0.008
Low time preference			-0.005	-0.005
Animal welfare			-0.035	-0.034
Like animal-based burger taste			0.000	-0.007
Like plant-based burger taste			-0.009	-0.007
N	2088		2088	
Log-likelihood	-1232.51		-1229.88	
AIC	2493.025		2503.761	
BIC	2572.04		2627.928	

Note: ***, **, * denotes significance at 1%, 5%, and 10% levels. Sd. indicates standard deviation. *Std.err* indicates standard errors. Model 1 allows the scale parameter to be drawn from a log-normal distribution. Gamma fixed at 0. [a] Coefficients are WTP. [b] A set of variables were used to predict the scale parameter in model 2.

Next, we ran model 2 by predicting the scale parameter using selected variables described in the method section. Model 2 in Table 2.4 presents the results. It yielded larger AIC and BIC than the model without the predictors (model 1). Furthermore, the intercept and the determinants of the scale parameter are not statistically significant. We, therefore, select the model without the predictor (model 1) as the preferred model.

Results in Table 2.4 suggest the existence of hypothetical bias and the potential of mitigation strategies to reduce or eliminate the HB for both products. However, we must test our hypothesis to confirm or reject such an observation. The following section presents the results of the hypotheses testing.

Hypothetical Bias

Table 2.5 provides the results of the hypotheses testing per product. We first conducted the hypotheses testing using the mean of individual WTP. Afterward, we performed the tests using the non-parametric combinatorial method of Poe, Giraud, and Loomis (2005).

The first hypothesis allows for identifying the presence of HB. Both t test and the Poe test yielded the rejection of the equality between WTP in the hypothetical control group (HCONT) and WTP in the non-hypothetical treatment (NH) for both products at the 1% significance level. Because we used a one-sided test, we can conclude the presence of HB, and that consumers overstated their WTP in the hypothetical setting.

Once HB was confirmed, we tested hypotheses H_2 to H_4 to confirm if the mitigation strategies reduced the HB. For the animal-based burger, the null hypothesis of equality between HCONT and each of the mitigation treatments (H_2 – H_4) is rejected for both tests. We, therefore, conclude that CT only, BR, and CT with BR mitigated the HB. For the plant-based burger, the null hypotheses H_2 and H_4 are rejected at the 1% significance level, suggesting that CT only and CT combined with BR mitigated the hypothetical bias. Contrary to the animal-based burger, we fail to reject the null hypothesis for the comparison between HCONT and HBR for the plant-based burger. This suggests that this approach did not affect the bias for the plant-based burger.

The main objective of all HB mitigation approaches is to eliminate the bias. We tested this by comparing WTP in HCT, HBR, and HCTBR, respectively, with WTP in NH. Hypotheses H_5 to H_7 show the comparisons. For the animal-based burger, WTP in HBR treatment was significantly larger than WTP in NH. Thus, the budget reminder did not eliminate the bias but only reduced it. We fail, on the other hand, to reject the null hypothesis of equality between HCT and NH (H_5) and HCTBR and NH (H_6) for both tests. Similar results are found for the plant-based burger. These results indicate that CT and CT with BR eliminated the bias.

Finally, we compared the effectiveness of the budget reminder method with cheap talk only and cheap talk with the budget reminder. HBR induced statically higher WTP than WTP in HCT and HCTBR for both burger types. It suggests that HCT and HCTBR were more effective than HBR. Given that CT and CTBR eliminated the bias, we checked if they have generated different WTP. For the animal-based burger, the t-test resulted in no difference in WTP between the two methods. However, the Poe test indicates that HCTBR recorded higher WTP than HCT. For the plant-based burger, both tests yielded the same result; there was no difference in WTP in the two treatments.

Table 2.5 Hypothesis Testing Results

	Hypotheses	Animal-based burger			Decision	Plant-based burger			Decision
		t stat	t p-value[a]	Poe p value[b]		t stat	t p-value	Poe p value	
H_1	HCONT vs NH	3.770	0.000	0.000	Reject	3.272	0.001	0.000	Reject
H_2	HCONT vs HCT	7.778	0.000	0.000	Reject	4.453	0.000	0.000	Reject
H_3	HCONT vs HBR	2.137	0.017	0.015	Reject	1.172	0.122	0.121	Accept
H_4	HCONT vs HCTBR	4.589	0.000	0.000	Reject	4.453	0.000	0.000	Reject
H_5	HCT vs NH	-0.887	0.811	0.189	Accept	-0.745	0.771	0.227	Accept
H_6	HBR vs NH	2.337	0.011	0.008	Reject	2.086	0.020	0.016	Reject
H_7	HCTBR VS NH	0.572	0.284	0.281	Accept	-1.024	0.846	0.153	Accept
H_8	HBR VS HCT	5.153	0.000	0.000	Reject	3.095	0.001	0.001	Reject
H_9	HBR vs HCTBR	2.475	0.008	0.006	Reject	3.202	0.001	0.001	Reject
H_{10}	HCT vs HCTBR[c]	-2.181	0.984	0.013	Reject	0.375	0.354	0.352	Accept

[a] One-sided test. Alternative hypothesis: WTP in the first group is larger than in the second group.
[b] P-values of Poe test are obtained from 10,000 bootstrapping.
[c] In the Poe test, the alternative is HCTBR is greater than HCT.

Figure 2.3 presents the WTP of each group computed using model 1. HCONT has the highest WTP and NH the lowest WTP for both products. Moreover, for the animal-based burger, there seems to be no remarkable difference between HCT and NH, and HCTBR and NH. For the plant-based burger, WTP appears identical between HCTBR and NH, while the WTP of HCT seems close to the WTP in NH.

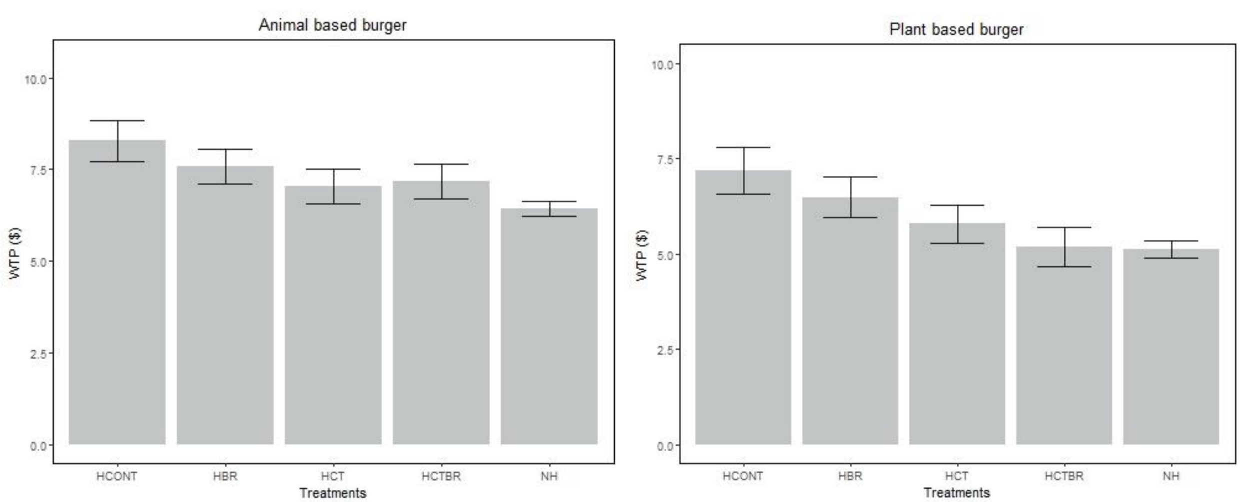

Figure 2.3 WTP for Animal-based and Plant-based Burger per Treatment.

Figure 2.4 presents the simulated market share for both burger products per treatment using the price range of the experiment. It shows how sensitive respondents' preferences were across treatments. For the animal-based burger, the non-hypothetical treatment generated the lowest market shares. One noticeable fact is that the order of the market share trend shifts significantly after \$6. For example, the market shares of HCONT were lower than the market shares of HCTBR for the price range of \$3-\$6. However, the market shares of the latter became smaller for price levels above \$6. For the plant-based burger, the market shares of the non-hypothetical treatment are below the market shares of all other groups except HCTBR at all price levels. However, the gap between the market shares of NH and HCTBR significantly reduce for prices greater than \$6. The two lines get closer, the more the price levels are greater than \$6. Furthermore, market shares of all groups decline markedly for price levels higher than \$6 for the plant-based burger.

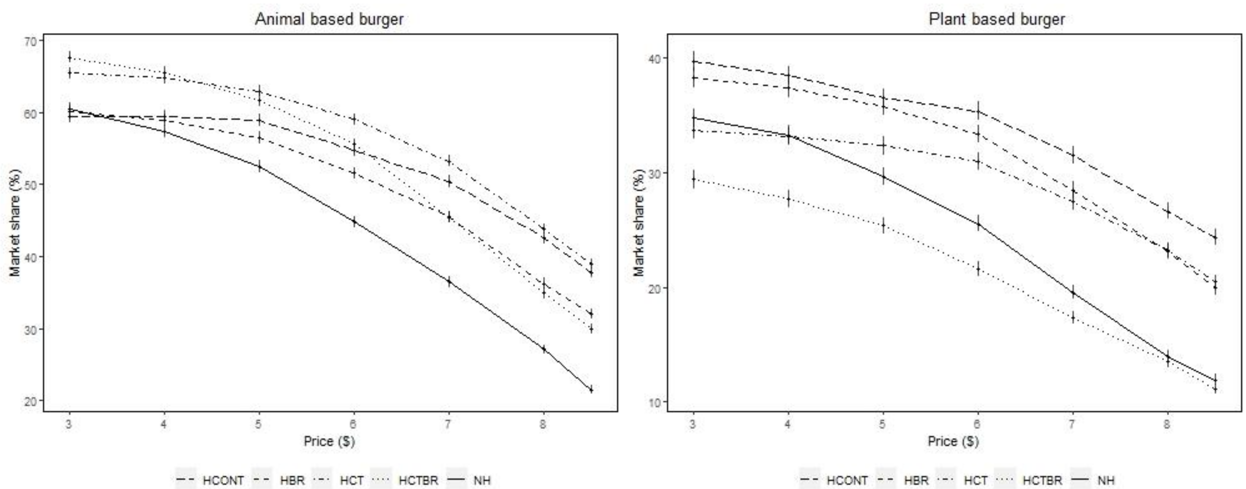

Figure 2.4 Simulated Market Share at Different Price Levels per Product and Treatment

Notes: vertical bars indicate 95% confidence interval at each price level.

Discussion

The choice characterization reveals that respondents tend to choose the no buy option more frequently in the hypothetical setting than in the non-hypothetical condition. This result aligns with previous findings by Bazzani et al. (2017) and Lusk and Schroeder (2004). Our new finding is that adding an explicit uncertainty option in the choice scenario can lead to choosing the no-buy option less. This is evidence of respondents' uncertainty, and it indicates that future DCE may need to incorporate uncertainty adjustment. Our percentage of uncertain answers is low compared to the 31.4% reported by Alberini, Boyle, and Welsh (2003) using contingent valuation. The gap could be explained by the types of products, the sample size, the labeled design, and the stated preference method used. However, the cause of the uncertainty is unknown. Further econometric analysis is needed to establish if adding the uncertainty choice in the choice sets affects welfare estimates and hypothetical bias.

Our results related to hypothetical bias add to the extensive literature on the existence of HB. We found that in the hypothetical control group, respondents overstated their WTP by a factor of 1.29 and 1.40 for the animal-based burger and the plant-based burgers, respectively, relative to the WTP in the non-hypothetical treatment. These ratios are close to mean ratios in Silva et al. (2011), who reported a ratio of 1.2 and Murphy et al. (2005), who found a ratio of 1.35 in their meta-analysis. Moreover, our ratios are in the range reported in the meta-analysis conducted by List and Gallet (2001) and Penn and Hu (2018). These studies have proved that decision-makers exhibit hypothetical bias by overstating their WTP by a factor of up to three times the WTP in an actual market. Our results support findings by Fang et al. (2021) that consumers overstate WTP in hypothetical setting while making food choices.

The price sensibility analysis of the market shares indicates that market shares noticeably declined for both products across treatments for price levels above $6. One possible explanation is that respondents used a reference price while making their choice. Reference price use occurs when consumers compare price encountered in the experiment to other prices encountered in a recent time frame (Miljkovic and Effertz 2010). In fact, consumers participate in experiments with some levels of prior knowledge related to the goods investigated (Caputo, Lusk, and Nayga 2018). Price is arguably one of the most common factors consumers have some prior knowledge on, especially for frequently purchased goods. There is strong empirical evidence that reference prices affect consumers' decisions (e.g., Mazumdar, Raj, and Sinha 2005; Caputo, Lusk, and Nayga 2020; Caputo, Lusk, and Nayga 2018). For instance, Caputo, Lusk, and Nayga (2020) reported that consumers are more price sensitive above a reference price than below, as according to Weaver and Frederick (2012), consumers perceive a deal above this price as a "bad deal".

The magnitude of the hypothetical bias is higher for the plant-based burger than the animal-based burger, even though the absolute WTP for the latter is greater. As explained by Chowdhury et al. (2011) , who found a similar result for different varieties of sweet potatoes in Uganda, this difference in HB may be due to a difference in familiarity with the two products. Lusk (2003) and Aadland and Caplan (2003) also pointed out the significant role of subjects' background related to the products investigated. Chowdhury et al. (2011) also found that subjects exhibit more hypothetical bias for unfamiliar products. In our study, the familiarity scores were 3.79 and 2.46 out of 5, respectively, for the animal-based burger the plant-based burger, suggesting that respondents were less familiar with the plant-based burger.

To investigate whether respondents who were more familiar with the plant-based burger exhibited less HB, we ran model 1 in the subsample of respondents who were more familiar with the plant-based burger[7]. The results indicated no presence of HB for the plant-based burger. However, knowledgeable consumers in the non-hypothetical treatment were willing to pay more for the plant-based burger, \$6.23/5 Oz, compared to the \$5.12/5 Oz in the entire sample. Likewise, we estimated WTP for respondents who were more familiar with the animal-based burger and found no statistical evidence of hypothetical bias for the animal-based burger for this subsample[8]. As for the plant-based burger, knowledgeable consumers in the non-hypothetical treatment were willing to pay more for the animal-based burger (\$7.41/5 Oz). Appendix C.2 reports the results of the subsample analysis. Even though it may be appropriate to define hypothetical bias as a "deviation", our findings support that the direction of the hypothetical bias in food choices is likely to be upward.

The HB mitigation approaches significantly affected the magnitude of the bias. Unlike Loomis, Gonzalez-Caban, and Gregory (1994) , who found that the budget reminder has no effect on WTP estimated using contingent valuation, our result suggests that the method is effective at reducing the hypothetical bias in choice experiment. Our hypotheses testing showed that WTP in HCONT was higher than WTP in HBR, and WTP in HBR were greater than WTP in NH for the animal-based burger. Consistent with Loomis, Gonzalez-Caban, and Gregory (1994), the method was ineffective at mitigating the HB for the plant-based burger, suggesting

---

[7] The subsample includes those with a familiarity score above the sample mean. Score above the mean are scores greater or equal to three.
[8] The subsample includes those with a familiarity score above the sample mean. Score above the mean are scores 4 and 5.

that this approach's effectiveness may depend on the product. Furthermore, BR did not eliminate the bias.

Moreover, we found that CT only and CT combined with BR eliminate the HB for both products. This result is in accord with the findings by Cummings and Taylor (1999), who were the first to investigate the effectiveness of cheap talk at eliminating the HB. Many studies have since investigated the effectiveness of CT or CT combined with BR. Landry and List (2007) , Silva et al. (2011), Aadland and Caplan (2003),and Morrison and Brown (2009) found that WTP with CT were not different from WTP in the non-hypothetical treatment, suggesting that this method eliminated the hypothetical bias. However, other studies found evidence to the contrary. For example, Lusk (2003) and Broadbent (2014) found no evidence of CT eliminating HB. Many studies reported that CT only reduces the bias, and does not completely eliminate it. Penn and Hu (2019) performed a meta-analysis using 67 studies and demonstrated that CT reduces the HB by only 20%. Penn and Hu (2019) and Gschwandtner and Burton (2020) reported that the method performs better while combined with other approaches. The effectiveness of sole CT to eliminate the bias implies that mentioning budget constraints or reminder as in previous studies is not useful or needed. This result is crucial given the popularity of including budget constraints in cheap talk scripts. Penn and Hu (2019) reported that 78% out of 298 cheap talk scripts included some sort of budget reminder or substitute[9]. Cheap talk without any budget constraint reminder is shorter than cheap talk with budget reminder, and it will contribute to reduce surveys duration. There are two main explanations for our results. First, the payment levels we used may have played a significant role. According to Brown, Ajzen, and Hrubes (2003), CT is ineffective for low price levels ($1 and $3), but it is effective for higher price levels ($5 and $8). An alternative

[9] The authors analyzed 238 cheap talk scripts from 67 studies.

explanation is that respondents' familiarity with the products is a potential factor. It is likely that respondents with familiarity score lower than the sample average mainly drove the results. This explanation was confirmed by the subsample analysis results in appendix D.2. Model 1 estimated in the subsample of those who were less familiar with each of the burgers indicate a strong presence of hypothetical bias.

As a robustness check, we estimated the G-MNL-I and the model in which gamma is random. Appendix E.2 presents the results. WTP for both products and in each treatment were similar to those reported above. However, WTP generated by the model with random gamma was closer to the results reported in Table 2.4 for both products (Table 2.6). Gamma is not significant in the model with random gamma, which suggests that the G-MNL-II model is appropriate. Furthermore, G-MNL-II performed better in terms of AIC and BIC than the two alternative models.

Table 2.6 Comparison of WTP ($) for Animal-based and Plant-based Burger Across Alternative Model Specifications

Treatments	Animal-based burger			Plant-based burger		
	G-MNL-II	Gamma random	G-MNL-I	G-MNL-II	Gamma random	G-MNL-I
HCONT	8.280	8.489	8.254	7.191	7.036	6.933
HCT	7.048	7.133	6.844	5.794	5.793	5.246
HBR	7.589	7.546	7.152	6.489	6.446	6.504
HCTBR	7.174	7.123	6.575	5.182	5.201	5.101
NH	6.431	6.278	7.011	5.120	5.038	5.793

Conclusion

Applied economists extensively use discrete choice experiments for different purposes, including evaluating the acceptability and valuation for new products. The method is often used in a hypothetical setting given that new products are not yet available, or policy interventions have not been implemented. One major constraint of using hypothetical discrete choice experiment is the hypothetical bias which can compromise the validity of estimates from such stated preference method. It is, therefore, paramount to investigate approaches to eliminate or at least significantly reduce this bias. This study investigates, for the first time, the effectiveness of budget reminder to eliminate or mitigate the hypothetical bias and compare its effectiveness to cheap talk and cheap talk with budget reminder in a discrete choice experiment. We also explore the effect of a new approach we refer to as "choice task uncertainty adjustment" on choices in a discrete choice experiment. We conducted a laboratory experiment by eliciting respondents' willingness to pay for 5-Oz animal-based and plant-based burgers. The study used a between-subject design with a hypothetical control group, four hypothetical treatment groups, and one non-hypothetical treatment.

The study yields four main findings. First, subjects are likely to choose one of the two proposed products in the hypothetical control group. The choice task uncertainty adjustment reduces the likelihood of choosing the no-buy option. The implication of this finding relates to the need in DCEs to incorporate uncertainty adjustment.

Second, hypothetical bias is present for both products with an overstated factor of 1.29 and 1.40 for the animal-based burger and the plant based-burger, respectively. Third, the budget reminder approach reduces the hypothetical bias for the animal-based burger, but not for the plant-based burger. Lastly, both cheap talk and cheap talk with budget reminder eliminate the hypothetical bias and perform better than the budget reminder by itself. This finding implies that cheap talk scripts designed to eliminate HB may not need a budget reminder to be effective. Given that our study represents only one study, more research is needed to further test the robustness of this finding. Moreover, future research is needed to assess the effect of the choice task uncertainty adjustment on welfare estimates and hypothetical bias.

We add to the growing literature related to hypothetical bias and consumers' acceptance and willingness to pay for plant-based food products. Future research should explore the effectiveness of budget reminders using an unlabeled discrete choice experiment, and also investigate different versions of budget reminder scripts or designs.

APPENDICES

Appendix A.2: Treatment scripts

Cheap talk only script

Studies show that people tend to act differently when they face hypothetical decisions. In other words, they say one thing and do something different. For example, some people state a price they would pay for an item, but when this item becomes available in a grocery store, they will not pay the price they said they would pay.

There can be several reasons for this different behavior. One possibility is that it might be difficult to visualize themselves getting the product from a grocery store shelf and paying for it. Do you understand what I am talking about?

We want you to behave in the same way that you would if you really had to pay for the product and take it home. Please consider how much you really want the product, as opposed to other alternatives that you like or any other constraints that might make you change your behavior, such as taste. Please try to really put yourself in a realistic situation.

Budget reminder script

We would like you to think about your budget, and how your choice could affect your ability to buy other goods.

If you spend more on the burger, you will have less money left for other goods that you could buy given your budget constraint.

Uncertainty script

In making your choices among the burgers products and other alternatives presented in each choice scenario, please choose Unsure/I don't know if you are uncertain which of the two proposed burgers along with the price you would choose. This implies that you would like to buy the product, but do not know which one to choose.

Non-hypothetical script

In the upcoming eight choice tasks, you will choose between two burger products and a "neither of the two options". At the end of the experiment, one of the eight choice tasks will be randomly selected to be the binding choice task. Each choice task has the same probability to be selected as the binding choice task. In the binding choice task, you will be getting the product you chose and will pay the price corresponding to this product (and the cost of the product will be deducted from your participation fee). If your choice in the binding choice task was "neither of the two options", you will not get any product and will not pay anything.

Appendix B.2: Principal components analysis (PCA) of Time preference

We employed a principal component analysis to reduce the 14 time-preference variables into two main components. First, we performed three tests to ensure that it is appropriate to implement a PCA using the data. The Bartlett's test of sphericity is highly significant ($\chi 2 = 894.63$, $p < 0.000$), and it indicates that the correlations are high enough for the PCA. The Kaiser–Meyer–Olkin measure is acceptably high (0.81). Finally, the determinant of the correlation matrix (0.02) indicates no presence of multicollinearity. Eigenvalues are used to select the number of components. As in previous studies (Joireman et al. 2012; De Marchi et al. 2016), the exploratory phase show three eigen values greater than one. Nevertheless, the scree plot (Figure

2.5) below clear shows two components. We, therefore, follow (De Marchi et al. 2016) and (Joireman et al. 2012) and selected the first two components.

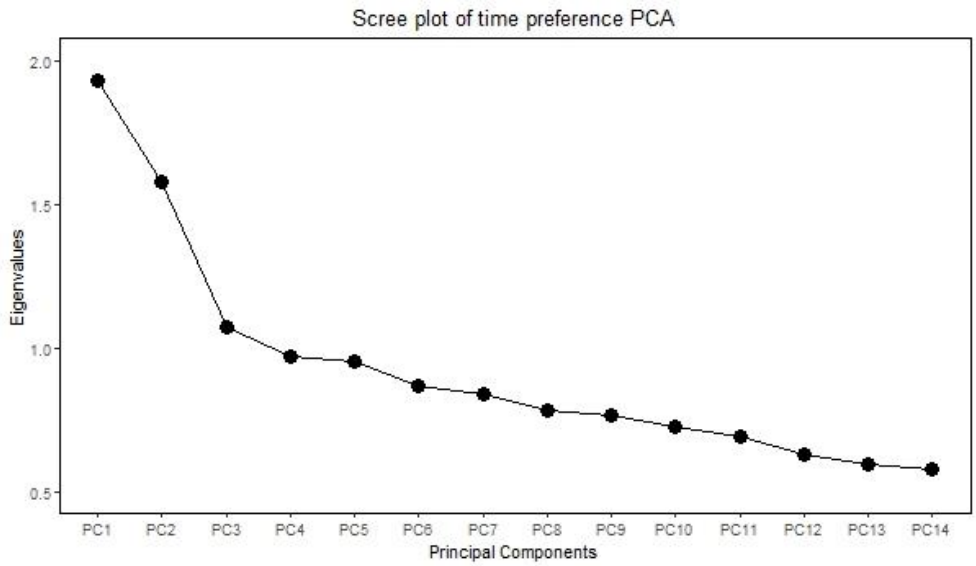

Figure 2.5. Scree Plot of Time Preference Principal Component Analysis

Table 1.7 presents the rotated factor loadings of the rotated component matrix. It is generated using the orthogonal rotation method. Results show that each time preference variable loaded to its expected factors.

Table 2.7 Rotated Component Matrix

Items	CFC-I factor	CFC-F factor
CFC_11 (I)	0.79	-0.2
CFC_5 (I)	0.73	-0.2
CFC_3 (I)	0.7	-0.13
CFC_9 (I)	0.68	-0.02
CFC_10 (I)	0.62	-0.09
CFC_12 (I)	0.61	0.11
CFC_4 (I)	0.53	0.11
CFC_8 (F)	0.23	0.5
CFC_2 (F)	0.05	0.57
CFC_6 (F)	-0.04	0.53
CFC_14 (F)	-0.09	0.79
CFC_13 (F)	-0.12	0.71
CFC_7 (F)	-0.13	0.63
CFC_1 (F)	-0.2	0.65

Appendix C.2: Additional Econometrics Analysis

Table 2.8 Estimates of WTP ($) Using Model 1 in the Subsample of Respondents who are More Familiar with the Animal-based and the Plant-based Burgers

Parameter	Subsample more familiar with animal-based burger[b]		Subsample more familiar with plant-based burger[c]	
	Coefficients[a]	Std.err.	Coefficients	Std.err.
Animal-based burger				
ASC	7.420***	0.257	7.128***	0.337
sd. (ASC)	2.122***	0.126	2.355***	0.179
ASC*HCONT	0.017	0.373	1.241**	0.553
ASC*HCT	-0.927***	0.296	-0.309	0.393
ASC*HBR	-0.273	0.298	-0.001	0.386
ASC*HCTBR	-0.464	0.325	0.806*	0.448
Plant-based burger				
ASC	5.056***	0.283	6.228***	0.331
sd. (ASC)	2.338***	0.101	2.336***	0.135
ASC*HCONT	1.421***	0.450	0.603	0.594
ASC*HCT	0.335	0.337	-0.529	0.401
ASC*HBR	1.158***	0.326	-0.393	0.402
ASC*HCTBR	-0.244	0.366	0.488	0.469
Consumer heterogeneity				
Intercept	1.415***	0.388	1.064***	0.388
Tau	1.299***	0.215	-1.165***	0.229
N	1312		848	
Log-likelihood	-719.042		-470.584	
AIC	1466.084		969.1689	
BIC	1538.594		1035.569	

Note: ***, **, * denotes significance at 1%, 5%, and 10% levels. Sd. indicates standard deviation. *Std.err* indicates standard errors. [a] Coefficients are WTP. Gamma is fixed at 0
[b] Respondents with a familiarity score for animal-based burger above the average were included in this sample. The sample included those with score 4 and 5.
[c] Respondents with a familiarity score for plant-based burger above the average were included in this sample. The sample included those with scores 3,4, and 5.

Appendix D.2: Additional Econometrics Analysis

Table 2.9 Estimates of WTP ($) Using Model 1 in the Subsample of Respondents who were Less Familiar with the Animal-based and the Plant-based Burgers

Parameter	Subsample less familiar with animal-based burger[b]		Subsample less familiar with plant-based burger[c]	
	Coefficients[a]	Std.err.	Coefficients	Std.err.
Animal-based burger				
ASC	4.860***	0.325	6.136***	0.250
sd.(ASC)	2.382***	0.239	2.380***	0.143
ASC*HCONT	4.256***	0.709	2.149***	0.362
ASC*HCT	2.709***	0.477	0.626*	0.346
ASC*HBR	2.479***	0.631	1.861***	0.427
ASC*HCTBR	2.610***	0.528	0.854**	0.345
Plant-based burger				
ASC	5.172***	0.322	4.481***	0.267
sd.(ASC)	2.724***	0.212	2.393***	0.138
ASC*HCONT	1.559**	0.689	2.753***	0.405
ASC*HCT	1.331***	0.481	1.003***	0.353
ASC*HBR	1.382**	0.636	2.373***	0.452
ASC*HCTBR	0.350	0.527	-0.161	0.377
Consumer heterogeneity				
Intercept	0.367	0.250	0.665***	0.243
Tau	-0.864***	0.192	1.010***	0.166
N	776		1240	
Log-likelihood	-500.6303		-752.2605	
AIC	1029.2606		1532.5209	
BIC	1094.4187		1604.241	

Note: ***, **, * denotes significance at 1%, 5%, and 10% levels. Sd. indicates standard deviation. *Std.err* indicates standard errors. [a] Coefficients are WTP. Gamma is fixed at 0

[b] Respondents with a familiarity score for animal-based burger below the average were included in this sample. The sample included those with score 1,2, and 3.

[c] Respondents with a familiarity score for plant-based burger above the average were included in this sample. The sample included those with scores 1 and 2.

Appendix E.2: Alternative Model Specifications

Table 2.10. WTP ($) Estimates Using Alternatives Specifications for Model 1 Using Different Values for Gamma

Parameter	Gamma is random		G-MNL-I[b]	
	Coefficients[a]	Std.err.	Coefficients	Std.err.
Animal-based burger				
ASC	6.278***	0.198	7.011***	0.260
sd.(ASC)	2.194***	0.127	2.283***	0.170
ASC*HCONT	2.211***	0.311	1.243***	0.415
ASC*HCT	0.855***	0.264	-0.167	0.323
ASC*HBR	1.268***	0.278	0.141	0.349
ASC*HCTBR	0.845***	0.262	-0.436	0.323
Plant-based burger				
ASC	5.038***	0.221	5.793***	0.283
sd.(ASC)	2.631***	0.134	2.598***	0.193
ASC*HCONT	1.998***	0.351	1.140**	0.456
ASC*HCT	0.755***	0.288	-0.546	0.348
ASC*HBR	1.408***	0.311	0.711*	0.378
ASC*HCTBR	0.162	0.290	-0.692*	0.366
Consumer heterogeneity				
Intercept	0.815***	0.215	0.241***	0.075
Tau	1.065***	0.141	0.756***	0.071
Gamma	0.068	0.072	-	-
N	2088		2088	
Log-likelihood	-1232.5		-1246.52	
AIC	2494.999		2521.045	
BIC	2579.658		2600.061	

Note: ***, **, * denotes significance at 1%, 5%, and 10% levels. Sd. indicates standard deviation. *Std.err* indicates standard errors. [a] Coefficients are WTP.
[b] gamma is fixed at 1

Appendix F.2: IRB Approval for the Burger Study

UNIVERSITY OF ARKANSAS

To:	Arsene Agossadou
From:	Justin R Chimka, Chair IRB Expedited Review
Date:	11/24/2021
Action:	**Exemption Granted**
Action Date:	11/24/2021
Protocol #:	2107344220
Study Title:	Consumers choices for burgers under uncertainty

The above-referenced protocol has been determined to be exempt.

If you wish to make any modifications in the approved protocol that may affect the level of risk to your participants, you must seek approval prior to implementing those changes. All modifications must provide sufficient detail to assess the impact of the change.

If you have any questions or need any assistance from the IRB, please contact the IRB Coordinator at 109 MLKG Building, 5-2208, or irb@uark.edu.

cc: Di Fang, Investigator
Wei Yang, Investigator

Page 1 of 1

REFERENCES

Aadland, David, and Arthur J. Caplan. 2003. "Willingness to Pay for Curbside Recycling with Detection and Mitigation of Hypothetical Bias." *American Journal of Agricultural Economics* 85 (2):492-502.

Alberini, Anna, Kevin Boyle, and Michael Welsh. 2003. "Analysis of contingent valuation data with multiple bids and response options allowing respondents to express uncertainty." *Journal of Environmental Economics and Management* 45 (1):40-62. doi: https://doi.org/10.1016/S0095-0696(02)00010-4.

Alemu, M. H., and S. B. Olsen. 2018. "Can a Repeated Opt-Out Reminder mitigate hypothetical bias in discrete choice experiments? An application to consumer valuation of novel food products." *European Review of Agricultural Economics* 45 (5):749-782. doi: 10.1093/erae/jby009.

Alemu, Mohammed Hussen, Morten Raun Mørkbak, Søren Bøye Olsen, and Carsten Lynge Jensen. 2013. "Attending to the Reasons for Attribute Non-Attendance in Choice Experiments." *Environmental and Resource Economics* 54 (3):333–359.

Alós-Ferrer, Carlos, Ernst Fehr, and Nick Netzer. 2021. "Time Will Tell: Recovering Preferences When Choices Are Noisy." *Journal of Political Economy* 129 (6):1828–1877. doi: 10.1086/713732.

Andor, M. A., M. Frondel, and C. Vance. 2017. "Mitigating Hypothetical Bias: Evidence on the Effects of Correctives from a Large Field Study." *Environmental & Resource Economics* 68 (3):777-796. doi: 10.1007/s10640-016-0047-x.

Balcombe, K., and I. Fraser. 2011. "A general treatment of 'don't know' responses from choice experiments." *European Review Of Agricultural Economics* 38 (2):171-191. doi: 10.1093/erae/jbr010.

Baskin, Barbara H., and Karen C. Harris. 1995. "Heard Any Good Books Lately? The Case for Audiobooks in the Secondary Classroom." *Journal of Reading* 38 (5):372-376.

Bateman, Ian J., Brett H. Day, Andrew P. Jones, and Simon Jude. 2009. "Reducing gain–loss asymmetry: A virtual reality choice experiment valuing land use change." *Journal of Environmental Economics and Management* 58 (1):106-118. doi: https://doi.org/10.1016/j.jeem.2008.05.003.

Bazzani, C., V. Caputo, R. M. Nayga, and M. Canavari. 2017. "Testing Commitment Cost Theory In Choice Experiments." *Economic Inquiry* 55 (1):383-396. doi: 10.1111/ecin.12377.

Beck, Matthew J, John M Rose, and David A Hensher. 2013. "Consistently inconsistent: The role of certainty, acceptability and scale in choice." *Transportation Research Part E: Logistics and Transportation Review* 56:81-93.

Bello, Muhammad, and Awudu Abdulai 2016. "Impact of Ex-Ante Hypothetical Bias Mitigation Methods on Attribute Non-Attendance in Choice Experiments." *American Journal of Agricultural Economics* 98 (5):1486-1506. doi: https://doi.org/10.1093/ajae/aav098.

Bliemer, Michiel C. J., John M. Rose, and Caspar G. Chorus. 2017. "Detecting dominance in stated choice data and accounting for dominance-based scale differences in logit models." *Transportation Research Part B: Methodological* 102:83-104. doi: https://doi.org/10.1016/j.trb.2017.05.005.

Bobinac, Ana. 2019. "Mitigating hypothetical bias in willingness to pay studies: Post-estimation uncertainty and anchoring on irrelevant information." *The European Journal of Health Economics* 20 (1):75-82.

Bohm, Peter. 1972. "Estimating demand for public goods: An experiment." *European Economic Review* 3 (2):111-130. doi: https://doi.org/10.1016/0014-2921(72)90001-3.

Broadbent, C. D. 2014. "Evaluating mitigation and calibration techniques for hypothetical bias in choice experiments." *Journal of Environmental Planning and Management* 57 (12):1831-1848. doi: 10.1080/09640568.2013.839447.

Brown, Thomas C, Icek Ajzen, and Daniel Hrubes. 2003. "Further tests of entreaties to avoid hypothetical bias in referendum contingent valuation." *Journal of environmental Economics and Management* 46 (2):353-361.

Campbell, Danny, and Victoria S Lorimer. 2009. "Accommodating Attribute Processing Strategies in Stated Choice Analysis: Do Respondents Do What They Say They Do." Comunicación Presentada a la XVII Conferencia Anual de la Asociación Europea de Economía Ambiental y de los Recursos Naturales, Amsterdam.

Caputo, V., E. J. Van Loo, R. Scarpa, R. M. Nayga, and W. Verbeke. 2018a. "Comparing Serial, and Choice Task Stated and Inferred Attribute Non-Attendance Methods in Food Choice Experiments." *Journal of Agricultural Economics* 69 (1):35-57. doi: 10.1111/1477-9552.12246.

Caputo, Vincenzina, Jayson L Lusk, and Rodolfo M Nayga. 2020. "Am I Getting a Good Deal? Reference-DependentDecision Making When the Reference Price Is Uncertain." *American Journal of Agricultural Economics* 102 (1):132-153.

Caputo, Vincenzina, Jayson L. Lusk, and Rodolfo M. Nayga. 2018. "Choice experiments are not conducted in a vacuum: The effects of external price information on choice behavior." *Journal of Economic Behavior & Organization* 145:335-351. doi: https://doi.org/10.1016/j.jebo.2017.11.018.

Caputo, Vincenzina, Giovanni Sogari, and Ellen J. Van Loo. 2022. "Do plant-based and blend meat alternatives taste like meat? A combined sensory and choice experiment study." *Applied Economic Perspectives and Policy* n/a (n/a). doi: https://doi.org/10.1002/aepp.13247.

Caputo, Vincenzina, Ellen J. Van Loo, Riccardo Scarpa, Rodolfo M. Nayga, and Wim Verbeke. 2018b. "Comparing Serial, and Choice Task Stated and Inferred Attribute Non-Attendance Methods in Food Choice Experiments." *Journal of Agricultural Economics* 69 (1):35–57. doi: 10.1111/1477-9552.12246.

Carlsson, Fredrik, Dinky Daruvala, and Henrik Jaldell. 2010. "Do you do what you say or do you do what you say others do?" *Journal of Choice Modelling* 3 (2):113-133.

Carlsson, Fredrik, Morten Raun Mørkbak, and Søren Bøye Olsen. 2012. "The first time is the hardest: A test of ordering effects in choice experiments." *Journal of Choice Modelling* 5 (2):19-37. doi: https://doi.org/10.1016/S1755-5345(13)70051-4.

Carson, Richard T, Theodore Groves, and John A List. 2014. "Consequentiality: A theoretical and experimental exploration of a single binary choice." *Journal of the Association of Environmental and Resource Economists* 1 (1/2):171-207.

Caussade, Sebastián, Juan de Dios Ortúzar, Luis I. Rizzi, and David A. Hensher. 2005. "Assessing the influence of design dimensions on stated choice experiment estimates." *Transportation Research Part B: Methodological* 39 (7):621-640. doi: https://doi.org/10.1016/j.trb.2004.07.006.

Cembalo, Luigi, Francesco Caracciolo, Alessia Lombardi, Teresa Del Giudice, Klaus G Grunert, and Gianni Cicia. 2016. "Determinants of individual attitudes toward animal welfare-friendly food products." *Journal of Agricultural and Environmental Ethics* 29 (2):237-254.

Chang, Anna C. S. 2009. "Gains to L2 listeners from reading while listening vs. listening only in comprehending short stories." *System* 37 (4):652-663. doi: https://doi.org/10.1016/j.system.2009.09.009.

Chang, Anna C. S., and Sonia Millett. 2015. "Improving reading rates and comprehension through audio-assisted extensive reading for beginner learners." *System* 52:91-102. doi: https://doi.org/10.1016/j.system.2015.05.003.

Channa, Hira, Amy Z. Chen, Patricia Pina, Jacob Ricker-Gilbert, and Daniel Stein. 2019. "What drives smallholder farmers' willingness to pay for a new farm technology? Evidence from an experimental auction in Kenya." *Food Policy* 85:64-71. doi: https://doi.org/10.1016/j.foodpol.2019.03.005.

Chowdhury, Shyamal, JV Meenakshi, Keith I Tomlins, and Constance Owori. 2011. "Are consumers in developing countries willing to pay more for micronutrient-dense biofortified foods? Evidence from a field experiment in Uganda." *American Journal of Agricultural Economics* 93 (1):83-97.

Chung, Chanjin, Tracy Boyer, and Sungill Han. 2011. "How many choice sets and alternatives are optimal? Consistency in choice experiments." *Agribusiness* 27 (1):114-125. doi: https://doi.org/10.1002/agr.20252.

Cook, Joseph, Marc Jeuland, Brian Maskery, and Dale Whittington. 2012. "Giving stated preference respondents “time to think”: results from four countries." *Environmental and Resource Economics* 51 (4):473-496.

Cook, Joseph, Dale Whittington, Do Gia Canh, F Reed Johnson, and Andrew Nyamete. 2007. "Reliability of stated preferences for cholera and typhoid vaccines with time to think in Hue, Vietnam." *Economic inquiry* 45 (1):100-114.

Cummings, Ronald G., and Laura O. Taylor. 1999. "Unbiased Value Estimates for Environmental Goods: A Cheap Talk Design for the Contingent Valuation Method." *American Economic Review* 89 (3):649-665. doi: 10.1257/aer.89.3.649.

Daniel, David B., and William Douglas Woody. 2010. "They Hear, but Do Not Listen: Retention for Podcasted Material in a Classroom Context." *Teaching of Psychology* 37 (3):199-203. doi: 10.1080/00986283.2010.488542.

De Marchi, Elisa, Vincenzina Caputo, Rodolfo M. Nayga, and Alessandro Banterle. 2016. "Time preferences and food choices: Evidence from a choice experiment." *Food Policy* 62:99-109. doi: https://doi.org/10.1016/j.foodpol.2016.05.004.

De-Magistris, T., F. Akaichi, and K. Ben Youssef. 2016. "Testing the effectiveness of the oath script in reducing the hypothetical bias in the Contingent Valuation Method." *Agricultural Economics-Zemedelska Ekonomika* 62 (8):378-384. doi: 10.17221/174/2015-agricecon.

de-Magistris, T., and S. Pascucci. 2014. "The effect of the solemn oath script in hypothetical choice experiment survey: A pilot study." *Economics Letters* 123 (2):252-255. doi: 10.1016/j.econlet.2014.02.016.

de-Magistris, Tiziana, Azucena Gracia, and Rodolfo M. Nayga Jr. 2013. "On the Use of Honesty Priming Tasks to Mitigate Hypothetical Bias in Choice Experiments." *American Journal of Agricultural Economics* 95 (5):1136-1154. doi: https://doi.org/10.1093/ajae/aat052.

Dekker, Thijs, Stephane Hess, Roy Brouwer, and Marjan Hofkes. 2016. "Decision uncertainty in multi-attribute stated preference studies." *Resource and Energy Economics* 43:57-73.

Dellaert, B. G. C., B. Donkers, and A. van Soest. 2012. "Complexity Effects in Choice Experiment-Based Models." *Journal of Marketing Research* 49 (3):424-434. doi: 10.1509/jmr.09.0315.

DeLong, K. L., K. G. Syrengelas, C. Grebitus, and R. M. Nayga. 2021. "Visual versus Text Attribute Representation in Choice Experiments." *Journal of Behavioral and Experimental Economics* 94. doi: 10.1016/j.socec.2021.101729.

Diao, Yali, and John Sweller. 2007. "Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations." *Learning and Instruction* 17 (1):78-88. doi: https://doi.org/10.1016/j.learninstruc.2006.11.007.

Ding, Min, Rajdeep Grewal, and John Liechty. 2005. "Incentive-aligned conjoint analysis." *Journal of marketing research* 42 (1):67-82.

Dohmen, Thomas, Armin Falk, David Huffman, Uwe Sunde, Jürgen Schupp, and Gert G. Wagner. 2011. "Individual Risk Attitudes: Measurement, Determinants, And Behavioral Consequences." *Journal of the European Economic Association* 9 (3):522-550. doi: https://doi.org/10.1111/j.1542-4774.2011.01015.x.

Drichoutis, Andreas C, Achilleas Vassilopoulos, Jayson L Lusk, and Rodolfo M Nayga. 2017. "Consumer preferences for fair labour certification." *European Review of Agricultural Economics* 44 (3):455-474.

Dunlap, REVL, K Val Liere, Angela Mertig, and Robert Emmet Jones. 2000. "Measuring endorsement of the new ecological paradigm: A revised NEP scale." *Journal of social issues* 56 (3):425-442.

Fang, Di, Rodolfo M. Nayga Jr, Grant H. West, Claudia Bazzani, Wei Yang, Benjamin C. Lok, Charles E. Levy, and Heather A. Snell. 2021. "On the Use of Virtual Reality in Mitigating Hypothetical Bias in Choice Experiments." *American Journal of Agricultural Economics* 103 (1):142-161. doi: https://doi.org/10.1111/ajae.12118.

Fenichel, Eli P, Frank Lupi, John P Hoehn, and Michael D Kaplowitz. 2009. "Split-sample tests of “no opinion” responses in an attribute-based choice model." *Land Economics* 85 (2):348-362.

Fiebig, Denzil G, Michael P Keane, Jordan Louviere, and Nada Wasi. 2010. "The generalized multinomial logit model: accounting for scale and coefficient heterogeneity." *Marketing Science* 29 (3):393-421.

Fox, J. A., J. F. Shogren, D. J. Hayes, and J. B. Kliebenstein. 1998. "CVM-X: Calibrating contingent values with experimental auction markets." *American Journal of Agricultural Economics* 80 (3):455-465. doi: 10.2307/1244548.

Greene, William H, and David A Hensher. 2010. "Does scale heterogeneity across individuals matter? An empirical assessment of alternative logit models." *Transportation* 37 (3):413-428.

Gschwandtner, A., and M. Burton. 2020. "Comparing treatments to reduce hypothetical bias in choice experiments regarding organic food." *European Review of Agricultural Economics* 47 (3):1302-1337. doi: 10.1093/erae/jbz047.

Gschwandtner, A., C. Jang, and R. McManus. 2020. "Improving Drinking Water Quality in South Korea: A Choice Experiment with Hypothetical Bias Treatments." *Water* 12 (9). doi: 10.3390/w12092569.

Haab, Tim, and Kenneth E. McConnell. 2002. *Valuing Environmental and Natural Resources: The Econometrics of NonMarket Valuation*. Cheltenham, UK and Northampton, MA: Edward Elgar Publishing.

Haaland, Ingar, Christopher Roth, and Johannes Wohlfart. 2022. "Designing Information Provision Experiments." *Journal of Economic Literature*.

Haghani, Milad, Michiel C. J. Bliemer, John M. Rose, Harmen Oppewal, and Emily Lancsar. 2021. "Hypothetical bias in stated choice experiments: Part II. Conceptualisation of external validity, sources and explanations of bias and effectiveness of mitigation methods." *Journal of Choice Modelling* 41:100322. doi: https://doi.org/10.1016/j.jocm.2021.100322.

Handel, Benjamin, and Joshua Schwartzstein. 2018. "Frictions or mental gaps: what's behind the information we (don't) use and when do we care?" *Journal of Economic Perspectives* 32 (1):155-78.

Harrison, Glenn W. 2006. "Experimental evidence on alternative environmental valuation methods." *Environmental and Resource Economics* 34 (1):125-162.

Hensher, David A, John M Rose, and William H Greene. 2005. *Applied choice analysis: a primer*: Cambridge university press.

Hensher, David A., John M. Rose, and William H. Greene. 2012. "Inferring Attribute Non-Attendance from Stated Choice Data: Implications for Willingness to Pay Estimates and a Warning for Stated Choice Experiment Design." *Transportation* 39 (2):235–245. doi: 10.1007/s11116-011-9347-8.

Herriges, Joseph, Catherine Kling, Chih-Chen Liu, and Justin Tobias. 2010. "What are the consequences of consequentiality?" *Journal of Environmental Economics and Management* 59 (1):67-81. doi: https://doi.org/10.1016/j.jeem.2009.03.004.

Hess, Stephane, and David A. Hensher. 2010. "Using conditioning on observed choices to retrieve individual-specific attribute processing strategies." *Transportation Research Part B: Methodological* 44 (6):781-790. doi: https://doi.org/10.1016/j.trb.2009.12.001.

Hoke, O., B. Campbell, M. Brand, and T. Hau. 2017. "Impact of Information on Northeastern US Consumer Willingness to Pay for Aronia Berries." *Hortscience* 52 (3):395-400. doi: 10.21273/hortsci11376-16.

Hole, A. R., J. R. Kolstad, and D. Gyrd-Hansen. 2013. "Inferred vs. stated attribute non-attendance in choice experiments: A study of doctors' prescription behaviour." *Journal of Economic Behavior & Organization* 96:21-31. doi: 10.1016/j.jebo.2013.09.009.

Howard, Gregory and Roe Brian E. and Nisbet Erik C. and Martin Jay F. 2017. "Hypothetical Bias Mitigation Techniques in Choice Experiments: Do Cheap Talk and Honesty Priming Effects Fade with Repeated Choices?" *Journal of the Association of Environmental and Resource Economists* 4 (2):543-573. doi: 10.1086/691593.

Hoyer, W. D., and N. E. Stokburger-Sauer. 2012. "The role of aesthetic taste in consumer behavior." *Journal Of The Academy Of Marketing Science* 40 (1):167-180. doi: 10.1007/s11747-011-0269-y.

Jacquemet, Nicolas, Robert-Vincent Joule, Stéphane Luchini, and Jason F Shogren. 2013. "Preference elicitation under oath." *Journal of Environmental Economics and Management* 65 (1):110-132.

Joireman, Jeff, Monte J Shaffer, Daniel Balliet, and Alan Strathman. 2012. "Promotion orientation explains why future-oriented people exercise and eat healthy: Evidence from the two-factor consideration of future consequences-14 scale." *Personality and Social Psychology Bulletin* 38 (10):1272-1287.

Kragt, Marit E. 2013. "Stated and Inferred Attribute Attendance Models: A Comparison with Environmental Choice Experiments." *Journal of Agricultural Economics* 64 (3):719-736. doi: 10.1111/1477-9552.12032.

Kragt, Marit E., and Jeffrey W. Bennett. 2012. "Attribute Framing in Choice Experiments: How Do Attribute Level Descriptions Affect Value Estimates?" *Environmental and Resource Economics* 51 (1):43-59. doi: 10.1007/s10640-011-9487-5.

Krinsky, Itzhak, and A. Leslie Robb. 1986. "On Approximating the Statistical Properties of Elasticities." *The Review of Economics and Statistics* 68 (4):715-719. doi: 10.2307/1924536.

Ladenburg, J., and S. B. Olsen. 2014. "Augmenting short Cheap Talk scripts with a repeated Opt-Out Reminder in Choice Experiment surveys." *Resource and Energy Economics* 37:39-63. doi: 10.1016/j.reseneeco.2014.05.002.

Lancaster, Kelvin J. C. 1966. "A New Approach to Consumer Theory." *Journal of Political Economy* 74 (2):132-157.

Landry, Craig E., and John A. List. 2007. "Using Ex Ante Approaches to Obtain Credible Signals for Value in Contingent Markets: Evidence from the Field." *American Journal of Agricultural Economics* 89 (2):420-429. doi: https://doi.org/10.1111/j.1467-8276.2007.01017.x.

Lawton, R. N., S. Mourato, D. Fujiwara, and H. Bakhshi. 2020. "Comparing the effect of oath commitments and cheap talk entreaties in contingent valuation surveys: a randomised field experiment." *Journal of Environmental Economics and Policy* 9 (3):338-354. doi: 10.1080/21606544.2019.1689174.

Lewis, K. E., C. Grebitus, and R. M. Nayga. 2016. "The Importance of taste in experimental auctions: consumers' valuation of calorie and sweetener labeling of soft drinks." *Agricultural Economics* 47 (1):47-57. doi: 10.1111/agec.12208.

List, John A., and Craig A. Gallet. 2001. "What Experimental Protocol Influence Disparities Between Actual and Hypothetical Stated Values?" *Environmental and Resource Economics* 20 (3):241-254. doi: 10.1023/A:1012791822804.

Lombardi, Alessia, Riccardo Vecchio, Massimiliano Borrello, Francesco Caracciolo, and Luigi Cembalo. 2019. "Willingness to pay for insect-based food: The role of information and carrier." *Food Quality and Preference* 72:177-187. doi: https://doi.org/10.1016/j.foodqual.2018.10.001.

Loomis, John, Armando Gonzalez-Caban, and Robin Gregory. 1994. "Do reminders of substitutes and budget constraints influence contingent valuation estimates?" *Land Economics*:499-506.

Luckstead, Jeff, Rodolfo M Nayga Jr, and Heather Snell. 2022. "US workers' willingness to accept meatpacking jobs amid the COVID-19 pandemic." *Journal of the Agricultural and Applied Economics Association*.

Luckstead, Jeff, Heather A Snell, Lawton Lanier Nalley, Rodolfo M Nayga Jr, and Joshua Sarpaning. 2022. "A multi-country study on consumers' valuation for child-labor-free chocolate: Implications for child labor in cocoa production." *Applied Economic Perspectives and Policy* 44 (2):1021-1048.

Lund, Randall J. C. 1991. "A Comparison of Second Language Listening and Reading Comprehension." *The Modern Language Journal* 75 (2):196-204. doi: 10.2307/328827.

Lundhede, Thomas Hedemark, Søren Bøye Olsen, Jette Bredahl Jacobsen, and Bo Jellesmark Thorsen. 2009. "Handling respondent uncertainty in choice experiments: Evaluating recoding approaches against explicit modelling of uncertainty." *Journal of Choice Modelling* 2 (2):118-147.

Lusk, Jayson L, and F Bailey Norwood. 2009a. "An inferred valuation method." *Land Economics* 85 (3):500-514.

Lusk, Jayson L, and F Bailey Norwood. 2009b. "Bridging the gap between laboratory experiments and naturally occurring markets: an inferred valuation method." *Journal of Environmental Economics and Management* 58 (2):236-250.

Lusk, Jayson L, and Ted C Schroeder. 2004. "Are choice experiments incentive compatible? A test with quality differentiated beef steaks." *American journal of agricultural economics* 86 (2):467-482.

Lusk, Jayson L. 2003. "Effects of Cheap Talk on Consumer Willingness-to-Pay for Golden Rice." *American Journal of Agricultural Economics* 85 (4):840-856. doi: https://doi.org/10.1111/1467-8276.00492.

Mamkhezri, J., J. A. Thacher, J. M. Chermak, and R. P. Berrens. 2020. "Does the solemn oath lower WTP responses in a discrete choice experiment application to solar energy?" *Journal of Environmental Economics and Policy* 9 (4):447-473. doi: 10.1080/21606544.2020.1738276.

Matthews, Yvonne, Riccardo Scarpa, and Dan Marsh. 2017. "Using virtual environments to improve the realism of choice experiments: A case study about coastal erosion management." *Journal of Environmental Economics and Management* 81:193-208. doi: https://doi.org/10.1016/j.jeem.2016.08.001.

Mazumdar, Tridib, S. P. Raj, and Indrajit Sinha. 2005. "Reference Price Research: Review and Propositions." *Journal of Marketing* 69 (4):84-102. doi: 10.1509/jmkg.2005.69.4.84.

McFadden, Daniel. 1973. "Conditional Logit Analysis of Qualitative Choice Behavior." In *Frontiers in Econometrics*, edited by Paul Zarembka, 105–142. New York, NY: Academic Press.

Miljkovic, Dragan, and Cary Effertz. 2010. "Consumer behavior in food consumption: reference price approach." *British Food Journal*.

Mokas, Ilias, Sebastien Lizin, Tom Brijs, Nele Witters, and Robert Malina. 2021. "Can immersive virtual reality increase respondents' certainty in discrete choice experiments? A comparison with traditional presentation formats." *Journal of Environmental Economics and Management* 109:102509. doi: https://doi.org/10.1016/j.jeem.2021.102509.

Moreno, Roxana, and Richard E. Mayer. 2002. "Verbal redundancy in multimedia learning: When reading helps listening." *Journal of Educational Psychology* 94 (1):156-163. doi: 10.1037/0022-0663.94.1.156.

Morrison, Mark, and Thomas C Brown. 2009. "Testing the effectiveness of certainty scales, cheap talk, and dissonance-minimization in reducing hypothetical bias in contingent valuation studies." *Environmental and Resource Economics* 44 (3):307-326.

Murphy, James J., P. Geoffrey Allen, Thomas H. Stevens, and Darryl Weatherhead. 2005. "A Meta-analysis of Hypothetical Bias in Stated Preference Valuation." *Environmental and Resource Economics* 30 (3):313-325. doi: 10.1007/s10640-004-3332-z.

Nayga, R. M., W. Aiew, and J. P. Nichols. 2005. "Information effects on consumers' willingness to purchase irradiated food products." *Review of Agricultural Economics* 27 (1):37-48. doi: 10.1111/j.1467-9353.2004.00206.x.

Norwood, F Bailey, and Jayson L Lusk. 2011. "Social desirability bias in real, hypothetical, and inferred valuation experiments." *American Journal of Agricultural Economics* 93 (2):528-534.

Olsen, S. B., T. H. Lundhede, J. B. Jacobsen, and B. J. Thorsen. 2011. "Tough and Easy Choices: Testing the Influence of Utility Difference on Stated Certainty-in-Choice in Choice Experiments." *Environmental & Resource Economics* 49 (4):491-510. doi: 10.1007/s10640-010-9443-9.

Papoutsi, Georgia S., Rodolfo M. Nayga, Panagiotis Lazaridis, and Andreas C. Drichoutis. 2015. "Fat tax, subsidy or both? The role of information and children's pester power in food choice." *Journal of Economic Behavior & Organization* 117:196-208. doi: https://doi.org/10.1016/j.jebo.2015.06.011.

Penn, J., and W. Y. Hu. 2019. "Cheap talk efficacy under potential and actual Hypothetical Bias: A meta-analysis." *Journal of Environmental Economics and Management* 96:22-35. doi: 10.1016/j.jeem.2019.02.005.

Penn, J., and W. Y. Hu. 2021. "Mitigating hypothetical bias by defaulting to opt-out in an online choice." *Applied Economics* 53 (3):315-328. doi: 10.1080/00036846.2020.1808172.

Penn, Jerrod M., and Wuyang Hu. 2018. "Understanding Hypothetical Bias: An Enhanced Meta-Analysis." *American Journal of Agricultural Economics* 100 (4):1186-1206. doi: https://doi.org/10.1093/ajae/aay021.

Pimentel, David, and Marcia Pimentel. 2003. "Sustainability of meat-based and plant-based diets and the environment." *The American Journal of Clinical Nutrition* 78 (3):660S-663S. doi: 10.1093/ajcn/78.3.660S.

Poe, Gregory L, Kelly L Giraud, and John B Loomis. 2005. "Computational Methods for Measuring the Difference of Empirical Distributions." *American Journal of Agricultural Economics* 87 (2):353–365.

Revelt, David, and Kenneth Train. 1999. "Customer-specific taste parameters and mixed logit." *University of California, Berkeley*.

Rogowsky, B. A., B. M. Calhoun, and P. Tallal. 2016. "Does Modality Matter? The Effects of Reading, Listening, and Dual Modality on Comprehension." *Sage Open* 6 (3). doi: 10.1177/2158244016669550.

Rose, John M, Michiel CJ Bliemer, David A Hensher, and Andrew T Collins. 2008. "Designing efficient stated choice experiments in the presence of reference alternatives." *Transportation Research Part B: Methodological* 42 (4):395-406.

Sarrias, Mauricio, and Ricardo Daziano. 2017. "Multinomial Logit Models with Continuous and Discrete Individual Heterogeneity in R: The gmnl Package." *Journal of Statistical Software* 79 (2):1 - 46. doi: 10.18637/jss.v079.i02.

Savage, Scott J., and Donald M. Waldman. 2008. "Learning and Fatigue during Choice Experiments: A Comparison of Online and Mail Survey Modes." *Journal of Applied Econometrics* 23 (3):351-371.

Scarpa, Riccardo, and John M Rose. 2008. "Design efficiency for non-market valuation with choice modelling: how to measure it, what to report and why." *Australian journal of agricultural and resource economics* 52 (3):253-282.

Scarpa, Riccardo, Mara Thiene, and Kenneth Train. 2008. "Utility in willingness to pay space: a tool to address confounding random scale effects in destination choice to the Alps." *American Journal of Agricultural Economics* 90 (4):994-1010.

Scarpa, Riccardo, Kenneth G Willis, and Melinda Acutt. 2007. "Valuing externalities from water supply: Status quo, choice complexity and individual random effects in panel kernel logit analysis of choice experiments." *Journal of Environmental Planning and Management* 50 (4):449-466.

Scarpa, Riccardo, Raffaele Zanoli, Viola Bruschi, and Simona Naspetti. 2013. "Inferred and Stated Attribute Non-attendance in Food Choice Experiments." *American Journal of Agricultural Economics* 95 (1):165–180. doi: 10.1093/ajae/aas073.

Schulze, William D, Ralph C d'Arge, and David S Brookshire. 1981. "Valuing environmental commodities: some recent experiments." *Land Economics* 57 (2):151-172.

Shr, Yau-Huo Jimmy, Richard Ready, Brian Orland, and Stuart Echols. 2019. "How do visual representations influence survey responses? Evidence from a choice experiment on landscape attributes of green infrastructure." *Ecological Economics* 156:375-386.

Silva, Andres, Rodolfo M Nayga Jr, Benjamin L Campbell, and John L Park. 2011. "Revisiting cheap talk with new evidence from a field experiment." *Journal of Agricultural and Resource Economics*:280-291.

Slade, Peter. 2018. "If you build it, will they eat it? Consumer preferences for plant-based and cultured meat burgers." *Appetite* 125:428-437. doi: https://doi.org/10.1016/j.appet.2018.02.030.

Tarfasa, Solomon, Roy Brouwer, Oleg Sheremet, and Jetske Bouma. 2017. "Informing water harvesting technology contract design using choice experiments." *Water Resources Research* 53 (10):8211-8225.

Train, Kenneth E. 2009. *Discrete choice methods with simulation*. Second edition ed. Cambridge: Cambridge university press.

Turner, Karynne L., and Mona V. Makhija. 2012. "The role of individuals in the information processing perspective." *Strategic Management Journal* 33 (6):661-680. doi: https://doi.org/10.1002/smj.1970.

Van Loo, E. J., V. Caputo, R. M. Nayga, H. S. Seo, B. Y. Zhang, and W. Verbeke. 2015. "Sustainability labels on coffee: Consumer preferences, willingness-to-pay and visual attention to attributes." *Ecological Economics* 118:215-225. doi: 10.1016/j.ecolecon.2015.07.011.

Van Loo, Ellen, Vincenzina Caputo, Jr Rodolfo M. Nayga, Jean-Francois Meullenet, Philip G. Crandall, and Steven C. Ricke. 2010. "Effect of Organic Poultry Purchase Frequency on Consumer Attitudes Toward Organic Poultry Meat." *Journal of Food Science* 75 (7):S384-S397. doi: https://doi.org/10.1111/j.1750-3841.2010.01775.x.

Vossler, Christian A, Maurice Doyon, and Daniel Rondeau. 2012. "Truth in consequentiality: theory and field evidence on discrete choice experiments." *American Economic Journal: Microeconomics* 4 (4):145-71.

Walker, Joan L, Yanqiao Wang, Mikkel Thorhauge, and Moshe Ben-Akiva. 2018. "D-efficient or deficient? A robustness analysis of stated choice experimental designs." *Theory and Decision* 84 (2):215-238.

Weaver, Ray, and Shane Frederick. 2012. "A reference price theory of the endowment effect." *Journal of Marketing Research* 49 (5):696-707.

Weir, Michael J., Hirotsugu Uchida, and Maya Vadiveloo. 2021. "Quantifying the effect of market information on demand for genetically modified salmon." *Aquaculture Economics & Management* 25 (1):1-26. doi: 10.1080/13657305.2020.1803447.

Wuepper, D., P. Wree, and G. Ardali. 2019. "Does information change German consumers' attitudes about genetically modified food?" *European Review of Agricultural Economics* 46 (1):53-78. doi: 10.1093/erae/jby018.

Zellner, Arnold. 1962. "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias." *Journal of the American statistical Association* 57 (298):348-368.

Zossou, Espérance, Rose Fiamohe, Simple Davo Vodouhe, and Matty Demont. 2022. "Experimental auctions with exogenous and endogenous information treatment: Willingness to pay for improved parboiled rice in Benin." *Journal of Agricultural Economics* n/a (n/a). doi: https://doi.org/10.1111/1477-9552.12482.