Who is asking: Impact of Gender and Ethnicity on Contingent Valuation Estimates

Abdelaziz Lawani[1*], Ph.D., Anna Liisa Ihuhwa[2]

[1] Tennessee State University, [2] Vanderbilt University, alawani@tnstate.edu

Selected Paper prepared for presentation at the 2024 Agricultural & Applied Economics Association Annual Meeting, New Orleans, LA; July 28-30, 2024

Who is asking: Impact of Gender and Ethnicity on Contingent Valuation Estimates

Abstract

Contingent valuation (CV) is a widely utilized method for assessing the economic value of non-market goods and services, yet debates persist regarding its validity and reliability. While previous research has explored various factors influencing CV estimates, including survey design and respondent characteristics, the influence of the interviewer, particularly their gender and ethnicity, remains understudied. This study addresses this gap by investigating how the gender and ethnicity of the interviewer affect CV estimates. Employing a double-bounded dichotomous choice survey with 513 respondents representative of the U.S. population, the study examines willingness to pay for child-labor-free chocolate and hemp-based clothing. Findings indicate significant effects of interviewer gender and ethnicity on willingness to pay for these goods.

Keywords: Experimental Design, contingent valuation, stated preference method, gender, ethnicity, willingness to pay.

JEL Codes: C9, Q51, C83

Who is asking: Impact of Gender and Ethnicity on Contingent Valuation Estimates

1. Introduction

Contingent valuation (CV) is one of the most widely used methods to assess the economic value of non-market goods and services. More than 10,000 CV studies have been undertaken worldwide (Haab et al., 2020) and the method has been applied across various fields such as economics, environmental and natural resources management, healthcare, infrastructure planning, education, parks, and recreation. Given that these goods and services are not typically traded in economic markets CV is a stated preference (SP) method relying on individual preferences expressed in surveys that simulate hypothetical markets (Cuccia, 2020).

As the methodology relies on asking people questions, as opposed to observing their actual behavior, it has become a subject of significant debates in economic literature (Adamowicz et al., 1994). Controversies surrounding CV methods often center on issues of validity and reliability (Bishop et al., 2019). Validity relates to bias or the degree to which the method measures the theoretical construct under investigation, while reliability is about variance, the variability of the estimates (Perni et al., 2021). The predominant body of research in economics has studied factors such as survey design, respondent characteristics, and model specifications to understand the origins and solutions to bias and variability in CV estimates. Yet, the impact of the individual administering the survey, including factors related to gender and ethnicity, remains a relatively understudied aspect in the CV literature even though they have been shown to influence data and estimates in diverse literature (Davis et al., 2010; Durant et al., 2010; West et al., 2017). This research aims to contribute to the literature on the effects of human interviewers in CV studies. It examines the extent to which gender and ethnicity affect the validity and reliability of CV estimates.

This research uses a double-bounded dichotomous choice survey of a total of 513 respondents. The respondents are representative of the U.S. population and were surveyed on their willingness to pay for child-labor-free dark chocolate and hemp-based white T-shirts for a robustness check. The respondents were randomly assigned to one of five main treatment categories, distinguished by the perceived race (black or white) and gender (male or female) of the researcher conducting the study. The survey is introduced using the recorded voice of the researcher and the hypothetical scenarios are also narrated using the voice of the researcher. Once the hypothetical scenarios were described to them, the respondents were asked about their willingness to pay for the child-labor-free chocolate, the hemp-based T-shirt, and individual characteristics. The results of the study show that gender and ethnicity affect the willingness to pay for child-labor-free chocolate and hemp-based clothing.

The remainder of the paper is organized as follows. Section 2 describes the survey and the empirical strategy of the study. Section 3 discusses the results and section 4 concludes with the policy implications of the study.

2. Methodology

To determine the effect of the interviewer's ethnicity and gender on CV estimates, we designed an experiment where respondents to two CV questions were randomly assigned to one of five primary treatment groups categorized by the perceived race and gender of the interviewing conducting the study. Respondents coming from a representative sample of U.S. population have been purchased from Qualtrics. They will be randomly assigned to one of the treatments groups. To manipulate race and gender, we follow Bertrand and Mullainathan (2004) Carpusor et al., (2006), and Edelman et al., (2017) by randomly assigning to the respondent an interviewer that has either African American male, African American female, or white male, or white female sounding name. The control group has no name and no recording. A typical treatment consists of having the respondent listen to a recording of a text describing the survey and if they prefer, they can read the corresponding text in addition to having listened to the recording. In the recording and text, the name of the human interviewer is mentioned. The recordings have been made by hired voice artists with tones corresponding to each group. To validate our list of names and the recordings, a short survey was conducted, and participants were asked to categorize each name and recording as white male, white female, African American male, or African American female within three seconds. According to Edelman et al. (2017), the three-second time limit to categorize the names ensures that the respondents had limited time to ponder beyond an instinctive response.

It is important to stress that we did not use an experiment where the interviewer introduced the survey to the respondent in person. Indeed, due to the inherent variability among human interviewers, who differ in numerous ways beyond the specific variable of interest here (gender and ethnicity), establishing a clear causal relationship in an in-person experiment can be challenging. Such an experiment would also be susceptible to experimenter bias, as it cannot be conducted in a fully blinded manner, and there is a risk that the interviewer anticipating a particular form of bias might inadvertently influence its occurrence.

After the treatment, the respondent answers a typical CV question. Here we use two products: a child-labor-free chocolate bar and a T-shirt made out of hemp. We use the following hypothetical scenarios for the child-labor-free dark chocolate and the hemp-based T-shirt:

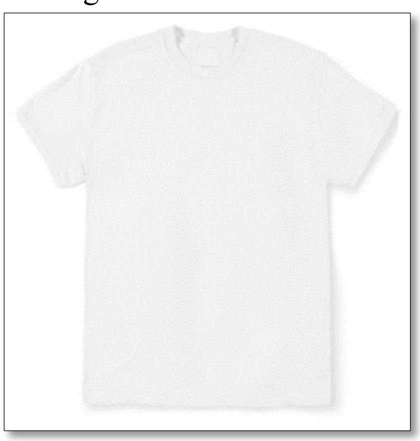
Hypothetical Scenario for child-labor free chocolate.
Chocolate is produced using coca beans and child labor is common in cocoa production. Côte d'Ivoire and Ghana are the two major cocoa producers, contributing to more than 60% of global cocoa bean production. Using the International Labor Organization's definitions of child labor, a study conducted by Tulane University shows that in 2014, 2.12 million children were working in child labor in cocoa production in Côte d'Ivoire and Ghana. The child labor practices included the use of machetes for weeding, carrying heavy loads exceeding permissible weight limits, working long hours on school days, withdrawal from school during cocoa season for farm work, exposure to agrochemicals, forest clearing, tree climbing, and various other activities.
The prevalence of such child labor practices raised concerns about the well-being and safety of young workers in the cocoa industry. Chocolate can also be produced using cocoa beans cultivated without child labor. In this scenario, we are only focused on child-labor-free 70% dark chocolate. More specifically, consider the following hypothetical 70% dark chocolate bar produced without child labor. The chocolate has the same size, quality, taste, and other characteristics as the 70% dark chocolate you usually buy when shopping. It is safe and will have no side effects on those consuming it.

Hypothetical Scenario for hemp-based clothing

Hemp is a plant belonging to the Cannabis sativa species. Hemp differentiates itself from other cannabis strains, like marijuana, by its significantly lower THC cannabinoid content (less than 0.3 percent) and higher cannabidiol (CBD) content. The tetrahydrocannabinol (THC) is the psychoactive compound responsible for the feeling "high" effects, such as euphoria or intoxication, commonly associated with recreational or medical cannabis. Hemp does not provide the psychoactive effects associated with marijuana. Consequently, hemp is utilized for various purposes distinct from those of its counterparts in the cannabis family.

Hemp plants are grown and used for several products such as building materials (roofing, flooring, paint), biofuel, beauty products (lotions, body wash, shampoo), foods (milk, protein powder, drinks), clothing (jeans, t-shirts, shoes, belts), etc.

In this scenario, we are only focused on the use of hemp in clothing. More specifically, consider the following hypothetical unisex T-shirt made out of hemp. The T-shirt has the same property as the usual T-shirt you usually buy when shopping, it is safe and will have no side-effects on those wearing it.



The hypothetical scenarios were narrated using the voices that corresponds to the treatment.

Respondents also have the option to read the scenario after the narration if they want.

Various alternative contingent valuation methods are available for assessing Willingness to Pay (WTP). In this study, we employ the double-bounded dichotomous choice method (DBDC), which is simple, incentive-compatible, and more efficient compared to its counterpart, the single-bounded dichotomous choice (SBDC) (Hanemann, Loomis and Kanninen 1991; Hoehn and Randall 1989). Unlike the SBDC, which provides respondents with a simple YES or NO decision at a single bid price, DBDC presents respondents with YES or NO decisions across multiple bid prices (Brago et al. 2022). In SBDC, respondents are asked whether they are willing to pay a specific premium for the marketed product. The double-bound method involves a two-step procedure. Initially, respondents are asked if they would be willing to pay a predetermined amount for the product after hypothetically knowing the outcome. Subsequently, in the second stage, a lower or higher bid is presented based on the response to the initial bid.

We have recruited a representative sample of the U.S. population through Qualtrics. In the survey, we include several attention-check questions. Some ask respondents to leave the answer blank, others ask to perform a simple computation (what is 2+2), and finally, the last type of attention check invites respondents to choose a specific answer. Only the respondents who passed all attention check questions were included in this study. Out of a total of the 4062 respondents who took part in this survey, only 513 respondents passed all the attention check questions and were included in the data analysis. The strict consideration of the responses of participants who pass the attention check questions is due to the sensitivity of the research question. Indeed, since we are using the voices as treatments, it is important we ensure that the respondents were not distracted during the survey.

Table 1 compares the socio-economic characteristics of the respondents to the characteristics of the U.S. population. We notice very minor differences between the two.

Table 1: Socio-economic characteristics of participants in the experiments and the U.S. population

		U.S. Population	Experimental sample
Number of participants			513
Age	18-34	0.30	0.29
	35-54	0.32	0.33
	55+	0.38	0.39
Gender	Male	0.48	0.48
	Female	0.52	0.52
Race	White	0.75	0.76
	Black	0.13	0.12
	Others	0.12	0.12
Household Income	Less than $50K	0.35	0.35
	$50K-100K	0.35	0.35
	100K+	0.30	0.30
Education	No college degree	0.65	0.28
	4-year degree or higher	0.35	0.72

Price ranges, for both products, were selected based on prior literature and store prices. To mitigate initial bid bias, for each product, three different starting bids were selected and equally randomly shown to the respondents. Table 2 summarizes the initial, lower, and higher bids used in this experiment.

Table 2: Different bids (in $) used in the experiment for each product.

Product	Initial bid B_i ($)	Lower bid, B_i^L, if the initial bid is declined ($)	Higher bid B_i^U, if the initial bid is accepted ($)
Child-labor-free chocolate	3.26	1.20	5.20
	5.32	3.20	7.20
	7.38	5.20	9.20
Hemp-based T-Shirt	10.57	5.14	20.46
	20.46	10.14	40.16
	30.60	15.14	60.16

For the initial bid, B_i, respondents will respond with either "yes" or "no." If the first bid is accepted, they are presented with the second higher bid, B_i^U; conversely, if the first bid is declined, they are presented with the second lower bid, B_i^L. Consequently, four potential

Willingness to Pay (WTP) outcome pairs may be derived for each product: "Yes-Yes," "Yes-No," "No-Yes," and "No-No."

Following Aikoh et al. (2020) we represent the utility of a respondent k in treatment T by the following:

$$U_{kT}^{i} = V_{kT}^{i} + \varepsilon_{kT}^{i}$$

where V_{kT}^{i} is the deterministic term of the utility and ε_{kT}^{i} is its stochastic term. If the respondent k answers Yes to the bid, i will take the value Y. Contrary, it will take the value N if he answers No to the bid. Considering treatment T, U_{kT}^{Y} is the utility he derives from answering Yes to the bid and U_{kT}^{N} is the corresponding utility when he answers No.

If the respondent answers Yes, then

$$U_{kT}^{Y} > U_{kT}^{N}$$

and the corresponding probability that the respondent answers Yes can be written as:

$$P_{kT}^{Y} = \Pr(U_{kT}^{Y} > U_{kT}^{N}) = \Pr(V_{kT}^{Y} + \varepsilon_{kT}^{Y} > V_{kT}^{N} + \varepsilon_{kT}^{N})$$

Given a logistic distribution, and ΔV the utility difference function, P_{kT}^{Y} can be rewritten as:

$$P_{kT}^{Y} = \{1 + \exp(-\Delta V)\}^{-1}$$

and $\Delta V = \alpha + \beta ln B_i + \gamma X_k$ where B_i is the initial bid and $X_k$the set of independent variables such as gender, age, income, and ethnicity. We can extend the first step decision by the respondent to develop the double-bounded dichotomous choice model.

If respondent k in treatment T answers no to both B_i, and B_i^L then

$$P_{kT}^{NN}(B_i, B_i^L) = \Pr(B_i > \ WTP \ and \ B_i^L > WTP) = D(B_i^L;\ \theta)$$

where D is a distribution function and θ the parameter vector. We can also derive P_{kT}^{NY}, P_{kT}^{YN}, and P_{kT}^{YY} as follows:

$$\begin{cases} P_{kT}^{NY}(B_i, B_i^L) = \Pr(B_i \geq \ WTP \ \geq B_i^L) = D(B_i;\ \theta) - D(B_i^L;\ \theta) \\ P_{kT}^{YN}\left(B_i, B_i^U\right) = \Pr\left(B_i \leq \ WTP \ \leq B_i^U\right) = D\left(B_i^U;\ \theta\right) - \ D(B_i;\ \theta) \\ P_{kT}^{YY}\left(B_i, B_i^U\right) = \Pr\left(B_i \leq \ WTP \ and \ B_i^U \leq WTP\right) = 1 - D\left(B_i^U;\ \theta\right) \end{cases}$$

The parameters can be estimated using the maximum likelihood method. Assuming a log-linear function for the distribution D, and α and β respectively the constant term and the parameter of logarithmic value of the bid for θ:

$$D(B) = \{1 + \exp\{-(\alpha + \ \beta ln B_i)\})\}^{-1}$$

Other factors (X_k) can also be included in the model.

Using the estimated parameters α and β we can calculate the median WTP which is the amount that the probability that the respondent will answer Yes to the bid is 50% (Hanemann 1984). It is calculated as follows:

$$Median\ WTP = \exp\left(-\frac{\alpha}{\beta}\right)$$

To test our hypothesis and compare the impact of gender and voice on WTP, we could also include the treatments in the model. However, doing so leads to non-convergence of the model. Instead, we computed the median WTP for each treatment and tested if there is a significant difference between them. Given that the data does not meet the normality assumptions of ANOVA, we use the non-parametric Kruskal-Wallis's test and perform pairwise comparisons with Dunn's post hoc test to determine which specific treatments differ from each other.

The analyses for this study were conducted using the R package DCchoice (Aizaki et al. 2022).

3. Results

3.1. Implied demand curve for child-labor-free chocolate and hemp-based T-shirt

Table 3 summarizes the survey response summary statistics for the bid values.

Table 3: Survey response summary statistics

	Child-labor-free Chocolate			Hemp-based T-shirt		
	3.26	5.32	7.38	10.57	20.46	30.6
0	18	32	47	14	67	80
1	92	119	80	102	78	47
Total	110	151	127	116	145	127

There are 110 survey responses for a bid price of $3.26. of the 110 responses, 18 said no they wouldn't be willing to pay $3.26 for child-labor-free chocolate and 92 said yes, they would be willing to pay. Also, off the 151 survey responses for $5.32, 32 said no, they wouldn't be willing to pay while 119 said yes, they would pay. Finally, for the price of $7.38, 47 said no, and 80 said yes.

Similarly, of the 116 responses for the price of $10.57 for hemp-based T-shirts, 14 said no, and 102 said yes. 67 said no to $20.46 and 78 said yes. And to the price of $30.6, 80 said no and 47 said yes. When observing the number of people who said yes for the hemp-based T-shirt, their number decreases as the bid amount increases (102,78, and 47). However, for child-labor-free chocolate, the number of yes does not follow the same pattern. It increases first (from 92 to 119) before decreasing (from 119 to 80). Figure 1 presents the implied demand curve for each product which is the proportion of respondents saying yes to each bid price. It shows that the law of demand holds. As the bid price increases, the proportion of people that are willing to pay for the bid decreases.

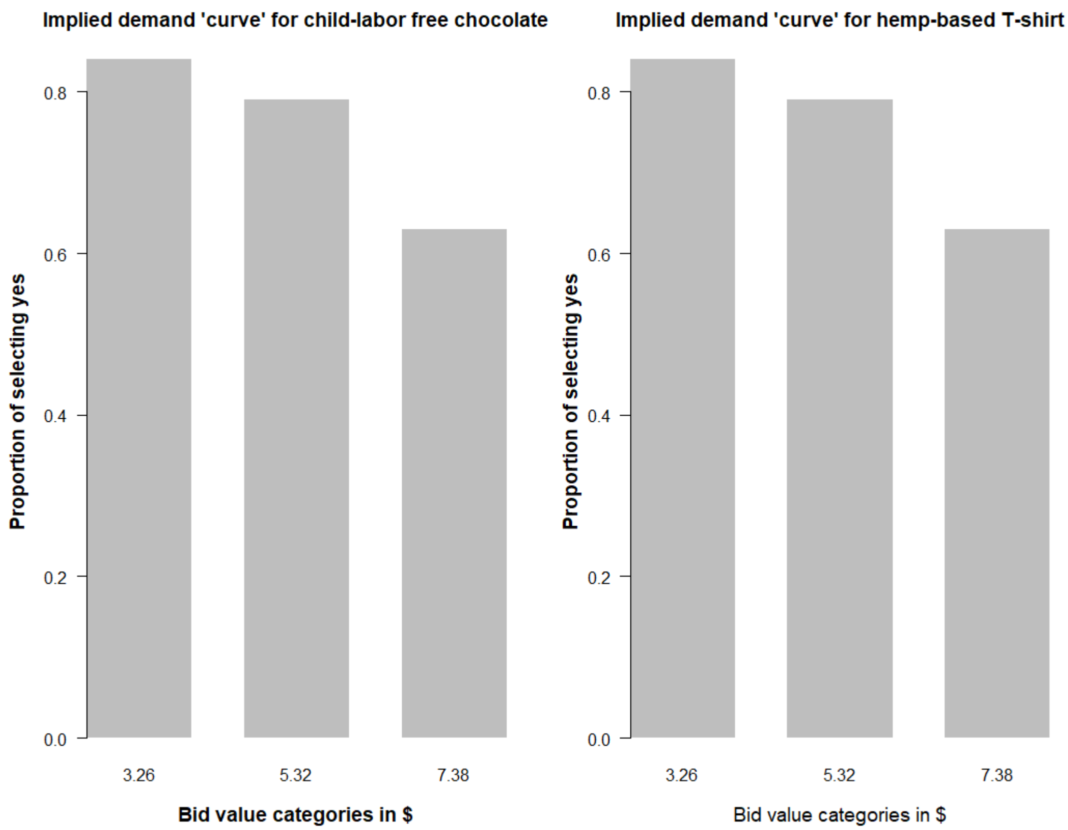


Figure 1: Proportion of people saying yes for each bid amount for child-labor-free chocolate and hemp-based T-shirt.

3.2. Comparison across treatment groups

Table 4 displays the median willingness to pay (WTP) for child-labor-free chocolate and hemp-based T-shirts across treatment and control groups. The analysis reveals several key findings.

First, respondents are willing to pay a price premium to eliminate child labor in chocolate production. This result is consistent with Luckstead et al. (2022) findings. Similarly, respondents are also willing to pay a premium for hemp-based clothing.

Second, including the voice of the researcher increases the respondents' WTP compared to the control group. This result holds for all group across the two products except for child-labor free chocolate when the respondents hear a black female voice. In this case the results show that respondents are willing to pay less compared to the control group.

Third, when comparing the different gender and ethnicity, for hemp-based clothing, the voice of white female interviewer leads to the highest WTP ($25.95). This is followed by the voice of white male interviewer ($25.46) and the voice of black male interviewer ($24.93). the voice of black female interviewer leads the lowest WTP ($21.72) which is closest to the control group ($21.51). it is as if the voice of black female interviewer does not induce any additional WTP.

For child-labor-free chocolate, black male voice induces the highest WTP ($7.39), followed by white female voice ($7.17), and white male voice ($7.06). Black female voice also lead to the lowest WTP ($6.21) which is even smaller than the WTP of the control group ($6.89).

These results show that the voice of the interviewer does impact WTP and some ethnicities and genders have a higher impact compared to when the respondents hear no voice

Table 4: Median WTP for child-labor-free chocolate and hemp-based T-shirt

	Child-labor-free chocolate					Hemp-based T-shirt				
	Control: no voice	Black male voice	Black female voice	White male voice	White female voice	Control: no voice	Black male voice	Black female voice	White male voice	White female voice
WTP	6.89***	7.39***	6.21***	7.06***	7.17***	21.51***	24.93***	21.72***	25.46***	25.95***
95%CI	(6.80,8.26)	(6.478,8.53)	(6.44,8.93)	(6.62,8.27)	(6.31,8.21)	(22.52,29.85)	(19.90,27.33)	(19.06,27.68	(20.46,27.05	(19.50,26.82)
L.L	-158.46	-71.99	-66.72	-96.20	-87.09	-149.69	-62.10	-56.23	-98.34	-85.11
AIC	332.93	159.96	149.44	208.41	190.19	315.39	140.22	128.45	212.68	186.22
Obs	129	61	50	81	67	129	61	50	81	67

*, **, and *** represent respectively 0.1, 0.05, and 0.01 significance levels; Std error is the standard error; L.L. is the log-likelihood; AIC is the Akaike Information Criteria; Obs are the number of observation in the estimation.

3.3. Within and inter-gender and race comparisons

Figure 2 presents the median WTP for a hemp-based T-shirt for specific groups. We examined the median WTP for each gender and ethnic group when they heard the voice of a researcher from their group or from another gender and or ethnic group.

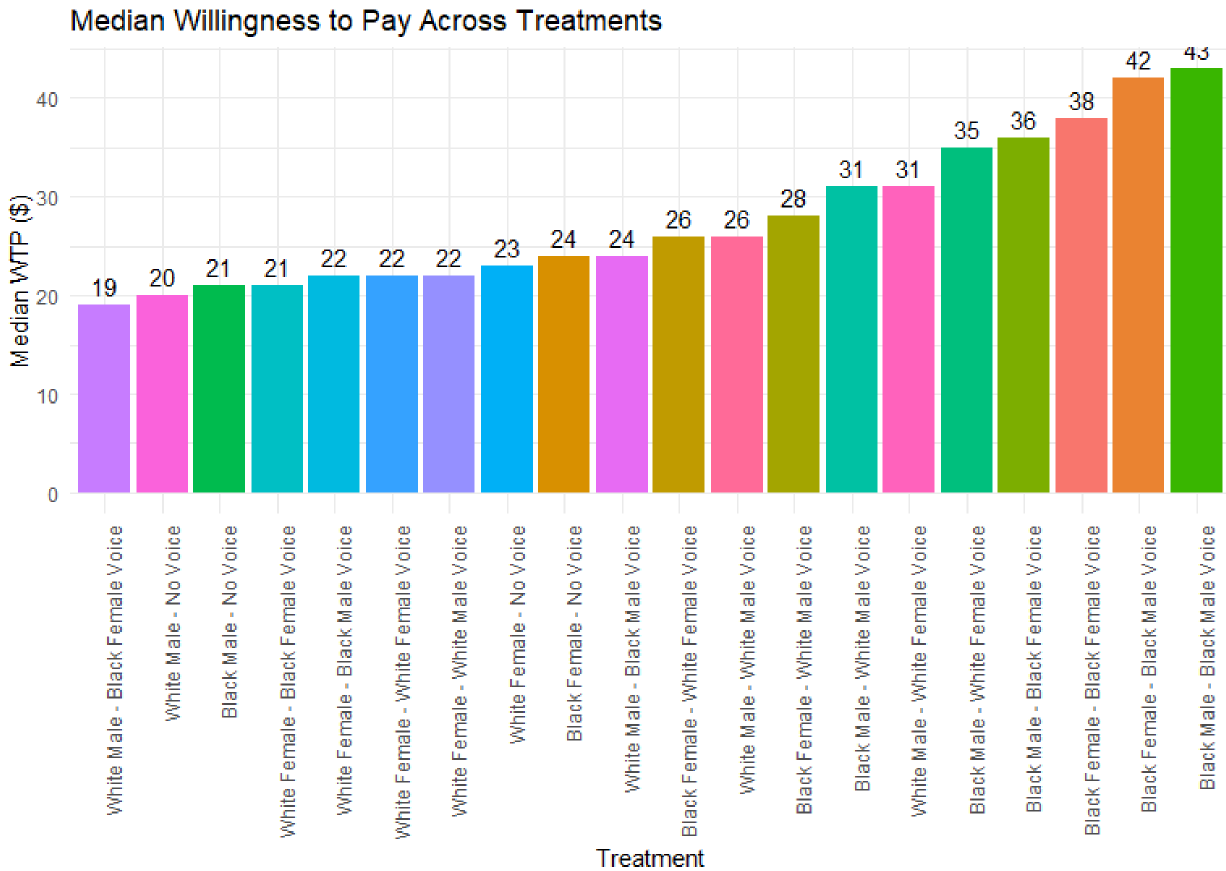


Figure 2: Impact of own and cross-gender and ethnic on WTP for hemp-based T-shirt.

The result shows that the highest WTPs are achieved when black male respondents hear the voice of a black male researcher followed by black females hearing a black male's researcher voice, black females hearing a black female researcher's voice, and black males hearing a black female researcher's voice. These estimations are almost two-fold higher than the WTP when these groups hear no voice.

Own and cross-gender WTPs for whites are also relatively higher compared to the WTPs of this ethnic group when they hear no researcher's voice. The lowest WTP is achieved when voice when white males hear the voice of a black female researcher.

We use the Kruskal-Wallis's test to determine if there are significant differences in median willingness to pay across the different treatments and perform pairwise comparisons with Dunn's post-hoc test. Figure 3 plots the results of the statistical differences between the median WTPs.

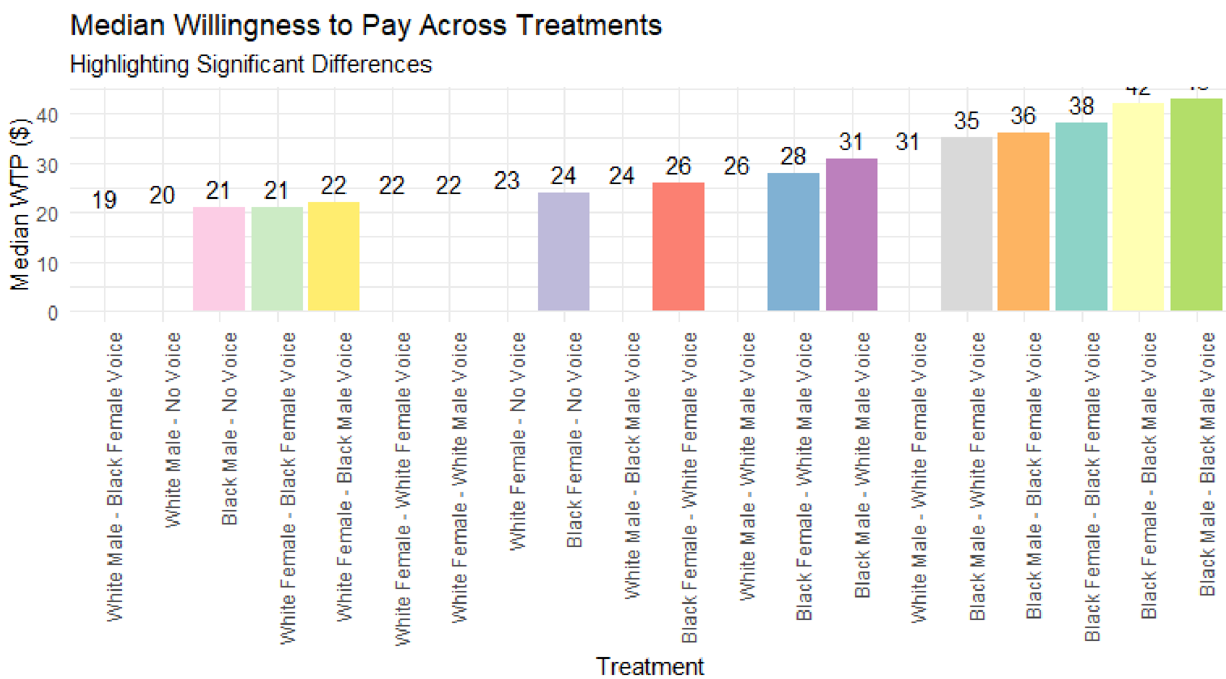


Figure 3: Significant median WTP across treatments.

When examining the significant differences between the WTPs of each group across the different treatments, the majority of the results hold.

Note: *an extended version of this paper will be presented during the 2024 AAEA annual meeting.*

Selected Works Cited

Aikoh, T., Y. Shoji, T. Tsuge, S. Shibasaki, and K. Yamamoto. 2020. "Application of the double-bounded dichotomous choice model to the estimation of crowding acceptability in natural recreation areas." *Journal of Outdoor Recreation and Tourism* 32:100195.

Aizaki, H., T. Nakatani, K. Sato, and J. Fogarty. 2022. "R package DCchoice for dichotomous choice contingent valuation: a contribution to open scientific software and its impact." *Japanese Journal of Statistics and Data Science* 5(2):871–884.

Brago, P., G. Danso-Abbeam, A.A. Ogundeji, J. Abankwa, D.S. Ehiakpor, J.A. Awuni, A.P. Aduku, and G. Dagunga. 2022. "Consumers' willingness to pay for quail products in Tamale metropolis, Ghana." *Journal of Agriculture and Food Research* 10:100445.

Fogarty J. and Aizaki H. (2018) An Illustrative Example of Contingent Valuation. NMVR Project Team, Non-Market Valuation with R. URL http://lab.agr.hokudai.ac.jp/nmvr/ (accessed on 04/04/2024).

Hanemann, M., J. Loomis, and B. Kanninen. 1991. "Statistical efficiency of double-bounded dichotomous choice contingent valuation." *American journal of agricultural economics* 73(4):1255–1263.

Hanemann, W.M. 1984. "Welfare evaluations in contingent valuation experiments with discrete responses." *American journal of agricultural economics* 66(3):332–341.

Hoehn, J.P., and A. Randall. 1989. "Too many proposals pass the benefit cost test." *The American Economic Review* 79(3):544–551.

Luckstead, J., H.A. Snell, L.L. Nalley, R.M. Nayga Jr, and J. 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.