Journal of Choice Modelling 44 (2022) 100369

Contents lists available at ScienceDirect

Journal of Choice Modelling

journal homepage: www.elsevier.com/locate/jocm

Matching reality: A basket and expenditure based choice experiment with sensory preferences☆

Clinton L. Neill [a,b,*], Jacob Lahne [c]

[a] Department of Population Medicine and Diagnostic Sciences, Cornell University, United States of America
[b] Dyson School of Applied Economics and Management, Cornell University, United States of America
[c] Department of Food Science and Technology, Virginia Tech, United States of America

ARTICLE INFO

JEL classification:
C90
Q13
D12

Keywords:
Consumer choice
Sensory
Basket and expenditure based choice experiment
Vegetables

ABSTRACT

This article introduces a basket and expenditure based choice experiment design to elicit consumer preferences for multiple products. This design is utilized to imitate a more realistic shopping scenario for consumers when choosing among many different products simultaneously. This approach allows participants to choose both multiple items, in this case vegetables, and related quantities/expenditures to place in a basket of goods. We provide an application of the experimental design to a vegetable choice experiment. This is done in conjunction with a sensory experiment to provide a contextual component to the experiment and econometric model. This type of experiment lends itself to the use of the Multiple Discrete-Continuous Extreme Value (MDCEV) class of models. More specifically, we use the extended version of the MDCEV model proposed by Palma and Hess (2020) that relaxes the need for a budget while also accounting for substitution and complementarity among products. We find that the proposed design and class of econometric methods present a flexible way to analyze consumer choice when the desire is to elicit preferences for a basket of goods rather than simple discrete alternatives or attributes.

1. Introduction

The concept of utility has long been the framework from which economists have estimated demand for goods and services. In its simplest economic form, one observes utility as the quantities and prices of the goods placed in a basket. However, much of the literature on consumer choice has centered on the attributes of two similar goods. While discrete trade-offs are an important facet of understanding consumer behavior, the traditional experimental designs and econometric approaches are narrow views of a much larger choice problem. In reality, consumers choose several items simultaneously based on their experience with each good, their preference for various attributes, and prices.

Traditional demand systems have been used to estimate complements and substitutes for inter-related products. However, traditional models have relied on aggregate data to examine these relationships and can suffer from endogeneity issues (Wang and Bessler, 2006). Discrete choice experiments (DCEs) offer relief from the endogeneity as prices are exogenous, but are less useful in estimating relative values across different categories of goods. Caputo and Lusk (2022) introduced the basket-based choice

☆ **Acknowledgments**: We want to thank Dr. Alex White for allowing us to recruit from his class of students, Alex Hessler for providing the fresh vegetables, and Katherine Phetxumphou for her help in coordinating the sensory studies. We would also like to thank the attendees of the 2020 Emerging Scholar presentation at the Southern Agricultural Economic Association Annual Meetings in Louisville, KY, the attendees of the Virginia Tech Food Science and Technology Seminar series, and two anonymous reviewers for their insightful comments and suggestions.

* Correspondence to: Department of Population Medicine and Diagnostic Sciences, Cornell University, Ithaca, NY 14853, United States of America.
E-mail address: cln64@cornell.edu (C.L. Neill).

https://doi.org/10.1016/j.jocm.2022.100369
Received 10 August 2021; Received in revised form 21 June 2022; Accepted 22 June 2022
Available online 24 June 2022
1755-5345/

experiment (BBCE) that combines the advantages of both approaches (demand systems and DCEs) and allows for a more realistic shopping analog. Within the BBCE, participants are able to choose as many alternative options to place in a basket, but they are only able to choose which products to place in the basket. This limits the consumer and researcher to only analyze basket diversity and only look at discrete complementarity/substitution. The BBCE does allows for a more robust set of approaches to analyze demand and can provide more meaningful policy implications. Unfortunately, choices are still discrete/binary which limits the realistic nature of the stated preference experiment. Moreover, demographic and other observable variables are only able to enter the econometric model through interaction terms rather than having explicit impacts on choices.

A similar type of choice experiment is the open-ended choice experiment (OECE) (Corrigan et al., 2009; Maynard et al., 2004; Pilon, 1998; Gabor et al., 1970). These experiments allow participants to choose any non-negative, discrete quantities (counts) of one to two options. Specifically, participants typically have two options in which they choose the desired quantity of each product. The products are often similar in nature and the goal is to look at the trade off in quantities between such goods. While some studies allowed for price variation and budget constraints, consumers are inherently forced into substitution patterns due to limited, within-product basket measures. The nature of the data (e.g. counts), again, limits the econometric analysis. In order to simulate the basket shopping experience an experimental design must allow for the multiple discrete and continuous choices inherent in the buying process.

This study presents a choice experiment combining aspects of the BBCE and OECE that allows for the discrete and continuous aspects of consumer decision making — the Basket and Expenditure Based Choice Experiment (BEBCE) design. We follow the work of Palma and Hess (2020) and extend the multiple discrete-continuous extreme value (MDCEV) model of Bhat (2005) to allow for substitution and complementarity without the need for a budget. We apply the experimental design and extended MDCEV model to a consumer choice experiment with a classroom of students on vegetable choice in conjunction with a sensory experiment. This approach allows us to directly estimate the effect of sensory experiences and demographic characteristics. We find that satiation effects increase when participants have recently experienced (tasted) the products. In addition, most vegetables in our experiment are complements or independent of each other rather than substitutes.

As such, this article contributes to the literature by introducing a flexible basket based choice experiment design and analogous model. In addition, we provide an application to demonstrate the experiment in practice. While the results may not be generalizable due to a small, limited sample the goal is to provide an example for future implementation of larger experiments. The remainder of this article is structured as follows: a conceptual model of the consumer choice problem; the experimental design of the BEBCE and sensory experiments; a discussion of the extended MDCEV model; followed by results, discussion, conclusions, and a discussion of future work that could take advantage of the proposed design and method.

2. Conceptualizing the consumer choice problem

Demand for multiple goods describes many, if not most, consumer choice problems (Bhat, 2008). This is especially true when it concerns food choices. However, much of the food choice literature is based on the random utility models of McFadden (1974) and focus on single discrete choices. Several methods have been developed to address the discrete-continuous nature of choices based on utility theory (see Bonnet and Richards (2016) for a discussion of the various models), but few approaches are available to analyze the multiple discrete-continuous problem inherent in actual consumer choice. Bhat (2005) developed the MDCEV model to address this type of choice scenario and remains consistent with utility theory via the Kuhn–Tucker first order conditions (Wales and Woodland, 1983). In addition, this approach allows for corner solutions that arise from non-consumption of some alternatives (Richards and Mancino, 2014).

For illustrative purposes, imagine a consumer's reservation price for a good is solely determined by the utility they gain from consuming said good. Further, all consumers in the market with a reservation price at or below the market price would purchase the product. The problem becomes increasingly complex with the addition of other goods within the choice set, which creates another issue for consideration — demand for a basket of goods (Ariely and Levav, 2000). A consumer can choose a non-negative quantity (which includes zero) of each good, and as the number of products increases the potential combination of basket diversity increases in the following fashion:

$$\sum_{r=0}^{K}\binom{K}{r} \tag{1}$$

where K is the size of the total number of goods available to the consumer and r is the number of different goods chosen from the choice set. Thus, the choice problem now becomes a choice between various combinations of goods. The simplistic choice model is clearly more complex, yet capturing all of these issues within econometric models is challenging.

In general, the consumer objective function of utility (u) maximization that represents the idea that a consumer, i, must choose which food products, j, to purchase from a set of available alternatives (Palma and Hess, 2020), such that

$$Max_{x_j} u_0(x_{i0}) + \sum_{j=1}^{J} u_j(x_{ij}) + \sum_{j=1}^{J-1}\sum_{l=j+1}^{J} u_{jl}(x_{ij}, x_{il}) \tag{2}$$

where x_{ij} denotes the level of consumption for consumer i in alternative j, and x_{i0} is an outside or numeraire good. As usual, a consumer is subject to a budget constraint, M, such as

$$M_i = x_{i0}p_{i0} + \sum_{j=1}^{J} x_{ij}p_{ij} \tag{3}$$

where $p_{io} = 1$ for the outside good and p_{ij} denotes the price of each inside good, j.

To account for the multiple-discrete nature of the consumer choice problem, Bhat (2008) proposes a generalized variant of the translated constant elasticity of substitution (CES) utility function. This form has been modified to generalize the utility equation and provides more flexible functional forms such as in Richards et al. (2012) and Richards and Mancino (2014). In our formulation, we follow that of Palma and Hess (2020) where utility of the inside goods are represented as

$$u_j(x_{ij}) = \psi_{ij}\gamma_j \log\left(\frac{x_{ij}}{\gamma_j} + 1\right) \tag{4}$$

where γ_j is the satiation parameter which indicates that when good j is chosen, higher values of γ_j corresponds with increases in consumption. The ψ_{ij} follows that of Bhat (2008) and is the base utility. Since ψ_{ij} represents marginal utility of an alternative at zero consumption, we ensure the parameters are positive by the following formulation:

$$\psi_{ij} = e^{\beta_j z_{ij} + \epsilon_{ij}} \tag{5}$$

where β_j are vectors of parameters representing attribute weights of alternative j, z_{ij} are attributes of the alternatives, and ϵ_{ij} is the random error term. As one can note, the utility function is flexible enough to incorporate product-specific attributes, which leads to the idea that traditional discrete choice experiments could be expanded within this framework. How to design such an experiment will be discussed in a later section, but if the goal is to allow for a more realistic experiment in consumer choice research, this approach is a potential alternative. For our study, this is where the experience/sensory parameters will enter the consumer's utility problem.

The functional form for the utility of the outside good is assumed to be linear. As such, we define the outside good utility function as

$$u_0(x_{i0}) = \psi_{i0} x_{i0} \tag{6}$$

$$\psi_{i0} = e^{\alpha z_{i0}} \tag{7}$$

where ψ_{i0} is the marginal utility of the outside good, α is a vector of parameters representing weights of characteristics, z_{i0}, of the outside good. Constructing utility in this way is a more accurate representation of the consumer choice problem as it incorporates product characteristics, consumer demographics, and the ability to conceptualize how buying multiple different products simultaneously affects utility.

In the traditional MDCEV econometric approach, a budget is assumed (or known in the case of time allocation examples). The difference in the budget and inside good expenditure is used for the outside good expenditure/quantity estimates. With the linear utility function for the outside good we are able to drop the outside good consumption, x_{i0}, and the budget, M, from the final model formulation. Instead, the outside good parameters can be used to estimate other characteristics (e.g. socio-demographics) that would act as global parameters that affect consumption of all inside goods. As Palma and Hess (2020) note, this functional form should be considered an approximation in cases that the expenditure on the inside goods in relatively small compared to the overall budget.[1] In the case of food expenditures in developed countries, the percentage of income devoted to food is relatively small and was about 9.5% for U.S. consumers in 2019 for food-at-home (FAH) and food-away-from-home (FAFH) combined (USDA Economic Research Service, 2020). FAH expenditure was about 4.8%. This suggests that this functional form for the outside good is appropriate for choice experiments surrounding food choice, as in our sample.

The final component of the utility function, $u_{jl}(x_{ij}, x_{il})$, is another addition within the extended utility formulation that expresses the complementarity and substitution effects of the inside goods. Bhat's (2005, 2008) original approach does not account for these interaction effects within the utility function. While several approaches have been developed to address the issue of complementarity and substitution (see Bhat et al. (2015), Vasquez Lavin and Hanemann (2008) and Richards et al. (2015), among others.), Palma and Hess (2020) is the only approach that also relaxes the need for a budget and is a flexible method that can be applied to various types of data. The generalized MDCEV formulation is perhaps the closest alternative to the proposed methods and utility, but still requires a budget assumption (Richards et al., 2015). Because the opportunity to misspecify budgets in food expenditure is so high (not known *a priori*, or because self reported budgets are often noisy) a formulation of the utility model without the need for a budget is ideal. To account for substitution/complementarity effects the functional form for u_{jl} is as follows:

$$u_{jl}(x_{ij}, x_{il}) = \delta_{ij}\left(1 - e^{-\delta_j x_{ij}}\right)\left(1 - e^{-\delta_l x_{il}}\right) \tag{8}$$

where $\delta_{ij} > 0$ denotes that the pair of goods, k and l, are complements. Likewise, $\delta_{ij} < 0$ denotes the pair of goods as substitutes. If $\delta_{ij} = 0$ then the pair of goods are considered independent of one another. Each of the δ_j and δ_l parameters determine the curvature of the effect.

The proposed formulation of utility here, as adapted from Palma and Hess (2020), increases the efficiency of the derivation as there is no need for the budget and accounts for complementarity and substitution effects. This avoids many theoretical issues in previous research. However, until recently, there has been no experimental choice design that matches this type of utility function. Specifically, there are no choice designs that account for both discrete and continuous choices simultaneously. Thus, our desire to develop such a design that is theoretically sound and empirically feasible.

[1] This is corroborated in Palma and Hess (2020) via simulation and through numerous applications.

Table 1
Vegetable Basket and Expenditure Based Choice Experiment price levels ($/lb).

	Low	Medium	High
Peppers	$1.80	$3.40	$5.20
Turnips	$1.00	$2.00	$3.00
Carrots	$0.60	$1.00	$1.60
Tomatoes	$0.60	$1.20	$1.80
Kale	$1.40	$3.00	$4.40
Lettuce	$0.80	$1.60	$2.20

Table 2
Summary statistics from sensory and economic experiments.

Variable	N	Mean/Percent	Std Dev
Female			
Pooled	286	58.00%	0.49
Sensory	107	68.00%	0.47
Non-sensory	179	53.00%	0.49
Household size	286	2.89	1.32
Vegetable consumption	286	2.17	1.06
Never (0)	7	2.45%	
1–2 days a week (1)	82	28.67%	
3–4 days a week (2)	92	32.17%	
4–6 days a week (3)	66	23.08%	
Everyday (4)	39	13.64%	
State resident	286	78.00%	0.42
Weekly food budget	286	$71.63	$45.42
Vegetable budget ($)	286	13.82	10.89
Expenditure within choice questions ($)	286	8.94	8.83

3. The basket and expenditure based choice experiment design

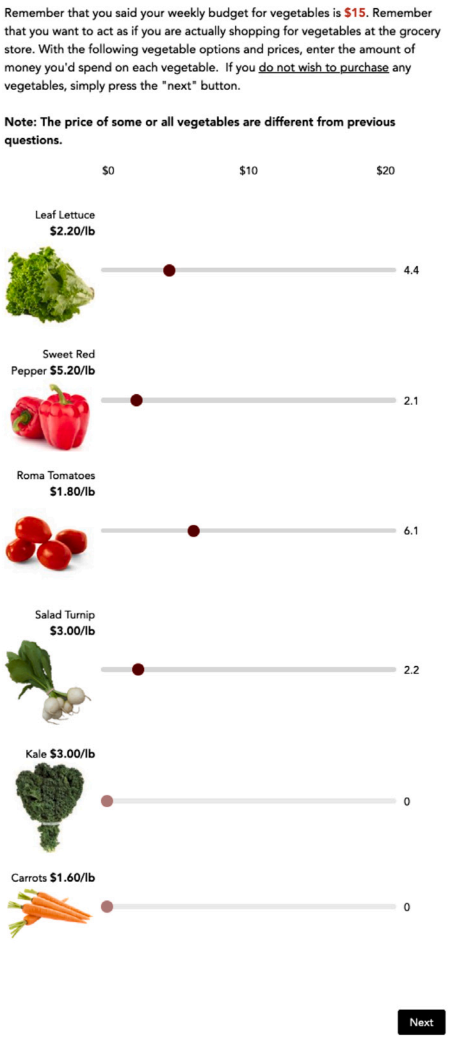
Discrete choice experiments have exploded in popularity over the past two decades, but as noted in recent advancements on experimental design (Caputo and Lusk, 2022), they lack the ability to effectively account for substitution and complementarity among various sets of products. Discrete choice experiments are useful in providing a value to specific attributes, but have less relevance for understanding a consumer's basket of goods. The basket-based and open-ended (quantity) approaches alleviate some of these concerns, yet still have their weaknesses (Caputo and Lusk, 2022; Corrigan et al., 2009). A basket-based discrete choice design only allows for binary choices of alternatives, while an open-ended design is usually limited to two alternatives and is more akin to contingent valuation methods where a control product does not vary in price.

In our experiment we employ a basket-and-expenditure-based choice experiment (BEBCE) that uses the strength of both previously mentioned designs. A BEBCE allows participants to choose any number of alternatives presented (discrete) and the amount of expenditure of their budget to spend on each alternative (continuous). This design is also reminiscent of the consumer expenditure problem of Deaton and Muellbauer (1980). Moreover, this design matches the utility theory proposed in Bhat (2005). Our application of the design was conducted along with a sensory experiment to test the effects of how experience with the vegetable and overall liking affect consumer choice. We recognize that none of the choices are binding and potentially suffer from hypothetical bias. Not all of our sample participated in the sensory experiment, so it was impossible to provide a binding choice to everyone. However, we did incentivize participants to participate and provide realistic responses. The procedures for the BEBCE is as follows:

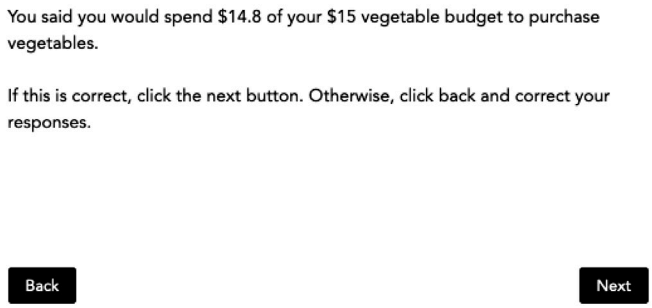
Step 1 Construct the choice questions. As in a typical choice experiment, prices for each vegetable option are varied at three different levels. We use a main-effect fractional factorial design. Our experiment had six vegetables at 3 price levels which resulted in 13 choice questions. The price levels for each vegetable are presented in Table 1. The medium prices chosen for the experiment were based on average prices for each vegetable in the local area. From there, the low and high prices were constructed to represent the range of prices from seasonal variation in fresh produce markets observed locally. Every choice question was identical except for the change in prices. All six vegetable options are available to place in the consumer's hypothetical basket in every choice scenario.

Step 1a Prior to the choice questions, we elicit an individual's weekly food and vegetable budget. The stated vegetable budget is presented in each of the choice questions to serve as a reminder and to suggest some cognitive budget constraint. Again, we do not use this budget constraint in the extended MDCEV model in order to reduce misspecification and noise in the data. The budget is solely used to create a cognitive reminder to treat the experiment as real rather than hypothetical.

Step 2 Within each choice question respondents are asked to determine how much they are willing to expend on each vegetable at a given price level (Fig. 1a). If they do not want to purchase any of the vegetables, students could choose to expend zero dollars across all vegetables. The question is framed to ask participants to purchase the amount of each vegetable they would want to consume for a given week. Note that we elicit expenditure using a slider tool. The slider allowed for continuous levels of expenditure to be selected at a precision of $0.10. As noted later, we translate expenditures to quantities, which preserves a continuous measure.



(a) Vegetable Basket and Expenditure Based Choice Question



(b) Vegetable Basket and Expenditure Based Followup Question

Fig. 1. Example of vegetable basket and expenditure based questions about choice expenditure.

Step 3 As a followup to each choice question, respondents are shown how much they expended relative to their stated weekly food budget. They are then given the opportunity to revise their expenditure choices (Fig. 1b). This budget reminder was only weakly binding as they are not forced to spend less than or equal to their budget (e.g. they could overspend). While not theoretically ideal, this still simulates a realistic shopping scenario for the participant.

Step 4 The experiment is repeated for all 13 choice questions. The choice scenarios are randomized to prevent order effects. In each scenario the participant chooses the amount of their budget to expend on each vegetable and is asked to follow-up question in Step 3 to align with each individual's budget constraint.

3.1. Sample collection

In order to test whether experience attributes affect consumer expenditure we use a large classroom of students to employ a sensory-economic approach to evaluating vegetable preferences. The students are a part of a general education course with students at all levels of undergraduate college education (freshman through senior). Students in the class are first asked to complete a sensory evaluation of six different vegetables: sweet red peppers, carrots, tomatoes, salad turnips, kale, and leaf lettuce. These vegetables were chosen as they could all be eaten raw and provide potential substitution and complementary patterns. All of these vegetables were also in season at the time of the experiment in October 2019 and obtained locally.

The sensory experiment was conducted in the Virginia Tech Sensory Evaluation Laboratory according to standard good practices (Lawless et al., 2010). Participants were seated in individual tasting booths and completed two evaluations of the same, six vegetables. First, participants were presented with the samples one at a time in randomized order, counterbalanced for presentation-order effects. They were instructed to smell and taste each vegetable sample. Participants rated samples for overall flavor and aroma liking on an unstructured, continuous line scale translated to ratings between 0 and 10 in 0.1 unit increments. In-between samples, subjects were given saltines and filtered water as palate cleansers and a 30-s wait time was enforced. Once they had completed the first task, the panelists received all six samples simultaneously and ranked them in order of their overall, holistic liking.

In total, 286 students participated in the BEBCE, with 107 also completing the sensory experiment. This leaves 179 participants who only completed the BEBCE and not the sensory experiment. Each student was incentivized to participate by being offered extra credit for each component completed. Again, the choices were not binding so there is potential for hypothetical bias in our results. An alternative to reducing hypothetical bias for future work would be to conduct the economic choice experiment and randomly choosing one of their choices as binding and the participant would then be presented with a plate of the vegetables chosen in proportion to the amounts chosen.[2]

After the conclusion of the sensory experiment, all students in the class are asked to complete the BEBCE that includes the same 6 vegetable options. As previously mentioned, each student is asked to report their estimated total weekly food budget and the amount they think they spend on vegetables. Their vegetable budget is then weakly binding throughout the BEBCE. This allows for a individually customized choice experiment for every participant. In addition, demographic and location (whether they are in- or out-of-state students) was gathered through the survey. Note that our results cannot be generalized to the larger population. However, we conduct sample diagnostics and present those results in Appendix.

4. Econometric models

In the following derivation, we leave the meaning of x_{ij} to generally represent consumption of goods (e.g. quantity or expenditure), but we specifically use quantities as is standard in the literature. Moreover, a BEBCE design lends itself to a variety of analyses that give a more detailed view of the consumer demand problem. This design allows for a deeper understanding of basket size/diversity, which has largely been ignored in individual consumer demand literature. Thus we also model basket diversity via negative-binomial regressions.

It is worth noting that some studies have incorporated sensory components into econometric analysis (Tozer et al., 2015; Waldrop and McCluskey, 2019). The idea of including sensory/experience variables when consumers make choices is one of the reasons many consumer studies are performed inside grocery stores (Lusk et al., 2001). In addition, beliefs about personal health, healthfulness of certain foods, and other social/environmental preferences are becoming popular additions to many consumer choice studies (Axsen et al., 2013; Lusk et al., 2014; Neill and Williams, 2016; Neill and Holcomb, 2019; Howard et al., 2020). Including such contextual factors in demand models is paramount as consumer behavior toward prices and resulting purchase quantity are not experienced in a vacuum. More importantly, neither are surveys/experiments. As previously mentioned, a subset of subjects participated in a sensory experiment along with the BEBCE to determine whether actual familiarity with the good influences individual preferences and basket size. We examine the heterogeneity of these individuals by comparing pooled and sensory specific models. For ease of notation, this is not distinguished in the below derivations.

4.1. The extended multiple discrete-continuous extreme value model

In this section, we follow the derivation of the extended MDCEV model set forth by Palma and Hess (2020). We begin by deriving the Lagrangian and associated KT conditions:

$$\mathcal{L}(x_i) = u_0(x_{i0}) + \sum_{j=1}^{J} u_j(x_{ij}) + \sum_{j=1}^{J-1}\sum_{l=j+1}^{J} u_{jl}(x_{ij}, x_{il}) - \lambda\left(x_{i0}p_{i0} + \sum_{j=1}^{J} x_{ij}p_{ij} - M_i\right) \tag{9}$$

$$\frac{\partial\mathcal{L}}{\partial x_{i0}} = 0 : \psi_{i0} = \lambda p_0 \tag{10}$$

$$\frac{\partial\mathcal{L}}{\partial x_{ij}} = 0 : \frac{\Psi_{ij}}{\frac{x_{ij}}{\gamma_j}+1} + \delta_j e^{-\delta_j x_{ij}} \sum_{l\neq j} \delta_{jl}\left(1 - e^{-\delta_l x_{il}}\right) \leq \lambda p_{ij} \tag{11}$$

where Eq. (11) will be an equality when alternative j is consumed as the marginal utility of a chosen vegetable at the optimum level of consumption will be λ scaled by its own price, p_{ij}. If vegetable j is not chosen in a particular choice scenario, then the marginal utility is lower than this scaled value. By combining the partial derivatives and replacing ψ_{i0} and ψ_{ij} with Eqs. (7) and (5), respectively, and isolating the random error term we have the following inequality:

$$\epsilon_{ij} \leq -\left(z_{ij}\beta_j - log\left(\frac{x_{ij}}{\gamma_j}+1\right) - log\left((e^{\alpha z_{i0}})\frac{p_{ij}}{p_{i0}} - \delta_j e^{-\delta_j x_{ij}} \sum_{l\neq j}\delta_{jl}\left(1 - e^{-\delta_l x_{il}}\right)\right)\right) \tag{12}$$

[2] We thank the reviewers of this paper for this suggestion on future research design.

where ϵ_{ij} is assumed to be independent and identically distributed via a Gumbel distribution with mean zero and scale σ to be estimated. Thus, the likelihood function is given by (Palma and Hess, 2020):

$$L(x_{ij}) = |Jac|\frac{1}{\sigma^T}\frac{\prod_{j=1}^{T_i} e^{\frac{W_{ij}}{\sigma}}}{\prod_{j=1}^{J} e^{-e^{\frac{W_{ij}}{\sigma}}}} \tag{13}$$

where the consumed alternatives are reordered so that they hold the indexes $j = 1...T_i$ and the non-consumed alternatives hold indexes $j = (T_i + 1)...J$; W_{ij} represents the right side of the inequality in Eq. (12):

$$W_{ij} = z_{ij}\beta_j - log\left(\frac{x_{ij}}{\gamma_j} + 1\right) - log\left((e^{\alpha z_{i0}})\frac{p_{ij}}{p_{i0}} - \delta_j e^{-\delta_j x_{ij}} \sum_{l\neq j} \delta_{jl}\left(1 - e^{-\delta_l x_{il}}\right)\right) \tag{14}$$

and $|Jac|$ is the determinant of the Jacobian of W_{ij} with the elements defined as

$$Jac_{nn} = \frac{1}{x_n + \gamma_n} + \frac{\delta_n E_n}{\psi_0 \frac{p_j}{p_0} - E_n} \tag{15}$$

$$Jac_{nm} = \frac{-\delta_{nm}\delta_n\delta_m e^{-\delta_n x_n} e^{-\delta_m x_m}}{\psi_0 \frac{p_j}{p_0} - E_n} \tag{16}$$

$$E_n = \delta_n e^{-\delta_n x_n} \sum_{l\neq n} \delta_{nl}(1 - e^{-\delta_l x_l}) \tag{17}$$

where n indexes row, m indexes columns, and the i index has been dropped for clarity. Once the model is estimated, we forecast the predication's based on solving the optimization problem via an iterative approach. We forgo presenting the procedure here for brevity and it is well documented in Palma and Hess (2020).

There are several aspects of the model that are important to note as it relates to the identification of the parameters. First, all consumers should be included in the estimation — even those that do not consume any alternatives. These consumers provide valuable information to the value of the utility parameter for outside goods. Additionally, there should be no intercept term in the outside good utility parameter (ψ_{i0}). Third, in order to aid in estimation of the model, we follow the advice of Palma and Hess (2020) and fix the δ_0 parameter be equal to $-\frac{log(1-\sqrt{\rho})}{q}$, where ρ is a percentage and q is the quantile associated with ρ in the consumption vector of all inside goods in the entire dataset, excluding zeros. The purpose of fixing this parameter is that as the value of δ_0 gets smaller it is harder to identify in addition to identifying δ_{jl}, δ_j, and δ_l. We set $\rho = 0.95$ as suggested (Palma and Hess, 2020). Finally, the delta parameters could also be capturing income effects since the implicit budget is restrictive and can lead to increased demand for an inexpensive product while decreasing the demand for an expensive one. However, this issue is partially alleviated if expenditure on inside goods is small captured to the consumer's overall budget (Palma and Hess, 2020). Since we are examining vegetable expenditures, we believe that income effects are small in this case. However, we test the issue by including the total food budget in the outside good equation and find no significant changes in the δ_{jl} parameters presented below.

4.2. Welfare analysis

To further evaluate the results from the BEBCE, elasticities can be constructed via (Palma and Hess, 2020) forecasting algorithm. In order to forecast the quantities of each variable given model parameters we must use the previously defined Lagrangian and KT conditions. As such, we obtain

$$x_{ij} = h(x_{ij}) = \gamma_j \left(\frac{\psi_{ij}}{\psi_{i0}\frac{p_{ij}}{p_{i0}} - E_n}\right) \tag{18}$$

where E_n is defined in Eq. (17) and depends on all values of x_n. Eq. (18) is a fixed point problem and as such at least one solution to the problem exists over the interval $[0, \frac{B_n}{p_{ij}}]$. However, we cannot ensure that the solution is unique, so Palma and Hess (2020) suggest an iterative approach which we do not present here. Given that we set the price in the budget equations, we are able to re-estimate the model under various price points and obtain the forecasted values. From these point estimates we can then calculate the own and cross-price elasticities of each vegetable and associated pair.

4.3. Modeling basket diversity with contextual parameters

One of the advantages of a basket based experiment is the ability to better understand consumer desire for diversity in their basket/choice set. Analyzing food basket diversity can assist in marketing, nutritional choices, and macroeconomic long-run growth and development — though each of these issues address different interests/questions (Lancaster, 1990; Thiele and Weiss, 2003). In this experiment, basket diversity is a concern for reasons of market segmentation and understanding health-related choices. Basket diversity in terms of vegetable choice may provide a better look into which consumers prefer a more diverse basket, and may also give insight into food substitution. For each consumer choice question, the number of different options placed in the basket are analyzed.

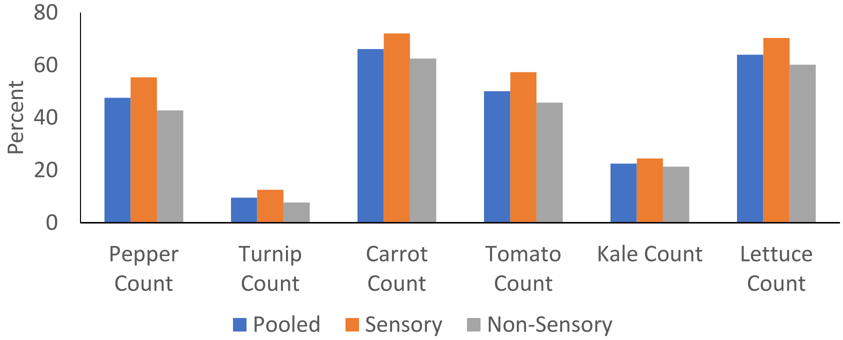


Fig. 2. Histogram of how often each vegetable was chosen across all choice questions..

Since there are six vegetable options to choose from, a consumer could choose $n = 0, 1, \ldots, 6$ total options to place in their cart. This type of dependent variable takes on a non-negative interval and must be modeled using a Poisson or negative binomial regression. In this study, we use a negative binomial to model the probability of choosing a particular number of vegetables. The probability function is represented as a Poisson-gamma mixture distribution

$$h = P(Y = y_i|\mu_i, \tau) = \frac{\Gamma(y_i + \tau^{-1})}{\Gamma(y_i + 1)\Gamma(\tau^{-1})}\left(\frac{\tau^{-1}}{\tau^{-1} + \mu_i}\right)^{\tau^{-1}}\left(\frac{\mu_i}{\tau^{-1} + \mu_i}\right)^{y_i} \tag{19}$$

$$\mu_i = t_i\mu \tag{20}$$

$$\tau = \frac{1}{\upsilon} \tag{21}$$

where y_i is the number of different vegetable options placed in the basket, μ_i is the mean incidence rate of y per unit of exposure in choice question t, and υ is a scale parameter (Cameron and Trivedi, 2005). Since we actually have a censoring issue from the fact that the largest number of vegetables that can be chosen is six, a right-truncated negative binomial is needed. The probability distribution is then represented as (Gurmu and Trivedi, 1992)

$$P(Y = y_i|Y \leq c) = \frac{h(y_i, \mu_i, \tau)}{H(c, \mu_i, \tau)} \tag{22}$$

where c is the censoring value, which in the case of this experiment is six. The regression coefficients are used to determine μ as follows:

$$log(\mu_i) = \eta_0 + \sum_{i=1}^{6}\gamma_i x_i + \sum_{k=1}^{s}\delta_k D_k \tag{23}$$

where x_i and D are defined as demographic and sensory values, respectively.

5. Results

The subjects of this study are entirely undergraduates simultaneously taking a large general education course. As noted in Table 2, the majority of students that completed the BEBCE and sensory experiments are female. Most participants live with one to two other people for an average household size of almost three total. The most common level of vegetable consumption is 3–4 days a week, with a very small percentage of participants claiming to never consume vegetables during a week. Unsurprisingly, a large majority of the students are in-state residents at the university in which they attended. The average stated vegetable budget for a week is \$13.82, though the average expenditure for each choice question is \$8.94. Eliciting a weekly vegetable budget is essential to give participants their own frame of reference within the choice questions. Since the vegetable options are limited in the experiment, this phenomenon of under-spending is likely due to preferences for vegetables outside of the presented choice sets. Since the experiment also has a sensory component, this was to be expected as we could not offer all possible options in a tractable way. Under-spending, as opposed to over-spending, also indicates that the follow-up questions likely helped the majority of participants remain below their specified budgets, though there is no way to know for certain.

Within the BEBCE, participants choose multiple options within the choice question which allows us to observe co-occurrence between the vegetable options. As noted in Caputo and Lusk (2022), the joint product selection allows for a better understanding of complementarity. In Table 3 the most common joint product selection is carrot–lettuce, followed by tomato–lettuce, carrot–tomato and carrot–pepper, respectively. The least common combination is turnip–kale followed by pepper–turnip, turnip–carrot and turnip–tomato, respectively. These joint product selections show the potential substitution and complementary patterns. For example, carrots are a likely complement to many of the other vegetables, but a likely substitute for turnips which are also a root vegetable. To better understand how joint product selection compares to overall selection of each vegetable, Fig. 2 represents how often each vegetable is placed in a participant's basket across all choice questions. Carrots and lettuce are the most popular individual choices, and turnips and kale are the least popular choices.

Table 3
Co-occurrence of vegetable choices across all survey participants (excluding no buy/purchase choices).

	Pepper	Turnip	Carrot	Tomato	Kale	Lettuce
Pepper	.					
Turnip	6.6% (7.7%) [5.5%]	.				
Carrot	35.0% (40.6%) [29.1%]	7.9% (9.1%) [6.6%]	.			
Tomato	29.9% (34.7%) [24.9%]	7.5% (8.7%) [6.2%]	37.3% (43.2%) [31.0%]	.		
Kale	14.7% (17.0%) [12.2%]	5.7% (6.6%) [4.8%]	17.0% (19.7%) [14.2%]	14.7% (17.0%) [12.2%]	.	
Lettuce	33.9% (39.3%) [28.2%]	7.5% (8.7%) [6.3%]	48.1% (55.9%) [40.1%]	37.9% (44.0%) [31.6%]	16.7% (19.4%) [13.9%]	.

Note: Top number is overall probability of joint choice; Number in ( ) is probability of buying row product conditional on buying the column product; Number in [ ] is probability of buying column product conditional on buying the row product.

5.1. Extended MDCEV results

As previously mentioned, the extended MDCEV model was run for the pooled sample and the various sub-samples — those that participated in the sensory experiment and those that did not. Results are presented in Table 4. Liking scores for each of the vegetables was included in the ψ_j parameter estimates for sensory participants. Since we have non-sensory participants in the pooled model we mean center the liking variable. Specifically, for those that completed the sensory portion of the experiment we subtract the average liking score for each variable from each individual liking score for all vegetables. We then code the non-sensory participants liking scores to be zero. The demographic variables included in the linear outside good (ψ_0) are gender, household size, age, and whether the student participant was considered an in-state resident at the university.

Across all three models, the demographic characteristics were not significant. This is likely due to the fact that the sample is college students and has less variation than a national sample. However, to provide an understanding of interpretation, a positive α value like Household Size in the pooled and non-sensory models indicates that larger households consume more outside goods than inside. Similarly, an $\alpha < 0$ like female participants indicates that women consume more of the inside vegetables and less of the outside goods.

The base utility parameters, β_j, all have the same sign except for β_{Carrot} and β_{Tomato} in the sensory model. Since all of the base utility parameters are interacted with liking scores, $\beta_{j,Liking}$ one cannot directly say which vegetable is most preferred. However, the relative magnitudes across the different sub-samples are consistent. Interestingly, the β_{Turnip} and β_{Tomato} for the sensory participants along with the β_{Carrot} parameter for non-sensory and pooled participants was not significant. While there is some endogeneity in the fact that we presented the participants in the sensory trial with quality carrots (e.g. not spoiled or rotten), the act of tasting can alleviate some of the hypothetical nature of choice experiments. As expected, all of the liking variables contribute to positive gains in base utility. Only kale liking scores had no statistically significant effect in the pooled and sensory models. One thing to note about including sensory scores in the base utility parameters is a potential for endogeneity as they are most likely correlated with the error term. While we do not attempt to address this issue in this manuscript, as it is outside the scope of the paper, one potential solution would be to use a hybrid model similar to that of a hybrid-choice model. This would require two indicators of liking for each alternative and is something to consider for future research.

The satiation parameters, γ_j, have interesting variation in relative ranking across the samples. Note that higher values of γ_j indicate that when a vegetable is placed in the basket, a higher amount is consumed. In the pooled model, tomatoes have the highest level of satiation and peppers have the least. For the sensory participants, carrots have the highest satiation levels with tomatoes being the second highest. Otherwise, all other vegetables maintain the same relative ranking of satiation. Perhaps not surprising, sensory participants indicated lower relative magnitudes in satiation for almost all vegetables, except carrots. This may be an indicator of hypothetical bias among the non-sensory participants.

The δ_{ij} parameters for the pooled model indicate that all vegetable pairs are either complements or independent of one another based on statistical significance. In the design of the experiment, our intent was to induce some substitution patterns such as lettuce and kale both being leafy greens. However, this was not the case even when we added a budget variable to the outside good to capture income effects.[3] Within the pooled model peppers–tomatoes, peppers–kale, peppers–lettuce, turnips–tomatoes, turnips–kale,

[3] The coefficient of the budget variable was small (<0.0001), statistically insignificant, and did not change the complementarity patterns when included.

Table 4
Extended MDCEV model with substitution and complementarity effect estimates by sample group.

Variable	Pooled		Sensory		Non-sensory	
	Estimate	Robust t-ratio	Estimate	Robust t-ratio	Estimate	Robust t-ratio
α_{Female}	−0.008	−0.46	0.000	0.01	−0.008	−0.41
α_{HHSize}	0.005	0.80	−0.003	−0.38	0.011	1.35
$\alpha_{AgeCategory}$	0.003	0.44	0.006	0.65	0.002	0.29
$\alpha_{StateResident}$	−0.006	−0.24	0.029	0.68	−0.022	−0.85
γ_{Pepper}	5.126*	4.73	5.712*	2.08	5.742*	3.42
γ_{Turnip}	20.189*	4.40	15.131*	2.20	22.129*	3.69
γ_{Carrot}	14.204*	4.74	16.778*	2.00	15.613*	3.32
γ_{Tomato}	17.516*	4.16	13.347	1.94	21.708*	3.13
γ_{Kale}	10.142*	4.97	10.672*	2.00	11.572*	3.51
$\gamma_{Lettuce}$	10.623*	4.47	11.099	1.93	12.659*	3.16
β_{Pepper}	1.136*	29.16	1.068*	13.10	1.150*	27.08
X $\beta_{PepperLiking}$	0.022*	3.40	0.036*	2.25		
β_{Turnip}	0.190*	2.22	0.183	1.09	0.255*	2.41
X $\beta_{TurnipLiking}$	0.034*	3.34	0.047*	2.33		
β_{Carrot}	0.039	0.90	−0.211	−2.49	0.065	1.26
X $\beta_{CarrotLiking}$	0.021*	3.21	0.062*	2.48		
β_{Tomato}	0.088*	2.22	−0.024	−0.34	0.116*	2.66
X $\beta_{TomatoLiking}$	0.028*	4.37	0.057*	2.71		
β_{Kale}	0.803*	14.64	0.752*	6.87	0.844*	13
X $\beta_{KaleLiking}$	0.015	1.77	0.031	1.94		
$\beta_{Lettuce}$	0.484*	11.58	0.410*	5.16	0.484*	10.6
X $\beta_{LettuceLiking}$	0.015*	2.33	0.033*	2.19		
$\delta_{Pepper-Turnip}$	0.013	0.67	−0.011	−0.38	0.026	1.24
$\delta_{Pepper-Carrot}$	0.012	0.72	0.025	1.38	0.007	0.37
$\delta_{Pepper-Tomato}$	0.041*	2.55	0.023	0.97	0.052*	2.53
$\delta_{Pepper-Kale}$	0.146*	4.22	0.156*	3.37	0.129*	2.59
$\delta_{Pepper-Lettuce}$	0.059*	2.60	0.060*	2.13	0.055	1.86
$\delta_{Turnip-Carrot}$	0.029	1.14	0.007	0.19	0.031	1.06
$\delta_{Turnip-Tomato}$	0.065*	2.38	0.073	1.34	0.039	1.2
$\delta_{Turnip-Kale}$	0.173*	5.62	0.201*	3.07	0.145*	4.1
$\delta_{Turnip-Lettuce}$	0.016	0.57	0.039	1.02	−0.006	−0.16
$\delta_{Carrot-Tomato}$	−0.009	−0.33	0.031	0.84	−0.026	−0.92
$\delta_{Carrot-Kale}$	0.007	0.29	0.032	0.71	−0.005	−0.2
$\delta_{Carrot-Lettuce}$	0.108*	4.46	0.145*	3.35	0.096*	3.24
$\delta_{Tomato-Kale}$	0.021	0.90	0.039	1.10	0.009	0.3
$\delta_{Tomato-Lettuce}$	0.076*	3.24	0.055	1.75	0.082*	2.66
$\delta_{Kale-Lettuce}$	0.034	1.09	0.009	0.20	0.044	1.12
σ	0.289*	6.46	0.255*	2.94	0.270*	4.42
d0	0.401		0.418		0.384	
Log Likelihood	−26,827.67		−9800.50		−16,717.38	
Observations	272		100		172	
Choice Questions	3197		1167		2030	

Note:
*Denotes statistical significance at 0.05 level or lower.

carrots–lettuce, and tomatoes–lettuce were complements. For the sensory participants peppers–kale, peppers–lettuce, turnips–kale, and carrots–lettuce were complements. In the non-sensory model peppers–tomatoes, peppers–kale, turnips–kale, carrots–lettuce, and tomatoes–lettuce were found to be complements. The variation across sub-samples indicates that fully hypothetical scenarios with no tasting/experience can influence the complementarity/substitution patterns of goods. Even though $\delta_{Kale-Lettuce}$ in the sensory model was not significant, it did maintain the expected negative sign for substitutes — though all we can say is that they are independent goods. Similarly, tomatoes and peppers are independent (not significant) in the sensory model but indicated as complements in the non-sensory model. This suggests that hypothetical choice experiments need to be wary of perceived substitution and complementarity. In addition, the flexibility of the extended MDCEV allows for researchers to examine these effects and could be an alternative to examining other basket based choices.

5.2. Welfare results

From the pooled model we use the forecasting procedure in Palma and Hess (2020) to predict the number of times each vegetable is placed within a basket of goods across all of the choice questions. From these predictions, own and cross price elasticities can be calculated by changing the prices used in the budget equation. To demonstrate, we predict the quantities for the most and least commonly chosen vegetables – carrots and turnips, respectively – and present the cross-price elasticity for the most common pair of goods chosen, which are indicated as complements from the resulting model — carrots and lettuce. As seen in Fig. 4a, we find is that carrots have an own price elasticity of about −1.259 using the most and least expensive prices. For turnips (Fig. 4b), we find

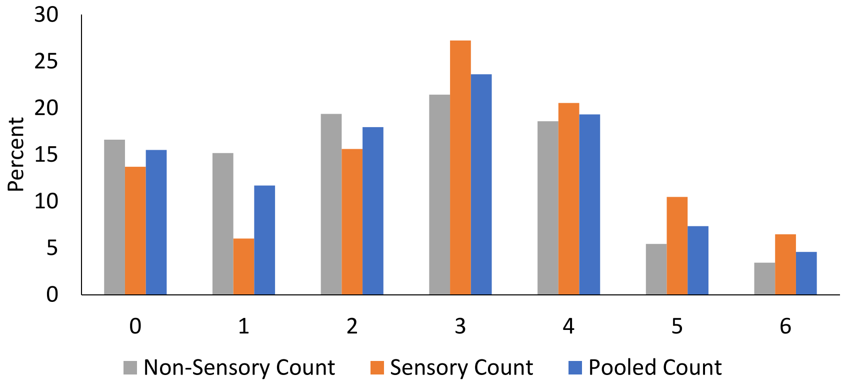


Fig. 3. Histogram of the number of different vegetables chosen across all choice questions.

an own price elasticity of −7.090. For the cross-price elasticity we examine the effect of lettuce price changes on the number of times that carrots are placed in a basket (Fig. 4c). As expected, we calculate a negative cross-price elasticity of about −0.238. Such results and further welfare analysis leads further credence to having a flexible design that captures the reality of consumer shopping baskets.

5.3. Basket diversity results

As noted before, experimentally modeling basket diversity allows for many insights. Moreover, the idea of basket diversity also allows for insights into how rigid a person's food preferences are when sensory/experience variables are added. As before, we analyze three model specifications based on sample composition/aggregation. Fig. 3 represents the varying basket diversity/size from every choice question. The most common basket size for sensory and non-sensory participants was three vegetables. Interestingly, a larger percentage of those who participated in the sensory experiment selected three or more vegetables than those who did not perform the sensory evaluation. This suggest that the act of sensory evaluation does directly influence basket diversity.

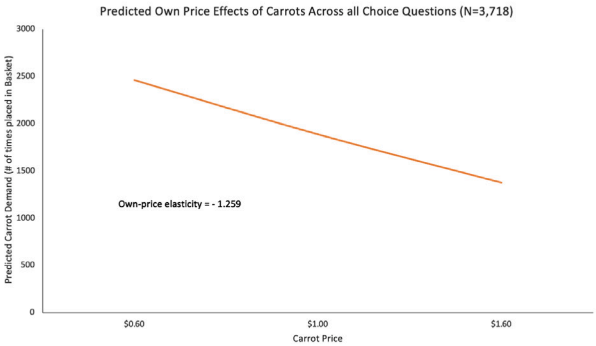
When examining the parameter estimates within each sample specification (Table 5), the parameters are robust except for weekly vegetable consumption and various age fixed effects in the sensory and non-sensory samples. Across samples there are distinct heterogeneous effects. For example, females and household size have no significant effect on basket diversity for sensory participants. Household size, however, has a negative and significant effect for the non-sensory participants. Also, the age fixed effects for the sensory sample are much smaller than the non-sensory sample.[4] As for the liking variables, the sensory model shows turnip, tomato, and kale liking have significant effects on increasing diversity. In the pooled model, increasing in carrot liking decreases diversity. These values are the least robust to sample aggregation and should be interpreted with caution.

6. Conclusions

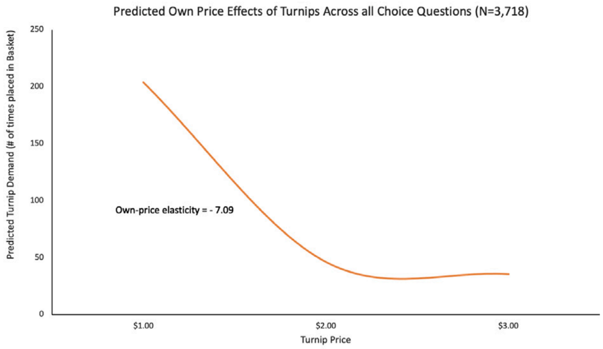
Many consumer choices are predominately characterized by an individual's preferences and experiences for many different goods. This study adds to previous literature by introducing a choice design that allows for a more realistic shopping experience, proposes a class of econometric models to analyze the data, and demonstrates the need to better understand consumer demand from multiple perspectives. This type of experiment would be applied to better understand the continuous variations in a goods attributes or a broader collection of goods across many different categories. The flexibility of the design and the ability to move beyond discrete choices is desirable across many choice problems.

Our proposed design combines aspects of the BBCE and OECE to elicit choices under price variation, quantities, expenditures, and product specific attributes. While we did not utilize the design to include product attributes (i.e. organic, local, etc.) the design is certainly flexible enough to allow for such modifications. Also, we test this design in a hypothetical (and semi-hypothetical with the sensory experiment) situation. Testing the design in non-hypothetical situations, such as a laboratory setting with real exchange of goods and money, is warranted. Moreover, a direct comparison of the approach is needed to validate many of the features that discrete choice experiments maintain in regards to incentive compatibility, anchoring effects, etc. Our goal with this design is to offer an alternative to study a wider range of problems without the need to purchase scanner data and other forms of secondary data. Note that the application in this study is likely not generalizable as the sample is constructed of a specific age range and of much smaller sample size than what is used in many previous studies. However, that was not the goal of the application, but rather to provide an example of how to use such an experimental design. As such, future research should focus on broader categories of spending or include many more options to provide more precise estimates to generalize the results of our application, if desired. The

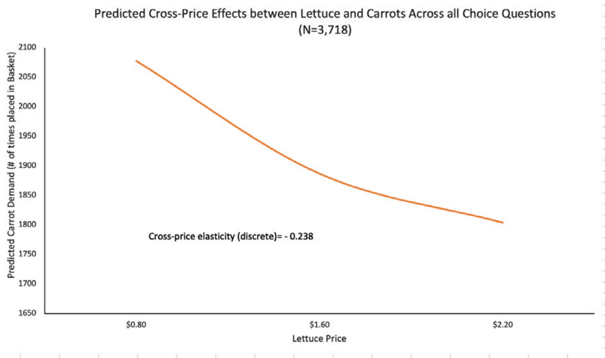
[4] Note that the and the non-sensory sample had no participants who were exactly 24 years of age, so 22 years old was used as the reference. This could partly influence the results.



(a) Own price Demand for Carrots



(b) Own Price Demand for Turnips



(c) Cross-Price Demand for Carrots under Lettuce Prices

Fig. 4. Welfare/forecasting results from the pooled extended MDCEV model.

Table 5
Basket diversity (Negative Binomial) parameter estimates for each model specification.

Parameter	Pooled		Sensory		Non-sensory	
	Estimate	S.E.	Estimate	S.E.	Estimate	S.E.
Intercept	0.91*	0.18	0.77*	0.2	1.17*	0.09
Female	−0.02	0.03	−0.03	0.05	−0.01	0.03
Household size	−0.05	0.01	−0.01	0.02	−0.06*	0.01
Vegetable consumption	0.11*	0.01	0.03	0.02	0.17*	0.02
Age = 18	0.08	0.17	0.16	0.19	−0.28*	0.07
Age = 19	0.08	0.17	0.27	0.19	−0.39*	0.07
Age = 20	−0.07	0.17	0.16	0.19	−0.47*	0.07
Age = 21	0.01	0.17	0.12	0.19	−0.38*	0.07
Age = 22[b]	0.27	0.18	0.30	0.20	.	.
Age = 24[a]	.	.	.	.	.	.
State Resident	−0.07*	0.03	−0.09	0.06	−0.06	0.04
Pepper Liking	0.01	0.01	0.01	0.01		
Turnip Liking	0.01	0.01	0.03*	0.01		
Carrot Liking	−0.02*	0.01	−0.01	0.01		
Tomato Liking	0.02*	0.01	0.02*	0.01		
Kale Liking	0.01	0.01	0.02*	0.01		
Lettuce Liking	0.01	0.01	0.01	0.01		
Scale	1.20E−12	0.00	4.09E−06	0.02	2.90E−05	0.00
−2 Log Likelihood	9438.10		3803.60		5587.80	
AICC	9472.30		3838.20		5608.00	
Observations	272		100		172	

Note: there were no participants aged 23 or over 24 in the sample.
[a]Denotes that the reference age for the pooled and sensory groups is 24.
[b]Denotes that the reference age for the nonsensory group is 22 as there were no participants aged 24.
*Denotes statistical significance at 0.05 level or lower.

Table A.1
Probit model for probability of overspending stated vegetable budget (N = 272).

Parameter	Estimate	S.E.
Intercept	−0.89*	0.14
Female	−0.01*	0.06
HHSize	0.01	0.02
Veg Consumption	−0.05*	0.03
Reference Age = 24		
Age = 18	0.11	0.10
Age = 19	0.13	0.10
Age = 20	0.01	0.10
Age = 21	−0.11	0.11
Age = 22	−0.25*	0.13
State Resident	−0.13	0.08
Sensory Participant	0.07	0.06
AIC	2891.11	
−2 LL	2889.11	

Note:
All values rounded up to nearest 0.01 value.
*Denotes statistical significance at 0.05 level or lower.

only limitation to adding more goods is the size of the design. As with many discrete choice experiments, blocking and alternative designs can alleviate this concern.

We also present an extended version of a common model to examine multiple discrete and continuous data that is produced by such an experimental design. Specifically, the extended MDCEV model of Palma and Hess (2020) allows researchers to examine complementarity and substitution without the need for a budget. Moreover, the model allows for demographic variables and product attributes to enter the model directly. The BEBCE design also allows for additional analysis such as our basket diversity model that takes into account the count data the results from the multiple discrete choice. While there are other econometric approaches that attempt to also account for many of the factors as Palma and Hess (2020), we opted for this model as it imposed fewer restrictions on the parameters.

Within our application, we find that sensory participants often have lower levels of satiation for all vegetables in the study as opposed to the non-sensory participants. We also find that all of the vegetables are considered to be complements or independent of one another rather than substitutes. As previously mentioned, this could be the result of a limited and narrow sample and requires further testing, but proves that forcing alternatives to be substitutes (as a result of model specifications) is not appropriate. The limited sample is also the likely cause of why we do not find any significant effects from the demographic variables. There is also a difference in base utility ranking for vegetables between different sub-samples. Our basket diversity results reveal that age and vegetable consumption levels are important factors as there is much variation in effect sizes across sub-samples. Experience/sensory variables have are much less important in basket diversity as compared to utility/demand, but still has an effect in some cases.

As was initially defined by Thurstone (1927), the concept of utility is embedded in how consumers choose goods and that creates psychological stimuli. This is why sensory experiences and directly measuring raw preferences/liking is critical to better estimation of food demand. In a similar vein of research, Durlauf and Young (2001) and Brock and Durlauf (2003) argue for increased efforts to incorporate sociological ideas into economic models and reasoning. Consumer choices are not made in a vacuum and neither are choice experiments. Designing experiments/surveys to extract the intricacies of consumer choices will aid in better consumer demand estimation.

CRediT authorship contribution statement

Clinton L. Neill: Conceptualization, Methodology, Writing – original draft. **Jacob Lahne:** Conceptualization, Writing – reviewing and editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The authors do not have permission to share data.

Appendix. Appendix of sample diagnostics

Considering the sample is small relative to many of the choice experiments performed in recent literature, it is prudent to determine how representative the student expenditure and budgets are with previous samples. Comparing the stated weekly food budget to findings from Neill and Holcomb (2019) for the same age range from a national survey, we see that the averages are quite similar. This suggests that the sample is, within reason, similar to a random national sample.

Another diagnostic measure we utilize is determining the probability of overspending on stated vegetable budgets. While average expenditure is below the average stated vegetable budget (Table 2), better understanding the probability of overspending will indicate if participants perform consistently with theoretical assumptions. Approximately 85.2% of the basket choices are less than or equal to the stated vegetable budget. In order to determine if there are any commonalities between participants or choice questions that increase the probability of overspending, a probit equation is utilized. We include demographic parameters as explanatory variables. The table of parameters are presented in Table A.1. The results indicate that female participants, vegetable consumption and those aged 22 (as compared to 24 year old participants) all decrease the probability of overspending one's stated budget.

References

Ariely, D., Levav, J., 2000. Sequential choice in group settings: Taking the road less traveled and less enjoyed. J. Consum. Res. 27 (3), 279–290.

Axsen, J., Orlebar, C., Skippon, S., 2013. Social influence and consumer preference formation for pro-environmental technology: The case of a UK workplace electric-vehicle study. Ecol. Econom. 95, 96–107.

Bhat, C.R., 2005. A multiple discrete–continuous extreme value model: formulation and application to discretionary time-use decisions. Transp. Res. B 39 (8), 679–707.

Bhat, C.R., 2008. The multiple discrete-continuous extreme value (MDCEV) model: role of utility function parameters, identification considerations, and model extensions. Transp. Res. B 42 (3), 274–303.

Bhat, C.R., Castro, M., Pinjari, A.R., 2015. Allowing for complementarity and rich substitution patterns in multiple discrete–continuous models. Transp. Res. B 81, 59–77.

Bonnet, C., Richards, T.J., 2016. Models of consumer demand for differentiated products. pp.

Brock, W.A., Durlauf, S.N., 2003. Multinomial choice with social interactions. pp.

Cameron, A.C., Trivedi, P.K., 2005. Microeconometrics: Methods and Applications. Cambridge University Press.

Caputo, V., Lusk, J.L., 2022. The basket-based choice experiment: a method for food demand policy analysis. Food Policy 109, 102252.

Corrigan, J.R., Depositario, D.P.T., Nayga, Jr., R.M., Wu, X., Laude, T.P., 2009. Comparing open-ended choice experiments and experimental auctions: an application to golden rice. Amer. J. Agric. Econ. 91 (3), 837–853.

Deaton, A., Muellbauer, J., 1980. An almost ideal demand system. Amer. Econ. Rev. 70 (3), 312–326.

Durlauf, S.N., Young, H.P., 2001. The new social economics. Soc. Dyn. 1–14.

Gabor, A., Granger, C.W., Sowter, A.P., 1970. Real and hypothetical shop situations in market research. J. Mar. Res. 7 (3), 355–359.

Gurmu, S., Trivedi, P.K., 1992. Overdispersion tests for truncated Poisson regression models. J. Econometrics 54 (1–3), 347–370.

Howard, G., Roe, B.E., Interis, M.G., Martin, J., 2020. Addressing attribute value substitution in discrete choice experiments to avoid unintended consequences. Environ. Resour. Econ. 77 (4), 813–838.

Lancaster, K., 1990. The economics of product variety: A survey. Mark. Sci. 9 (3), 189–206.

Lawless, H.T., Heymann, H., et al., 2010. Sensory Evaluation of Food: Principles and Practices, Vol. 2. Springer.

Lusk, J.L., Fox, J.A., Schroeder, T.C., Mintert, J., Koohmaraie, M., 2001. In-store valuation of steak tenderness. Amer. J. Agric. Econ. 83 (3), 539–550.

Lusk, J.L., Schroeder, T.C., Tonsor, G.T., 2014. Distinguishing beliefs from preferences in food choice. Eur. Rev. Agric. Econ. 41 (4), 627–655.

Maynard, L.J., Hartell, J.G., Meyer, A.L., Hao, J., 2004. An experimental approach to valuing new differentiated products. Agric. Econ. 31 (2–3), 317–325.

McFadden, D., 1974. Conditional logit analysis of qualitative choice behavior. Front. Econom. 105–142.

Neill, C.L., Holcomb, R.B., 2019. Does a food safety label matter? Consumer heterogeneity and fresh produce risk perceptions under the food safety modernization act. Food Policy 85 (C), 7–14.

Neill, C.L., Williams, R.B., 2016. Consumer preference for alternative milk packaging: the case of an inferred environmental attribute. J. Agric. Appl. Econ. 48 (3), 241–256.

Palma, D., Hess, S., 2020. Some adaptations of multiple discrete-continuous extreme value (MDCEV) models for a computationally tractable treatment of complementarity and substitution effects, and reduced influence of budget assumptions. Inst. Transp. Stud. Choice Model. Centre Univ. Leeds pp.

Pilon, T.L., 1998. Extensions to the analysis of choice studies. Sawtooth Softw. Res. Pap. Ser. pp.

Richards, T.J., Gómez, M.I., Pofahl, G., 2012. A multiple-discrete/continuous model of price promotion. J. Retail. 88 (2), 206–225.

Richards, T.J., Mancino, L., 2014. Demand for food-away-from-home: a multiple-discrete–continuous extreme value model. Eur. Rev. Agric. Econ. 41 (1), 111–133.

Richards, T., Yonezawa, K., Winter, S., 2015. Cross-category effects and private labels. Eur. Rev. Agric. Econ. 42 (2), 187–216.

Thiele, S., Weiss, C., 2003. Consumer demand for food diversity: evidence for Germany. Food Policy 28 (2), 99–115.

Thurstone, L.L., 1927. A law of comparative judgment.. Psychol. Rev. 34 (4), 273.

Tozer, P.R., Galinato, S.P., Ross, C.F., Miles, C.A., McCluskey, J.J., 2015. Sensory analysis and willingness to pay for craft cider. J. Wine Econ. 10 (3), 314–328.

USDA Economic Research Service, 2020. Food expenditure series. https://www.ers.usda.gov/data-products/food-expenditure-series/, Accessed: 2021-1-15.

Vasquez Lavin, F., Hanemann, W.M., 2008. Functional forms in discrete/continuous choice models with general corner solution. pp.

Waldrop, M.E., McCluskey, J.J., 2019. Does information about organic status affect consumer sensory liking and willingness to pay for beer? Agribusiness 35 (2), 149–167.

Wales, T.J., Woodland, A.D., 1983. Estimation of consumer demand systems with binding non-negativity constraints. J. Econometrics 21 (3), 263–285.

Wang, Z., Bessler, D.A., 2006. Price and quantity endogeneity in demand analysis: evidence from directed acyclic graphs. Agric. Econ. 34 (1), 87–95.