Vol. 9, No. 3, 2020

BAE

Bio-based and Applied Economics

Editorial

Full Research Articles

Poste Italiane spa - Tassa pagata - Piego di libro
Aut. n. 072/DCB/FI1/VF del 31.03.2005

FIRENZE UNIVERSITY PRESS

ISSN 2280-6180
www.fupress.net

Bio-based and Applied Economics is the official journal of the Italian Association of Agricultural and Applied Economics – AIEAA (www.aieaa.org; E-mail: info@aieaa.org).

Journal Contact: Meri Raggi, via Belle Arti 41, Dept. of Statistics, 40124 Bologna; Email: meri.raggi@unibo.it

Available online at: http://www.fupress.com/bae

Università degli Studi di Firenze
Firenze University Press
via Cittadella, 7 - 50144 Firenze, Italy
www.fupress.com/
E-mail: journals@fupress.com

ISSN 2280-6180 (print)
ISSN 2280-6172 (online)
Direttore Responsabile: Corrado Giacomini
Registrata al n. 5873 in data 10 maggio 2012
del Tribunale di Firenze

Bio-based and Applied Economics 9(3): 223, 2020
DOI: 10.13128/bae-10443

Editorial

Eliciting preferences using stated choice experiments

Mara Thiene[1], Jürgen Meyerhoff[2]

[1] University of Padova, Padova
[2] Technische Universität, Berlin

Stated choice experiments are nowadays widely used in different fields ranging from marketing and transport economics to health and environmental economics. They are a survey-based preference assessment technique that presents respondents with mutually exclusive alternatives described by attributes and their levels and asks them to choose the most preferred of those alternatives. Subsequently, the choices recorded enable estimates of the trade-offs among attribute levels respondents are willing to make, giving insights into their preferences. If one of the attributes is a cost variable, marginal willingness to pay estimates can be calculated, representing peoples' preferences for different attributes on the same monetary unit. Due to the comprehensive information choice experiments can provide such as marginal and non-marginal welfare measures, they have recently become a favored method to evaluate individual preferences. However, stated choice experiments are by no means a method that can be employed by simply following standard recipes from a cookbook. Understanding participants responses to the designed choice tasks presented in surveys and their adequate analysis still requires further research to achieve validity and reliability of the requested results such as welfare estimates.

This special issue wants to contribute to the development of choice experiments by presenting a number of selected papers that present results from methodological investigations as well as from policy-oriented applications of choice experiments in the area of environmental and agricultural economics. The authors are mainly members of a group of academics who have met regularly over the last decade as members of the ENVECHO network, which is a scientific network of researchers using discrete choice modelling in the field of environmental valuation (www.envecho.com).

We wish to thank all authors and reviewers and, of course, the publisher for giving us the opportunity to publish this special issue in the journal Bio-based and Applied Economics.

ISSN 2280-6180 (print)
ISSN 2280-6172 (online)

www.fupress.com/bae

Bio-based and Applied Economics 9(3): 225-240, 2020
DOI: 10.13128/bae-7768

Full Research Article

Investigating determinants of choice and predicting market shares of renewable-based heating systems under alternative policy scenarios

Cristiano Franceschinis*, Mara Thiene

University of Padova, Italy

Abstract. Fostering the uptake of heating technologies based on renewable resources is an important part of the EU energy policy. Yet, despite efforts to promote their diffusion, heating systems based on fossil fuels are still predominant. In order to better tailor energy policies to citizens preferences, it is crucial to collect accurate information on their determinants of heating choices. At this purpose, we adopted a choice experiment and a latent class model to analyze preferences of householders in the Veneto region (North-East Italy) for different heating systems and their key features. We focused on three devices based on biomass and three on fossil fuels, and accounted for technical, economic and environmental characteristics of such systems. Model estimates highlight the presence of substantial preference heterogeneity among the population, which can be partially explained by citizens socio-demographics. We also use model outputs to simulate market shares for heating systems under alternative policy scenarios. Results provide interesting suggestions to inform the design of policies aimed at fostering the adoption of biomass-based heating systems.

Keywords. Ambient heating systems choice, Latent class Model, Market shares, Willingness to pay.

JEL Codes. C01, Q42, Q47.

1. Introduction

Developing a strategy to increase the sustainability of the heating sector is a priority for the European Union, in order to reduce energy imports and dependency and meet the greenhouse gas emission target established under the Paris Agreement. Currently, heating and cooling account for half of the EU energy consumption and 75% of the fuel used in this sector comes from non-renewable resources (European Commission, 2016).

To tackle such issues, the 2030 Climate and Energy Policy Framework adopted by the European Council in 2014 includes three key targets for 2030: i) a 40% reduction of green-

*Corresponding author. E-mail: cristiano.franceschinis@unipd.it

Editor: Meri Raggi.

ISSN 2280-6180 (print)
ISSN 2280-6172 (online)

house gas (GHG) emissions compared to 1990 levels, ii) a 27% share of renewable energy in gross final energy demand and iii) a 27% increase in energy efficiency (European Council, 2014). The targets for renewables and energy efficiency were revised upwards in 2018 at 32% and 32.5% respectively. Member States are obliged to adopt integrated National Climate and Energy Plans (NECPs) for the period 2021-2030 to define how they plan to achieve such goals. Member States submitted their draft plans in 2018 and final plans must be submitted by the end of 2019. Italy, in its draft plan, set the targets of a 33% reduction of GHG emissions compared to 2005 levels and a 30% share of renewable energy on final consumption to be achieved by 2030 (Italian Government, 2018). In 2017, the values for the two targets were 18% and 17% respectively. Thus, as emphasized in the EU Country Report 2019 (European Commission, 2019), further efforts are needed to ensure the achievement of 2030 objectives. Among the specific targets set by the plan, there is an annual increase of 1.3% of renewables share in the sector of residential heating and cooling. To achieve such target – among other measures – the plan aims to promote an active role by citizens on the energy demand market and the uptake of micro-generation technologies based on renewables. As such, installation of renewable based residential heating systems in new buildings and replacement of fossil fuel technologies in existing ones plays a crucial role in the energy system transition. To entice the active participation of citizens, it is important to collect information on their heating preferences, in order to retrieve determinants of heating choices.

Information about heating preferences can be collected via choice experiment, an increasingly popular method for stated preferences analysis. For example, Rommel and Sagebiel (2017) investigated preferences of German homeowners for micro-cogeneration units for residential use. Their results suggest how householders have a strong interest in adopting such technologies, with willingness to pay (WTP) values ranging from 11.000 to 23.000 Euros. Features of micro-cogeneration products, as well as socio-demographics characteristics of houseowners, were found to substantially affect their WTP. Scarpa and Willis (2010) investigated WTP for the adoption of different renewable micro-generation technologies in the UK. Specifically, they focused on solar photovoltaic, micro-wind, solar thermal, heat pumps, biomass boilers and pellet stoves. Their results suggest that householders are willing to adopt such technologies, but for most of them WTP values do not cover capital cost. A similar study was carried out by Su et al. (2018) in Lithuania. Authors found householders to prefer solar energy-based technologies over the other renewable based ones. Claudy et al. (2010) also estimated consumers' WTP for different microgeneration technologies, namely micro wind turbines, wood pellet boilers, solar panels and solar water heaters. The study showed how WTPs vary substantially among different technologies and how consumers attitudes and beliefs about the technologies significantly influence their WTPs. Rouvinen and Matero (2013) focused on preferences towards different types of heating systems (based on fuel used) and examined the role of system features on householders' choices in Finland. Investment cost was identified as the most impactful attribute on householders' decisions, but non-monetary attributes played a significant role as well. Results also provided evidence of preference heterogeneity, partially linked to individuals' characteristics. Similarly, Michelsen and Madlener (2012) analyzed the influence of sensitivity to different heating systems' attributes on homeowners' adoption decision. Their findings suggest that importance attached to different attributes affects technological features choice: for example, people focused on energy saving are more likely to adopt condensing boilers with thermal support, while consumers attaching a strong value to use of renewables prefer pellet-fired boilers. Furthermore, they found socio-demographics and spatial factors to affect preferences. Ruokamo (2016) explored homeowners' attitudes towards innovative hybrid home heating systems, described in terms of fuel used and key features, such as costs, comfort of use and environmental impact. The author found that such technologies are generally well accepted by houseowners and that their preferences are strongly affected by socio-demographic characteristics. Yoon et al. (2015) compared householders' WTP for district heating and individual heating. While they found citizens to be generally willing to pay more for district heating, substantial differences emerged when accounting for preference heterogeneity: consumers with higher income and education were found to prefer district heating, while those more concerned about costs were willing to pay more for the individual one.

In this paper we present the results of a choice experiment aimed at eliciting householders' preferences towards different heating systems in the Veneto region (Italy). Specifically, we focus on six different heating systems, three based on renewables (chip wood, firewood and wood pellet) and three on fossil fuels (methane, oil and LPG).

Veneto is a fairly populated region (almost five million residents) characterized by air pollution mainly related to high road traffic in all main cities and by the presence of large industrial districts in several sectors, such as tanning, cement production and furniture manufacturing. When accounting specifically for carbon dioxide emission, the residential impact is substantial as well, around the 20% of total emissions (ARPAV, 2015).

To decrease the negative impact of residential sector on the production of greenhouse gases, since 2014 the regional authority supports the purchase of biomass-based heating systems by annually allocating financial subsidies (up to €1,600 for stoves and €5,000 for boilers). Such policy, however, seems to only marginally meet the expectations of the population, in terms of fostering the adoption of such technologies. For example, in 2018 only 29 citizens applied for the funding and 25 requests were approved, for a total of around €55.000, out of €500,000 allocated policy budget. In 2019 the number of requests was higher (76, of which 66 approved) but again most of the policy budget was not used (around €120,000 allocated out of a budget of €500,000)[1]. Thus, a better understanding of underlying factors motivating householders to stick with a fossil fuel system or to switch to a renewable one, is of crucial to reach the goals of the energy transformation process in the region.

Building on the evidence provided by the literature on how preferences towards different heating systems are highly heterogeneous and on how socio-demographics play an important role in such variability, we adopt a Latent Class approach and use socio-demographic characteristics of respondents to predict probability to belong to different classes. This approach allows to: i) identify different segments of the population according to socio-demographic characteristics; ii) explore how preferences towards different heating systems and their features vary across segments. We then use the estimates of such model to predict market shares for alternative heating systems within two policy scenarios, considering/based on a reduction of: i) investment costs for biomass fueled heating systems; ii) operating costs for biomass-based technologies. Both simulated scenarios are in line with the policies implemented by the Veneto region, the idea being that our empirical results may become useful

[1] Data retrieved from https://www.regione.veneto.it/web/ambiente-e-territorio/rottamazione-stufe-bando-2019.

to better tailor the features of such policies to the population of the region. Reduction of investment cost is a commonly adopted policy to foster use of renewable based technologies, such as the subsidies provided by the Veneto region. Reduction of operating costs is a possible effect of targeted policies as well (e.g. subsidies on fuel purchase).

The objective of our study is twofold: on one hand to investigate how socio-demographic characteristics influence citizens choice of heating systems, in order to gain insights on the determinants of adoption of renewable based technologies; on the other, to identify which, among a set of possible policy interventions, can be more effective in terms of fostering the diffusion of such technologies among the population.

The remainder of the paper is organized as follows: section 2 describes data collection (sampling procedure, survey design and administration); section 3 formally describes the econometric approach; section 4 reports the results of our study and section 5 draws its conclusions.

2. Data collection and survey

This section reports a succinct description of sampling procedure and survey. Further details can be found in Franceschinis et al. 2016, 2017.

Data were collected with the support of a market research firm via a web-based survey addressed to a sample of householders of the Veneto region. We used a random sample of householders, stratified on the main socio-demographics (age, education, gender, place of residence). A total of 1,557 questionnaires were collected, out of which 1,451 were complete and used for the analysis. The questionnaire was structured in five sections. The first focused on heating system and energy resources currently used by respondents. The following section included the choice experiment, which is described in detail below. The third section included follow-up questions about the choices made in the previous section. The fourth section presented attitudinal questions related to respondents' psychological traits. The last collected socio-demographic information.

The choice experiment involved a hypothetical scenario in which respondents were asked to select the heating system they would adopt if they had to renovate their current one among a set of alternative options. The heating systems presented to respondents were six, three based on biomass (firewood, chip wood and wood pellet) and three on fossil fuels (methane, LPG and oil). Each alternative system was described in terms of six attributes: i) investment cost, ii) investment duration, iii) annual operating cost, iv) CO_2 emissions, v) fine particle emissions and vi) required own work. The respective levels were system-specific and are reported in Table 1. Investment cost is the cost for purchasing and installing the heating system. Possible public incentives were not accounted for in defining the levels of the attribute. Investment duration refers to the lifespan of the heating system, from purchase to dismantling. Operating costs include fuel price, maintenance costs and electricity costs for those systems that need it to work. CO_2 emissions and fine particles emissions refer to the quantity of CO_2 and fine particles released by the fuel combustion processes. To facilitate the evaluation of CO_2 emissions levels, respondents were informed that 1,000 kg of CO_2 corresponds to the emissions from driving 6,000 km in a new generation car. To illustrate fine particles health impacts, respondents were informed that "it has been estimated that if annual fine particle emissions for one house are 2,000 g, then

Table 1. Choice Experiment attributes and levels.

Attributes	Firewood	Chip wood	Wood Pellet	Methane	Oil	LP Gas
Investment cost (€)	9,500; 11,000; 12,500	11,500; 13,000; 14,500	13,000; 15,000; 17,000	4,000 4,800; 5,600	4,500; 5,500; 6,500	4,000; 5,000; 6,000
Investment duration (y)	15; 17; 19	17; 20; 23	16; 19; 22	16; 18; 20	16; 18; 20	14; 17; 20
Operating cost (€/y)	1,200; 2,000; 2,800	2,000; 2,800; 3,600	2,500; 3,750; 5,000	4,000; 5,500; 7,000	6,000; 8,000; 10,000	9,000; 12,500; 16,000
CO_2 Emissions (kg/y)	150; 225; 300	300; 375; 450	375; 450; 525	3,000; 3,750; 4,500	3,900; 4,575; 5,250	3,525; 4,125; 4,725
Fine particle emissions (g/y)	4,500; 6,000; 7,500	2,250; 3,750; 5,250	750; 1,500; 2,250	15; 30; 45	150; 450; 750	15; 30; 45
Required own work (h/m)	5; 10; 15	1; 2; 3	1; 2; 3	-	0.5; 1; 1.5	0.5; 1; 1.5

the total emissions of 10,000 similar houses cause one premature death per year". Finally, required own work refers to the time required to ensure the faultless operation of the heating system (e.g., cleaning and handling fuel loads). The choice of attributes and their levels was based on earlier studies and on feedback from experts. The annual operating cost and CO_2 and fine particle emissions were computed using as reference the energy consumption of an average detached house with a living area of 120 m^2.

The experimental design adopted in the study was an efficient availability design (Rose et al., 2013), according to which only three alternatives were shown in each choice task. The combination of levels that appeared in each scenario was defined with three different sub designs, namely near orthogonal, D-efficient (Scarpa and Rose, 2008; Rose and Bliemer, 2009) and serial designs. For the latter, an orthogonal design was used for the first respondent. After the choice sequence was completed, a multinomial logit model was estimated in the background and statistically significant parameters were used as priors to generate an efficient design. This process continued after each respondent and priors were continuously updated to generate a gradually more efficient design. Overall, the design generated 60 choice scenarios blocked in six groups, so that each respondent faced 10 of them. The sample was split so to have the same number of respondents assigned to the three different sub designs. An example of choice scenario is reported in Table 2.

3. Econometric approach

In our study we estimated a latent class model to investigate variation of tastes towards heating systems types and their features among the householders of the Veneto

Table 2. Example of choice scenario.

Attributes	Wood Pellet	LP Gas	Firewood
Investment Duration (years)	19	20	19
Fine particles emissions (g/year)	2,250	15	7,500
CO_2 emission (kg/year)	375	3,525	150
Required own work (hours/month)	1	1	15
Investment cost (€)	17,000	5,000	12,500
Operative cost (€)	3,750	9,000	1,200
Your choice	○	○	○

Region. The model is based on the Random Utility Theory (Luce, 1959; McFadden, 1974), according to which a respondent n facing a set of J mutually exclusive alternatives has utility U_i for alternative i as a function of attributes X_k, so that:

$$U_{ni}=\beta x_{ni}+\varepsilon_{ni} \tag{1}$$

where ε_{ni} is the unobserved error assumed to be i.i.d. extreme value type I.

To account for heterogeneity in sensitivity to attributes X_k, we adopted a latent class model. Such model assumes the existence of C classes of respondents, where C is exogenously defined by the analyst, based on information criteria indexes. Preference vary across classes but are homogeneous within them. As the classes are latent, an equation explaining the probabilistic assignment of individual n into class c needs to be defined. Using a logit formulation for the class allocation model, with Z_n being a vector of socio-economic variables and θ_c a vector of estimated coefficients, the probability that individual n belongs to segment c is given by (Bhat, 1997):

$$\pi_{nc} = \frac{\exp(\theta_c' Z_n)}{\sum_{c=1}^{c=C} \exp(\theta_c' Z_n)} \tag{2}$$

Specifically, the variables we used in Z vector are: i) age, ii) education, iii) income, iv) currently owning a biomass-based heating system.

Then, the probability that individual n chooses alternative i, conditional on belonging to class c, takes the logit form (Hensher and Greene, 2003):

$$\pi_{ni|c} = \frac{\exp(\beta_{nc}' x_i)}{\sum_{j=1}^{j=J} \exp(\beta_{nc}' x_j)} \tag{3}$$

Where X_i represents the vector of attributes associated with each alternative and β_{nc} the vector of estimated coefficients for class c.

The estimated parameters of the latent class model were used to simulate the market shares of different heating systems under different policy scenarios. Specifically, the scenarios involve reductions of investment cost (ranging from none to 50% reduction) and operational costs (same levels as previous case) for biomass-based heating systems. We computed choice probabilities in each scenario with the logit formula described in Equation 3, by including in it estimated coefficients β_{nc} and by varying the levels of investment and operational costs according to the reduction scenarios.

4. Results

This section reports the results of our study. In the first part of the section the estimates of the latent class model are presented, while the second part focuses on the policy scenarios.

4.1 LC model estimates

The first step of our modelling approach involves the identification of the optimal number of classes. As suggested by the literature (Hurvich and Tsai, 1989; Nylund et al., 2007), we referred to the AIC and BIC information criteria, which both favour a specification with 4 classes (Table 3). Class membership probabilities are 23% for class 1, 36% for class 2, 16% for class 3 and 25% for class 4 (Table 4). Results of latent class model with four classes are reported in Table 2[2]. The table also reports WTP values for heating systems features, which were computed with respect to the investment cost.

To class 1 are more likely to belong older individuals with low income and education, who currently do not own a biomass-based heating system. Such class exhibits a strong preference towards methane-fuelled technologies with seemingly no interest in biomass-based ones. As it concerns the attributes, it can be noticed how members of this class are very sensitive to installation and operational costs, which is consistent with their feature of individuals with low income. This class seems also sensitive to technical features of heating systems, and it shows a preference for systems with a long duration (WTP value of €0.38 for each additional year of duration) and which require low amount of time for maintenance (€0.27 to avoid an hour of work per month). Emissions, instead, do not seem to affect choices of members of this class.

Moving to class 2, to this class are more likely to belong younger individuals with high education and income who currently do not possess a biomass-based technology.

Table 3. Information criteria for alternative model specifications.

Number of classes	Number of parameters	LL	AIC	BIC
1 (MNL)	11	-15,713	31,448	31,506
2	28	-15,081	30,218	30,367
3	44	-15,017	30,122	30,356
4	60	-14,894	29,908	30,227
5	76	-14,886	29,924	30,328

[2] A part of these results was included in the report “Veneto 100% Rinnovabile: fotografia e prospettive” by the Interdepartmental Centre Giorgio Levi Cases for Energy Economics and Technology (University of Padova), available at http://levicases.unipd.it/wp-content/uploads/2019/11/Relazione-finale.pdf

Table 4. LC model results.

Class size	Class 1 23%		Class 2 36%		Class 3 16%		Class 4 25%	
	Estimate	WTP	Estimate	WTP	Estimate	WTP	Estimate	WTP
Class membership function								
Intercept	**0.16**		0.24		-0.11		-0.07	
Age	**0.31**		-0.31		0.12			
Degree	**-0.22**		0.43		-0.23			
Income	**-0.15**		0.32		0.02			
Owning a biomass fuelled system	**-0.22**		-0.11		0.19			
Heating system features								
Investment cost	**-0.41**		-0.89		-0.18		-0.49	
Operational cost	**-0.38**		-0.95		-0.31		-0.46	
Investment duration	**0.16**	**0.38**	0.34	0.38	**0.28**	**1.58**	**0.17**	**0.35**
Required own work	**-0.11**	**-0.27**	-0.24	-0.27	0.09	0.50	**-1.41**	**-2.87**
CO_2 emissions	-0.01	-0.03	**-0.73**	**-0.82**	**-0.06**	**-0.36**	**-0.02**	**-0.04**
Fine particle emissions	0.04	0.10	**-0.28**	**-0.31**	-0.05	-0.25	**0.01**	**-0.01**
Heating system type								
Firewood	2.00		4.99		1.89		1.55	
Chip wood	-3.96		1.94		0.99		1.21	
Wood pellet	0.42		10.88		0.46		4.20	
Methane	**4.81**		7.65		0.19		6.29	
Oil	**-0.39**		2.14		-0.04		-1.26	

Note: Coefficients statistically significant at 95% in bold. WTP values were computed with respect to the investment cost.

As it concerns preferences towards different types of heating system, it can be noticed how members of such class show a strong interest in biomass fuelled system, especially those based on wood pellet. This suggests that such class has a strong potential in terms of switching from a fossil fuel-based system to a biomass one. This seems corroborated by the high sensitivity to carbon dioxide and fine particles emissions of members of this class, which are those willing to pay the most to avoid them among all classes (€0.73/kg/year for CO_2 and €0.28/g/year for fine particles). At the opposite, in this class there seems to be no concern about technical features of heating systems.

To class 3, instead, are more likely to belong older individuals who currently own biomass-based heating system. As in the previous class, there seems to be a strong interest in biomass technologies, but in this case, firewood is the preferred fuel. As it concerns heating systems features, members of this class seem to strongly appreciate technologies with long investment duration (€1.58 for each additional year, largest value among all classes) and low emissions of carbon dioxide (€0.36 to avoid a kilogram per year).

Finally, members of class 4 (the baseline class) seem to be interested mainly in methane-based systems and in those fuelled by wood pellet. They also have a strong aversion to oil-based technologies. As it concerns systems' features, they seem interested in low maintenance requirements and to a lesser degree to avoiding emissions, with low WTP values (€0.04/kg/year to avoid CO_2 and €0.01/g/year to avoid fine particles).

4.2 Market shares for different heating systems in alternative policy scenarios

In this sub-section we report results of market shares simulations for different heating systems in two sets of policy scenarios: i) reduction of investment cost for biomass fuelled heating systems, specifically none, 10%, 20%, 30%, 40% and 50%; ii) reduction of operating costs for biomass based technologies (same range as above). In the first part of the sub-section (4.2.1) we present the average market shares weighted by class size for different heating systems; in the second (4.2.2) we report the shares for biomass technologies within each class.

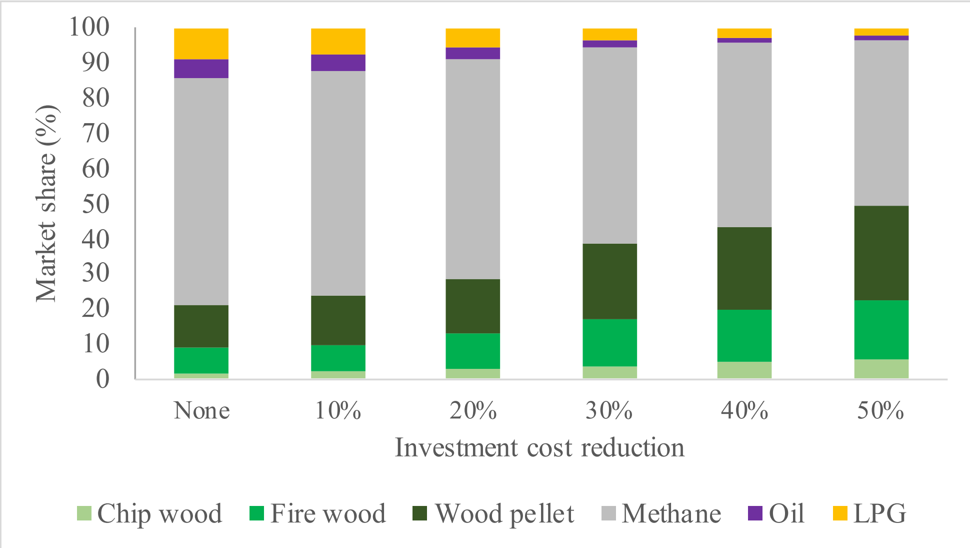
4.2.1 Average market shares in the investment cost reduction scenarios

Table 5 and Figure 1 illustrate weighted average market shares for different heating systems under the investment cost reduction scenario. In the baseline scenario, i.e. no investment cost reduction, most of the population would choose a methane heating system to replace the current one (64.40%), followed by wood pellet (12.61%), LPG (8.82%), firewood (7.35%), oil (5.39%) and chip wood (1.23%). Moving to the 10% reduction scenario, it can be noticed how shares for biomass fuelled technologies slightly increase (around 1% for each system). A 20% reduction seems to trigger a stronger response, with an increase of around 3% for wood pellet and 1.5% for the other biomass-based systems compared to the 10% reduction scenario. In the 30% reduction scenario the share for biomass-based systems further increases, in particular for wood pellet technologies, with a share of around 21% compared to the 12.61% of the baseline scenario. The fourth scenario (40% reduction), instead, does not show substantial differences compared to the previous one. Finally, in the last scenario (50% reduction) around 30% of citizens would choose a wood pellet fired system, around the 17% a firewood one and around the 5% a chip wood one. Overall, in such scenario, nearly half of the population would choose to adopt a bio-

Table 5. Average market shares under the investment cost reduction scenarios.

	Investment cost reduction					
Heating system	None (baseline)	10%	20%	30%	40%	50%
Chip wood	1.23	1.83	2.69	3.21	4.55	5.49
Firewood	7.35	7.90	10.32	13.86	14.93	16.59
Wood pellet	12.61	13.94	15.21	21.36	23.79	26.91
Biomass total	21.19	23.67	28.22	38.43	43.27	49.00
Methane	64.60	63.66	62.49	55.91	52.32	47.33
Oil	5.39	4.84	3.53	2.31	1.70	1.21
LPG	8.82	7.84	5.76	3.36	2.71	2.46
Total	100.00	100.00	100.00	100.00	100.00	100.00

Figure 1. Average market shares under the investment cost reduction scenarios.



mass-based system. The overall increases of the share of biomass-based systems compared to the baseline scenario are: 2.5%, 7%, 17.2%, 22% and 28% for the alternative investment cost reductions.

4.2.2 Average market shares in the operational costs reduction scenarios

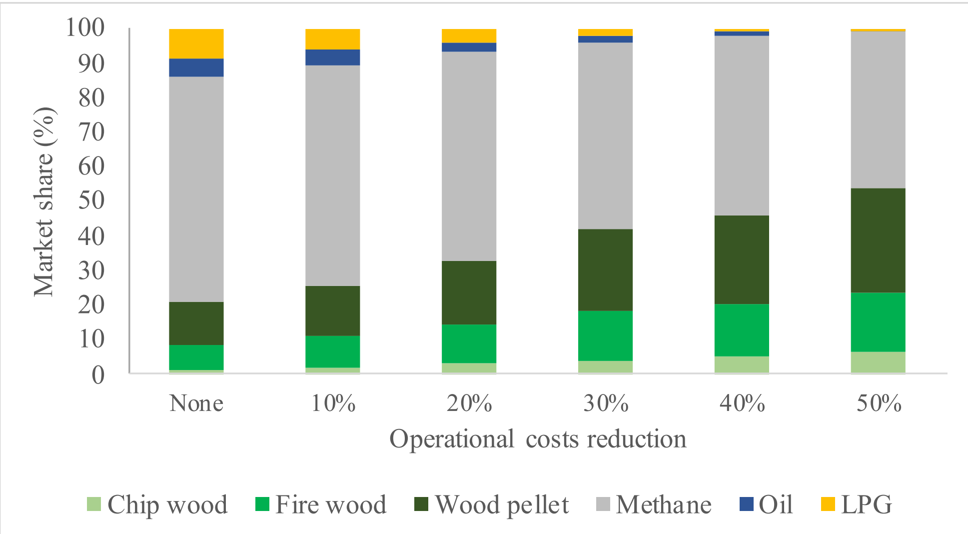
Table 6 and Figure 2 report the estimated average shares under the alternative operational cost reduction scenarios. Firstly, it is of interest to notice how reducing operational costs seems to have a stronger effect in terms of increasing biomass systems market shares compared to the reduction of investment cost. This seems true in each scenario (i.e. for each magnitude of the reduction) and is particularly evident in the 50% reduction scenario, under which the overall biomass technologies share is around 54% for operational costs reduction and 49% for investment cost reduction. In terms of fostering diffusion of biomass fired systems, this seems to suggest how policies aimed at decreasing operational costs for citizens may be more effective than those providing a reduction of the investment cost.

By looking more closely at the operational costs reduction scenarios, it can be noticed that, similarly to the previous scenario, a 10% reduction does not lead to a substantial increase in the shares of biomass technologies (between 1% and 2% increase for each system). A 20% reduction has an only slightly stronger effect, with an increase of about the 3% for biomass systems shares compared to the previous scenario. A similar relative increase is also shown for the 30% and 40% reduction scenarios. Finally, in the last scenario there is a substantial increase in the shares for biomass technologies, especially as it concerns wood pellet, which would be chosen by around the 30% of the population. Overall, it seems that a reduction of the operational costs would strongly favour the dif-

Table 6. Average market shares under the operational costs reduction scenarios.

	Operational costs reduction					
Heating system	None (baseline)	10%	20%	30%	40%	50%
Chip wood	1.23	2.11	3.37	3.88	4.91	6.11
Firewood	7.35	8.65	11.21	14.65	15.32	17.71
Wood pellet	12.61	14.99	18.07	23.16	25.89	29.94
Biomass total	21.19	25.75	32.65	41.69	46.12	53.76
Methane	64.60	63.19	60.66	54.08	51.80	45.19
Oil	5.39	4.65	2.60	1.71	0.98	0.46
LPG	8.82	6.41	4.09	2.52	1.10	0.59
Total	100.00	100.00	100.00	100.00	100.00	100.00

Figure 2. Average market shares for under the operational costs reduction scenarios.



fusion of wood pellet systems and only to a lesser degree the diffusion of other biomass-based systems. This may be due to higher operational costs of pellet fired heating systems, compared to other biomass ones.

4.2.3 Market shares in the investment cost reduction scenarios within each class

In this section we move from the population-level picture to a class-specific analysis, to explore how preference heterogeneity influences the diffusion of biomass heating systems. Specifically, we report and discuss probabilities to choose a biomass-based technology as replacement of the current one within each class in different policy scenarios.

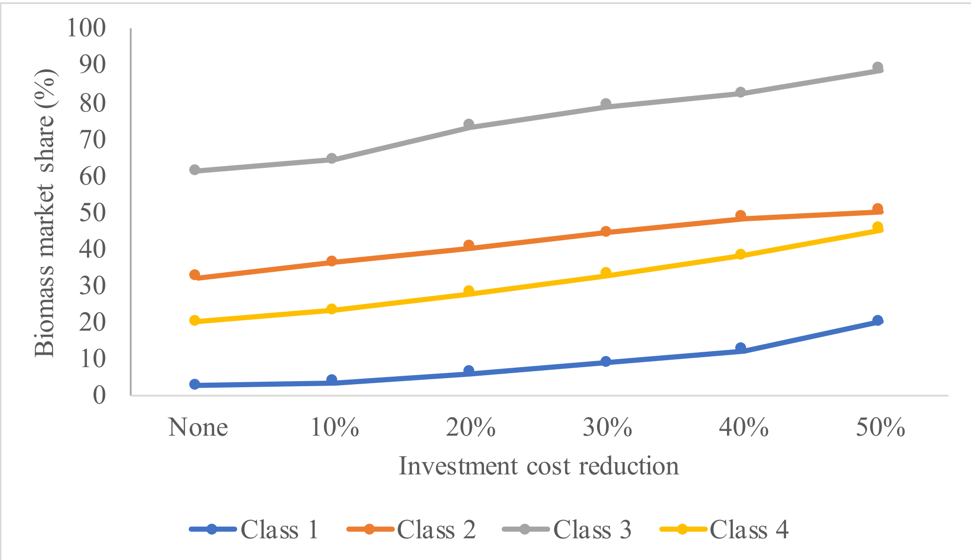
Starting from class 1, Table 7 and Figure 3 show how in class 1 the market share for biomass devices in the baseline scenario is extremely low (2.60%). Such value is consistent with the profile illustrated in Section 4.1, which highlighted how members of this class are characterized by little interest in biomass technologies and absence of sensitivity to carbon emissions. Moving to the cost reduction scenarios, their effect on adoption probability seems limited. Only in the 50% reduction scenario there seems to be a substantial increase in the biomass share (8% increase compared to the 40% reduction scenario). It seems that a very strong incentive is needed to foster diffusion of renewable based systems in this class.

Moving to class 2, the share for biomass devices in the baseline scenario equals 32.11%. Such results – together with the LC estimates – suggest that this class includes individuals who currently own a fossil fuel fired heating system and around one third of them would switch to a biomass fuelled one, even with no cost reduction. This seems to corroborate the potential of this class in terms of increasing the diffusion of renew-

Table 7. Class-specific biomass systems market shares under the investment cost reduction scenarios.

Class	Investment cost reduction					
	None (baseline)	10%	20%	30%	40%	50%
Class 1	2.60	3.39	5.91	8.66	12.11	19.99
Class 2	32.11	36.11	40.18	44.18	48.12	50.08
Class 3	61.16	64.12	73.18	78.81	82.18	88.88
Class 4	19.98	23.11	27.61	32.81	38.11	45.18

Figure 3. Class-specific biomass systems market shares under the investment cost reduction scenarios.



able based technologies across the population. As for the previous class, the effect of the investment cost reduction seems limited. In this case, however, the low effect may be linked to the high income of its members, that could make them less sensitive to costs.

Class 3 exhibits the highest biomass system adoption probability in the baseline scenario (61.6%). Overall, this class seems characterized by individuals that currently use a biomass system and show a high probability of choosing one of the same kind as replacement. Importantly, this class seems to be strongly affected by cost reduction, with an around 28% increase of the biomass devices share between the baseline and the 50% reduction scenario. This might be due to the low income of members of this class.

Finally, biomass systems share in class 4 equals 19.89% in the baseline scenario and 45.18% in the 50% reduction one, thereby suggesting a high sensitivity to investment cost reduction.

4.2.4 Market shares in the operational costs reduction scenarios within each class

Table 8 and Figure 4 report market shares within each class in the operational costs reduction scenarios. By comparing results with those reported in the previous section, it is interesting to notice how class 2 and 3 are affected more strongly by operational costs reduction, while classes 1 and 4 are affected more by investment cost reduction. This seems to be related to different sensitivity to the two costs highlighted in Section 4.1: classes 2 and 3 are more sensitive to operational costs, and as such reducing it has a stronger effect in such classes, while the opposite is true for classes 1 and 4. For example, in class 1, at a 50% reduction the share is around 15% for operational costs and 20% for investment cost. At the opposite, in class 3 the share in 5% higher in the case of operational cost reduction.

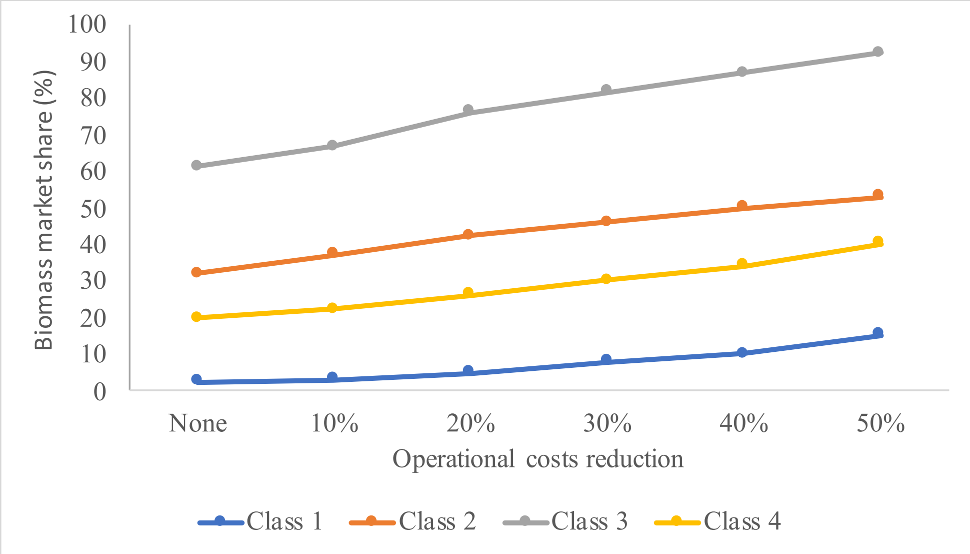
5. Discussion and conclusions

In the light of the importance of increasing the sustainability of the residential heating sector, it is crucial to inform energy policies with an accurate knowledge of the determinants of citizens heating choices. To this purpose, we designed a choice experiment aimed at investigating preferences towards heating systems and their features among the citizens of the Veneto region. We analysed choice data by means of a latent class model and we used the estimates to forecast market shares for different heating systems under alternative policies scenarios.

Table 8. Class-specific biomass systems market shares under the operational costs reduction scenarios.

Class	Operational costs reduction					
	None (baseline)	10%	20%	30%	40%	50%
Class 1	2.60	3.11	4.88	7.91	10.12	15.18
Class 2	32.11	37.14	42.24	46.11	49.88	53.11
Class 3	61.16	66.89	76.11	81.56	86.88	92.21
Class 4	19.98	22.41	26.22	30.33	34.16	40.11

Figure 4. Class-specific biomass systems market shares under the operational costs reduction scenarios.



The results of our study suggest how householders' preferences towards different heating systems and their features are strongly heterogeneous and how such variability can be partially explained by householders socio-demographic characteristics. Such findings support those of previous studies of the energy literature (e.g. Yoon et al., 2015, Ruokamo, 2016). Importantly, our estimates highlight the presence of population segments which seem to have a strong potential in terms of switching from a fossil fuel system to a renewable-based one. To this segment are more likely to belong individuals who already own a biomass-based heating systems and young individuals with high income and education level. At the opposite, our results highlight the existence of segments of the population with low interest towards the adoption of renewable-based technologies. Such segments are characterized by individuals with low income and education who currently own a fossil fuel system.

The simulations of policy scenarios allowed us to retrieve some important information about the efficiency of different policy measures in terms of fostering the diffusion on biomass technologies. Overall, we found that measures aimed at reducing operational costs for householders may induce a broader uptake of biomass appliances compared to those which target investment cost, even if the opposite is true in some segments of the population. This is particularly important in the context of the Veneto region, where subsidies for investment cost are currently in place and they seem to be only partially successful in nudging citizens towards the adoption of biomass-based appliances. We also found that only a large reduction of costs (i.e. 40% or 50% reduction) has a substantial effect on the increase of biomass systems shares, in classes with low interest in such technologies. This suggests that current incentives provided by authorities may not be enough to entice such segments of the population to switch from a fossil fuel to a biomass-based technology.

Acknowledgments

The research was funded by Interdepartmental Centre Giorgio Levi Cases for Energy Economics and Technology (University of Padova), "Sustainability of introduction of pellet based heating systems in a mountain area".

References

ARPAV, 2015. Inventario regionale delle emissioni in atmosfera.

Bhat, C.R. 1997. An endogenous segmentation mode choice model with an application to intercity travel. *Transportation Science*, 31(1):34–48.

Claudy, M.C., Michelsen, C., O'Driscoll, A., Mullen, M.R. 2010. Consumer awareness in the adoption of microgeneration technologies: an empirical investigation in the Republic of Ireland. *Renewable & Sustainable Energy Reviews*, 14(7):2154-2160.

European Commission. 2016. An EU Strategy on Heating and Cooling.
https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52016DC0051&from=EN

European Commission. 2019. Country Report Italy 2019.
https://ec.europa.eu/info/sites/info/files/file_import/2019-european-semester-country-report-italy_en.pdf

European Council. 2014.: Conclusions. European Council, General Secretariat of The Council. https://www.consilium.europa.eu/uedocs/cms_data/docs/pressdata/en/ec/145397.pdf

Franceschinis, C., Scarpa, R., Thiene, M., Rose, J., Moretto, M., Cavalli, R. 2016. Exploring the Spatial Heterogeneity of Individual Preferences for Ambient Heating Systems. *Energies*, 9, 407.

Franceschinis, C., Thiene, M., Scarpa, R., Rose, J., Moretto, M., Cavalli, R. 2017. Adoption of Renewable heating systems: an empirical test of the diffusion of innovation theory. *Energy*, 125:313-326.

Hensher, D. and Greene, W. (2003). The mixed logit model: the state of practice. *Transportation*, 30(2):133–176.

Hurvich, M., and Tsai, C. (1989). Regression and time series model selection in small samples. *Biometrika*, 76: 297–307.

Italian Government. 2018. Proposta Di Piano Nazionale Integrato Per L'energia E Il Clima.
https://www.mise.gov.it/images/stories/documenti/Proposta_di_Piano_Nazionale_Integrato_per_Energia_e_il_Clima_Italiano.pdf

Luce, D. 1959. Individual choice behavior: A theoretical analysis. John Wiley and Sons.

McFadden, D. 1974. Conditional logit analysis of qualitative choice behavior. In Zarembka P. (ed), Frontiers in econometrics. Academic Press, New York, pp 105–142.

Michelsen, C.C., Madlener, R. 2012. Homeowners' preferences for adopting innovative residential heating systems: A discrete choice analysis for Germany. *Energy Economics*, 34(5):1271-1283.

Nylund, K. L., Asparouhov, T. and Muthén, B.O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(4): 535-569.

Rommel, K., Sagebiel, J. 2017. Preferences for micro-cogeneration in Germany: Policy implications for grid expansion from a discrete choice experiment. *Applied Energy*, 206:612-622.

Rose, J.M., Bliemer, M.C.J. 2009. Constructing efficient stated choice experimental designs. *Transport Reviews*, 29(5):587-617.

Rose, J.M., Louviere, J.J., Bliemer, M.C.J. 2013. Efficient stated choice designs allowing for variable choice set sizes. In: International choice modelling conference, Sydney, Australia, 3rd-5th July 2013.

Rouvinen, S., Matero, J. 2013. Stated preferences of Finnish private homeowners for residential heating systems: A discrete choice experiment. *Biomass and Bioenergy*, 57:22–32.

Ruokamo, E. 2016. Household preferences of hybrid home heating systems – A choice experiment application. *Energy Policy*, 95:224-237.

Scarpa, R., Rose, J.M. 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*, 52:253-282.

Scarpa, R., Willis, K. 2010. Willingness-to-pay for renewable energy: Primary and discretionary choice of British households' for micro-generation technologies. *Energy Economics*, 32(1):129-136.

Su, W., Mengling, L., Zeng, S., Štreimikienė, D., Baležentis, T., Ališauskaitė-Šeškienėc, I. 2018. Valuating renewable microgeneration technologies in Lithuanian households: A study on willingness to pay. *Journal of Cleaner Production*, 191:318-329.

Yoon, T., Ma, Y., Rhodes, C. 2015. Individual Heating systems vs. District Heating systems: What will consumers pay for convenience? *Energy Policy*, 86:7.

Bio-based and Applied Economics 9(3): 241-262, 2020
DOI: 10.13128/bae-7928

Full Research Article

Multi-country stated preferences choice analysis for fresh tomatoes

Maria De Salvo[1,*], Riccardo Scarpa[2,3,4], Roberta Capitello[2], Diego Begalli[2]

[1] University of Catania, Department of Agriculture, Food and Environment, Italy
[2] University of Verona, Department of Business Administration, Italy
[3] Durham University, Business School, United Kingdom
[4] University of Waikato, Waikato Management School, New Zealand

Abstract. In this study we investigate consumers' preferences for fresh tomato attributes in four European countries by assessing and comparing Marginal Willingness-To-Pay (MWTP) estimates from panel Mixed Logit (MXL) models with utility specifications in the WTP-space. We performed an in-depth post-estimation inference to identify what attributes are the main determinants of fresh tomato purchases in each domestic market. We also assess the choice probabilities for tomatoes of various origins and types to illustrate how these post-estimation inference can be used to inform strategies designed to increase the market shares of Italian fresh tomato exports in new markets and to consolidate positions in markets where Italian fresh tomatoes are already appreciated by local consumers.

Keywords. Mixed Logit Model, Marginal Willingness-To-Pay, WTP space, preference space, fresh tomato.

JEL Codes. D12, Q13, Q18.

1. Introduction

Fresh tomato is one of the most commonly consumed vegetable in Europe. Over the last decade its consumption has remained stable at about 15 kg/year per capita, although stark changes have been observed concerning the range of quality consumers demand (European Union, 2018).

Italy is one of the major tomato producers in Europe (European Commission, 2020) with exports to German, Austrian, British, French and Romanian markets, where Italian fresh tomatoes are traditionally very appreciated. However, consumers' preferences gradually change, and year after year, the diversity of tomato types sold has increased everywhere to meet a rapidly evolving and diversifying demand. Health, convenience, taste and type of packaging are nowadays some of the most important product values for consumers. In the case of tomatoes, as for other foods, the market for 'specialties' is growing at

*Corresponding author. E-mail: mdesalvo@unict.it

Editor: Meri Raggi.

ISSN 2280-6180 (print)
ISSN 2280-6172 (online)

www.fupress.com/bae

a significant rate (Santeramo et al., 2018). New tomatoes varieties with attractive shapes, colours and tastes, innovative recyclable packaging, health claims and/or environmental certifications have been emerging as valuable product features that producers and retailers use to grow their market shares (Yue and Tong, 2009; Tonsor and Shupp, 2009; Alamanos et al., 2013; Oltman et al., 2014).

Nevertheless, the demand for tomatoes shows substantive differences across countries in terms of favorite shapes, packaging, origins and many other factors. Determining consumers' preferences and Willingness to Pay (WTP) for fresh tomato attributes is important to stakeholders in this industry (e.g., agricultural producers, intermediaries and retailers). It helps them determine which types of fresh tomato to grow and trade, how to manage the marketing mix, what communication content to emphasize in advertising campaigns, and how to apply fair prices along the supply chain. This information is particularly crucial for small-scale farmers who experience a strong competitive pressure from bigger companies of producers and importers. For them it is essential to correctly identify and characterize the market segments to supply, so as to define the assortment of tomatoes to produce the following season.

Against this background, the objective of this study is threefold. Firstly, this study aims at estimating consumers' willingness to pay for fresh tomato attributes across four key importing countries in Europe. Secondly, it aims at identifying the main determinants of tomato purchases across such markets. Thirdly, it aims at exploring how structural estimates of heterogeneous preferences can be used to inform marketing strategies which could guide the growth of Italian fresh tomato exports.

Alongside these research objectives, this paper also aims at achieving methodological and disseminative purposes. It will present and discuss the estimation strategies that could be implemented in a discrete choice cross-sectional analysis to face heterogeneity in preferences and take into account correlations between attributes. Frequently, in discrete choice applications, post estimation analyses are limited to the assessment of the Marginal Willingness-To-Pay (MWTP). But several additional results can be derived by the estimation of a discrete choice model with preference heterogeneity. In order to make concrete the methodological dissemination purposes of this study, the 'Rmarkdown' and 'markstat' codes we used in our analyses are made available to the reader.

The data collection took place in Germany, Russia, the UK and Norway. These countries were selected for different reasons. Two countries, Germany and the UK, are traditional export markets for Italian fresh tomatoes. In particular, Germany has been for several years the main European country for Italian fresh tomatoes exports. In 2015, Germany imported 28,188 tons of Italian fresh tomatoes, equivalent to 31% of the total fresh Italian tomato export. In 2015 the UK ranked third in terms of imported quantity from Italy, with 8,250 tons of fresh tomatoes.[1] The other two countries, Russia and Norway, instead, are marginal markets for Italian fresh tomato producers. Here Italian tomatoes compete with imports from other countries, such as the Netherlands, Spain, Egypt and Morocco. Nevertheless, the four markets under investigation in this study have all, to larger or smaller extent, the potential for future growth of Italian exports if producers will implement strategies aimed at meeting consumers preferences.

[1] Source: Trade Map (http://www.trademap.org).

To achieve the study objectives, the same choice experiment was administered to four representative samples of consumers, one for each country. In the data analysis, to account for heterogeneity in preferences, we used Mixed Logit models (henceforth MXL, see Train, 1998; 1999, 2009) with utility specified in WTP-space, as suggested by Train and Weeks (2005) and Scarpa et al. (2008). Despite the well-argued methodological advantages of this approach, when compared with the more conventional preference-space specification, applications in food choice experiments are still infrequent (Balcombe et al. 2010, Balogh et al. 2016, Caputo et al. 2016, and Caputo et al. 2018). Researchers have generally opted for the more traditional preference-space approach (Loureiro and Umberger, 2007; Ortega et al., 2011; Zanoli et al., 2013; Liu et al., 2019), even to assess the MWTP for tomato attributes (Onozaka and McFadden, 2011; Caputo et al., 2013; Carroll et al., 2013; Maples et al., 2014; Skreli et al., 2017).

In the post-estimation stage of our analysis, estimates of the Marginal Willingness-To-Pay (MWTP) for fresh tomato attributes were derived and compared across the four surveyed countries. From an empirical point of view, MWTPs provide producers with evidence to adjust their price strategy in line with market preferences. Further, we estimated full correlation matrices for random taste coefficients of tomato attributes. We used these to estimate market shares for combinations of tomato shapes and certifications. Signs and magnitudes of significant correlations between random attributes provide crucial information to producers and exporters. They are needed to define product profiles that meet consumers' demand and identify those combinations of tomato traits that consumers dislike. Moreover, to illustrate, probability choice functions were derived for selected product profiles. These functions are useful to predict consumer behavior, since differences in choice probabilities are dependent on tomato attributes. Finally, marginal changes in choice probabilities within samples and for the whole population were simulated. This type of analysis serves as a tool to predict changes in consumer behaviour specific to the different export markets.

The rest of the article is organized as follows. Section 2 discusses materials and method, while Section 3 illustrates the econometric analysis. Section 4 reports and discusses results. We conclude with some final remarks in Section 5.

2. Materials and method

Data used for this analysis were collected through a choice experiment designed to gather statements on hypothetical purchases of fresh tomato by consumers living in Germany, Russia, the UK and Norway. Preliminary focus groups and pilot surveys supported the final design of the questionnaire. Tomato attributes and levels were identified from previous studies (Yue and Tong, 2009; Onozaka and McFadden, 2011; Caputo et al., 2013; Carroll et al., 2013; Maple et al., 2014; Oltman et al., 2014; Meyerding, 2016; Skreli et al., 2017) and via discussion with experts in these export markets. Ten attributes were selected to profile fresh tomato characteristics. These were: tomato shape (which acts as a label for the product alternatives), colour, skin thickness, pulp type, packaging format, country of origin, production method, workers' health and safety certification[2], eco-sustainability

[2] In the questionnaire this attribute was explained as follows: "*tomatoes can be produced according to systems that ensure high health and safety standards to workers. The final product can have a label that certifies that these standards were implemented in production*".

Table 1. Attributes and levels.

Attribute	Attribute levels
Shape	**beef**, salad (*Salad*), vine (*Vine*), cherry (*Cherry*), date (*Date*)
Colour	red (*Red*), **not-red** (i.e., yellow, orange or variegated)
Skin	**thin**, thick (*Thick*)
Pulp	**juicy**, rich (*Rich*)
Packaging	**loose tomatoes**, net (*Net*), tray (*Tray*)
Origin	**Italy**, Netherlands (*NLD*), Spain (*ESP*), Morocco (*MAR*), Egypt (*EGY*), others (*OTH*)
Production method	**conventional**, low environmental impact (*Env*), organic (*Org*)
Workers' health and safety certified	**not present**, present (*Safety*)
Eco-sustainable certified	**not present**, present (*Eco*)
Price (euro/kg)	1.18, 1.58, 2.37, 2.76

Note: qualitative attributes were coded using dummy variables. The price attribute was coded using a continuous variable. In **bold** font the reference level. In brackets the variable name.

certification of production methods (including organic)[3] and price[4]. Table 1 reports the attributes and their levels.

To generate the alternatives, we used a fraction of the full factorial design, that was *D*-error minimizing within the sets that are orthogonal in the difference (refer to NGENE handbook for details). By using NGENE 12.0, 144 choice tasks were generated, blocked in twelve blocks of twelve each. Respondents were randomly assigned in a balanced rotation to one of the twelve blocks. Each was asked to complete the twelve randomized choice tasks in their assigned block. Given the complexity of the experimental design, the qualitative attribute named 'colour' was not directly included in the experimental design but was paired with tomato shape. Consequently, the combination between shape and colour was constant in each block, but combinations changed between blocks and assigned to different people. In this way, each respondent always visualized the same pictures for the five tomato alternatives in all choice scenarios under his/her scrutiny. The other attributes were presented in a textual form. Figure 1 illustrates a choice card.

The final questionnaire contained three sections. The first was designed to identify the respondent's socio-demographic profile; the second relates to food consumption habits, with specific reference to fresh tomatoes; the final section was dedicated to the choice experiment.

The survey was carried out in April 2016. The target population consisted of adult consumers that consumed fresh tomatoes in the last six months and were aware about the product characteristics. Country samples were selected to be representative of national populations in terms of age and gender. Interviews were conducted online and administered by Toluna (www.it.toluna.com), a market research company that deals with market

[3] In the questionnaire this attribute was explained as follows: "*tomatoes can be produced according to systems that ensure ecological sustainability and biodiversity protection. The final product can have a label that certifies that these standards were implemented in production*".

[4] In the questionnaire, prices were expressed in the national currency.

Figure 1. Example of a choice card.

Workers' health and safety certification	No	No	Yes	No	No
Packaging	Loose tomatoes	Net	Loose tomatoes	Loose tomatoes	Net
Eco-sustainable certified	Yes	No	No	Yes	Yes
Pulp	Juicy	Juicy	Juicy	Rich	Juicy
Production method	Low env. Impact	Organic	Low env. impact	Conventional	Organic
Skin	Thick	Thin	Thin	Thick	Thin
Origin	Morocco	Italy	Morocco	Netherlands	Netherlands
Price (€/kg)	2.76	2.37	1.58	1.18	1.58

analysis and has a world opt-in panel with over 9 million consumers. The company supplied the availability of high-quality Internet panels (i.e., ISO certification and application of international quality standards for market research) and guaranteed an incidence rate equal to 0.70 for each country.

The online questionnaire was completed by a total number of 2,600 respondents: 700 in Germany, Russia and the UK and 500 in Norway. The total choice observations generated were 31,200 (12 choice cards for 2,600 respondents). The number of products evaluated by respondents amounted to 156,000 (5 tomato shapes/scenarios for 2,600 respondents).

Table 2 reports the summary statistics at country level and for the whole sample.

Table 2. Summary statistics.

	Germany (n = 700)		Norway (n = 500)		Russia (n = 700)		UK (n = 700)		All (n = 2,600)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
White European ethnicity*	0.89	0.31	0.81	0.40	0.95	0.22	0.81	0.39	0.87	0.34
BMI	25.42	5.68	25.50	5.05	24.72	6.86	26.00	6.53	25.40	6.16
Vegetarian/vegan *	0.09	0.29	0.07	0.26	0.10	0.31	0.10	0.31	0.10	0.29
Female*	0.60	0.49	0.53	0.50	0.60	0.49	0.60	0.49	0.59	0.49
Age (in year)	39.01	12.09	40.34	16.74	38.70	11.09	39.23	12.40	39.24	12.96
Education**	0.29	0.45	0.39	0.49	0.75	0.43	0.43	0.50	0.47	0.50
Family size (n.)	2.54	1.23	2.47	1.39	3.23	1.17	2.84	1.34	2.79	1.31
Minor or dependent (n.)	0.68	0.92	0.61	1.02	0.93	0.94	0.79	1.04	0.77	0.99

*1 if yes; ** 1 if university graduate or post graduate.

3. Econometric model and inference

The choice data were analyzed by means of econometric models based on random utility maximization with heterogenous preference parameters (McFadden, 2001).

We assumed a linear and additive indirect utility function:

$$U_{njt}=-\alpha_n p_{njt}+\boldsymbol{\beta}_n^{'}\mathbf{x}_{njt}+\varepsilon_{njt} \tag{1}$$

where p_{njt} is the price attribute, $\mathbf{x}_{njt}$ represents the vector of non-price tomato attributes, and α_n and $\boldsymbol{\beta}_n$ are random parameters which represent n^{th} respondent's taste intensities for each attribute describing the tomato profile of each j^{th} alternative in the t^{th} choice occasion in the sequence. For the random component, we hypothesized that $\varepsilon_{njt}\sim$ i.i.d. Gumbel. Assumptions imply that, conditional on $\boldsymbol{\beta}_n$, the probability of observing a particular sequence of 12 choices for each n^{th} respondent ($y_n=y_{n1},y_{n2},\ldots,y_{n12}$) is the product of standard logit formulas:

$$L(y_{n1},y_{n2},\ldots,y_{n12}\,|\alpha_n,\boldsymbol{\beta}_n)=\prod_{t=1}^{12}\frac{\exp(-\alpha_n p_{nit}+\boldsymbol{\beta}_n'\mathbf{x}_{nit})}{\sum_{j=1}^{5}\exp(-\alpha_n p_{njt}+\boldsymbol{\beta}_n'\mathbf{x}_{njt})} \tag{2}$$

Unconditional probability was calculated as the integral of equation (2) weighted by the density function $g(\alpha_n,\boldsymbol{\beta}_n|\boldsymbol{\mu},\boldsymbol{\Omega})$:

$$P_n(y_n)=\int L(y_{n1},y_{n2},\ldots,y_{n12}|\alpha_n,\boldsymbol{\beta}_n)g(\alpha_n,\boldsymbol{\beta}_n|\boldsymbol{\mu},\boldsymbol{\Omega})d\alpha_n d\boldsymbol{\beta}_n \tag{3}$$

This integral was approximated through simulation, by: *i*) taking draws from the $g(.)$ function; *ii*) calculating the Likelihood function for each draw; and *iii*) averaging the results. The maximum simulated likelihood estimator is the value of the unknown parameters that maximizes the likelihood of the sample simulated in this manner.

Equation (3) represents the so-called panel mixed logit, which allowed us to use a mixed logit model specification in the context of repeated choices by respondents assuming specific taste distributions (Revelt and Train, 1998). To obtain a posterior distribution of $\alpha_n,\boldsymbol{\beta}_n$ for each respondent, the procedure described by Revelt and Train (2000) can be used.

Following Train and Weeks (2005) we specified the utility function in the *WTP space*[5]. With a Gumbel distributed unobservable component of utility, the error variance varies among respondents:

$$Var\left(\varepsilon_{njt}\right)=k_n^2\left(\pi^2/6\right) \tag{4}$$

where k_n represents a scale parameter for the n^{th} respondent. To allow for random scale parameter, Train and Weeks (2005) suggested to divide equation (1) by the scale parameter:

[5] Sonnier et al. (2007) called this model the "consumer's surplus model". It is also known in literature as "expenditure function space" model, "valuation function", or "money-space" (Thiene and Scarpa, 2009).

$$U_{njt} = -\left(\alpha_n/_{k_n}\right) p_{njt} + \left(\boldsymbol{\beta}_n/_{k_n}\right)' \mathbf{x}_{njt} + u_{njt} \quad (5)$$

As a consequence, in equation (5), u_{njt}~i.i.d. Gumbel, but with constant variance equal to $\pi^2/6$. Assuming that $\lambda_n = a_n/k_n$ and $\boldsymbol{c}_n = \boldsymbol{\beta}_n/k_n$, equation (5) becomes:

$$U_{njt} = -\lambda_n p_{njt} + \boldsymbol{c}_n' \mathbf{x}_{njt} + u_{njt} \quad (6)$$

where $\lambda_n = a_n/k_n$, $\boldsymbol{w}_n = \boldsymbol{c}'_n/\lambda_n$, $\boldsymbol{c}_n = \boldsymbol{\beta}_n/k_n$ and k_n represents the scale parameter for the n^{th} respondent.

Equation (6) is the so-called utility function in the *preference space*. Given that, by definition, the MWTP for an attribute is the ratio between the attribute's coefficient and the coefficient of the price attribute, equation (6) can be re-written as follows:

$$U_{njt} = -\lambda_n p_{njt} + (\lambda_n \boldsymbol{w}_n)' \mathbf{x}_{njt} + u_{njt} \quad (7)$$

where $\boldsymbol{w}_n = \boldsymbol{c}'_n/\lambda_n$. Equation (7) is the so-called utility function in *WTP space* (Train and Weeks, 2005).[6]

Through the direct choice of specific random WTP distributions, the WTP space approach prevents situations where the implied MWTP distributions from the random preference coefficients show excessively long tails. This is often the case in preference-space utility specifications (Scarpa et al., 2008). The literature reports controversial results on what approach produces a better fit to the empirical data. However, there is a general consensus on the ability of WTP space specifications to generate more reasonable and less disperse estimates of WTP distributions (Scarpa et al., 2008; Balcombe et al., 2009; Hensher an Greene, 2009; Rose and Masiero, 2009; Daly et al 2012; Owusu Coffie et al., 2016).

Our estimator was implemented in STATA 15.0 and employed the packages mixlogitwtp (Hole, 2007). We did not find significant evidence of heterogeneity in preliminary estimations for tomato colour, skin, pulp and country of origin. So, we assumed these to be fixed, meaning that we hypothesized homogeneous preferences for these features. Conversely, we obtained significant variance estimates in preliminary results for tomato shapes, packaging types and certifications. Hence, the associated random parameters were consequently assumed to be random and specifically distributed multi-variate normal with a full correlation matrix. The coefficient for the negative of price was assumed to have a log-normal distribution, to constrain the price coefficient to be always negative. Estimates were obtained with 1,000 Halton draws, which despite the high number of random parameters, can assure sufficiently low simulation variance of the maximum simulated likelihood estimator according to Zeng (2016) and Palma et al. (2018).

While all of the above is informative, it is also quite standard. In this study, however, we extended the range of inference in a more novel direction. We used the estimates of the vector of means $\boldsymbol{\mu}$ and their variance-covariance matrix $\boldsymbol{\Omega}=(\mathbf{LL}')$ for each country to infer the probabilistic choice behavior in the underlying population of consumers. Note

[6] Sonnier et al. (2007) called this model the "consumer's surplus model". It is also known in literature as "expenditure function space" model, "valuation function", or "money-space" (Thiene and Scarpa, 2009).

that estimates of the Cholesky decomposition $\mathbf{L}$ of the full variance-covariance matrix $\mathbf{\Omega}$ enabled us to derive the correlation matrix for the random α and $\boldsymbol{\beta}$. With this we identified patterns of covariation across taste parameters $\boldsymbol{\beta}$ that we then used in behavioral inference. For example, we used them in the derivation of probabilistic demand functions based on the simulation of distributions of preference values $\boldsymbol{\beta}$ in the population from which to infer market choice probabilities. We replicated this for selected tomato attributes (tomato profiles) and compared them across countries.

Another type of inference was conducted at the sample level. Here information on the observed choice sequence of each respondent was brought to bear by deriving individual specific means $\overline{\boldsymbol{\beta}}_n$ for marginal WTPs. These are graphically represented for the samples by kernel smoothing plots for the sample of each country. Hypothetical choice probabilities were also simulated at the population estimates by modifying the choice sets to evaluate shares for what-if scenarios. Scenarios simulated the introduction in the choice tasks of specific tomatoes profiles at a given price. We provide an illustration of the latter obtained with the post estimation commands in Stata. This required the modification of one or more attribute levels in the choice set and the re-computation of the in-sample predicted probabilities of choice to obtain changes in market shares following the introduction of new tomato profiles.

4. Results and discussion

Table 3 reports the coefficients estimates for each country. The model estimated for the pooled samples across countries are reported in the last columns to the right. The uninformed sequence of 12 choices between 5 alternatives has a log-likelihood of $\ln[(1/5)^{12}]=-19.31$, while the averages in our estimated model range between a maximum of -16.33 and a minimum of -16.67, respectively 0.85 and 0.86 percent of the uninformed probability. This implies a good explanatory power of the joint model.

Findings suggest that red color (baseline yellow/orange/variegated) and country of origin (baseline Italy) are key determinants of choice, while a thick tomato peel (baseline thin) and a rich pulp (baseline juicy) do not seem to be relevant. Preferences vary across the investigated markets, especially for the country of origin. Italian tomatoes are always preferred to those coming from other origins for Germans. These, for example, are willing to pay an average premium of 0.90 €/kg for Italian tomatoes in comparison to those coming from Morocco. Russians are generally indifferent to country of origin when tomatoes come from Egypt, Italy or Spain. However, they significantly dislike those produced in the Netherlands, Morocco or "other countries". For the latter, WTP is comparatively lower by about 0.19-0.21 €/kg. The UK consumers emerged as "origin-blind" as the country of origin never emerges as significant.

Country-level models' results further suggest that Norwegians appreciate juicy tomatoes (+ 0.15 €/kg in comparison to rich-pulp tomatoes), while Russians prefer to buy thin-skin tomatoes (+ 0.23 €/kg in comparison to thick-skin ones).

Interestingly, preferences for tomato shapes vary across countries and, at the same time, are also significantly heterogeneous within each country. For the coefficients of tomato shapes, standard deviations are significantly different from zero, with the exception of "cherry" and "date" shaped tomatoes in the UK and "date" shaped tomato in Nor-

Table 3. Mixed logit models' coefficients estimates.

Variables	Germany		Norway		Russia		UK		All	
	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation	Mean	Standard Deviation
	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.	Coef.
Red	3.77 *** (12.87)		2.26 *** (15.70)		1.39 *** (23.26)		3.95 *** (13.69)		2.52 *** (31.02)	
Thick	0.05 (0.56)		0.01 (0.12)		-0.23 *** (5.79)		-0.03 (0.35)		-0.06 (1.86)	
Rich	-0.07 (0.85)		-0.15 * (2.57)		0.05 (1.39)		-0.09 (1.17)		-0.03 (0.91)	
NLD	-0.39 ** (2.64)		-0.09 (0.93)		-0.19 ** (2.92)		-0.01 (0.10)		-0.18 ** (3.41)	
ESP	-0.65 *** (4.28)		-0.15 (1.43)		-0.12 (1.89)		0.09 (0.63)		-0.21 *** (3.94)	
MAR	-0.90 *** (5.50)		-0.19 (1.89)		-0.21 ** (3.15)		0.01 (0.06)		-0.29 *** (5.44)	
EGY	-0.86 *** (5.15)		-0.16 (1.51)		-0.09 (1.38)		-0.05 (0.39)		-0.25 *** (4.64)	
OTH	-0.62 *** (4.00)		-0.23 * (2.31)		-0.19 ** (2.87)		-0.27 (1.90)		-0.29 *** (5.34)	
Salad	-2.70 *** (10.29)	2.05 *** (8.80)	-1.35 *** (10.05)	1.17 *** (8.81)	-0.20 ** (2.78)	0.75 *** (10.12)	-2.36 *** (9.20)	2.48 *** (9.49)	-1.45 *** (18.64)	1.43 *** (18.20)
Vine	1.11 *** (7.58)	0.54 ** (3.47)	0.68 *** (6.41)	0.50 *** (5.30)	0.57 *** (9.21)	0.48 *** (9.45)	1.57 *** (8.56)	1.16 *** (7.58)	0.73 *** (14.76)	0.18 * (2.12)
Cherry	-2.00 *** (7.89)	1.81 *** (7.96)	-0.82 *** (6.46)	0.80 *** (6.48)	-0.95 *** (11.00)	0.41 *** (3.96)	-0.68 *** (4.47)	0.21 (1.17)	-1.12 *** (14.56)	0.52 ** (2.63)
Date	-3.62 *** (9.95)	0.83 ** (2.92)	-1.13 *** (7.57)	0.10 (0.77)	-1.21 *** (10.53)	0.71 *** (5.15)	-1.80 *** (7.81)	0.02 (0.14)	-1.75 *** (18.41)	0.07 (0.52)
Net	-0.25 * (1.97)	0.26 (0.76)	-0.10 (1.04)	0.38 * (2.37)	-0.14 * (2.08)	0.05 (0.53)	-0.10 (0.79)	0.53 (2.64)	-0.16 ** (3.38)	0.02 (0.21)
Tray	-0.72 *** (5.23)	0.34 (1.80)	-0.45 *** (4.66)	0.32 * (2.34)	-0.05 (0.80)	0.06 (0.81)	-0.64 *** (4.63)	0.05 (0.31)	-0.39 *** (7.83)	0.03 (0.41)
Env	0.18 (1.24)	0.25 (0.80)	-0.05 (0.50)	0.02 (0.14)	-0.22 ** (2.95)	0.38 *** (4.43)	0.23 (1.75)	0.15 (0.69)	0.01 (0.22)	0.08 (0.81)
Org	1.22 *** (6.23)	0.99 *** (4.23)	0.26 * (2.32)	0.80 *** (6.03)	-0.07 (0.93)	0.17 (1.90)	0.39 ** (2.65)	0.24 (1.02)	0.37 *** (6.58)	0.56 *** (3.73)
Safety	0.56 *** (4.81)	0.04 (0.17)	0.50 *** (5.36)	0.23 (1.66)	0.23 *** (3.85)	0.33 ** (3.28)	0.79 *** (6.26)	0.19 (0.97)	0.43 *** (10.12)	0.13 (0.81)
Eco	0.60 *** (4.69)	0.06 (0.30)	0.16 (1.82)	0.03 (0.25)	0.37 *** (5.56)	0.36 *** (3.85)	0.47 *** (3.98)	0.31 (1.49)	0.44 *** (9.36)	0.20 (1.94)
Ln-neg.p	-1.21 *** (15.69)	0.13 * (1.99)	-0.68 *** (11.21)	0.10 (0.83)	-0.48 *** (10.06)	0.04 (0.41)	-1.19 *** (16.06)	0.01 (0.19)	-0.85 *** (27.52)	0.17 ** (2.94)
Choices	42,000		30,000		42,000		42,000		156,000	
N	700		500		700		700		2,600	
ln-L/N	-16.335		-16.429		-16.488		-16.488		-16.674	
AIC	23,039.51		16,598.87		23,252.95		23,253.01		86,875.59	
BIC	23,774.37		17,305.14		23,987.81		23,987.87		87,721.99	

*p<0.05; **p<0.01; *** p<0.001. *z* statistic in parenthesis. Variables coding is reported in Table 1.

way. “Vine” tomatoes are always preferred to “beef” (the baseline) across all countries, while the latter are always preferred to “salad”, “cherry” and “date” tomatoes.

In general, consumers prefer to buy loose (the baseline) rather than packaged tomatoes. However, preferences for packaging types (in “nets” or “tray”) are not always significant and emerge as heterogenous at the country level.

Tomatoes with certified credence attributes are preferred to those without certification. In particular, Germans are willing to pay, on average, 1.22 €/kg for organic-certified tomatoes, even if the distribution is strongly dispersed in comparison to those in other countries. German consumers are also sensitive to certifications ensuring workers’ health and safety (+ 0.56 €/kg) and eco-sustainability (+ 0.60€/kg). However, the choices observed in the UK sample imply a higher willingness to pay for workers’ health and safety certification (+ 0.79 €/kg).

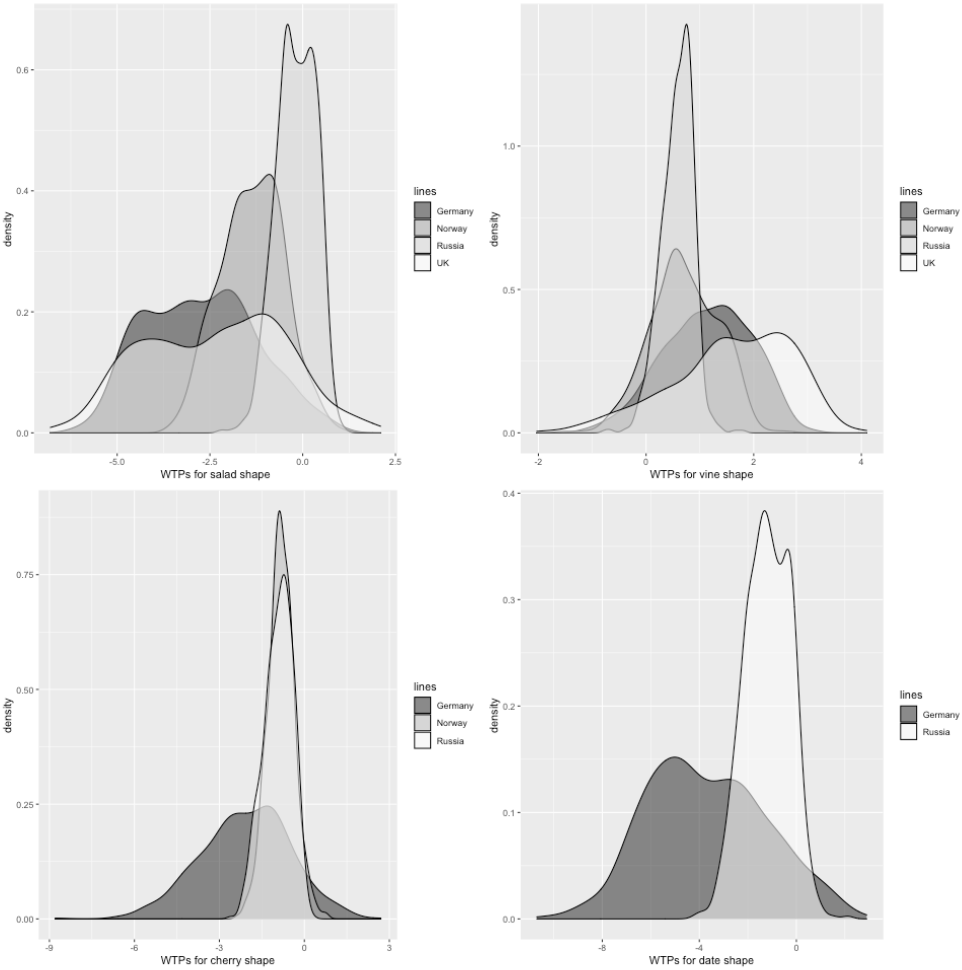
Preferences for certifications show a significant heterogeneity for organic products in Germany and Norway. Russians demonstrated significant heterogeneous preferences for low-environmental impact, workers’ health and safety, and eco-sustainable certified products. Instead, organic certification is not appreciated by Russians, who in turn are the only consumers with a significant positive appreciation for certification for low-environmental impact production methods.

Figure 2 displays the kernel smoothing of individual posterior means of MWTP sample distributions for each country for tomato shape, whose coefficients showed a significant heterogeneity in the majority of the investigated countries. Sample distributions are displayed only for those tomato shapes which have both significant mean and standard deviation estimates. Some distributions differ significant in terms of range, number of modes and relative positions in the WTP space. As pointed out earlier, German and UK consumers do not appreciate salad tomatoes in comparison to beef ones. Their MWTP distributions are prevalently located in the negative range and are multimodal in both cases. This implies that, everything else equal, for most consumers, salad tomatoes need to be sold at a lower target price compared to beef tomatoes to induce a purchase; how much lower is different in the two countries.

In contrast, MWTP for salad tomatoes in the other two markets are located to the right, especially for Russia, that present both positive and negative modal values. Negative means of MWTP are shown also for cherry and date tomatoes in comparison to beef tomatoes. However, preferences for cherry tomatoes seem to be more similar among Norwegian and Russian consumers than Germans; while MWTP distribution for date tomatoes are less dispersed for Russians and more variable for Germans. Vine tomatoes, in contrast, are preferred to beef ones in all markets. Modal value estimates are always positive, although the ranges of variation are extremely different between countries, with the widest one in the UK, where there is also the higher modal value.

Table 4 reports the estimates of correlation coefficients (lower triangle), variances (diagonal) and covariances (upper triangle) between random MWTPs in each country. For some pairs of random attributes, correlations are significant across all models and have concordant signs, even if they have different magnitude. Correlations between salad and date tomatoes, for instance, are always positive and significant in all markets, meaning that these kinds of tomato could be jointly sold in these countries through focussed advertising strategies exploiting the “drag” effect of a tomato type on the other. Converse-

Figure 2. Kernel density plots for conditional WTPs for tomato shapes.



ly, a negative correlation is estimated between salad and vine-shaped tomatoes. For Germany and Norway, in particular, the correlation coefficients are high and significant, -0.89 and -0.84 respectively. This finding is focal to support product marketing by the sellers: salad and vine tomatoes are antagonist in these markets and meet the preferences of different consumers and consequently separate market targets. This could suggest locating these products on different shelves or even separate shops when the locations of these are correlated with one type of buyers. Country-level preferences for type of packaging vary and they are correlated with the tomato's shape. In general, all consumers prefer to buy loose-packaged tomatos. However, Germans prefer to buy salad tomatoes that are tray-packaged (correlation coefficients: 0.77) and dislike trays for vine ones (-0.85); Norwegians like cherry tomatoes when packaged in a net (0.79). For the UK and Russia, some

Table 4. Estimates of correlation and covariance matrixes in each country.

Germany	Salad	Vine	Cherry	Date	Net	Tray	Env	Org	Safety	Eco	Np
Salad	**4.20**	**-2.19**	-0.38	**4.58**	0.47	**1.25**	0.14	0.46	**-0.58**	-0.14	0.06
Vine	***-0.89***	**1.44**	**1.05**	**-1.23**	-0.34	**-0.81**	-0.15	0.30	**0.46**	0.18	**-0.15**
Cherry	*-0.08*	***0.37***	**5.77**	**4.82**	0.10	-0.14	0.24	**-1.09**	**0.62**	-0.27	0.17
Date	***0.66***	***-0.30***	***0.60***	**11.30**	0.74	0.94	-0.30	-0.05	-0.08	-0.58	-0.15
Net	*0.38*	*-0.46*	*0.07*	*0.36*	0.38	0.19	-0.29	**-1.17**	-0.14	**-0.36**	0.10
Tray	***0.77***	***-0.85***	*-0.07*	*0.35*	*0.40*	**0.63**	0.29	-0.56	-0.12	-0.08	**0.20**
Env	*0.07*	*-0.13*	*0.11*	*-0.10*	*-0.50*	*0.39*	**0.86**	0.58	0.08	0.23	0.17
Org	*0.09*	*0.09*	***-0.17***	*-0.01*	***-0.72***	*-0.27*	*0.24*	**6.98**	-0.35	0.58	**-0.37**
Safety	***-0.37***	***0.50***	***0.34***	*-0.03*	*-0.29*	*-0.19*	*0.11*	*-0.10*	**0.59**	**0.61**	-0.05
Eco	*-0.06*	*0.13*	*-0.09*	*-0.15*	*-0.51*	*-0.09*	*0.21*	*0.26*	*0.68*	1.37	-0.21
-Price	*0.05*	***-0.22***	*0.12*	*-0.07*	*0.28*	***0.42***	*0.32*	***-0.08***	*-0.12*	***-0.31***	**0.34**

Norway	Salad	Vine	Cherry	Date	Net	Tray	Env	Org	Safety	Eco	Np
Salad	**1.37**	**-0.93**	-0.21	**1.21**	-0.14	0.14	-0.08	0.24	-0.13	0.07	0.07
Vine	***-0.84***	**0.88**	0.21	**-0.40**	0.16	-0.05	-0.08	**-0.33**	0.04	-0.16	**-0.23**
Cherry	*-0.21*	*0.27*	**0.69**	**0.43**	**0.41**	0.03	0.00	-0.06	-0.09	-0.06	0.04
Date	***0.71***	***-0.29***	***0.35***	**2.14**	0.26	0.23	-0.26	-0.01	**-0.34**	-0.16	-0.16
Net	*-0.19*	*0.26*	***0.79***	*0.29*	**0.39**	0.08	-0.01	0.00	0.19	-0.01	-0.06
Tray	*0.30*	*-0.13*	*0.08*	*0.39*	*0.33*	0.16	0.01	0.09	-0.07	0.01	0.01
Env	*-0.20*	*-0.24*	*0.02*	*-0.48*	*-0.05*	*0.06*	0.14	0.16	-0.11	0.01	**0.11**
Org	*0.20*	*-0.35*	*-0.07*	*-0.01*	*0.01*	*0.21*	*0.43*	**1.02**	0.03	**0.36**	0.10
Safety	*-0.12*	*0.05*	*-0.11*	***-0.24***	*0.31*	*-0.17*	*-0.31*	*0.03*	**0.93**	**0.27**	**-0.16**
Eco	*0.08*	*-0.24*	*-0.10*	*-0.15*	*-0.02*	*0.03*	*0.05*	***0.25***	***0.41***	**0.49**	0.01
-Price	*0.12*	***-0.46***	*0.09*	*-0.20*	*-0.17*	*0.07*	***0.54***	*0.05*	***-0.31***	*0.02*	**0.29**

Russia	Salad	Vine	Cherry	Date	Net	Tray	Env	Org	Safety	Eco	Np
Salad	**0.57**	-0.08	0.12	**0.71**	-0.09	**0.20**	0.09	-0.06	**-0.17**	**-0.12**	**-0.24**
Vine	*-0.23*	**0.24**	**0.28**	**0.25**	0.02	-0.02	-0.02	-0.06	-0.01	-0.01	**-0.17**
Cherry	*0.21*	***0.75***	**0.60**	**0.51**	-0.10	-0.08	-0.07	**-0.26**	-0.06	0.09	**-0.44**
Date	***0.67***	***0.35***	***0.47***	**1.98**	**-0.22**	**0.21**	0.14	-0.10	**-0.26**	-0.14	**-0.39**
Net	*-0.33*	*0.14*	*-0.36*	***-0.42***	**0.14**	0.10	0.05	**0.10**	0.00	**-0.14**	0.06
Tray	***0.53***	*-0.10*	*-0.22*	***0.29***	*0.55*	**0.25**	**0.14**	0.08	**-0.14**	**-0.28**	0.05
Env	*0.23*	*-0.09*	*-0.17*	*0.19*	*0.23*	*0.54*	0.28	0.16	-0.02	**-0.16**	0.14
Org	*-0.13*	*-0.19*	***-0.54***	*-0.11*	***0.43***	*0.24*	***0.48***	**0.40**	0.09	-0.05	0.12
Safety	***-0.36***	*-0.03*	*-0.13*	***-0.31***	*0.00*	***-0.45***	*-0.07*	*0.09*	**0.38**	**0.26**	0.03
Eco	***-0.21***	*-0.03*	*0.15*	*-0.13*	***-0.49***	***-0.73***	***-0.40***	*-0.07*	***0.55***	**0.58**	**-0.14**
-Price	***-0.39***	***-0.42***	***-0.71***	***-0.34***	*0.21*	*0.12*	***0.32***	*0.16*	*0.06*	***-0.23***	**0.64**

	UK										
	Salad	Vine	Cherry	Date	Net	Tray	Env	Org	Safety	Eco	Np
Salad	**6.14**	**-2.66**	**-0.69**	**5.70**	0.43	**0.92**	-0.30	-0.02	-0.58	-0.50	0.12
Vine	***-0.68***	**2.50**	-0.16	**-2.63**	-0.25	**-0.63**	**0.47**	**0.51**	0.29	0.04	**-0.62**
Cherry	***-0.53***	*-0.19*	0.27	-0.50	-0.04	-0.09	0.00	-0.15	0.11	0.14	0.14
Date	***0.98***	***-0.71***	*-0.41*	**5.48**	0.38	**0.73**	-0.16	-0.03	-0.42	-0.39	0.09
Net	*0.25*	*-0.23*	*-0.11*	*0.24*	**0.47**	**0.35**	0.10	-0.27	**-0.29**	**-0.41**	0.14
Tray	***0.49***	***-0.53***	*-0.23*	***0.41***	***0.68***	**0.58**	-0.04	-0.32	**-0.29**	**-0.40**	**0.28**
Env	*-0.18*	***0.44***	*0.01*	*-0.10*	*0.22*	*-0.07*	0.46	0.14	0.06	-0.22	**-0.15**
Org	*-0.01*	***0.28***	*-0.25*	*-0.01*	*-0.35*	*-0.38*	*0.19*	**1.28**	-0.07	0.26	**-0.20**
Safety	*-0.24*	*0.19*	*0.22*	*-0.18*	***-0.44***	***-0.39***	*0.09*	*-0.06*	**0.94**	**0.53**	-0.15
Eco	*-0.22*	*0.03*	*0.29*	*-0.18*	***-0.65***	***-0.57***	*-0.35*	*0.21*	***0.59***	**0.86**	**-0.15**
-Price	*0.08*	***-0.66***	*0.44*	*0.07*	*0.33*	***0.61***	***-0.37***	***-0.11***	*-0.26*	***-0.26***	**0.36**

Note: estimates of variances are reported in the diagonal. Covariance and correlation estimates are reported above and below the diagonal, respectively. Correlations are in *italic*. In **bold** the estimates which are significant with p<0.05.

estimates of correlation coefficients between tomato shapes and package types are significant, but their values are lower than 0.60, showing a low-to-moderate correlation.

Another interesting result concerns the relationship between certifications of workers' health and safety protection and eco-sustainability. Their correlation is always significant and positive, suggesting that consumers who are willing to pay for an eco-sustainable tomato are also willing to pay for a health and safety-certified tomato. However, the correlation coefficients are moderate ranging between 0.41 (for Norway) and 0.68 (for Germany).

Table 5 reports estimates of the market shares for combinations of tomato shape and certifications. In all investigated markets, vine tomato shows the higher shares. In all countries, this tomato type could increase its market share when certified as produced with methods that promote workers' health and safety, or eco-sustainability.

Figure 3 displays the estimated choice probability functions for selected product profiles along the price/kg dimension, when the baseline is a beef tomato without any additional attribute. For all graphs we adopted a price ranging from the lower level assumed in the choice experiment (1.81 €/kg) to three times the higher level (3 x 2.76 €/kg).

The top left plots of figure 3 show that as price increases the predicted purchase probability by Germans of red rich-pulp Italian tomatoes drops rapidly, so that at a price of 6€/kg it is basically zero except for vine tomato sold loose and certified.

The function for Norway investigates the simulated effects of eco-sustainable and organic certification of red Italian tomatoes, while the function for Russia is about the effect of tomato shape. Finally, the function for the UK investigates the role of certification for the same tomato profile and demonstrates that in this country consumers have the same reactions to price changes regardless the type of certification.

Figure 4 focuses on co-variation of preferences for tomato shapes, specifically salad and date, both of which tends to be disliked compared to the beef tomato baseline. It reports the iso-quantile plots for all the countries of bivariate kernel densities of MWTP

Table 5. Market shares for combinations of tomato shapes and certifications.

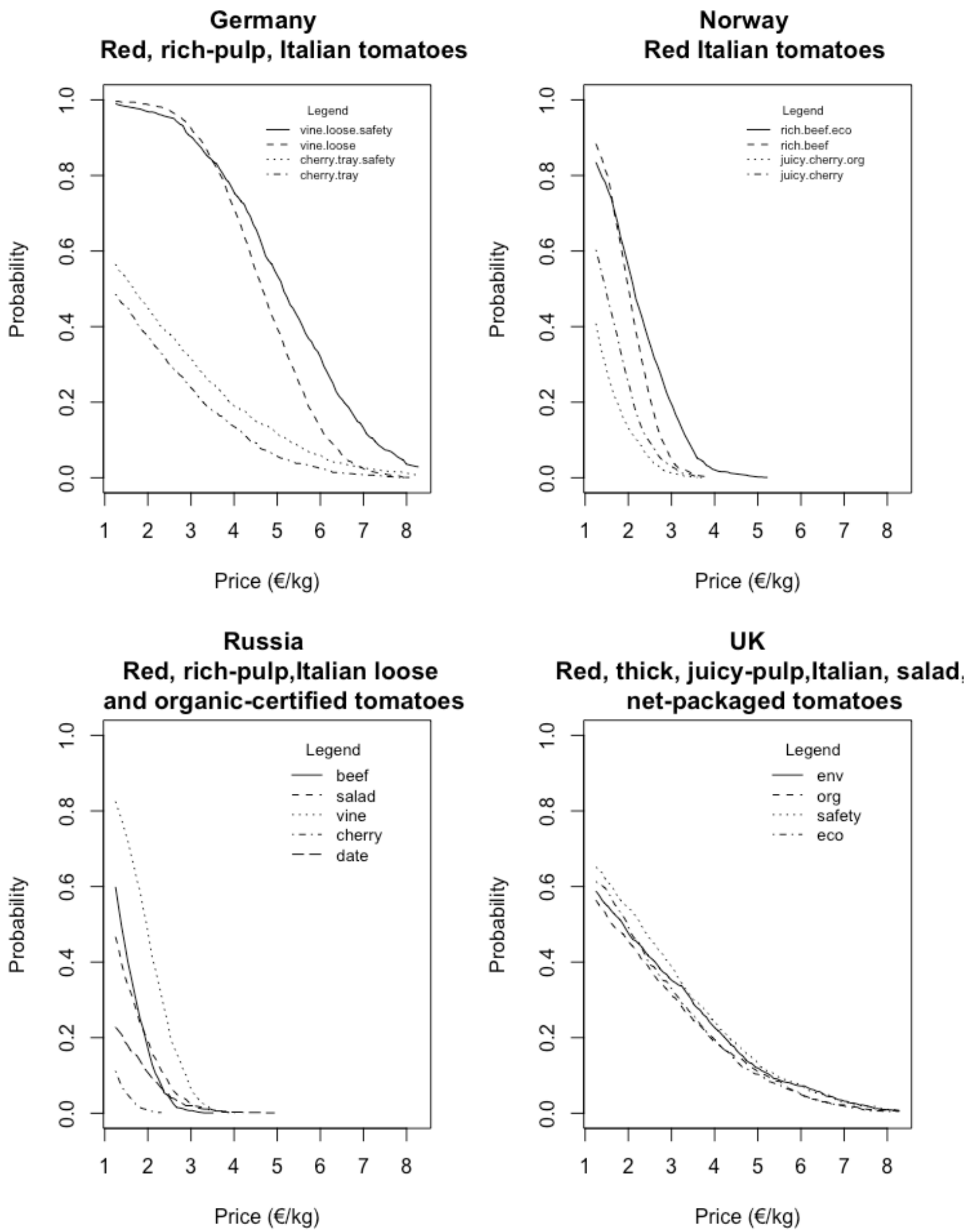
Shape	Certifications	Germany	Norway	Russia	UK
Salad	org-safety	4%	6%	11%	7%
	org-safety-eco	4%	4%	8%	5%
	org-eco	6%	6%	10%	7%
	safety-eco	5%	5%	17%	8%
	env-safety	4%	2%	10%	7%
	env-safety-eco	3%	1%	7%	4%
Vine	org-safety	45%	29%	28%	45%
	org-safety-eco	37%	22%	21%	36%
	org-eco	41%	28%	25%	40%
	safety-eco	54%	32%	46%	54%
	env-safety	39%	17%	20%	48%
	env-safety-eco	32%	11%	14%	34%
Cherry	org-safety	11%	5%	0%	4%
	org-safety-eco	8%	3%	0%	3%
	org-eco	8%	6%	1%	4%
	safety-eco	12%	5%	6%	8%
	env-safety	12%	3%	1%	6%
	env-safety-eco	8%	2%	1%	5%
Date	org-safety	7%	8%	5%	10%
	org-safety-eco	6%	6%	3%	7%
	org-eco	6%	8%	5%	9%
	safety-eco	8%	7%	8%	11%
	env-safety	6%	2%	5%	10%
	env-safety-eco	4%	2%	3%	6%

estimates in a range of change between -5 to +5 Euro/kg compared to the baseline product profile. The price change combinations along each iso-quantile curve represent the proportion of the population with the same probability of selecting tomatoes with one of the two shapes rather than the baseline beef tomato.

The isoquantiles highlight a positive correlation between salad and date-shaped tomatoes in the four countries, but the price set combinations with which they relate to the baseline are quite different. For German consumers, with a correlation estimate of 0.66, the curves cover a much larger set of MWTP values than in the Russian and Norwegian samples. For the UK consumers, the MWTP ranges are similar to those shown in Germany. However, because of the much stronger correlation of 0.98 between the shape attributes, the room for a differentiated pricing policy is much reduced. Norwegian and Russian consumers show quite similarly sets in terms of preferences and willingness to pay.

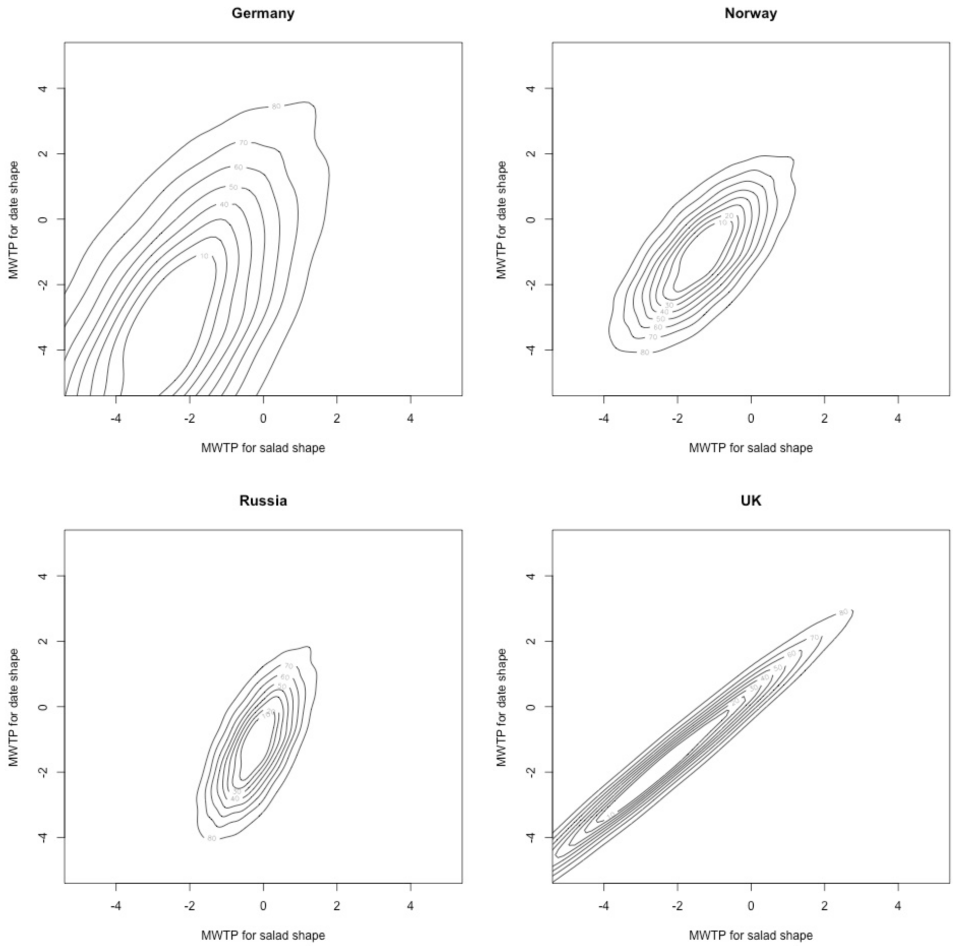
Finally, the estimated models can be used to simulate marginal changes in probability of choice within the samples rather than in the population. For example, what would the distribution of choice probability be, within the German sample, if all choice tasks including the baseline Italian tomato were offered with certification for workers' health and safe-

Figure 3. Country-level demand functions for some types of tomato.



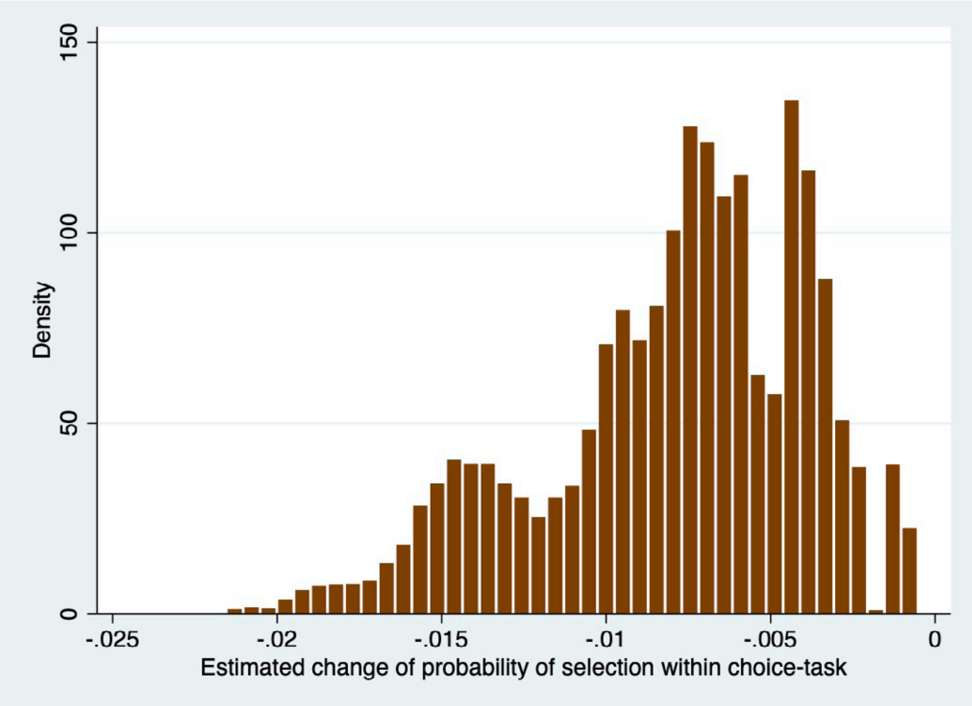
ty at a price increased by ten percent? This comes down to computing the choice probabilities for all five alternatives in each choice task, i.e. the probability vectors with the

Figure 4. Iso-Quantile Plots of Bivariate Kernel Density distributions for MWTP estimates for salad and date-shaped tomatoes in the four countries.



price increase for certitication ($\mathbf{p}^1$) and without such change ($\mathbf{p}^0$) for the baseline tomato. Then the difference between the two sets of predicted selection probabilities ($\mathbf{p}^1$ - $\mathbf{p}^0$) for the alternatives with the profile of interest is computed and the distribution of these values examined. In our case we have 13,071 choice sets containing the baseline profile in the German sample. An increase of ten percent would always result in a decreased selection probability, as shown in Figure 5. This suggests that either the price change should be lowered, or some additional positive features should be added, for example organic certification, that the German consumers seem to strongly appreciate. One can also envisage iterating this exercise at gradually lower price increases until a sufficient fraction of the within sample predicted choices show a positive value.

Figure 5. In-sample simulation of selection probabilities for workers' health and safety certification at 10% price increase.



5. Conclusions

We conduct identical surveys across four countries to estimate the marginal WTPs for a set of attributes of fresh tomatoes. Estimates were obtained in WTP-space, which several authors encourage practitioners to adopt to obtain more reliable, interpretable and plausible MWTP distributions. Specific differences in preferences across countries have been highlighted in terms of sign and magnitude of the coefficient estimates, conditional MWTPs, correlation coefficients and market shares. Further, simulations of choice purchase behaviour were inferred within-sample and at the population level. These were discussed with regards to their effects of price changes on tomato profiles in the four markets, to explore marketing implications of population distributions of marginal MWTPs and to exemplify the range of uses analysts can make of these model post-estimations.

The method can produce evidence that could be used to support the design of strategies aimed at consolidating the position of Italian tomatoes on traditional European markets, such as Germany and the UK; and at the same time, it could help Italian producers to identify what types of tomato produce to improve their share in Norwegian and Russian markets.

The tomato profile, which shows the highest probability to be purchased in all markets is vine, red and sold loose (unpackaged). However, some specific tomato profiles have been identified for each market. In Germany, where Italian tomatoes are preferred to those coming from other countries, consumers ask tomatoes whose quality is certified for workers' health and safety and eco-sustainability, but only within a restricted price range, as shown by the in-sample inference, where a ten percent increase was found too high. Salad-shaped tomatoes is more likely to be purchased when packaged in trays, while the use of this package should be avoided for vine-shaped tomatoes. In the UK, the same types of tomato certifications are also appreciated. However, the UK consumers seem to be not interested in the country of origin, unlike German consumers. Norwegian and Russian consumers adopt an intermediate behaviour. Consequently, tomatoes from Italy may not enjoy the same level of competitive advantage abroad, as it is generally assumed, and hence export penetration strategies should vary across countries. To sell more tomatoes in Norway, Italian producers should offer juicy-pulp tomatoes and certify their quality with organic and worker's health and safety labels. Cherry tomatoes are more appreciated in the UK market if are packaged in a net. Finally, Russians prefer thin-skin tomatoes and appreciate certifications for workers' health and safety and eco-sustainability, rather than for organic production.

Further research should address some of the limitations of our study in order to confirm or disconfirm our findings, which were only illustrative in their nature. In fact, we are aware about a number of limitations of our study. They arise from the choices we have been forced to make regarding the experimental design and the data analysis. Firstly, to assure that the survey respect international quality standards for market research in a cross-country context, we decided to collect data engaging a market research company. The use of such online survey has grown rapidly in social science and policy research in the last ten years (Lehdonvirta et al., 2020). However, it is well known that data generated in this manner could be affected by self-selection issues and non-random and non-representativeness of the samples, and these limitations should be taken into account in evaluating the external validity of our results. Further, to reduce the choice task complexity, we simulated a forced choice decision context, asking consumers to imagine they had to decide to buy one of the proposed options, without including an opt-out alternative. This decision has been supported by Dhar and Simonson (2003) who suggested that forced choice may generate more accurate and complete results in categories of familiar commodities in which the deferral option is available but rarely exercised. We assumed that this is the case of our research given that participants in our survey are consumers of fresh tomatoes, fresh tomatoes are characterized by high versatility in cooking and individual diets, and the expenditure of this product has a low impact on the individual/household budget. However, we are aware that this can be seen as a limitation of our study. Therefore, market shares estimates could be affected by the adopted choice design. This possibility must be taken into account by the reader. Moreover, each choice card includes several attributes and levels and, despite this well simulates the real-life scenario faced by consumers when purchasing fresh tomatoes, at the same time, respondents may not have attend to a certain number of attributes. An attribute-not-attendance phenomenon (Hensher, 2010) could consequently affect this survey as a limitation. We plan to analyse this eventuality through a further paper, given that it is not the focus in this one. Another limitation is related to the econometric approach. We chose to use Halton draws for simulations, despite the use of Scrambled Sobol draws could be more appropriate, as demonstrated by Czajkowski and Budziński (2019). Our choice stemmed from the fact that one of the aims of this paper is to provide the reader with estimation and post-estimation codes used in data analysis to facilitate dissemination. Further, it is worth observing that we took the exporter viewpoint, and consequently we did not adjust prices according to the national purchasing power given that results are mainly presented at a country level. Therefore, it is important to underscore that, in the case of a country comparison, the same tomato profile could be perceived as relatively cheap or expensive in countries with different purchase powers. These cases could affect choice probability estimates. Finally, we used maximum likelihood estimators, which suffer from the limitation of local optima, and assumed normal and log normal distributions of qualitative attributes and price, respectively, for the random parameters. Assumptions of unimodal symmetric distributions surely affect our estimates and the analysis might also have been conducted with more flexible semi-parametric mixtures (Train 2016, Caputo et al. 2018, Scarpa et al. 2020).

Despite these limitations, this study presented useful insights into consumer choices and their impact on market competitiveness for food producers. It demonstrated how the use of stated food choice experiments in a multi-country context is focal to support decision makers in determining which types of product to grow and promote, how to manage the marketing mix, what communication content to emphasize in advertising campaigns and how adopt price differentiation strategies in different markets to face consumers' demand.

References

Alamanos, E., Bourlakis, M., Tzimitra-Kalogianni, I. (2013). Segmenting Greek tomato consumers: policy and marketing insights towards a healthy diet. *British Food Journal* 115(4): 488-507.

Balcombe, K., Chalak, A., Fraser, I. (2009). Model selection for the mixed logit with Bayesian estimation. *Journal of Environmental Economics and Management* 57(2): 226-237.

Balcombe, K., Fraser, I., Di Falco, S. (2010). Traffic lights and food choice: A choice experiment examining the relationship between nutritional food labels and price. *Food policy* 35(3): 211-220.

Balogh, P., Békési, D., Gorton, M., Popp, J., Lengyel, P. (2016). Consumer willingness to pay for traditional food products. *Food Policy* 61: 176-184.

Caputo, V., Nayga Jr, R. M., Scarpa, R. (2013). Food miles or carbon emissions? Exploring labelling preference for food transport footprint with a stated choice study. *Australian Journal of Agricultural and Resource Economics* 57(4): 465-482.

Caputo, V., Scarpa, R., Nayga Jr, R. M. (2016). Cue versus independent food attributes: the effect of adding attributes in choice experiments. *European Review of Agricultural Economics* 44(2): 211-230.

Caputo, V., Scarpa, R., Nayga Jr, R. M., Ortega, D. L. (2018). Are preferences for food quality attributes really normally distributed? An analysis using flexible mixing distributions. *Journal of choice modelling* 28: 10-27.

Carroll, K. A., Bernard, J. C., Pesek Jr, J. D. (2013). Consumer preferences for tomatoes: the influence of local, organic, and state program promotions by purchasing venue. *Journal of Agricultural and Resource Economics* 38(3): 379-396.

Crastes Dit Sourd, R., Beaumais, O., Mahieu, P. A., Martinez-Camblor, P., Scarpa, R. (2020). A contingent valuation test for measuring the construct validity of willingness-to-pay estimates derived from choice-experiments. Land economics, accepted and forthcoming.

Daly, A., Hess, S., Train, K. (2012). Assuring finite moments for willingness to pay in random coefficient models. *Transportation* 39(1):19-31.

Dhar, R., & Simonson, I. (2003). The effect of forced choice on choice. *Journal of marketing research* 40(2): 146-160.

European Union (2018): EU agricultural outlook for markets and income, 2018-2030. European Commission, DG Agriculture and Rural Development, Brussels.

European commission (2020): DG agri-dashboard: tomatoes (https://ec.europa.eu/info/sites/info/files/food-farming-fisheries/farming/documents/tomato-dashboard_en.pdf).

Goett, A., K. Train, Hudson, K. (2000). Customer Choice Among Retail Energy Sup- pliers: The Willingness-to-Pay for Service Attributes. *Energy Journal* 21:1–28.

Hensher, D.A. (2010). Attribute processing, heuristics and preference construction in choice analysis. In: Hess, S., Daly, A.J. (Eds.), State-of Art and State-of Practice in Choice Modelling: Proceedings from the Inaugural International Choice Modelling Conference, Chap. 3, pp. 35–70. Emerald, Bingley, UK.

Hensher, D. A., Greene, W. H. (2011). Valuation of travel time savings in WTP and preference space in the presence of taste and scale heterogeneity. *Journal of Transport Economics and Policy (JTEP)* 45(3): 505-525.

Hole AR. (2007). Fitting mixed logit models by using maximum simulated likelihood. *Stata Journal* 7: 388-401.

Hole, A.R. Kolstad, J., (2012). Mixed logit estimation of willingness to pay distributions: a comparison of models in preference and WTP space using data from a health-related choice experiment. *Empirical Economics* 42(2): 445-469.

Layton, D.F., Brown. G. (2000). Heterogeneous Preferences Regarding Global Climate Change. *The Review of Economics and Statistics* 82(4): 616–24.

Lehdonvirta, V., Oksanen, A., Räsänen, P., & Blank, G. (2020). Social Media, Web, and Panel Surveys: Using Non-Probability Samples in Social and Policy Research. Policy & Internet.

Liu, R., Gao, Z., Nayga Jr, R. M., Snell, H. A., Ma, H. (2019). Consumers' valuation for food traceability in China: Does trust matter? *Food Policy* 88: 101768.

Loureiro, M. L., Umberger, W. J. (2007). A choice experiment model for beef: What US consumer responses tell us about relative preferences for food safety, country-of-origin labeling and traceability. *Food Policy* 32(4): 496-514.

Mabit, S., Caussade, S., Hess, S. (2006). Representation of taste heterogeneity in willingness to pay indicators using parameterization in willingness to pay space. In European Transport Conference 2006. European Transport Conference 2006.

Maples, M., Morgan, K. L., Harri, A., Hood, K., Interis, M. (2014). Consumer willingness to pay for environmental production attributes in tomatoes: a southeastern consumer survey, Selected paper prepared for presentation at the Southern Agricultural Economics Association 2014 Annual Meeting, Dallas, TX.

Meyerding, S. G. (2016). Consumer preferences for food labels on tomatoes in Germany–A comparison of a quasi-experiment and two stated preference approaches. Appetite 103: 105-112.

Meyerding, S. G., Trajer, N., Lehberger, M. (2019). What is local food? The case of consumer preferences for local food labeling of tomatoes in Germany. *Journal of Cleaner Production* 207: 30-43.

Morey, E., Rossmann, K. (2003). Using Stated- Preference Questions to Investigate Variations in Willingness to Pay for Preserving Marble Monuments: Classic Heterogeneity, Random Parameters, and Mixture Models. *Journal of Cultural Economics* 27(3/4): 215–29.

Oltman, A. E., Jervis, S. M., Drake, M. A. (2014). Consumer attitudes and preferences for fresh market tomatoes. *Journal of food science* 79(10): S2091-S2097.

Onozaka, Y., McFadden, D. T. (2011). Does local labeling complement or compete with other sustainable labels? A conjoint analysis of direct and joint values for fresh produce claim. *American Journal of Agricultural Economics* 93(3): 693-706.

Ortega, D. L., Wang, H. H., Wu, L., Olynk, N. J. (2011). Modeling heterogeneity in consumer preferences for select food safety attributes in China. *Food Policy* 36(2): 318-324.

Owusu Coffie, R., Burton, M. P., Gibson, F. L., Hailu, A. (2016). Choice of rice production practices in Ghana: a comparison of willingness to pay and preference space estimates. *Journal of Agricultural Economics* 67(3): 799-819.

Palma, M. A., Vedenov, D. V., and Bessler, D. (2018). The order of variables, simulation noise, and accuracy of mixed logit estimates. *Empirical Economics* 58: 2049–2083.

Revelt, D., Train, K. (1998). Mixed logit with repeated choices: households' choices of appliance efficiency level. *Review of economics and statistics* 80(4): 647-657.

Revelt, D., Train, K. (2000). Customer-specific taste parameters and mixed logit: Households' choice of electricity supplier, working paper No. E00-274, University of California at Berkley, Department of Economics, Berkley, California.

Rose, J. M., Masiero, L. (2010). A comparison of the Impacts of Aspects of Prospect Theory on WTP/WTA Estimated in Preference and WTP/WTA Space. *European Journal of Transport and Infrastructure Research* 10(4): 330-346.

Santeramo, F. G., Carlucci, D., De Devitiis, B., Seccia, A., Stasi, A., Viscecchia, R., Nardone, G. (2018). Emerging trends in European food, diets and food industry. *Food Research International* 104: 39-47.

Scarpa, R., Thiene, M., Train, K. (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: 94-1010.

Scarpa, R., Thiene, M., Marangon, F. (2007). The value of collective reputation for environmentally-friendly production methods: the case of Val di Gresta. *Journal of Agricultural & Food Industrial Organization* 5(1).

Scarpa, R., Franceschinis, C., Thiene, M. (2020). Logit mixed logit under asymmetry and multimodality of WTP: a Monte Carlo evaluation. *American Journal of Agricultural Economics*, accepted and forthcoming.

Skreli, E., Imami, D., Chan, C., Canavari, M., Zhllima, E., Pire, E. (2017). Assessing consumer preferences and willingness to pay for organic tomatoes in Albania: a conjoint choice experiment study. *Spanish journal of agricultural research* 15(3): 1-13.

Sonnier, G., Ainslie, A., Otter, T. (2007). Heterogeneity distributions of willingness-to-pay in choice models. *Quantitative Marketing Economics* 5(3): 313-331.

Street D.J., Burgess L., Louviere J.J. (2005). Quick and easy choice sets: constructing optimal and nearly optimal stated choice experiments. *International Journal of Research in Marketing* 22(4): 459-470.

Thiene, M., Scarpa, R., (2009). Deriving and testing efficient estimates of WTP distributions in destination choice models. *Environmental and Resource Economics* 44: 379–395.

Train, K.E. (1998). Recreation demand models with taste variation. *Land Economics* 74: 230–239.

Train, K. (1999). Mixed logit models for recreation demand, in J. Herriges & C. Kling, eds, 'Valuing Recreation and the Environment', Edward Elgar, Northampton, MA.

Train, K., Weeks, M. (2005). Discrete choice models in preference space and willingness-to-pay space. In Applications of simulation methods in environmental and resource economics (pp. 1-16). Springer, Dordrecht.

Train K., (2016). Mixed Logit with a Flexible Mixing Distribution. *Journal of Choice Modelling* 19: 40–53.

Tonsor, G. T., Shupp, R. (2009). Valuations of 'sustainably produced' labels on beef, tomato, and apple products. *Agricultural and Resource Economics Review* 38(3): 371-383.

Yue, C., Tong, C. (2009). Organic or local? Investigating consumer preference for fresh produce using a choice experiment with real economic incentives. *HortScience* 44(2): 366-371.

Zanoli, R., Scarpa, R., Napolitano, F., Piasentier, E., Naspetti, S., Bruschi, V. (2013). Organic label as an identifier of environmentally related quality: A consumer choice experiment on beef in Italy. *Renewable Agriculture and Food Systems* 28(1): 70-79.

Zeng, T. (2016). Using Halton sequences in random parameters logit models. *Journal of Statistical and Econometric Methods* 5(1): 59-86.

Bio-based and Applied Economics 9(3): 263-282, 2020
DOI: 10.13128/bae-7758

Full Research Article

Not my cup of coffee: Farmers' preferences for coffee variety traits – Lessons for crop breeding in the age of climate change[1]

Abrha Megos Meressa, Ståle Navrud*

School of Economics and Business, Norwegian University of Life Sciences, P.O. Box 5003, N-1432 Ås, Norway

Abstract. The advent of biotechnology and conservation of genetic resources hold promise to improve traits to meet the challenges to coffee growing from climate change. Developing new varieties by integrating traits in high demand by farmers could greatly increase farmers' adoption of new varieties. This study aims to inform breeding priority setting by examining farmers' preferences for coffee traits. A Discrete Choice Experiment was applied to smallholder farmers in northern Ethiopia to map their willingness-to-pay for improvements in four coffee traits: i) yield, ii) weather tolerance, iii) disease resistance, and iv) the maturity period. The traits are important to the farmers in their choice of coffee varieties. They prefer weather tolerant and disease resistant varieties; implying that they prefer yield stability over high yielding and early maturing varieties. Education level, access to irrigation and farmers' experience in coffee farming explain the preference heterogeneity across farmers. These results suggest that breeding programs should give priority to yield stability in order to increase farmers' adoption of new varieties, and secure *in situ* preservation of these traits. Thus, *ex situ* conservation programs are needed for early maturing and high yielding varieties, which farmers do not give priority to maintain in their own fields. This would improve climate resilience of coffee farming, and at the same time conserve the Arabica coffee genetic heritage of Ethiopia.

Keywords. Coffee, traits, crop breeding, climate change, discrete choice experiment, willingness-to-pay.

JEL Codes. Q18, Q51, Q55, Q57.

[1] We would like to acknowledge funding from the NORHED project through Capacity Building for Climate Smart Natural Resource management and Policy (CLINSRAP), a collaboration project between the Norwegian University of Life Sciences (NMBU) and Mekelle University, Ethiopia. We would also like to thank two anonymous referees for very detailed and constructive comments, which greatly helped us improve the paper. Any mistakes that remain are, of course, the sole responsibility of the authors.

*Corresponding author. E-mail: stale.navrud@nmbu.no

Editor: Meri Raggi.

ISSN 2280-6180 (print)
ISSN 2280-6172 (online)

1. Introduction

Coffee is grown by 20-25 million families in more than 80 tropical and subtropical countries (Bacon, 2005; Vega et al., 2003). Two main coffee species are grown; Arabica coffee (*coffea arabica)* and Robusta coffee (*coffea canephora*), with the former accounting for more than half of the world coffee production. Meeting the growing demand for coffee while safeguarding the genetic biodiversity of coffee is, however, a great challenge for policy makers. The advent of biotechnology and conservation of genetic resources hold promise to improve phenotypes of high economic importance and bring socially desirable outcomes.

Ethiopia is one of the world's largest coffee producing countries and known to harbor a wide range of coffee genetic diversity in a diverse array of coffee farming systems. There are more than 5,000 varieties of Arabica coffee in the country (Labouisse et al., 2008, Tsegaye et al., 2014), and they can still be found growing wild or semi-wild in the undergrowth of tropical highland forests. Ethiopian foreign exchange earnings largely depend on coffee export. There are four main coffee farming practices in Ethiopia: i) forest coffee, accounting for 8-10 % of the production, ii) semi-forest coffee (30-35 %), iii) garden coffee (50-57 %) and iv) plantations (5 %)(Kufa, 2012). Thus, 95 % of the total coffee produced can be attributed to smallholder farmers.

The productivity of forest coffee and semi-forest coffee farming is about 200-500 Kg per hectare, which is lower than the national average productivity (600 -700 Kg per hectare). The coffee species in the forests and farms vary in productivity per hectare, appearance and internal genetic structure (López-Gartner et al., 2009). The vast genetic variability in *Coffea arabica* genotypes of Ethiopia provides opportunities for creating coffee varieties, through selection and hybridization, with good yield performance, distinct quality characters, and resistance to major diseases.

The few common pests and coffee diseases include coffee berry disease (CBD) *(Colletotrichum kahawae*), coffee root-knot nematode (*Meloidogyne spp*.) and coffee rust (Muller et al., 2009; Dubale & Teketay, 2000). The threat of CBD remains prevalent in coffee growing regions despite research efforts and policy interventions encouraging planting of disease resistant coffee varieties and fungicide spraying. Pest and disease resistant cultivars yield economic benefits because they reduce yield losses and pesticide costs of coffee growers (Hein & Gatzweiler, 2006).

Previous studies and policies on annual crops narrowly focus on evaluating the benefits of high yielding varieties, but farmers' adoption of these improved varieties is low (e.g., Dalton, 2004; Shiferaw et al., 2014; Zeng et al., 2014). In addition, evidence from multi-attribute crop studies in developing countries show that farmers exhibit higher preferences for drought tolerant than high yielding crops (Asrat et al., 2010; Kassie et al., 2017). However, these studies examine farmers' preferences for crops such as teff (*Eragrostis abyssinica*) and maize. In contrast, coffee is arguably more robust to weather shocks than annual crops, but the practice of coffee farming is more challenging because of long-lasting effects of farming decision, less opportunities for inter-annual agronomic adjustments, as well as the ecological importance of preserving genetic diversity.

Farmers focus on their private economic benefits, and select and cultivate coffee varieties based on the benefits they obtain and/or expect to obtain from a particular trait

(Hein & Gatzweiler, 2006). However, farmers' emphasis on adoption of high yield coffee varieties could erode the genetic diversity of coffee in the forests and the semi-forest coffee farms. Fluctuating market price of coffee, coffee diseases, increased frequency of extreme weather events, and substitute cash crops like khat *(Catha edulis)* can also reduce the genetic diversity of coffee.

In coping with the environmental stressors, farmers' selection of coffee varieties to cultivate and maintain on their farm along with natural processes over generations of cultivation shapes the genetic structure of coffee (Baidu-Forson et al., 1997; Smale et al., 2001). Farmers' interest in increasing yield per hectare, reducing yield loss or shortening the waiting period to start harvesting a normal yield might motivate their decisions to cultivate new varieties and maintain them in their fields.

Climate change is threatening global coffee yields as changing temperatures and rainfall patterns affect plant growth. The changing climate may also be leaving coffee plants more vulnerable to diseases. Thus, in the age of climate change it is important to conserve the genetic diversity in Arabica coffee in countries like Ethiopia, as this genetic pool is likely to improve the possibilities for adapting coffee growing to future climates and secure the livelihood of smallholder coffee farmers in developing countries (FAO 2015)..

This paper aims at increasing our understanding of Ethiopian smallholder farmers' preferences for Arabica coffee traits. This knowledge can be used to construct breeding programs for coffee varieties farmers are likely to adopt, and thus conserve *in-situ.* For example, if farmers have strong preferences for high yield traits, they are more likely to maintain such varieties in their farmed fields. However, the farmers would then be less likely to cultivate or maintain other coffee varieties with lower yields, but with drought tolerance and other traits that could critically affect the future ability of coffee to adapt to climate change. In order to preserve these traits, *ex-situ* conservation efforts would be needed to supplement on the farm (*in situ*) conservation.

While previous studies of Ethiopian smallholder farmers have examined trait preferences for annual crops like teff and sorghum (Asrat et al., 2010), and found environmental adaptability and yield stability to be important, very little is known about the trait preferences of farmers for perennials like coffee.

This paper seeks to answer the following three research questions: 1) Which traits of Arabica coffee varieties do smallholder farmers prefer to cultivate? 2) Are there trait preference variations among the farmers? 3) Which sociodemographic factors explain the variations in farmers' preferences for coffee traits?

We employ a discrete choice experiment (DCE) to elicit farmers′ preferences and willingness-to-pay (WTP) for improvements in the following traits of Arabica coffee: i) yield per hectare, ii) weather tolerance, iii) diseases resistance, and iv) the maturity period. We also explore the preference heterogeneity among the smallholder farmers, and the sources of heterogeneity. The latter is found to be important for designing targeted communication programs, differentiated product offerings, and for identifying market segments and market niches (Allenby & Rossi, 1999). Thus, the results from this study can be used in the dissemination and adoption of improved coffee varieties.

2. Method and Data

2.1 Description of the Study Area

The study area is the Raya Alamata and Raya Azebo districts of the regional state of Tigray in northern Ethiopia. The study area is located about 600 km north of Addis Ababa, the capital of Ethiopia and 180 km south of Mekelle, the capital of the regional state of Tigray with about 4 million inhabitants. Most people in this rural area base their livelihood on rain-fed agriculture. The study area includes most of the Raya valley, which is one of the focal areas for agricultural expansion with its fertile soils and high agricultural potential. The Ethiopian Ministry of Water Resources initiated a hydrogeological study in the Raya valley in 2008 aiming to encourage farmers to adopt new technologies to improve productivity and ensure food security in the region (Ayenew et al., 2013). The study area, like the other regions in Ethiopia, has seen frequent variability in the weather; i.e. fewer normal years and more frequent droughts and flooding (Siam & Eltahir, 2017). Higher rainfall variability in the region has become a challenge for agriculture and environmental conservation as farmers have not adopted technologies that could mitigate crop yield losses.

Agriculture, being the main source of livelihood activity, involves a mixture of food and cash crop production. The main crops grown are maize, sorghum and teff; but also coffee and khat are found. Fruits are also grown as cash crops in the lowland areas. Although annual rainfall is moderate, ranging from 450 to 600 mm, the availability of farmland and fertile clay loam soils makes the area well suited to crop production. Since 2001/02, the regional government has made unsuccessfully efforts to get khat producers to convert to coffee production. The regional government has banned transportation, selling and buying of khat in the regional markets during the coronavirus pandemic state of emergency, and is planning to introduce new lasting laws to permanently prohibit the use and marketing of khat. One of the tentative measures proposed is to provide subsidies and other incentives to farmers that convert from khat to coffee farming. Thus. understanding farmers' preferences for coffee traits, and factors explaining potential preference heterogeneity among these farmers, could help us understand how effective alternative measures would be and improve their design.

2.2 Design of survey, choice experiment and attributes

2.2.1 Survey instrument

Discrete Choice Experiments (DCEs) enable us to study goods and attributes for which no market exists (Hanley et al 2001). We use DCE to evaluate farmers' preferences for the various traits of coffee varieties, as the other Stated Preference technique of Contingent Valuation is not able to value each individual trait. The DCE approach is based on a combination of Lancaster's household production theory (Lancaster, 1966), and McFadden's random utility theory (McFadden, 1973). Lancaster's household production theory states that the total utility of a good is derived from the characteristics or attributes of the good (Lancaster, 1966); while the random utility maximization (RUM) model is used for analyzing discrete choices, based on the assumption of utility maximizing behavior of individuals (McFadden, 1973). In DCEs, individuals are asked to make repeated hypothetical choices among alternatives in choice sets where the pre-specified levels of the different attributes vary.

The final survey instrument was designed in a stepwise process; including discussions with key informants and experts from Mekelle University, focus group discussions with the farmers; and a series of pretests of the survey instrument prior to the final survey; see table 1. We conducted pre-test surveys in April and May 2016 in four villages in the study area. In the first exploratory survey, we used a structured questionnaire, and carried out face-to-face interviews with informed village community members and local agriculture and development extension agents in the study area. The focus group discussants (N= 20, in five groups, each with four participants) and informant interviewees were used to determine the coffee attributes that were most important to them and to the community. In a pre-test survey we tested the questionnaire on a broad range of respondents in order to reflect the variation we expected to see in the final survey sample and checked whether respondents understood the questionnaire. We kept refining and clarifying the attributes and their levels using reports and opinions from discussants to make them easier for the respondents to understand.

Using information from the pre-testing, focus group discussions, key informants, model farmers and extension workers in the study area as well as discussions with experts, we selected five coffee attributes to define new coffee variety alternatives.

The questionnaire was translated into the local language (Tigrigna), and a pre-test face-to-face survey was conducted in May 2016. 36 farmers from the study area who were

Table 1. Description of process of developing the Discrete Choice Experiment (DCE) survey.

Stage	Research activity	Period	Description	Purpose
1.	Literature review and semi-structured interviews with key stakeholders in the area	March-April 2016	Identification of coffee attributes, and farming practices in the case study area	Identify relevant attributes to include in the DCE, and sociodemographic and other factors explaining farmers´ choices
2.	Focus groups (5 groups; each with four discussants; N=20), used both to explore and to pre-test a tentative version of the DCE	April-May 2016	Assess farmers' perception towards the coffee attributes and climate change	Identify and refine relevant attributes to include in the DCE exercise, and questions to map factors affecting respondents´ choices
3	Pre-test survey (N =36 face-to-face interviews)	May 2016	Test survey instrument and follow-up questions about the attributes and the credibility of the valuation scenarios/ choice cards	Check whether the choice cards and questions are found to be realistic, acceptable and understandable to the respondents
4.	Final Survey (N = 358 face-to-face interviews)	May-August 2016	Assess preferences of the local people towards different coffee attributes	Conduct the DCE exercise with the selected coffee attributes

engaged in farming activities (not only coffee production) at the time of the survey were randomly selected for the pre-test. During the pretest of the DCE, the choice sets included "quality" and "marketability" attributes, and each choice set had three alternatives and an opt-out option (i.e. none of the alternatives). Each alternative was characterized by five attributes. In the pretest, the respondents reported the choice sets to be too complex. Therefore, we changed each choice set in the final survey to include only two new alternatives and the opt-out option, where the alternatives included four non-monetary coffee attributes and a cost attribute.

Previous studies have shown that the use of labeled alternatives in DCE has a significant effect on individual choices, and could reduce respondents' attention to the actual attributes and make them look only at the labels of the alternatives (Jin, Jiang, Liu, & Klampfl, 2017). Since the goal of this study is to examine preferences for coffee traits, the choice sets comprised the unlabeled alternatives: "Alternative A" and "Alternative B"; and the opt-out alternative "Neither Alternative A nor Alternative B", having no additional cost.

The final survey was conducted from May to August 2016 by seven experienced interviewers who were trained for three days in survey techniques. They conducted face-to-face interviews of a random sample of 358 heads of farming households in the study area. During the interview, interviewers started by explaining the proposed breeding program and possible improvements in the coffee traits/attributes in order to help respondents to prepare for the choice cards. After addressing questions from the respondents, if any, the interviewers proceeded to the DCE. Afterwards, information about the sociodemographic characteristics of respondents were collected.

2.2.2 Design of attributes

The procedure in the final selection of attributes and definition of attribute levels is based on a review of previous studies (Asrat et al., 2010; Wale & Yalew, 2007), and examination of opinions expressed in the carefully crafted focus group discussions that include experienced and model farmers, ordinary farmers (mainly coffee breeders) and agricultural researchers as well as extension workers in the area. The experts on crop breeding and agricultural researchers have hands-on experience and practical knowledge about which coffee attributes are important. Similarly, the discussants reported that they considered the attributes as important for their selection of a particular coffee variety. The additional payment to fund the breeding program to improve the coffee attributes is presented as an extra cost of the seedlings for that particular coffee plant and is included along with the coffee attributes. Thus, the attributes included in the choice sets are: i) yield, ii) weather tolerance, iii) disease resistance, iv) maturity period, and v) extra cost of the seedling. Table 2 provides a description of the attributes and their levels.

Yield refers to the increase in average productivity of a coffee variety in quintal (1 quintal (Q) = 100 kg) per hectare. The improvement in yield has been emphasized by policy makers and development practitioners aiming at increasing farmers´ income and ensuring food security. The yield attribute has three levels: no change (the current yield per ha), and 1/4th (one fourth) and 1/3rd (one third) increase in productivity. The current yield per ha varies across different production systems and the coffee varieties. The average productivity in quintals per hectare (Q/ha) is 2-3 for forest coffee, 4-5 in semi-forest coffee, 7-8 for garden coffee and 9 for plantation coffee; and the national average is 6-7 Q/ha. The productiv-

Table 2. Attributes and attribute levels, including the "no change" levels of the opt-out option, used in the discrete choice experiment.

Attribute	Description	Attribute levels
Yield	Increased average productivity in terms of yield per hectare of a particular coffee variety	No change*, 1/4th increase, 1/3rd increase
Weather tolerance	Whether the coffee variety is tolerant to drought and frost and gives stable yield in the face of such weather stress factors.	No change*, Drought only tolerant, Drought and frost tolerant
Disease resistance	Whether the coffee variety gives stable yield despite the occurrences of coffee diseases or pest infections in scenarios of no drought and/or no cold weather.	No change*, Moderate disease resistant, Strong disease resistant
Maturity period	The time (in years) the coffee variety needs before giving its first normal yield.	No change*, 3 years, 5 years
Extra Cost per seedling	The additional payment, in Ethiopian Birr (ETB), an individual farmer is expected to pay per seedling	0, 7, 15, 20, 25 ETB

Notes: # ETB = Ethiopian birr; at the PPP conversion factor on 31 December 2016, 1 USD=8.68 ETB.
* "no change" in the opt-out option correspond to a maturity period of approximately 7 years for the traditional coffee varieties. No change to weather tolerance and diseases resistance traits are associated with a little drought and frost tolerance and a little disease resistance, respectively. The opt out traits/attribute levels are not included in constructing the hypothetical choice sets.

ity for selected varieties and hybrid varieties is in the range of 6-17 Q/ha and 15-24 Q/ha, respectively. Increased yield per hectare raises household income and is expected to have a positive effect on farmers´ willingness-to-pay (WTP) for seedlings of a coffee variety.

Weather tolerant and disease resistant traits are associated with the performance of the coffee variety in terms of giving a stable yield. *Weather tolerance* refers to the capacity of the coffee variety to withstand drought and frost, and to give a stable yield year after year. This attribute has three levels: no change (meaning little drought or frost tolerant), drought tolerant, and drought and frost tolerant. *Disease resistance* refers to the resilience and resistance of the coffee variety to diseases and pest infections when there is neither drought nor frost and it gives a stable yield year after year. The disease resistance attribute has three levels: no change (meaning little disease resistant), resistance only to common diseases, and high resistance to common and uncommon diseases. Increased weather tolerance and disease resistance are expected to increase farmers´ WTP for coffee traits.

Maturity period refers to the duration of time (in years) the coffee plant need to fully develop and start giving a normal yield. The maturity period attribute has two levels: five years and three years. An increase in the maturity period of the coffee is expected to have a negative effect on people's wellbeing and their preferences for the coffee variety. The *Cost* attribute is defined as *extra* costs per seedling. The average cost of a coffee seedling in the area at the time of the survey was approximately ETB 5-7.

2.2.3 Experimental design

This study employs an orthogonal main effect experimental design (OMED) to combine attribute levels and create choice sets. In creating the choice sets, we used the R soft-

Figure 1. Example of a choice card as it appeared in the questionnaire in the final survey. The "Neither A nor B" alternative to the right is the opt-out option.

Which of the following coffee varieties do you prefer? Alternative A and Alternative B would entail a cost to your household, while no payment would be required for the "Neither" option

	Alternative A	Alternative B	Neither Alternative A nor Alternative B: I prefer none of the new varieties
Yield	1/4th increase	1/3rd increase	
Weather tolerance	Drought and frost	Drought	
Disease resistance	Disease resistant	Disease resistant	
Maturity duration	3 years	5 years	
Cost per seedling	ETB 5	ETB 20	

I would prefer: Alternative A ____ Alternative B ___ Neither ___

Note: ETB = Ethiopian birr; 1 USD=8.68 ETB in terms of Purchase Power Parity (PPP) corrected exchange rate on December 31st, 2016.

ware version 3.3.2 and adopted the code by Aizaki (2012) to execute the experimental design and randomly assign the choice sets into two blocks. The experimental design creates 16 choice sets, and the two blocks include 8 choice sets each. Figure 1 shows a choice set as it was presented in a choice card to the respondents. The choice tasks put respondents in a hypothetical setting, offering them choice sets comprising two new alternative coffee varieties (presented as "Alternative A" and "Alternative B"), and an opt-out option ("Neither Alternative A nor B"). The two new coffee varieties come at an extra cost of the seedling in order to cover the costs of developing a new variety. The opt-out option has no extra cost of the seedlings as the farmers will then have the traditional coffee variety. The alternatives in the choice sets differ in one or more of the attribute levels.

The respondents are randomly assigned to the two blocks, and asked to choose his or her most preferred alternative in a sequence of eight choice sets. The respondents are subjected to only eight choice sets each, with the aim of attaining a balance between fatigue and learning (Caussade et al., 2005).

Similar to Meyerhoff and Liebe (2009), this study imposed restrictions to avoid unrealistic choice tasks by making the new alternatives have at least one higher attribute level than the opt-out alternative. This avoids new alternatives having inferior values to the opt-out option, but they can have higher extra costs. However, dominant choices created from the experimental design were also presented to the respondents as the removal of irrational or inferior preferences from the choice experiments could affect statistical efficiency (Lancsar & Louviere, 2006). Besides, the presence of new alternatives with higher/lower non-monetary attribute levels but less/equal cost (dominant/dominated alternatives) than other alternatives could help to examine whether respondents pay enough attention to and understand the choice task. Further, having generic alternatives such as "Alternative A" and "Alternative B" can make respondents focus on the attributes/traits rather than the labels we could have put on the alternatives/ coffee varieties.

2.3 Sample characteristics

In the final survey we interviewed 358 farmers residing in the rural areas of Raya Alamata and Raya Azebo districts of Tigray in northern Ethiopia. We applied proportional sampling to give larger quota to districts and villages with larger population and vice versa, and systematic random sampling to select farmers from household head name lists in sub-district offices. According to the most recent Ethiopian Central Statistical Agency census report (CSA, 2007), the total number of households in Raya Alamata and Raya Azebo was 20,532 and 32,360, respectively. Accordingly, the proportion of sampled household heads from the two districts was 60 percent from Raya Azebo and 40 percent from Raya Alamata. The sociodemographic characteristics of the farmers are presented in Table 3.

2.4 Model specification and estimation

The conditional logit model is commonly used to analyze consumer choice behavior based on random utility theory (McFadden, 1974). Conditional logit assumes the idiosyncratic errors to be independently and identically distributed (IID) extreme values, and the tastes for observed attributes to be homogeneous. Evidence shows that individuals exhibit significant heterogeneity in preferences for goods and services (see Alberini & Ščasný, 2013; Allenby & Rossi, 1999; Birol, Karousakis, & Koundouri, 2006). Mixed logit (MIXL) models relax the independence of irrelevant alternative (IIA) assumption of the more restrictive closed-form discrete choice models and allows for heterogeneity of preferences for observed attributes (Hensher & Greene, 2003; McFadden & Train, 2000). In this model, utility U is assumed to be latent, but observed only with the choice Y of alternative j (0, 1, 2) by individual i (i=1, … 358) in choice set t (t=1,2, … 8). A utility function given a choice set t with j alternatives for individual i can be written as;

$$U_{ijt}=\beta_i X_{ijt}+\varepsilon_{ijt}$$

Table 3. Description of sociodemographic variables used to explain the variations in farmers' preferences for the selected coffee traits.

Variable	Mean	Median	Std. Dev.	Definition
Age	43.2	40	13.6	Age of the household head; in years
Family size	5.6	6	2	Total number of family members in the household (including the respondent)
Education	1.8	0	3	Education level of household head; in years
Market	60	60	49.9	The distance to the main market from home; walking time in minutes
Farm size	2.9	3	1.9	The area of the farmed land the farmer owns; in Timad (1 hectare= 4 Timad)
Irrigable land	0.44	-	0.5	Whether the farmer owns irrigable land; 0=No; 1=Yes
Experience	0.28	-	0.47	Whether the farmer has ever managed a coffee farm (now or before); 0=No; 1=Yes

where X_{ijt} is a vector of observed explanatory variables including coffee attributes and sociodemographic characteristics, β_i is a vector of conformable parameters (unknown utility weights) the individual assigns to these variables; and ε_{ijt} is a random term that does not depend on underlying parameters or observed data, with zero mean and IID over alternatives. The utility weight (β_i) for a given attribute is given as;

$$\beta_i=\beta+\delta'_i\nu_{ij}$$

Where β is a vector of mean attribute utility weights in the population, δ is a diagonal matrix which contains the standard deviation (σ) of the distribution of the individual taste parameters (β_i) around the mean taste parameter (β), and ν_{ij} is the individual specific heterogeneity with mean equal to 0 and standard deviation of 1. The MIXL model permits random parameters to vary over individuals, and not observation, in order to measure interpersonal heterogeneity. The vector X_{ijt}, can include 0/1 terms to allow for alternative specific constant (ASC), where ASC takes the value 1 for "Alternative A" and "Alternative B" and 0 for the opt-out option. ASC accounts for the systematic differences in choice patterns between the alternatives. Behaviorally speaking, the ASC parameter reflects the average effect of various components such as endowment effect, status quo bias, omission bias, and the impacts of complexity such as fatigue effects and other unobserved attributes (Boxall et al, 2009; Meyerhoff & Liebe, 2009). The inclusion of an opt-out option can also reflect actual behavioral phenomena by avoiding forced demand, and hence improves the reliability of the welfare measures (Boxall et al., 2009; Veldwijk et al., 2014).

We set the parameters on yield, weather tolerance, disease resistance and maturity period attributes as random and with normal distribution, and the parameter on the cost attribute is set as fixed. A positive sign for significant coefficients of the attributes in the econometric estimation indicates a positive effect of the increase in the respective attribute on farmers' preferences, whereas a negative sign indicates a negative effect of the attribute on their preferences. Statistically significant coefficients on the attributes also enable the calculation of WTP for a change in the attribute. In a utility function linear in its parameters, the marginal WTP equals the negative ratio of the respective coefficient of non-monetary attribute and the coefficient of the monetary attribute (Hensher & Greene, 2011). The WTP estimates presented in Table 4 refer to a marginal, one level change in the attributes. The attributes levels included in this model are presented in Table 2, and the sociodemographic variables are defined in Table 3.

The coefficients in MIXL models are estimated with a simulated maximum likelihood estimation technique. This study used the gmnl-package by (Sarrias & Daziano, 2017) in R software version 3.3.2 to estimate the coefficients on alternative attributes and sociodemographic variables. Since the sociodemographic variables do not vary across choices/observations, their interaction with ASC are included to test whether they explain the observed taste variations across farmers or are random parameters across individuals. Akike information criteria (AIC), Bayesian information criteria (BIC) and likelihood ratio tests are used to compare the goodness of fit of the model and select the model with superior goodness of fit compared to other models. The inclusion of the sociodemographic variables in the MIXL model is used to uncover the factors explaining farmers' preference heterogeneity.

3. Results and Discussions

Standard multinomial logit (MNL) models were estimated first, before proceeding to MIXL models. Table 4 presents the results. Other models such as Scaled-multinomial logit model and generalized multinomial logit model were also estimated; see appendix A-1. The results from the MIXL models show superior fit to the data in this study. In the MIXL estimation, we set the coefficients on the attributes yield, weather tolerance, disease resistant and maturity duration to be random parameters with normal distribution, while the coefficient on the cost of seedlings is fixed in order to use it to compute WTP estimates. The maturity duration and cost of seedlings attributes are continuous variables; while the yield, weather tolerance and disease resistance attributes are categorical.

The coefficient on ASC is significant and positive, implying that farmers prefer the new alternative varieties at some additional cost to the existing varieties that come at no additional cost. Less than two percent of the respondents chose the opt-out option, but none of these respondents protested the proposed coffee variety development program and the changes in traits/attributes. Although the interviewers were trained to avoid experimenter demand effects (Zizzo, 2010), i.e. the respondent trying to please the interviewer by saying what they assumes the interviewer would like to hear, we cannot rule out that this effect might have contributed to the low opt-out percentage.

ASC captures the average effect of all relevant factors that are not included in the model. Thus, farmers´ choice of new improved varieties over the traditional ones seem to be motivated not only by coping with frequent weather changes and occurrence of coffee diseases, but also by the desire for high yield and early maturing traits.

Results from the MIXL model show that the estimated coefficients on yield, weather tolerance, disease resistance and maturity duration are all statistically significant. This implies that any developments in the specified coffee traits have significant effects on farmers' preferences for coffee varieties. The parameter on the yield attribute is interpreted in relation to an increase in productivity per hectare or an increase in farm income resulting from cultivating a coffee variety. The weather tolerance trait enhances resilience against drought and frost, while the disease resistance trait increases resilience against coffee diseases and pest infections occurring under "no drought" and "no frost" weather conditions. Thus, the coefficients on disease resistant and weather tolerant traits can be interpreted as farmers´ preferences for yield stability or resilience to risk of yield loss, and hence is also as an indicator of farmers` risk preferences. The parameter for the maturity period attribute reflects the time preference of farmers. The signs of the coefficients for all attributes/traits are consistent with standard economic theory as farmers prefer increased weather tolerance, higher disease resistance, and higher yield per hectare, but reduced duration of the maturity period and lower extra cost per seedling.

The significant and positive coefficient for the yield attribute implies that farmers prefer high yield coffee varieties to low yield coffee varieties, holding all other things constant. This implies that improvement in productivity per hectare of a coffee variety increases the farmer's preference for this variety. Previous DCE studies of annual crops (Asrat et al., 2010; Kassie et al., 2017) showed farmers to have similar positive preferences for the yield improvement attribute.

Table 4. Results of the MNL model and MIXL models without (MIXL1) and with (MIXL2) sociodemographic determinants of preferences heterogeneity.

	MNL model	MIXL1 model	MIXL2 model
ASC	4.621***	8.750***	6.825***
	(0.220)	(0.572)	(0.600)
Yield high	0.754***	1.078***	0.838***
	(0.065)	(0.117)	(0.231)
Weather tolerant	0.970***	1.292***	1.453***
	(0.067)	(0.135)	(0.284)
Disease resistant	0.929***	1.425***	2.713***
	(0.061)	(0.131)	(0.521)
Maturity duration	-0.452***	-0.548***	-0.665***
	(0.034)	(0.071)	(0.129)
Cost of seedling	-0.044***	-0.058***	-0.065***
	(0.005)	(0.006)	(0.009)
Yield high. Experience			0.028*
			(0.012)
Weather tolerant. Irrigation			-0.001*
			(0.001)
Disease resistant. Education			-0.018*
			(0.009)
Disease resistant. Age			-0.063
			(0.045)
Maturity duration. Education			0.051**
			(0.018)
Maturity duration. Market			-0.005
			(0.004)
Maturity duration. Experience			-0.001*
			(0.001)
N	2860	2860	1869
Log-likelihood	-1765.161	-1594.251	-1131.047
BIC (BIC/N)	3578.073 (1.251)	3315.839 (1.159)	2435.356 (1.303)
AIC (AIC/N)	3542.321 (1.239)	3220.502 (1.126)	2308.094 (1.235)

Note: Standard error in parentheses. ***, ** and * denote significant at the 1, 5 and 10 % level; respectively.

Weather tolerant and disease resistant attributes are associated with the ability of a coffee variety to withstand environmental stressors and to give stable yield. The estimated coefficients for these two attributes are consistently significant and positive. This could imply that farmers are willing to pay more for seedlings with these traits and are thus willing to give up part of their income in order to ensure stable yield. A DCE by Asrat et al. (2010) assessing the trait preferences of Ethiopian farmers for sorghum and teff crop varieties showed that farmers are willingly forego some income or yield to obtain a more stable and environmentally adaptable crop variety. The coefficient on the maturity period is significant and negative, indicating that farmers prefer early maturing coffee varieties over those coffee varieties that take longer to start giving normal yield. Similarly, experimental evidence on rice traits in western Africa showed farmers to be willing to pay for early maturing traits (Dalton, 2004) Note, however, that both Asrat et al. (2010) and Dalton (2004) looked at annual crops, while coffee is a perennial crop.

Policy makers often stress the importance of high yield varieties to meet the growing demand for food, but adoption of high yielding variety technologies is low. Our study shows that farmers are willing to pay more for improving traits associated with yield stability, such as weather tolerant and diseases resistant traits, than for increasing the yield per hectare or early maturity. The magnitude of the coefficients corresponds to the importance the farmers put on the traits. In a related study, Kassie et al. (2017) examined farmers' preferences for drought tolerant maize in rural Zimbabwe, and found that farmers are willing to pay five times more for a variety with a drought tolerance trait than for a variety providing an additional ton of yield per hectare. This implies that farmers are willing to forgo an increase in yield per hectare to get a stable yield on the farm. The subsistence nature of agriculture and escalated poverty in the area might restrain them from adopting a high yield cash crop variety technology with some risk and keep farmers trapped with a low yield and low cost variety technology.

Table 4 also reports the coefficients of sociodemographic factors that can explain preference heterogeneity among the farmers. Heterogeneity around the mean of the taste parameters is consistently apparent with respect to yield, weather tolerance, diseases resistance, and maturity duration traits. Therefore, we included age, education, experience with coffee farming, access to irrigation and distance to market in order to assess the observed sources of variation and to identify factors responsible for the heterogeneity. Note that the models in Table 4 are not directly comparable in the conventional model fit criteria of log likelihood, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC); as the number of observations in the model with the sociodemographic factors (MIXL2) is much smaller than in the models without these variables. Although BIC divided by number of observations (BIC/N) is higher in MIXL2, this is not the case for the AIC/N. Thus, we cannot conclude that the inclusion of these sociodemographic factors increases the model fit. We focus on the estimates from the MIXL model since the results demonstrate the presence of preference heterogeneity among the farmers. Education, access to irrigation, and experience of the farmer in coffee farming were found to be the factors that explain variation around the average level of taste preference for the traits. About 28% of the respondents reported having some experience in coffee farming activities, which explains preference variations for high yield and early maturing traits.

Considering the high yield trait, farmers with experience in coffee farming exhibit higher preferences for improvements of yield per hectare than farmers without experience. Some farmers in the study area are replacing low yield coffee varieties with improved coffee varieties, while others are shifting towards cultivation of other more lucrative cash crops such as khat. Farmers with relatively high levels of literacy are found to have lower preferences for disease resistant traits. This finding coincides with Gächter et al (2007) that found increased level of education to decrease loss aversion. On the other hand, farmers with better access to irrigation reveal lower preferences for weather tolerant coffee traits than the farmers who have no access to irrigation. This is as expected as farmers' lack of access to irrigation could increase their vulnerability to drought, and thus their risk aversion.

The coefficient on the maturity duration attribute is negative. A negative significant coefficient on maturity duration indicates that an increase in maturity duration of the coffee variety reduces farmers' preferences for that particular variety. Farmers' years of education reduces the negative effect of increasing maturity duration of late-maturing coffee varieties, whereas coffee farming experience increases the negative effect of increasing maturity duration. The could be explained by farmers′ private discount rate to increase with age and decrease with educational level and literacy, as observed by (Kirby et al., 2002). These days, almost the entire coffee farming area in the study area has been turned into production of khat and other cash crops. Thus, farmers with coffee farming experience tend to be older, and older farmers could have higher private discount rates and thus prefer early maturing traits.

In DCE analysis, the coefficients in themselves have no direct economic interpretation, but the negative ratio of the coefficients of the attribute to the cost coefficient give the marginal WTP estimate for the changes in the attributes (Hensher & Greene, 2003). Positive and negative marginal WTP estimates reflect utility and disutility of the attribute, respectively. The WTP for a change in an attribute level combined with the increment in the attribute level, leaves the deterministic part of the respondent's utility for a profile unchanged (Fiebig et al 2010) Table 5 presents the marginal WTP of the four coffee traits.

Observing the marginal WTP estimates (deferring the heterogeneity, i.e. the MIXL2 model), the farmers are willing to pay more for frost and drought tolerance as well as disease resistance traits, compared to increased yield. The premium is 2-3 times the amount they are willing to pay for a 1/3rd increase in the yield of 1 quintal/ha (1 quintal = 100 kg). This compares well with a similar study of farmers' preference for maize traits in Zimbabwe. Kassie et al. (2017) showed that the value farmers attach to drought tolerance is about five times higher than the WTP they attach to changing a variety. Our results also reflect the difficulties in making inter-annual adjustment in coffee farming practices. These results can explain the prevailing low adoption of high yield varieties by farmers in Ethiopia (Wale & Yalew, 2007).

The coefficient on the maturity period is significant and negative, which implies that an early maturity trait is more preferred to a late maturing trait. The negative sign implies

Table 5. Marginal WTP; in Ethiopian Birr (ETB) (1 USD=8.68 ETB in terms of Purchase Power Parity (PPP) corrected exchange rate on December 31st 2016).

Attributes	WTP Estimates from the MIXL1 model	WTP estimates from the MIXL2 model
ASC	150	105
Yield, high	18	13
Weather tolerant	22	22
Disease resistant	24	42
Maturity period	-9	-10

that farmers are willing to give up part of their income or yield to shorten the waiting period for the full development of the coffee plant and to start harvesting normal yield. In other words, farmers have disutility from a delay in the time it takes for the coffee seedling to give normal yield.

The significant and positive coefficient on ASC implies that other unobservable systematic factors also increase farmers' preferences for new alternative coffee variety over traditional varieties.

To summarize, the WTP results confirms that farmers prefer stable yield varieties (i.e. high disease resistant and weather tolerant traits) to high yield varieties or early maturing varieties, holding all other things constant.

4. Conclusion

Understanding farmers' preferences for coffee traits can help develop policies and breeding programs for new varieties that integrate traits in demand by the farmers, and thus increase farmers' adoption of new varieties. Using a discrete choice experiment, this paper examines farmers' preferences for increased yield, weather tolerance in terms of adaptation to drought and frost, disease resistance, and early maturing traits of Arabica coffee. The results show that farmers are willing to cultivate and pay more for weather tolerant and disease resistant coffee varieties than high yielding and early maturing ones. This indicates that farmers prefer improvements in yield stability traits to traits that maximize yields. Thus, crop-breeding programs aiming for larger uptake of new coffee varieties among farmers in order to increase coffee production should primarily develop weather tolerant and disease resistant varieties and combine them with high yield and early maturing traits.

The trait preferences of smallholder farmers also have implication for *in-situ* versus *ex-situ* conservation of coffee genetic diversity in Ethiopia. Smallholder farmers with no experience in coffee farming will not cultivate and maintain coffee varieties in their fields if yields are unstable, as they prefer the yield stability traits of weather tolerance and disease resistance. Thus, the uptake of varieties with high yield and early maturing traits will be low among farmers in regions without a history of coffee growing. *Ex-situ* conservation programs should therefore give priority to coffee varieties with these and other traits that are less preferred by farmers in order to preserve the full genetic heritage Ethiopian coffee.

Although farmers prefer stable yield to high yield traits, the mixed logit model results show heterogeneity in farmers' preferences for the coffee traits. Farmers with coffee farming experience exhibited higher preferences for high yielding and early maturing coffee traits than those that had no experience in coffee farming. In contrast, farmers with more years of education prefer maturing traits and disease resistant traits less than those with little education. Further, farmers with access to irrigable farmland exhibit lower preferences for weather tolerant traits. This implies that tailoring the improved coffee varieties to the preferences of these different groups of farmers would enhance farmers' adoption of the new varieties. This could make a significant contribution to improving coffee farmers' adaptation and resilience to climate change.

Future research is needed in order to test whether our findings on smallholder farmers' preferences can be generalized to other coffee growing regions in Ethiopia and around the world. Such new stated preference surveys should be based on best practice guidelines;

see Johnston et al (2017). Preferably, similar surveys should be carried out at the same time in different regions in order to better understand what measures are needed for coffee farmers to adapt to climate change impacts.

References

Aizaki, H. (2012). Basic functions for supporting an implementation of choice experiments in R. *Journal of Statistical Software, 50*, 1-24.

Alberini, A., & Ščasný, M. (2013). Exploring heterogeneity in the value of a statistical life: Cause of death v. risk perceptions. *Ecological Economics, 94*, 143-155.

Allenby, G. M., & Rossi, P. E. (1999). Marketing models of consumer heterogeneity. *Journal of Econometrics, 89*(1-2), 57-78.

Asrat, S., Yesuf, M., Carlsson, F., & Wale, E. (2010). Farmers' preferences for crop variety traits: Lessons for on-farm conservation and technology adoption. *Ecological Economics, 69*(12), 2394-2401.

Ayenew, T., GebreEgziabher, M., Kebede, S., & Mamo, S. (2013). Integrated assessment of hydrogeology and water quality for groundwater-based irrigation development in the Raya Valley, northern Ethiopia. *Water International, 38*(4), 480-492.

Bacon, C. (2005). Confronting the coffee crisis: can fair trade, organic, and specialty coffees reduce small-scale farmer vulnerability in northern Nicaragua? *World Development, 33*(3), 497-511.

Baidu-Forson, J., Ntare, B. R., & Waliyar, F. (1997). Utilizing conjoint analysis to design modern crop varieties: empirical example for groundnut in Niger. *Agricultural Economics, 16*(3), 219-226.

Bello, M., Abdoulaye, T., Abdulai, A., Wossen, T., & Menkir, A. (2019). Modeling farmers' willingness to pay for stress tolerance in maize in northern Nigeria: how does the states of gains and losses matter? 2019 Sixth International Conference, September 23-26, 2019, Abuja, Nigeria Paper 295797, African Association of Agricultural Economists (AAAE).

Bertrand, B., Etienne, H., Cilas, C., Charrier, A., & Baradat, P. (2005). Coffea arabica hybrid performance for yield, fertility and bean weight. *Euphytica, 141*(3), 255-262.

Birol, E., Karousakis, K., & Koundouri, P. (2006). Using a choice experiment to account for preference heterogeneity in wetland attributes: the case of Cheimaditida wetland in Greece. *Ecological Economics, 60*(1), 145-156.

Boxall, P., Adamowicz, W. L., & Moon, A. (2009). Complexity in choice experiments: choice of the status quo alternative and implications for welfare measurement. *Australian Journal of Agricultural and Resource Economics, 53*(4), 503-519.

Caussade, S., de Dios Ortúzar, J., Rizzi, L. I., & Hensher, D. A. (2005). Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B: Methodological, 39*(7), 621-640.

Coromaldi, M., Pallante, G., & Savastano, S. (2015). Adoption of modern varieties, farmers' welfare and crop biodiversity: Evidence from Uganda. *Ecological Economics, 119*, 346-358.

Dalton, T. J. (2004). A household hedonic model of rice traits: economic values from farmers in West Africa. *Agricultural Economics, 31*(2-3), 149-159.

Dubale, P., & Teketay, D. (2000). *The need for forest coffee germplasm conservation in Ethiopia and its significance in the control of coffee diseases.* Paper presented at the the proceedings of the workshop on control of coffee berry disease in Ethiopia.

FAO (2015). Voluntary guidelines to support the integration of genetic diversity into national climate change adaptation planning: Food and Agricultural Orgnasation of the United Nations, Rome, Italy. 50 pp. ISBN: 978-92-5-108882-1. http://www.fao.org/documents/card/en/c/290cd085-98f3-43df-99a9-250cec270867/

Fiebig, D. G., Keane, M. P., Louviere, J., & Wasi, N. (2010). The generalized multinomial logit model: accounting for scale and coefficient heterogeneity. *Marketing Science, 29*(3), 393-421.

Gächter, S., Johnson, E. J., & Herrmann, A. (2007). Individual-level loss aversion in riskless and risky choices. Discussion Paper no. 2007-02, The Centre for Decision Research and Experimental Economics, School of Economics, University of Nottingham.

Hein, L., & Gatzweiler, F. (2006). The economic value of coffee (Coffea arabica) genetic resources. *Ecological Economics, 60* (1), 176-185.

Hensher, D. A., & Greene, W. H. (2003). The mixed logit model: the state of practice. *Transportation, 30*(2), 133-176.

Hensher, D. A., & Greene, W. H. (2011). Valuation of travel time savings in WTP and preference space in the presence of taste and scale heterogeneity. *Journal of Transport Economics and Policy (JTEP), 45*(3), 505-525.

Jin, W., Jiang, H., Liu, Y., & Klampfl, E. (2017). Do labeled versus unlabeled treatments of alternatives' names influence stated choice outputs? Results from a mode choice study. *PloS one, 12*(8): e0178826. https://doi.org/10.1371/journal.pone.0178826

Johnston, R. J. K. J. Boyle, W. Adamowicz, J. Bennett, R. Brouwer, T. A. Cameron, W. M. Hanemann, N. Hanley, M. Ryan, R. Scarpa, R. Tourangeau, C. A. Vossler (2017). Contemporary guidance for stated preference studies. *Journal of the Association of Environmental and Resource Economists*, 4(2):319–405.

Kassie, G. T., Abdulai, A., Greene, W. H., Shiferaw, B., Abate, T., Tarekegne, A., & Sutcliffe, C. (2017). Modeling preference and willingness to pay for drought tolerance (DT) in maize in rural Zimbabwe. *World Development, 94*, 465-477.

Kirby, K. N., Godoy, R., Reyes-Garcıa, V., Byron, E., Apaza, L., Leonard, W., Pèrez, E, Vadez, V. & Wilkie, D. (2002). Correlates of delay-discount rates: Evidence from Tsimane' Amerindians of the Bolivian rain forest. *Journal of Economic Psychology, 23*(3), 291-316.

Kufa, T. (2012). *Recent coffee research development in Ethiopia.* Paper presented at the Presentation at the "Ethiopian Coffee Export Conference: Strengthening the Legacy of Our Coffee. November 8-9, 2012, Addis Ababa, Ethiopia.

Labouisse, J.-P., Bellachew, B., Kotecha, S., & Bertrand, B. (2008). Current status of coffee (Coffea arabica L.) genetic resources in Ethiopia: implications for conservation. *Genetic Resources and Crop Evolution, 55*(7), 1079-1093.

Lancaster, K. J. (1966). A new approach to consumer theory. *The journal of political economy*, 132-157.

Lancsar, E., & Louviere, J. (2006). Deleting 'irrational'responses from discrete choice experiments: a case of investigating or imposing preferences? *Health Economics, 15*(8), 797-811.

López-Gartner, G., Cortina, H., McCouch, S. R., & Moncada, M. D. P. (2009). Analysis of genetic structure in a sample of coffee (Coffea arabica L.) using fluorescent SSR markers. *Tree genetics & genomes, 5*(3), 435-446.

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior.

McFadden, D. (1974). Analysis of Qualitative Choice Behavior. Zarembka, P.(ed.): Frontiers in Econometrics. In: Academic Press. New York, NY.

McFadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Journal of applied Econometrics, 15*(5), 447-470.

Meyerhoff, J., & Liebe, U. (2009). Status quo effect in choice experiments: empirical evidence on attitudes and choice task complexity. *Land economics, 85*(3), 515-528.

Michler, J. D., Tjernström, E., Verkaart, S., & Mausch, K. (2019). Money matters: The role of yields and profits in agricultural technology adoption. *American Journal of Agricultural Economics, 101*(3), 710-731.

Muller, R. A., Berry, D., Avelino, J., & Bieysse, D. (2009). Coffee diseases. *Coffee: growing, processing, sustainable production. A guidebook for growers, processors, traders and researchers*, 495-549.

Ortega, D. L., Waldman, K. B., Richardson, R. B., Clay, D. C., & Snapp, S. (2016). Sustainable intensification and farmer preferences for crop system attributes: evidence from Malawi's central and southern regions. *World Development, 87*, 139-151.

Sarrias, M., & Daziano, R. (2017). Multinomial logit models with continuous and discrete individual heterogeneity in R: the gmnl package. *Journal of Statistical Software, 79*(2), 1-46.

Shiferaw, B., Kassie, M., Jaleta, M., & Yirga, C. (2014). Adoption of improved wheat varieties and impacts on household food security in Ethiopia. *Food Policy, 44*, 272-284.

Siam, M., & Eltahir, E. A. B. (2017). Climate change enhances interannual variability of the Nile river flow. *Nature Clim. Change, 7*(5), 350-354. doi:10.1038/nclimate3273

Smale, M., Bellon, M. R., & Gomez, J. A. A. (2001). Maize Diversity, Variety Attributes, and Farmers' Choices in Southeastern Guanajuato, Mexico. *Economic Development and Cultural Change, 50*(1), 201-225.

Teressa, A., Crouzillat, D., Petiard, V., & Brouhan, P. (2010). Genetic diversity of Arabica coffee (Coffea arabica L.) collections. *Ethiopian journal of applied sciences and technology, 1*(1), 63-79.

Tsegaye, B., Mohammed, A., Shimber, T., Getachew, Y., & Getachew, E. (2014). The Influence of sun drying methods and layer thickness on the quality of lowland Arabica coffee varieties at Gomma I, Southewesten Ethiopia. *Research Journal of Agriculture and Environmental Management. Vol, 3*(11), 547-554.

Vega, F. E., Rosenquist, E., & Collins, W. (2003). Global project needed to tackle coffee crisis. *Nature, 425*(6956), 343-343.

Veldwijk, J., Lambooij, M. S., de Bekker-Grob, E. W., Smit, H. A., & De Wit, G. A. (2014). The effect of including an opt-out option in discrete choice experiments. *PloS one, 9*(11), e111805.

Waldvogel, S. R. (2003). Caffeine - a drug with a surprise. *Angewandte Chemie International Edition, 42*(6), 604-605.

Wale, E., & Yalew, A. (2007). Farmers' Variety Attribute Preferences: Implications for Breeding Priority Setting and Agricultural Extension Policy in Ethiopia. *African Development Review, 19*(2), 379-396.

Ward, P. S., Ortega, D. L., Spielman, D. J., & Singh, V. (2014). Heterogeneous demand for drought-tolerant rice: Evidence from Bihar, India. *World Development, 64*, 125-139.

Zeng, D., Alwang, J., Norton, G. W., Shiferaw, B., Jaleta, M., & Yirga, C. (2014). *Agricultural technology adoption and child nutrition: Improved maize varieties in rural Ethiopia.* Paper presented at the Selected Paper prepared for presentation at the Agricultural and Applied Economics Association's 2014 AAEA Annual Meeting.

Zizzo, D. J. (2010). Experimenter demand effects in economic experiments. *Experimental Economics*, 13(1):75–98.

Appendices

Table A-1. Results from Multinomial logit (MNL), Scaled Multinomial logit (S-MNLl, Mixed logit model with correlated alternatives (MIXL), Mixed logit model without correlation (MIXL_U) and generalized multinomial logit (G-MNL) models.

	MNL	S-MNL	MIXL_U	MIXL	G-MNL
ASC	4.621***	25.530	8.512***	8.392***	9.636***
	(0.220)	(16.563)	(0.559)	(0.512)	(0.813)
Yield high	0.754***	1.907**	1.067***	1.041***	1.198***
	(0.065)	(0.701)	(0.108)	(0.113)	(0.143)
Weather tolerant	0.970***	2.309*	1.421***	1.252***	1.342***
	(0.067)	(0.965)	(0.125)	(0.122)	(0.145)
Disease resistant	0.929***	2.092**	1.366***	1.388***	1.631***
	(0.061)	(0.717)	(0.119)	(0.127)	(0.173)
Maturity duration	-0.452***	-1.406*	-0.734***	-0.493***	-0.593***
	(0.034)	(0.595)	(0.069)	(0.063)	(0.065)
Cost seedling	-0.044***	-0.112*	-0.064***	-0.056***	-0.065***
	(0.005)	(0.046)	(0.006)	(0.006)	(0.007)
Tau		1.410***			0.477***
		(0.323)			(0.091)
Gamma					-0.648
					(0.354)
N	2860	2860	2860	2860	2860
Log-likelihood	-1765.161	-1751.089	-1632.741	-1596.428	-1577.198
BIC	3578.073	3557.888	3345.067	3320.192	3297.651
AIC	3542.321	3516.178	3285.482	3224.855	3190.396

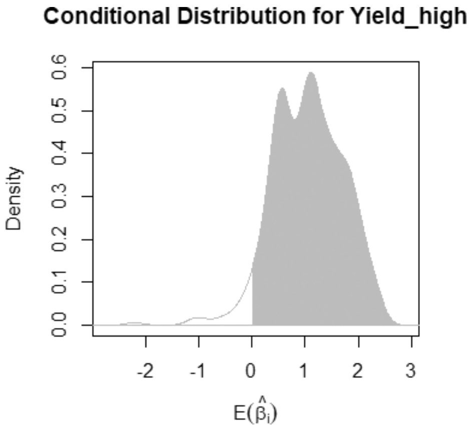
Notes: ***, ** and * denotes significant at the 1, 5 and 10 % level; respectively. Standard error in parentheses.

Table A-2. Standard deviations of the random parameters from mixed logit model results.

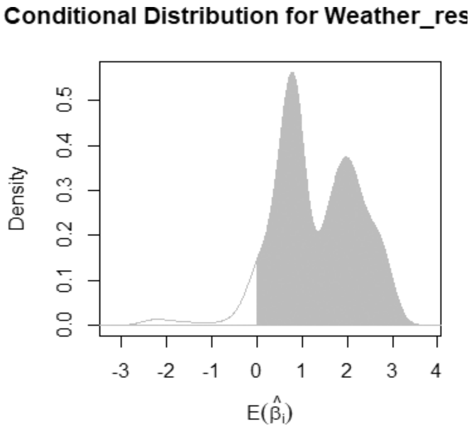
	Estimate	Std. Error	z-value	Pr(>|z|)
Yield high	1.0931	0.1985	5.51	3.7e-08 ***
Weather tolerant	1.3818	0.1906	7.25	4.2e-13 ***
Disease resistant	1.3674	0.2494	5.48	4.2e-08 ***
Maturity duration	0.6452	0.0964	6.69	2.2e-11 ***

Note: ***, ** and * denotes significant at the 1, 5 and 10 % level; respectively.

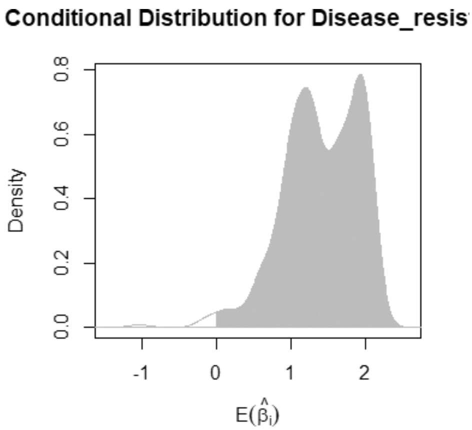
Figure A-1. Distribution of the individuals' conditional mean for the parameters of yield, weather tolerant, diseases resistant and maturity duration. The grey area displays the proportion of individual with positive conditional mean.



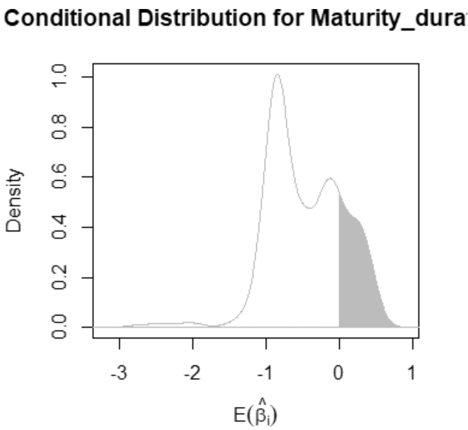
a) Kernel density for yield improvement



b) Kernel density for weather tolerant



c) Kernel density for Disease resistant



d) Kernel density for Maturity duration

Bio-based and Applied Economics 9(3): 283-304, 2020
DOI: 10.13128/bae-7764

Full Research Article

Does the place of residence affect land use preferences? Evidence from a choice experiment in Germany

Julian Sagebiel[1,*], Klaus Glenk[2], Jürgen Meyerhoff[3]

[1] Swedish University of Agricultural Sciences
[2] SRUC, Land Economy and Environment Research Group
[3] Technische Universität Berlin

Abstract. Discrete choice experiments can be used to inform policy makers on people's preferences for landscapes and cultural ecosystem services. Recent studies have shown that the spatial context influences preferences and related willingness to pay values. In this paper we investigate the effect of the landscape surrounding people's places of residence on their willingness to pay using data from a discrete choice experiment on local land-use changes and cultural ecosystem services throughout Germany. For analysis, we apply a latent class logit model and include landscape categories as explanatory variables for class membership. We find that the different landscapes people live in are correlated with preferences. Especially people from urban areas and farm- and grassland landscapes have larger willingness to pay values for improvements in cultural ecosystem services than people from forest landscapes and cultural landscapes. The results are important for policy makers as different willingness to pay values in different landscapes imply different welfare effects for land use changes. Taking this information into account can help in reaching more efficient resource allocations.

Keywords. Landscape preferences, latent class model, spatial heterogeneity, willingness to pay.

JEL Codes. Q51, Q57.

1. Introduction

Policy makers at different scales initiate land use changes to conform with subordinated laws and guidelines. Decisions should balance social and private costs and benefits for different stakeholder groups and the local population. Rigorous cost-benefit analysis is often difficult to conduct, as most regulating and cultural ecosystem services that are produced by landscapes are not traded in markets, making it impossible to directly observe societal demand for them. Benefit estimates of changes in ecosystem service provision need to be inferred through the use of non-market valuation techniques; in particular stated

*Corresponding author. E-mail: julian.sagebiel@slu.se

Editor: Meri Raggi.

ISSN 2280-6180 (print)
ISSN 2280-6172 (online)

preference methods, which allow estimation of willingness to pay through direct elicitation of preferences in hypothetical markets. In Europe, several non-market valuation studies assessing preferences for components and management of agrarian landscapes have been conducted, but they rarely accounted for spatial differences in preferences (Zanten et al. 2014; Glenk et al. 2019). The few studies that considered spatial heterogeneity in preferences found that the place of residence of respondents in stated preference surveys influences willingness to pay estimates (Campbell, Scarpa, and Hutchinson 2008; Campbell, Hutchinson, and Scarpa 2009; Brouwer, Martin-Ortega, and Berbel 2010; Broch et al. 2013; Garrod et al. 2012; Johnston and Ramachandran 2014). As land-use changes are often conducted locally, such information can significantly impact the results of cost-benefit analyses and may reveal insights on where a land-use change offers the largest benefits.

This paper contributes to the literature on spatial preference heterogeneity by investigating how preferences for policy-relevant landscape attributes differ across respondents residing in different landscapes. Spatially-driven differences in preferences for changes in landscape attributes can occur for two main reasons. First, it is well-established that individual preferences are affected by the current level of endowment (Glenk 2011; Hess, Rose, and Hensher 2008; Tversky and Kahneman 1981). Therefore, an increase or decrease in a good is valued relative to this status quo situation. Because the marginal value of a good or service may not be constant over levels of provision, individuals with different status quo situations may value additional changes in provision differently. In particular, economic theory suggests that the utility or value that is attributed to an additional unit of a good or service is higher if its scarcity increases. The concept of diminishing marginal utility suggests, for example, that people residing in a forest landscape are willing to pay less for additional forest area created than individuals living in farm- and grassland landscapes with little forest cover (Sagebiel, Glenk, and Meyerhoff 2017). Diminishing marginal utility may apply if more of a good or service is always preferred over less; however, this may not always apply to landscape attributes, where optimal shares of certain land use shares and landscape elements may exist. That is, an increase in land use share may be perceived beneficially up to a threshold, beyond which utility for an additional increase in provision decreases (Schmitz, Schmitz, and Wronka 2003).

Second, the overall composition of a landscape has a unique value that is qualitatively different from other landscapes and that is difficult if not impossible to describe in terms of separate landscape attributes. That is, residents have different perceptions of landscapes and of the role that specific elements play in achieving uniqueness. Consequently, preferences for changes in landscape attributes may differ across landscape types, either in a systematic fashion in case that subjective perceptions of landscape amenity and value are similar across individuals living in a particular landscape type, or in an unpredictable way if there is considerable heterogeneity in perceptions. For example, those individuals living in forest landscapes may have a systematically greater demand for enhancing biodiversity, whereas people living in farm- and grassland landscapes may prefer additional structural elements. Similarly, some people living in farm- and grassland landscapes may perceive their openness as a cultural heritage characteristic of a particular region, thus objecting structural change.

In the paper, we investigate the correlation between residing in different landscape categories (i.e., different status quo situations) and preferences for changes in landscape attributes, for example share of forest or levels of biodiversity. We use data from a web-based discrete choice experiment (DCE) survey in Germany to empirically test if differences in willingness to pay for landscape attributes exist; and if the 'status quo' landscape serves as a reference point for choices in the DCE with impacts on willingness to pay estimates.

The results are relevant for policy makers dealing with local land-use changes and researchers considering DCEs to assist cost-benefit analyses. For example, in Germany, there is a discussion about combating climate change by increasing the share of energy crops for renewable energy generation. A policy maker can set spatially varying incentives or other policy tools aiming at increasing or decreasing the share of corn on agricultural fields. Typically, such incentives are based on private benefits and ecological constraints, e.g. where gross margins are high. Social welfare impacts associated with landscape change are often not considered at all, or are not directly compared with private costs and benefits. Additionally, the importance of acceptance of the land use change by the local population is often neglected, and willingness to pay values, distinguished by landscape categories, can help to identify areas where such a land-use change is likely to find local support.

2. Survey and Data

2.1 Data Collection and Discrete Choice Experiment

The DCE is part of a German-wide, web-based survey conducted in March 2013. The respondents were recruited from an online panel of a large international market research institute. People 18 years or older who resided in Germany at the time of the study were eligible to participate. The survey consists of the six sub-samples with different DCEs, totalling around 10,000 respondents. The DCEs differ in their attributes and had different land-use foci. In all samples, the scenario was a local land use change within a radius of 15 km around the respondent's place of residence. The radius should represent a typical distance for everyday activities. We discussed the radius in focus groups and came up with 15 km being a widely accepted distance. Besides the DCE, the survey includes questions on leisure activities, perceptions and knowledge on land use and climate change as well as socio-demographic variables. Respondents were requested to provide their postal code or to use the integrated geo-tool which supplies the coordinates of the places identified by respondents such as their residence location.

In this paper, we use a sub-sample with attributes related to agricultural land-use changes. The DCE comprises five non-monetary attributes each having three levels, with zero indicating the status quo as today. Table 1 gives a description of all attributes of the used sample as well as the dummy codes used in the analysis.

The first attribute *Forest* refers to the share of forest. It takes the values *as today*, *10% less* and *10% more*. We assume that an increase in forest area increases utility with a decreasing rate (diminishing marginal utility). That implies that people living in forest rich areas gain less utility from an increase in forest than people living in areas with a low share of forest. The second attribute *Fieldsize* describes the average size of fields and forests. The levels include *as today*, *half the size of today* and *double the size of today*.

Table 1. Attribute description.

Attribute	Description	Levels	Dummy Code
Forest	Share of forest in %	as today 10% decrease 10% increase	omitted ForMinus10 ForPlus10
Fieldsize	Average size of forest and fields	as today half the size double the size	omitted FieldHalf FieldDouble
Biodiversity	Degree of biodiversity measured with bird indicator	as today (55 Points) slight increase (85 Points) strong increase (105 Points)	omitted Bio85 Bio105
Cornshare	Share of corn on agricultural fields	as today share of 30% share of 70%	omitted Corn30 Corn70
Meadows share	Share of meadows in %	as today share of 25% share of 50%	omitted Mead25 Mead50
Price	Annual payment to a local landscape fund in Euro	0, 10, 25, 50, 80, 110, 160	

Smaller field sizes imply a less monotonic landscape and more structural elements, which are assumed to be more attractive in terms of visual amenity (Zanten et al. 2014). On the other hand, larger forests can lead to better forest connectivity which may have positive implications for biodiversity and recreation. We therefore have no clear expectation for this attribute. The third attribute *Biodiversity* is described with a bird indicator as a proxy for biodiversity. Bird indicators are used in several countries as headline indicators for biodiversity (Gregory et al. 2003; Butchart et al. 2010). The bird indicator, developed by the German Federal Agency for Nature Conservation, provides information on the suitability of the area for birds, where 100 points describe the state in the year 1975 in Germany (Doerpinghaus and Ludwig 2005). For Germany as a whole, the bird indicator is currently estimated to lie at about 55 points. The levels used in the DCE are *as today (55 Points)*, *slight increase (85 Points)* and *strong increase (105 Points)*. We expect that utility increases with increasing points, as it has been found in other DCE studies (Shoyama, Managi, and Yamagata 2013). The fourth attribute *Cornshare* is the share of corn on agricultural fields. The levels are *as today*, *30%* and *70%* on the agricultural fields in the surrounding. In the focus group discussions conducted prior to the survey, corn was often described as having a negative impact on landscape. We expect that a larger share of corn leads to a decrease in utility. *Meadowsshare*, the fifth attribute, refers to the share of meadows and grassland used for grazing. It takes the levels *as today*, *25% of the area*, *50% of the area*. In the focus group discussions, most participants linked a high share of meadows to a more natural landscape. We thus expect a utility increase from an increase in the share. Note that some attribute levels imply a reduction in the endowment compared to the status quo. This is explicit for Forest and Fieldsize and implicit for Cornshare and Meadowsshare. In the former case, we expect that some respondents have preferences for a

Figure 1. Example of a choice set.

If only the following options were available for the future development of the landscape within a radius of up to 15 kilometers around your place of residence, which one would you choose? If you live in a large city, please consider the surrounding area of the city.

	Landscape A	Landscape B	Landscape C
Share of forest	As today	Increase by 10%	As today
Field size	As today	Twice the size	As today
Biodiversity on agricultural fields	Strong increase	As today	As today
Share of maize on arable land	max. 70% of fields	max. 30% of fields	As today
Share of grassland on agricultural fields	25%	25%	As today
Financial contribution to fund per year	110 €	80 €	0 €
I CHOOSE ☑	☐	☐	☐

reduction. For example, in forest rich areas, people may prefer a reduction in forest share (Sagebiel, Glenk, and Meyerhoff 2017). To account for such preferences, we used a positive and a negative level. In the case of Cornshare and Meadowsshare, the direction of the change (reduction or increase) depends on the respondent's current situation. However, absolute percentage values are useful as, in practice, land use changes are often announced in such values. We expected that people understand an absolute percentage value better than a relative change. Thus we used absolute percentage values for these attributes, taking into account that the change people value varies between respondents. Finally, the price attribute is framed as an annual payment to a newly introduced landscape fund per person for an unspecified period of time. We explained to the respondents that all residents who are affected by the land use change will have to contribute to the fund (i.e. a compulsory payment) and that the money in the fund was to be exclusively used to finance and maintain the land use changes. The exact description of the payment vehicle was informed by focus group discussions. The framing of the payment vehicle as a fund was preferred to other possible payment vehicles and regarded as credible. Tax payments were not regarded as credible, because the land use change was local while taxes are usually collected at least at county level and often used for multiple purposes. The levels of the fund range from 10 to 160 Euro and is set to szero in the status quo alternative.

Each choice set consists of three unlabelled landscape alternatives, where *landscape 3* represents the status quo (Figure 1). The experimental design was created with the software package NGene, maximizing C-efficiency, which relates to the minimization of variance of willingness to pay estimates. The design was optimized for a multinominal logit model with linearity in utility and priors close to zero. It consisted of 18 choice sets divided into two blocks. Each respondent answered nine choice sets. The order of the choice sets was randomized across respondents.

2.2 Landscape categories and socio-demographics

The German Federal Agency for Nature Conservation has developed a system to classify landscapes within Germany. The intention behind this approach is to provide a basis for effective conservation and development of cultural landscapes along the objectives of the European Landscape Convention. Overall, the German land surface was divided in 858 landscapes including 59 urban conglomerations. The system comprises overall 24 landscape types that are assigned to the following six main categories (Gharadjedaghi et al. 2004):[1]

1. Coastal landscapes: This type is characterized by landscapes near the German coast of the North Sea and the Baltic Sea.
2. Forest landscapes: These landscapes have a large share of forests between 40% and 70%.
3. Cultural landscapes: These landscapes have a share of forest between 20% and 40% and a high share of one of the following items: water bodies, meadows and grassland, wine-growing, glaciers and rocks, orchards, wetlands, a combination of the items.
4. Farm- and grassland dominated landscapes: In contrast to the cultural landscapes, they have a share of forest that is less than 20%. They are further characterized by a large share of grassland and arable land.
5. Mining areas: Landscapes with more than 10 percent of the land surface under open cast mining.
6. Urban agglomerations: These landscapes comprise cities and areas with a high density of settlements and infrastructure.

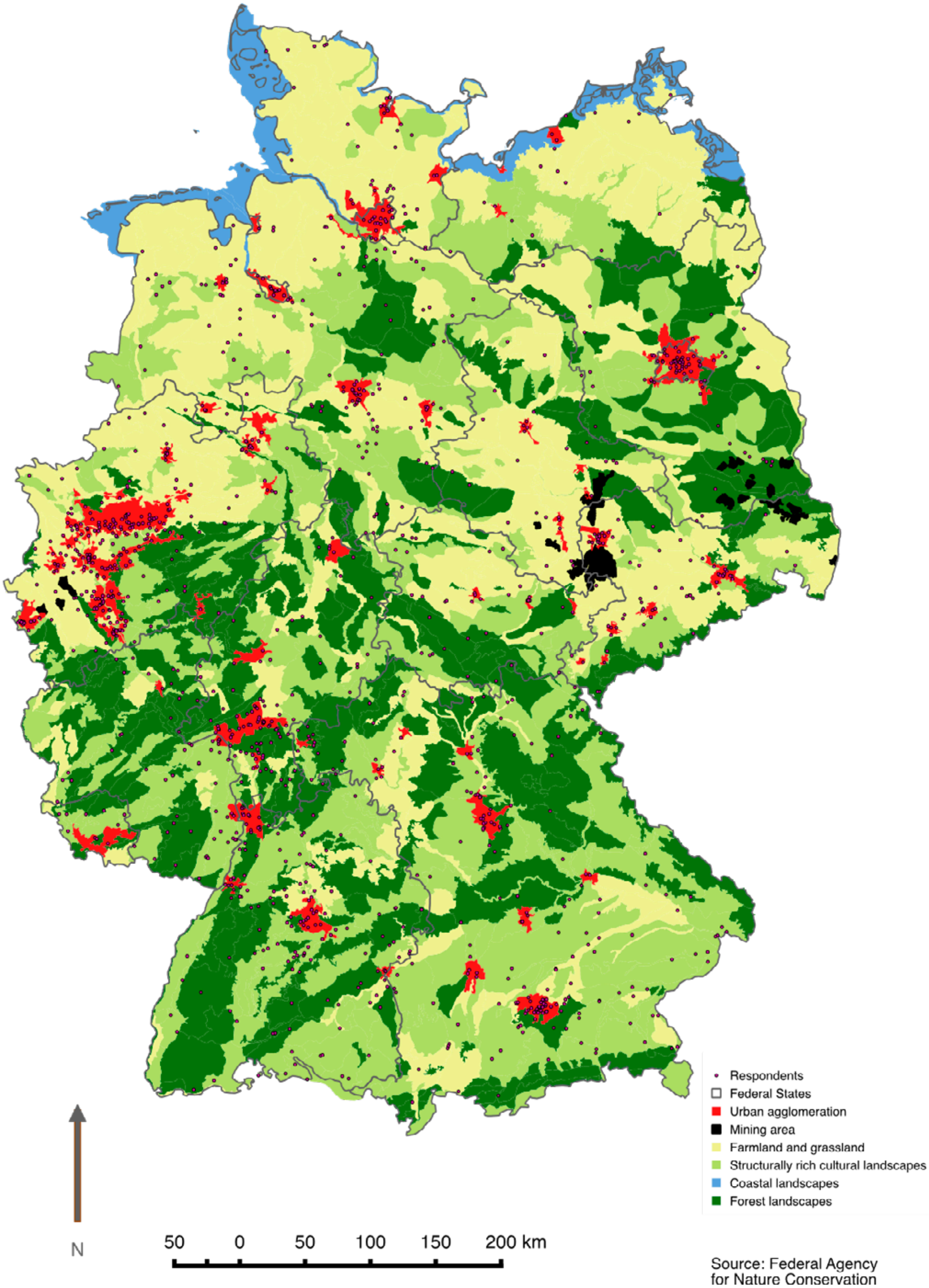
Table 2 summarizes the distribution of the respondents according to the landscapes. Each respondent is uniquely allocated to one of the categories. In this process, the actual place of residence was used to determine the landscape category rather than the percentage share of landscape categories surrounding the place of residence. Figure 2 maps both the landscape categories and the respondents' locations. We exclude five respondents from coastal landscapes and mining areas from the analysis as these categories are too small. The final sample size is 1409.

Table 2. Distribution of landscape categories.

Landscape Category	No.	%
Coastal landscapes	3	0.2
Forest landscapes	204	14.4
Cultural landscapes	326	23.1
Farm- and grassland landscapes	309	21.9
Mining areas	2	0.1
Urban Agglomerations	570	40.3
Total	1414	100.0

[1] See https://www.bfn.de/en/activities/protecting-habitats-and-landscapes/landscapes-of-conservation-importance/landscape-types.html for a brief description of the 24 landscape types.

Figure 2. Spatial distribution of sample.



2.3 Hypotheses and empirical strategy

The geo-referenced respondents are distinguished by the landscape categories described in Table 2. The main aim is to find out whether respondents from different landscapes exhibit different preferences. Hence, the main hypothesis is: preferences and willingness to pay values for landscape attributes correlate with the landscape in which a respondent lives.

We expect decreasing marginal utility, i.e. marginal willingness to pay is lower in landscapes where the status quo levels of defining attributes are already high. For example, marginal willingness to pay for more forest is lower in forest landscapes than in the other landscape categories. Additionally, we expect some kind of place attachment for attributes that dominate a landscape (Scannell and Gifford 2010). For example, a respondent living in a forest rich area is not willing to give up forest as it is a dominant characteristic of the landscape. In contrast, a respondent living in an area with a medium share of forest is more interested in gaining forest but also less averse against a loss in forests. Table 4 shows that in farm- and grassland landscapes, fieldsize is higher than in the other categories, where it is rather similar. Hence, the hypothesis is that the willingness to pay for half the size differs between farm- and grassland landscapes and the other landscapes. Corn share is highest in the two cultural landscapes and lowest in urban agglomerations. As a high corn share is expected to be perceived negatively, and for most respondents the first level already implies an increase over the status quo, we expect negative willingness to pay values. These would be highest in cultural landscapes and lowest in urban agglomerations. Therefore, we focus on the second level of this attribute, i.e. an increase to 70%. The average share of meadows is relatively similar in all landscapes, so that large differences in willingness to pay may not be present.

3. Econometric approach

In the analysis, we use a latent class logit model to investigate the effects of the landscape categories on preferences and willingness to pay. The model is consistent with microeconomic theory, assuming rational individuals who maximize a utility function under constraints. An individual i chooses in t choice situations between a given set of alternatives n – each described by a conditional indirect utility function U_{int} – the alternative that provides the maximum amount of utility. Each alternative is characterized by k attributes that have levels A_{iknt}. We assume the utility functions for alternatives to be linear and additive in the attributes, and add an error term e_{int} which is Extreme Value Type I distributed to the random utility model. A utility function can be written as

$$U_{int}=beta_1 A_{i1nt}+\beta_2 A_{i2nt}+\ldots+\beta_k A_{iknt}+e_{int} \tag{1}$$

where the β_ks are the corresponding utility coefficients. The probability of an individual choosing alternative n can be written as a conditional logit model:

$$Pr_{imt} = \frac{exp(\beta_1 A_{i1mt} + \beta_2 A_{i2mt}+\ldots+\beta_k A_{ikmt})}{\sum_{n=1}^{N} e\,xp(\beta_1 A_{i1nt} + \beta_2 A_{i2nt}+\ldots+\beta_k A_{iknt})} \tag{2}$$

This model has a closed form and can be estimated using maximum likelihood.

In order to incorporate preference heterogeneity, we apply a latent class logit model. We assume that a given number of preference classes S, differing in their utility parameters $<\beta_{k|1},\beta_{k|2},\ldots,\beta_{k|s}>$, exists. Each individual has probabilities $<h_1,h_2,\ldots,h_s>$ to be member of the preference classes. The probabilities h_s can be estimated with a multinomial logit model

$$h_s = \frac{exp(\zeta_s X_i)}{\sum_{s=1}^{S} e\,xp(\zeta_s X_i)} \tag{3}$$

where X_i are explanatory variables, in this case the landscape categories, and ζ_s are the coefficients. The unconditional choice probability to choose alternative m is given as

$$Pr_{imt} = \sum_{s=1}^{S} h_s Pr_{imt|s} \tag{4}$$

The latent class logit model as described in equation 4 introduces preference heterogeneity between classes. Within a class, preferences are fixed. To relax this assumption without introducing a large amount of new parameters, we extend the model to a scale-adjusted latent class model (Magidson and Vermunt 2008). In this model, each preference class s is separated by a constant which can be interpreted as a scale parameter. The scale parameter merely states that preferences for all attributes are higher in the one scale class than in the other scale class. Whether the differences between respondents are caused by different preferences (all very high, vs. all very low) or by differences in the error variances (more random vs. less random choices) cannot be answered empirically (Hess and Train 2017). Still, the introduction of this parameter captures another dimension of heterogeneity, which can improve model fit significantly. As the scale classes are restricted in a way that all preference parameters differ similarly, willingness to pay values between scale classes are not affected. Technically, the scale parameter is estimated by another multinomial logit model, and each respondent has a probability g to belong to scale class r – similar to the preference classes. The unconditional choice probability in equation 4 becomes

$$Pr_{imt} = \sum_{s=1}^{S} \sum_{r=1}^{R} g_r h_s Pr_{imt|sr} \tag{5}$$

If an earlier analysis has already identified some respondents belonging to a specific class, one can add a known-class parameter τ_r. This parameter is zero if a respondent cannot be assigned a priori to a certain class, leading to

$$Pr_{imt} = \sum_{s=1}^{S} \sum_{r=1}^{R} g_r h_s \tau_r Pr_{imt|sr} \tag{6}$$

In this study, we use the known class indicator to classify all respondents who have always chosen the status quo option into class 1. To determine the number of preference classes *S*, one can use statistical measures of fit such as the Bayesian Information Criterion (BIC), or the corrected Akaike Information Criterion (cAIC). Both BIC and cAIC penalize for more parameters and are therefore preferred over other information criteria. Additional to the statistical criteria, one can rely on own judgment concerning reasonable parameter estimates and knowledge gained from earlier analyses (Boxall and Adamowicz 2002; Scarpa and Thiene 2005).

To calculate willingness to pay values for each class individually, the respective class preference parameter is divided by the class cost parameter. Confidence intervals of willingness to pay are calculated with the delta method.

4. Results

4.1 Descriptive statistics of landscape categories

We first analyze the relationship between socio-demographic variables and landscape categories. This step is important to understand whether and how potential differences in preferences could arise from differences in socio-demographics rather than the landscape respondents are living in.

Most differences are found between urban agglomerations and the other landscapes (Table 3). Respondents from urban agglomerations are more educated and have fewer children. We use Kruskall-Wallis and t-tests to test for overall differences between the landscape categories. Statistically significant differences on a 5% level are present for all variables except personal income and sex. Although there are differences in socio-demographics between landscape categories (especially between urban areas and all other areas), we will not investigate those here. We acknowledge that the differences in preferences may be driven by socio-demographics rather than landscape categories, but this is not relevant for the policy question of how land use changes are perceived in different landscapes. Our analysis thus only provides correlations.

Using data from the German Federal Agency for Cartography and Geodesy (BKG) and the German Federal Institute of Research on Building, Urban Affairs and Spatial Development (BBSR), we investigated the actual status quo attribute levels of the respondents. Table 4 summarizes the actual status quo in the 15km radius by landscape categories. In most cases, there are relatively large differences between the landscape categories. For the sake of parsimony, we will not investigate the actual status quo and possible effects any further. Sagebiel, Glenk, and Meyerhoff (2017) conduct a detailed investigation of the actual status quo and its effects on willingness to pay.

4.2 Latent class analysis

We estimate the latent class model described in section 3 using the software package LatentGold Choice 4.5 with the Syntax module. To select a specific number of classes we compared BIC and cAIC for two to eight class models, in the absence and presence of a scale class. We choose a model with five preference classes and two scale classes. This model turned out have the lowest BIC and cAIC values and offered plausible parameter values.

Table 3. Frequencies and column percentages (in parentheses) of socio-demographic variables.

	Forest	Cultural	Farm- and grassland	Urban	Total
Education					
Secondary or less	83 (40.9)	122 (37.4)	121 (39.4)	141 (24.7)	467 (33.2)
Higher education	46 (22.7)	86 (26.4)	77 (25.1)	155 (27.2)	364 (25.9)
University	74 (36.5)	118 (36.2)	109 (35.5)	274 (48.1)	575 (40.9)
Sex					
Male	100 (49.0)	168 (51.5)	175 (56.6)	308 (54.0)	751 (53.3)
Female	104 (51.0)	158 (48.5)	134 (43.4)	262 (46.0)	658 (46.7)
Children in household					
Yes	68 (33.3)	135 (41.4)	102 (33.0)	134 (23.5)	439 (31.2)
No	136 (66.7)	191 (58.6)	207 (67.0)	436 (76.5)	970 (68.8)
Income					
Less than 1500 Euros	81 (39.7)	128 (39.3)	120 (38.8)	227 (39.8)	556 (39.5)
1500 to 2600 Euros	58 (28.4)	91 (27.9)	74 (23.9)	153 (26.8)	376 (26.7)
More than 2600 Euros	65 (31.9)	107 (32.8)	115 (37.2)	190 (33.3)	477 (33.9)
Age					
19 to 29	32 (15.7)	67 (20.6)	60 (19.4)	132 (23.2)	291 (20.7)
30 to 39	50 (24.5)	62 (19.0)	58 (18.8)	110 (19.3)	280 (19.9)
40 to 49	39 (19.1)	89 (27.3)	92 (29.8)	144 (25.3)	364 (25.8)
50 to 59	40 (19.6)	64 (19.6)	55 (17.8)	102 (17.9)	261 (18.5)
Older than 60	43 (21.1)	44 (13.5)	44 (14.2)	82 (14.4)	213 (15.1)

All attributes except price were dummy coded with the status quo level *as today* as the reference. The landscape categories entered the class membership function as dummy coded variables with forest landscapes as the reference category. We did not include any socio-demographic variables as these are correlated with the landscape categories,

Table 4. Mean and standard deviation (in parenthesis) of actual status quo by landscape categories.

	Forest	Cultural	Farm- and grassland	Urban	Total
Forest Share	41.7	29.8	17.5	18.4	24.2
	(12.2)	(11.2)	(9.7)	(10.0)	(13.7)
Field Size	17.7	17.5	25.9	17.0	19.2
	(7.5)	(6.8)	(12.2)	(6.7)	(9.1)
Corn Share	14.9	20.9	19.9	10.5	15.5
	(10.3)	(14.8)	(15.8)	(10.0)	(13.5)
Meadows Share	15.6	17.9	18.1	12.6	15.5
	(6.1)	(9.0)	(12.1)	(7.1)	(9.1)

potentially causing multicollinearity. 23% of all respondents chose the status quo alternative in all choice situations and were assigned to class 1 with a probability of 1. As several respondents seemed to have ignored the price attribute, we fixed the price parameter to zero in class 3 to capture price non-attendance. In models without this restriction, at least one class is characterized by willingness to pay values three times as high as the highest price level of 160 Euro, which we consider implausible.

In a first step, we describe the five classes in terms of estimated utility parameters and willingness to pay values. Then, we investigate the relationship between class membership and landscape categories. Table 5 shows the estimation results and Table 6 its willingness to pay values.

The overall model is highly significant. The statistically significant coefficient for the scale class of -0.302 translates to scale class probabilities of 57.5% and 42.5% for scale classes 1 and 2, respectively, indicating that additional heterogeneity and correlation patterns are present. In Class 1, price, ForMinus10, FieldHalf, FieldDouble, Corn70 and Mead50 are highly significant and negative. Willingness to pay values range between -88 and -35 Euro, i.e. people are opting against all land use changes and would need to be compensated. The positive and significant ASCsq means that Class 1 is characterized by preferences towards the status quo. Class 2 has a negative and significant ASCsq, indicating preferences for land use changes. ForMinus10, ForPlus10, FieldDouble, Bio105, Corn70, Mead50 and price are significant with the expected signs. The willingness to pay for ForMinus10 and ForPlus10 is -165 Euro and 64 Euro, respectively. People are willing to pay for increases in forest, but would need to be compensated nearly three times as much for decreases in forest. For a reduction of field sizes (FieldHalf), willingness to pay is nearly 20 Euro while a doubling of field sizes would need to be compensated with 45 Euro. Willingness to pay for increases in biodiversity is 32 Euro for an increase to 85 points and 51 Euro for 105 points. A share of corn of 30% is not significant but a share of 70% requires a compensation of 61 Euro. Willingness to pay for a share of meadows of 25% is positive (42 Euro) while a share of 50% is not significant and close to zero. In summary, Class 2 is characterized by large positive and negative willingness to pay values for land use changes. Class 3 is the price non-attendance class. Respondents who disregard the cost attribute are likely choosing a land use change scenario over the status quo if they

Table 5. Latent class model with five classes.

	Class 1	Class 2	Class 3	Class 4	Class 5
ASCsq	0.760	-1.281***	-23.897**	-3.993***	-3.535***
ForMinus10	-3.151***	-4.169***	-17.829*	-1.287***	-0.062
ForPlus10	-0.316	1.624***	0.926***	0.317*	1.108**
FieldHalf	-2.678***	0.474	-0.075	-0.961***	-1.305***
FieldDouble	-2.120***	-1.132***	-0.275**	0.390**	-0.361
Bio85	0.774	0.814**	-5.791**	0.444	-0.455
Bio105	0.431	1.295***	1.233***	0.103	0.964*
Corn30	-0.647	0.592	1.055***	0.038	0.351
Corn70	-5.326***	-1.563***	0.214	-1.195***	-0.763
Mead25	-1.173	1.056***	-0.035	-1.120***	-1.503***
Mead50	-2.525***	0.090	0.773***	-0.802***	-0.586
price	-0.060***	-0.025***	0.000	-0.013***	-0.078***
Covariates of membership function					
Forest	ref	ref	ref	ref	ref
Cultural	ref	0.172	0.877**	0.041	0.924*
Grass/Farm	ref	0.655**	1.245***	0.2462	1.167**
Urban	ref	0.198	1.326***	0.438	1.325***
Scale classes					
	Scale Class 1	Scale Class 2			
Constant	ref	-0.302**			
Log-Likelihood	-8976.153				
Observations	12681				
Respondents	1409				

* p < 0.10, ** p < 0.05, *** p < 0.01 , ref = reference category with parameter fixed to zero

have a positive attitude towards policy change. This is reflected in the very large and negative ASCsq. Similarly, the very large and negative coefficient for decreases in forest share can be explained by this phenomenon. Nearly all coefficients of the remaining attributes are significant and have the expected signs. Bio85 is significant and negative which could imply that members of this class have already a high degree of biodiversity and regard 85 points as a deterioration. Similarly, the positive coefficient of Corn30 implies that people have already high shares of corn and regard a 30% share as an improvement. Finally, Mead25 is not significant while Mead50 is significant and positive. Class 4 is characterized by comparatively large negative willingness to pay values to avoid decreases in forest share, field size and a corn share of 70%. Interestingly, the willingness to pay for meadows share is negative for both 25% and 50%. In Class 5, positive willingness to pay values are significant and positive only for ForPlus10 (14 Euro) and Bio105 (12 Euro) and negative for FieldHalf (-16 Euro) and Mead25 (-19 Euro). This class comprises small or no utility gains from land use changes.

The landscape categories have a significant impact on the probability to be member of a class. Forest landscapes and Class 1 are the reference categories, the parameters in the membership function are interpreted relative to them. Class 1 is the largest class with a share of about 40%. Classes 2 to 4 have a share between 16% and 19%. Class 5 is the smallest class with a share of 9%. Note that Classes 1 and 5 are characterized by no or low willingness to pay values and make up nearly 50% of class membership.

Table 7 shows class membership probabilities calculated for each landscape category separately. Differences in class membership between landscape categories are present in Classes 1, 3 and 5. Membership probabilities are rather homogeneous for Classes 2 and 4. Respondents from forest landscapes are more likely to be member of Class 1 compared to the other categories with a share of nearly 51% (against the class average of 39%) and less likely member of Classes 3 and 5 with shares of only 8% and 4% (compared to the class averages of 18% and 9%). Respondents from cultural landscapes are slightly more likely to be member of class 1 (42%) and less or equally likely in the other classes. Respondents from grass- and farmlands are less likely to be member of Class 1 (34% against 39%) and Class 4 (14% against 16%) and more likely to be member of Class 2 (24% against 19%). Finally, respondents from urban agglomerations are less likely to be member of Class 1 (35% against 39%) and 2 (16% against 19%) and more likely in Classes 3 (21% against 18%), 4 (17% against 16%) and 5 (11% against 9%).

The results are partly in line with our expectations. Forest landscapes and cultural landscapes have high shares of forest and are relatively bio-diverse, with many structural landscape elements. Such landscapes are generally associated with high recreational values. Respondents from these categories are more likely to be member of Class 1 which is characterized by status quo choices and strong opposition against reductions of forest share, increases in corn and changes in field size. This aligns with our expectation of place attachment. The zero willingness to pay for increases of forest indicates diminishing marginal utility. People from forest landscapes are also less likely to be members of Class 3, which is characterized by a strong tendency towards land use changes and cost non-attendance, and of Class 5, which is characterized by low willingness to pay, implying some heterogeneity within this landscape category. About 55% are allocated to Classes 1 and 5 (low willingness to pay), while the remaining share belongs to the other classes which are characterized by high willingness to pay and strong preferences for land-use changes.

Farm- and grasslands are dominated by agriculture and monotonic landscapes with low shares of forest. Respondents from farm- and grasslands are more likely to be members of Class 2, which is characterized by rather large willingness to pay values. This result fits to our expectations of marginal diminishing utility. People living in this landscape have a low endowment of forest, biodiversity and meadows and are thus more willing to pay for an additional unit.

Finally, respondents from urban agglomerations are more likely to be member of Class 3, i.e. are more likely to not attend to costs. While we have no expectation here, this result may be explained by hypothetical bias. The choice scenario is less realistic for people in urban areas and they are less used to the landscapes. They may have ignored the price attribute more often, while at the same time exhibit strong preferences for land use changes. It should be noted that our results indicate preference heterogeneity within landscape categories. We do observe deterministic patterns of distinct preferences between landscape categories. Each landscape category is present in each class with a probability close to the average group probability. Class 1 is the largest class for all landscape catego-

Table 6. Willingness to pay values.

Attribute	Class 1	Class 2	Class 3	Class 4	Class 5
ForMinus10	-52.52***	-165.19***	-	-100.67***	-0.7954
ForPlus10	-5.27	64.33***	-	24.75*	14.24***
FieldHalf	-44.64***	18.77*	-	-75.17***	-16.76***
FieldDouble	-35.34**	-44.83**	-	30.51**	-4.64
Bio85	12.90	32.24**	-	34.75	-5.84
Bio105	7.17	51.32***	-	8.04	12.38**
Corn30	-10.78	23.44	-	2.96	4.50
Corn70	-88.78***	-61.93*	-	-93.47**	-9.80
Mead25	-19.55	41.82**	-	-87.58***	-19.31**
Mead50	-42.09**	3.57	-	-62.74**	-7.53

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

Table 7. Class probabilities by landscape categories.

Landscape	Class 1	Class 2	Class 3	Class 4	Class 5
Forest	0.51	0.19	0.08	0.16	0.04
Cultural	0.42	0.19	0.17	0.14	0.09
Grass/Farm	0.34	0.24	0.19	0.14	0.09
Urban	0.35	0.16	0.21	0.17	0.11
Overall	0.39	0.19	0.18	0.16	0.09

ries and Class 5 is the smallest class for all landscape categories. The effects that we identified should be interpreted as tendencies.

Additional to the latent class analysis, we have estimated separate conditional logit models by landscape categories and used Poe et al. tests (Poe, Giraud, and Loomis 2005) to test for differences in willingness to pay between landscape categories. While exact quantitative results differ, the key findings are similar irrespective of the approach used. The appendix provides more details on the conditional logit models, willingness to pay values and the Poe et al. test results.

5. Conclusion and policy implications

This paper investigated preferences for land-use changes and compared willingness to pay values between different landscape categories in Germany. The data came from a discrete choice experiment inferring preferences for forest share, average size of forest and fields, degree of biodiversity, share of corn and share of meadows within the 15 kilometer radius of the respondents' places of residence. The radius was chosen to represent a typical distance for everyday activities. As the places of residence were geo-referenced, we could combine the data with landscape categories compiled by the German Federal Agency for Nature Conservation. The categories comprised forest landscapes, cultural landscapes, farm- and grassland landscapes and urban agglomerations. The aim of the study was to test whether preferences for land-use changes are correlated with these landscape categories. To do so, we estimated a five-class latent class model and used the landscape categories as explanatory variables in the class membership function. The classes can be distinguished by different willingness to pay values. It turned out that people from forest landscapes and cultural landscapes were less willing to pay for land-use changes and showed a preference towards the status quo situation. Further, people from urban agglomerations and farm- and grassland have high probabilities to be member of classes with large willingness to pay values.

In summary, the results showed that the preferences do differ among landscape categories, but not as systematically as we had expected. Although we find systematic differences in preferences between landscape categories, all landscape categories are relatively evenly distributed across classes. As the latent class analysis has shown, preference heterogeneity exists also within the landscape categories. That is, each respondent, independent of which landscape category the respondent is from, has a probability of at least 8% to be member of any class.

The analysis has implications for policy makers. Our study provides evidence that there are differences in preferences determined by the place of residence. Integrating such differences in landscape planning and cost-benefit analyses may help to improve decisions and induce land-use changes to areas where people appreciate them most or are least reluctant towards a change. A relevant example is the share of corn among agricultural fields. While an increase in the production of energy corn can potentially help to reduce carbon dioxide emissions, it is largely regarded as a disfigurement of the landscape. Our study revealed that opposition to corn is generally large, but stronger in forest and cultural landscapes than in other landscapes. Similarly, increases in forest share should take place in areas with limited forests and near urban agglomerations. Areas characterized by high recreational values such as forest and cultural landscapes should be preserved. Here, people tend more towards the status quo and changes are less appreciated by residents. There is limited interest in increases in forest shares or biodiversity, and at the same time a large resistance against reductions. In contrast, respondents from urban agglomerations and farm- and grasslands are more likely to benefit from increases in forest shares and biodiversity. Here, significant welfare effects of such measures are more likely. Our findings may also be used to inform the design of agri-environmental schemes. For example, compensation may be higher for measures to increase agro-biodiversity in a rather monotonic landscape or near urban areas, because benefits of measures are greater.

Similar studies have investigated land use changes on a broader scale. In their meta-analysis, van Zanten et al. (2014) have found preferences for various landscape elements such as smaller field sizes, but no spatial determinants of preferences. Garrod et al. (2012) have found that preferences for improving ecosystem services depend on the landscape where they are present. This result is in line with our findings, yet our study differs as the proposed land use change always took place at the person's place of residence. In Garrod et al. (2012), this was not the case. To our knowledge, our study is the first study that identifies spatial differentiated preferences for local land use changes.

This study is limited by the fact that we did not investigate the underlying sources for the differences. The landscape categories differ in the status quo of the investigated attributes and in socio-demographic variables. Thus, we are not able to identify the causal effect of living in a certain landscape on preferences. Yet, the study insights provide correlation patterns which are sufficient to foster an understanding of the variation of preferences and willingness to pay between qualitatively different regions.

Acknowledgments

We thank Henry Wuestemann for support in the processing of the GIS data and Dr. Roland Goetzke and Raphael Knevels from the Federal Institute of Research on Building, Urban Affairs and Spatial Development (BBSR) for preparing the landscape structure data. Financial support from the German Ministry of Education and Research within the project CC-LandStraD (Grant Number: 01LL0909C) is gratefully acknowledged.

References

Boxall, Peter C, and Wiktor L Adamowicz. 2002. "Understanding Heterogeneous Preferences in Random Utility Models: A Latent Class Approach." *Environmental and Resource Economics* 23 (4): 421–46.

Broch, Stine Wamberg, Niels Strange, Jette B. Jacobsen, and Kerrie A. Wilson. 2013. "Farmers' Willingness to Provide Ecosystem Services and Effects of Their Spatial Distribution." *Ecological Economics* 92: 78–86.

Brouwer, Roy, Julia Martin-Ortega, and Julio Berbel. 2010. "Spatial Preference Heterogeneity: A Choice Experiment." *Land Economics* 86 (3): 552–68.

Butchart, Stuart HM, Matt Walpole, Ben Collen, Arco van Strien, Jörn PW Scharlemann, Rosamunde EA Almond, Jonathan EM Baillie, et al. 2010. "Global Biodiversity: Indicators of Recent Declines." *Science* 328 (5982): 1164–8.

Campbell, Danny, W George Hutchinson, and Riccardo Scarpa. 2009. "Using Choice Experiments to Explore the Spatial Distribution of Willingness to Pay for Rural Landscape Improvements." *Environment and Planning A* 41: 97–111.

Campbell, Danny, Riccardo Scarpa, and W George Hutchinson. 2008. "Assessing the Spatial Dependence of Welfare Estimates Obtained from Discrete Choice Experiments." *Letters in Spatial and Resource Sciences* 1 (2-3): 117–26.

Doerpinghaus, Annette, and Gerhard Ludwig. 2005. *Methoden Zur Erfassung von Arten Der Anhänge Iv Und V Der Fauna-Flora-Habitat-Richtlinie*. Bundesamt für Naturschutz.

Garrod, Guy, Eric Ruto, Ken Willis, and Neil Powe. 2012. "Heterogeneity of Preferences for the Benefits of Environmental Stewardship: A Latent-Class Approach." *Ecological Economics* 76: 104–11.

Gharadjedaghi, B, R Heimann, K Lenz, Ch Martin, V Pieper, A Schulz, A Vahabzadeh, P Finck, and U Riecken. 2004. "Verbreitung Und Gefährdung Schutzwürdiger Landschaften in Deutschland." *Natur Und Landschaft* 79 (2): 71–81.

Glenk, Klaus. 2011. "Using Local Knowledge to Model Asymmetric Preference Formation in Willingness to Pay for Environmental Services." *Journal of Environmental Management* 92 (3): 531–41.

Glenk, Klaus, Robert J Johnston, Jürgen Meyerhoff, and Julian Sagebiel. 2019. "Spatial Dimensions of Stated Preference Valuation in Environmental and Resource Economics: Methods, Trends and Challenges." *Environmental and Resource Economics*, 1–28.

Gregory, RD, D Noble, R Field, J Marchant, M Raven, and DW Gibbons. 2003. "Using Birds as Indicators of Biodiversity." *Ornis Hungarica* 12 (13): 11–24.

Hess, Stephane, John M Rose, and David A Hensher. 2008. "Asymmetric Preference Formation in Willingness to Pay Estimates in Discrete Choice Models." *Transportation Research Part E: Logistics and Transportation Review* 44 (5): 847–63.

Hess, Stephane, and Kenneth Train. 2017. "Correlation and Scale in Mixed Logit Models." *Journal of Choice Modelling* 23: 1–8.

Johnston, Robert J, and Mahesh Ramachandran. 2014. "Modeling Spatial Patchiness and Hot Spots in Stated Preference Willingness to Pay." *Environmental and Resource Economics* 59 (3): 363–87.

Magidson, Jeroen K., and Jeroen K. Vermunt. 2008. "Removing the Scale Factor Confound in Multinomial Logit Choice Models to Obtain Better Estimates of Preference." In *Sawtooth Software Conference Proceedings*, 139–55. Sawtooth Software.

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): 353–65.

Sagebiel, Julian, Klaus Glenk, and Jürgen Meyerhoff. 2017. "Spatially Explicit Demand for Afforestation." *Forest Policy and Economics* 78: 190–99.

Scannell, Leila, and Robert Gifford. 2010. "Defining Place Attachment: A Tripartite Organizing Framework." *Journal of Environmental Psychology* 30 (1): 1–10.

Scarpa, Riccardo, and Mara Thiene. 2005. "Destination Choice Models for Rock Climbing in the Northeastern Alps: A Latent-Class Approach Based on Intensity of Preferences." *Land Economics* 81 (3): 426–44.

Schmitz, K, PM Schmitz, and TC Wronka. 2003. "Valuation of Landscape Functions Using Choice Experiments." *German Journal of Agricultural Economics* 52 (8): 379–89.

Shoyama, Kikuko, Shunsuke Managi, and Yoshiki Yamagata. 2013. "Public Preferences for Biodiversity Conservation and Climate-Change Mitigation: A Choice Experiment Using Ecosystem Services Indicators." *Land Use Policy* 34: 282–93.

Tversky, Amos, and Daniel Kahneman. 1981. "The Framing of Decisions and the Psychology of Choice." *Science* 211 (4481): 453–58.

van Zanten, Boris T, Peter H Verburg, Mark J Koetse, and Pieter JH van Beukering. 2014. "Preferences for European Agrarian Landscapes: A Meta-Analysis of Case Studies." *Landscape and Urban Planning* 132: 89–101.

Appendix

In order to further investigate differences between landscape categories, we estimate separate conditional logit models for the landscape categories. Table 8 provides the estimation results.

Table 8. Conditional logit models by landscape.

	(1) Forest	(2) Cultural	(3) Farm- and Grasslands	(4) Urban
ASCsq	0.101	0.0171	-0.0959	0.0436
	(0.201)	(0.155)	(0.154)	(0.112)
ForMinus10	-0.647***	-0.594***	-0.532***	-0.458***
	(0.135)	(0.103)	(0.106)	(0.0778)
ForPlus10	0.139	0.284***	0.384***	0.401***
	(0.104)	(0.0789)	(0.0769)	(0.0562)
FieldHalf	-0.152	-0.314***	-0.283***	-0.221***
	(0.119)	(0.0905)	(0.0906)	(0.0670)
FieldDouble	-0.302***	-0.345***	-0.267***	-0.0984*
	(0.107)	(0.0791)	(0.0773)	(0.0561)
Bio85	0.0381	0.129	0.0650	0.204***
	(0.127)	(0.0994)	(0.102)	(0.0753)
Bio105	0.0126	0.246***	0.327***	0.417***
	(0.118)	(0.0884)	(0.0843)	(0.0614)
Corn30	-0.0983	0.176*	0.000266	-0.0231
	(0.125)	(0.0972)	(0.0965)	(0.0706)
Corn70	-0.691***	-0.390***	-0.550***	-0.548***
	(0.127)	(0.0930)	(0.0919)	(0.0670)
Mead25	0.0596	0.0768	-0.131	0.0247
	(0.136)	(0.103)	(0.105)	(0.0761)
Mead50	-0.234*	-0.151	-0.182*	-0.125*
	(0.130)	(0.0939)	(0.0938)	(0.0683)
price	-0.00615***	-0.00755***	-0.00520***	-0.00583***
	(0.00115)	(0.000892)	(0.000858)	(0.000630)
N	5508	8802	8343	15390
pseudo R^2	0.171	0.117	0.087	0.081
AIC	3366.8	5717.0	5604.7	10379.3
BIC	3446.1	5802.0	5689.0	10471.0
χ^2	691.3	753.6	529.8	916.5
Log-Likelihood (NULL)	-2017.1	-3223.3	-3055.2	-5635.9
Log-Likelihood	-1671.4	-2846.5	-2790.3	-5177.6

Standard errors in parentheses * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

All models are highly significant and differences between the landscape categories are visible. In forest landscapes, ForPlus10 is not significant, according to the hypothesis that respondents living in areas with a lot of forests have a limited preference for an increase in the share of forests. An increase in biodiversity to 85 points is only significant in urban agglomerations, where people are characterized by a low degree of biodiversity. Hence, an increase to 85 points has already a positive effect on utility. In the other categories, biodiversity is significant only at the 105 point level. In order to better understand the differences, Table 9 displays the estimated willingness to pay values for the different categories and Figure 3 gives a graphical overview of the willingness to pay values and corresponding 95% confidence intervals. Finally, Table 10 provides the p-values of the Poe test. If the p-value is larger than 0.95 or smaller than 0.05, the willingness to pay values are significantly different. The Poe test has to be interpreted with care. Significant differences will only appear when confidence intervals are small enough. Hence, if the test does not reject the hypotheses that the willingness to pay values are similar, it does not necessarily mean that they are not. It rather means that we cannot show that they are.

Differences in willingness to pay are significant for ForPlus10, Bio105, Corn30 and Corn70. ForPlus10 is not significant for forest landscapes and is significantly higher in open cultural landscapes and urban agglomerations. An increase in biodiversity is valued most in open cultural landscapes and in urban agglomerations and is significantly higher than in forest landscapes. An increase in corn share to 70% has the highest negative willingness to pay in forest landscapes and in open cultural landscapes. There are very few differences between open cultural landscapes and urban agglomerations and no significant differences for Meadows Share and Bio85, which however maybe due to the large confidence intervals. FieldHalf and FieldDouble are nearly always significant, but again, no significant willingness to pay differences exist. Thus, preferences for this attribute are relatively similar.

The results from the Poe test are corresponding to the findings from the latent class analysis. In both exercises, people from open cultural landscapes and urban agglomerations seem to have relatively equal preferences. Similarly, people from forest landscapes and from structurally rich cultural landscapes exhibit similar preferences. The main hypotheses of decreasing marginal utility seem partly confirmed. For example, people in forest landscapes have no willingness to pay for an increase, but a strong willingness to pay against a decrease. However, not in all cases, the results correspond to our expectations.

Table 9. Willingness to pay for different landscape models.

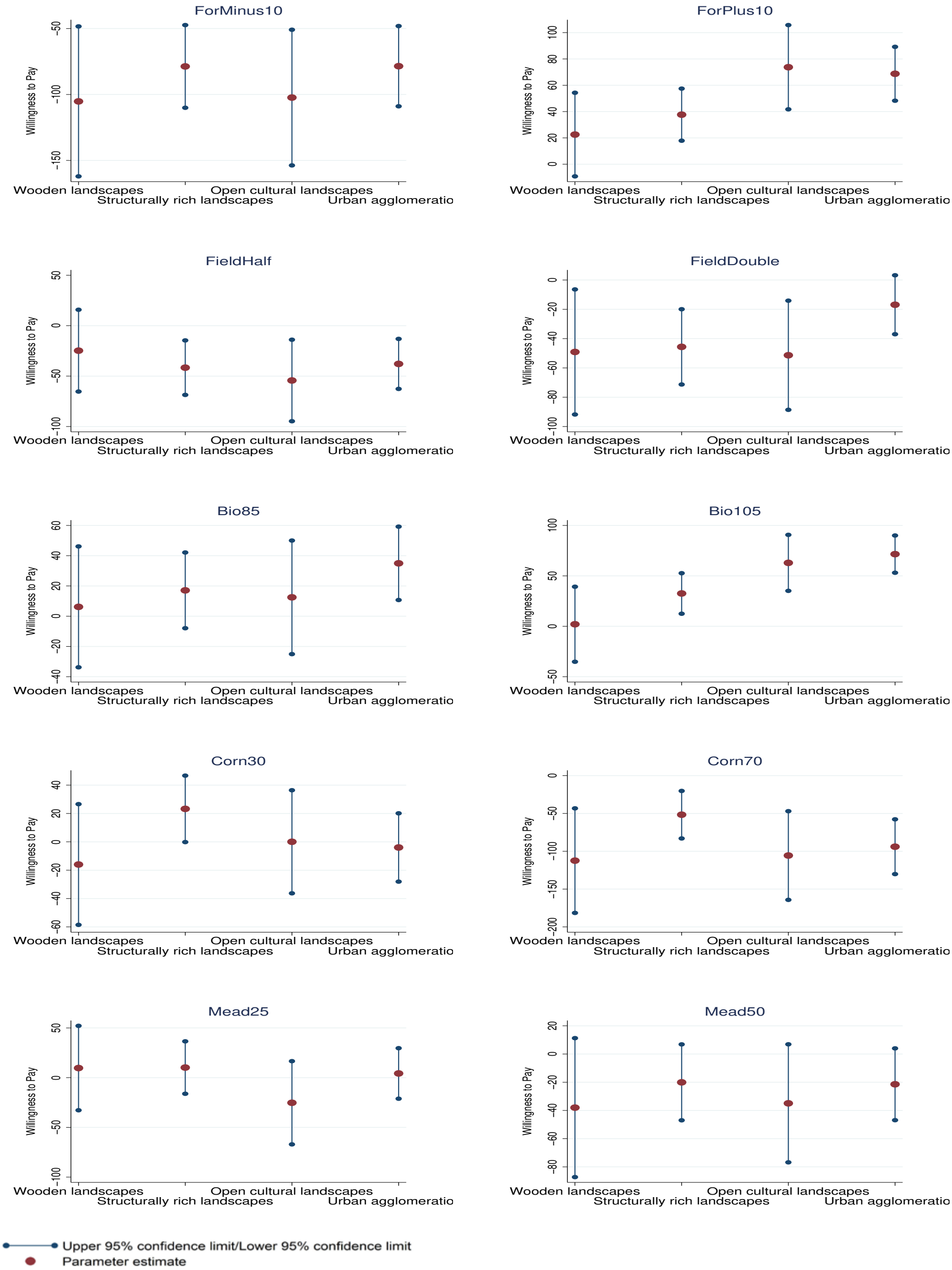
	(1) Forest	(2) Cultural	(3) Farm- and grassland	(4) Urban
ForMinus10	-105.2***	-78.72***	-102.3***	-78.52***
	(29.01)	(16.00)	(26.24)	(15.52)
ForPlus10	22.54	37.66***	73.74***	68.74***
	(16.21)	(10.10)	(16.33)	(10.44)
FieldHalf	-24.75	-41.61***	-54.31***	-37.89***
	(20.69)	(13.79)	(20.61)	(12.65)
FieldDouble	-49.10**	-45.62***	-51.34***	-16.88
	(21.77)	(13.10)	(19.00)	(10.26)
Bio85	6.188	17.10	12.49	34.97***
	(20.38)	(12.77)	(19.15)	(12.37)
Bio105	2.044	32.57***	62.91***	71.57***
	(18.98)	(10.26)	(14.17)	(9.419)
Corn30	-15.98	23.26*	0.0511	-3.970
	(21.72)	(11.95)	(18.54)	(12.29)
Corn70	-112.4***	-51.64***	-105.6***	-93.96***
	(35.27)	(16.03)	(29.91)	(18.47)
Mead25	9.688	10.17	-25.24	4.244
	(21.66)	(13.45)	(21.35)	(12.97)

Standard errors in parentheses.
* p < 0.10, ** p < 0.05, *** p < 0.01.

Table 10. Poe test results.

	ForMinus10	ForPlus10	FieldHalf	FieldDouble	Bio85	Bio105	Corn30	Corn70	Mead25	Mead50
Forest vs. Cultural	0.845	0.905	0.251	0.481	0.879	0.987	0.992	0.965	0.696	0.558
Forest vs. Farm	0.412	0.998	0.174	0.529	0.644	0.993	0.805	0.589	0.327	0.517
Forest vs. Urban	0.795	0.997	0.318	0.711	0.877	1.000	0.837	0.683	0.675	0.492
Cultural vs. Farm	0.074	0.986	0.33	0.55	0.185	0.775	0.056	0.046	0.14	0.449
Cultural vs. Urban	0.4	0.975	0.597	0.779	0.486	0.908	0.025	0.041	0.474	0.403
Farm vs. Urban	0.889	0.321	0.733	0.683	0.802	0.623	0.505	0.587	0.867	0.458

Figure 3. Willingness to pay confidence intervals by sample.



Bio-based and Applied Economics 9(3): 305-324, 2020
DOI: 10.13128/bae-8087

Full Research Article

The use of latent variable models in policy: A road fraught with peril?

Danny Campbell*, Erlend Dancke Sandorf

University of Stirling, Stirling Management School, Economics Division

Abstract. This paper explores the potential usefulness and possible pitfalls of using integrated choice and latent variable models (hybrid choice models) on stated choice data to inform policy. Using a series of Monte-Carlo simulations, we consider how model selection depends on the strength of relationship between the latent variable and preferences and the strength of relationship between the latent variable and the indicator. Our findings show that integrated choice and latent variable models are difficult to estimate, even when the data generating process is known. Ultimately, we show that their use should be driven by the analyst's belief about the strength of correlations between preferences, the latent variable and indicator. We discuss the implications of our results for policy.

Keywords. Stated preferences, choice modelling, integrated choice and latent variables, hybrid choice model.

JEL codes. C25, H41, Q51.

1. Introduction

Many policies affect the natural environment: e.g. a new hydro-electric dam will provide clean renewable energy and jobs, but may cause damage to the local river; a new motorway will reduce travel time, but may be built in a vulnerable natural area; and, a new conservation area will protect a number of vulnerable species, but possibly displace existing and future industrial activity and development. Policy makers are routinely faced with these decisions and trade-offs, and in many countries they are required to undertake cost-benefit analyses or assessments. Problematically, many of these costs and benefits are not traded in markets and policy makers have no information on society's preferences for these non-market goods and services. Stated choice experiments, where people are asked to make a choice between competing policy alternatives, are a way to elicit people's preferences for non-market goods and services.

Economists have long recognized that people's choices are affected by a multitude of observable (e.g. gender, age and income) and unobservable (e.g. attitudes and beliefs)

*Corresponding author. E-mail: danny.campbell@stir.ac.uk

Editor: Meri Raggi.

ISSN 2280-6180 (print)
ISSN 2280-6172 (online)

individual characteristics in addition to the characteristics of the options amongst which they choose. For example, when asked to choose whether to support a policy to protect a river from hydropower development, people's decision will likely depend on their income and where they live in relation to the river, but also their attitudes towards development, clean energy and conservation. Testing whether choices are different between high and low income people is trivial and straightforward, but how do we test for differences in attitudes and beliefs? How do we incorporate and consider them in our models? The most obvious, and perhaps most intuitive, way to test for the marginal effect of an attitude or belief is to use an interaction term the same way we would when exploring the marginal effects of age, gender or income. However, unlike age, gender and income, attitudes and beliefs are likely correlated with unobserved factors affecting choice (i.e. the error term) and indicators of attitudes and beliefs (e.g. Likert scale survey questions) are themselves imperfect measures of the true underlying attitude or belief. If either of these are true, then the model will be misspecified and the estimated parameters may be biased (endogeneity bias and measurement error) (Ben-Akiva et al., 2002; Hess, 2012).

Recently, the integrated choice and latent variable (ICLV), or hybrid choice model (Ben-Akiva et al., 2002; McFadden, 1986), popularized in transport (Bhat et al., 2015; Hess and Stathopoulos, 2013), has gained traction in environmental economics (Alemu and Olsen, 2019; Hoyos et al., 2015; Kassahun et al., 2016; Mariel and Meyerhoff, 2016; Taye et al., 2018; Zawojska et al., 2019). An ICLV model combines structural equation modelling with discrete choice modelling. In this modelling framework, we assume that (unobserved) character traits, such as pro-environmental attitudes, can be captured by one or more latent variables defined as functions of observable characteristics and measures intended to capture such attitudes, e.g. Likert scale questions. These latent variables can be included directly in our choice models to capture the effect of (latent) attitudes and beliefs on the probabilities of choice (Ben-Akiva et al., 2002). The popularity of ICLV models stems from claims that the inclusion of attitudes and beliefs through latent variables leads to improved forecasts (Vij and Walker, 2016; Yáñez et al., 2010), that it sheds more light on preference heterogeneity (Kassahun et al., 2016; Mariel and Meyerhoff, 2016), and that it allows for the inclusion of attitudinal variables and beliefs while avoiding issues with measurement error and possible endogeneity bias (Ben-Akiva et al., 2002; Guevara and Ben-Akiva, 2010).[1] The latter is only true under specific conditions (Vij and Walker, 2016).

Measurement error and endogeneity bias aside, the interpretability of the parameters in ICLV models remain a challenge, especially if we seek to use the model results to influence policy. In an ICLV model, indicators only affect choice indirectly through the latent variable. The latent variable is, by definition, unknown and has no direct interpretability. As such, the indicators can only be interpreted in relation to their directional impact on the latent variable and its directional impact on utility. For examples from environmental economics, see Kassahun et al. (2016) who study farmers' marginal willingness to pay (MWTP) to adopt irrigation methods, Taye et al. (2018) who study how people's environmental attitudes affect their MWTP for forest management options, Alemu and Olsen (2019) who try to understand how people's food choice motives affect their MWTP for

[1] For an overview of the historical development of hybrid discrete choice models, we refer the reader to (Bahamonde-Birke and Ortúzar, 2017).

insect based food products or Lundhede et al. (2015) who look at how perceived uncertainty about policy outcomes affect bird conservation under climate change. To aid interpretability of the latent variable and to gain a better understanding of what drives heterogeneity in welfare measures, Hoyos et al. (2015), Mariel and Meyerhoff (2016) and Mariel et al. (2018) argue in a series of papers, all in environmental economics, that practitioners should use exploratory factor analysis to identify which indicators are appropriate for each latent variable. This approach can also be helpful in model estimation, because more appropriate indicators should make estimation of the model easier. An alternative, or perhaps complement, to the exploratory analysis is to use already validated scales to elicit attitudes or personality traits (Alemu and Olsen, 2019; Boyce et al., 2019; Hoyos et al., 2015; Taye et al., 2018). That said, Vij and Walker (2016) show that a reduced form model without latent variables may fit the data at least as well as a latent variable model if the observable explanatory variables are good predictors of the latent variables, which is a specific case of the general result provided by (McFadden and Train, 2000). Chorus and Kroesen (2014) caution that using the results of an ICLV model to inform policies that seek to influence choice by targeting the latent variable is inappropriate given the cross-sectional nature of the data (i.e. only between-individual comparisons based on differences in the latent variable can be accommodated, rather than within-individual comparisons based on changes in the latent variable) and the possibly endogenous relationship between the latent variable and choice. It is also important to keep in mind that as the complexity of our models – and our ability to capture more heterogeneity – increase, we need to be careful that we do not tailor our model too close to the sample data. This may compromise our ability to generalize our model and results beyond the existing dataset and limit the usefulness to policy makers. While end users will often want to establish the relationship between the dependent variable(s) and a relatively small number of key independent variables, increasing model complexity is justified only if it produces reasonably more accurate results. While a familiar aphorism among econometricians is that "all models are wrong", some models are more wrong than others, and to be of practical use there is a need to ensure that our results are understandable and meaningful. That responsibility lies with us.

So what then, is the additional benefit of developing an ICLV model? We argue in this paper that while model fit is obviously important, it is not the be all and end all of model selection; and that while using hybrid models to suggest polices that target the latent variable itself is inappropriate (Chorus and Kroesen, 2014; Kroesen et al., 2017; Kroesen and Chorus, 2018), these models can provide rich insight into behaviour (Hess, 2012), help de-bias estimates (Vij and Walker, 2016), offer improvements in prediction in certain contexts (Vij and Walker, 2016) and reveal additional layers of heterogeneity (Hess, 2012; Mariel and Meyerhoff, 2016; Taye et al., 2018). However, we show that retrieving the true parameters of ICLV models can be challenging, and that the benefits of developing and using them are not always clear-cut.

This paper is a practical illustration of the points outlined above, and can work as a clarification for practitioners and policy makers alike. Using Monte Carlo simulations, we show the important role that correlation between the attributes, indicators and latent variables play in model selection and that the econometricians belief about the strength of this correlation is the main thing to consider when trying to decide whether an ICLV model is appropriate. Furthermore, we show that the bias of not accounting for these correlations in parameters and MWTP is generally increasing with the strength of the correlations. The practical implication is that the strength of the endogeneity bias from including the indicator directly in the choice model is related to the strength of the correlation between the indicator and the latent variable. For low degrees of correlation, omitting the latent variable or using a reduced form model does not lead to substantial bias in MWTP, but for high degrees of correlation between the indicator and the latent variable, the bias in MWTP is less than the model without indicators or latent variables. As such, our results can be viewed as an illustration of Vij and Walker (2016) and Kroesen and Chorus (2018).

The rest of the paper is outlined as follows: Section 2 outlines our econometric approach, Section 3 details the Monte-Carlo data generation processes, Section 4 presents the results from the simulation study, and Section 5 discusses the implications of our results for the use of ICLV models for policy and concludes the paper.

2. Econometric approach

To illustrate our point and substantiate our conclusions, we use a straightforward stated choice data setup. We generate synthetic datasets and show through Monte-Carlo simulation how misspecification of the model can lead to bias and under which circumstances this may not be the case. In the following, we assume that the reader is somewhat familiar with discrete choice modelling. To introduce notation, and to save space, we start with a standard random parameters mixed logit model where the probability of observing the sequence of T_n choices y_n made by individual n is a K dimensional integral of the logit formula over all possible values of $\widehat{\boldsymbol{\beta}}_n$:[2]

$$\Pr\left(\mathrm{y}_n \mid \mathrm{X}_n, \widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\gamma}}, \mathrm{z}_n, \widehat{\boldsymbol{\Sigma}}\right) = \underbrace{\int \int \dots \int}_{K} \prod_{t=1}^{T_n} \frac{\exp(\widehat{\boldsymbol{\beta}}_n \mathrm{x}_{njt})}{\sum_{j=1}^{J} \exp(\widehat{\boldsymbol{\beta}}_n \mathrm{x}_{njt})} f\left(\widehat{\boldsymbol{\beta}}_n \mid \widehat{\boldsymbol{\mu}}, \widehat{\boldsymbol{\gamma}}, \mathrm{z}_n, \widehat{\boldsymbol{\Sigma}}\right) d\left(\widehat{\boldsymbol{\beta}}_n\right), \tag{1}$$

where x_{njt} is a column vector of attribute levels and the joint density of the row vector $\widehat{\boldsymbol{\beta}}_n$ of marginal utilities is given by $f(\widehat{\boldsymbol{\beta}}_n|\cdot)$. A key consideration when specifying random parameters is the assumption regarding their distribution. In this paper, we express the individual marginal utility parameter for attribute k, $\hat{\beta}_{nk}$, as follows:

$$\hat{\beta}_{nk} = \hat{\mu}_k + \widehat{\boldsymbol{\gamma}}_k \mathrm{z}_n + \epsilon_{nk}, \tag{2}$$

where $\hat{\mu}_k$ is the mean of the distribution for attribute k, z_n is a column vector of regressors relating to individual-specific characteristics, e.g. age, gender, attitudinal responses or latent variables, $\widehat{\boldsymbol{\gamma}}_k$ is a conformable row vector of estimated mean shifter parameters and ε_{nk} is a deviate from a multivariate normal distribution with zero mean and covariance $\widehat{\boldsymbol{\Sigma}}$. Introducing individual specific characteristics, e.g. responses to an attitudinal question, allows us to assess and interpret the marginal effect of the attitudinal response on margin-

[2] The ICLV model can be specified with other choice kernels as well, e.g. multinomial logit or latent class, but throughout this paper, whenever we refer to the ICLV model it is one specified with a random parameters mixed logit kernel.

al utility in the same way as we would for age, gender or income. However, as discussed above, by including attitudinal measures directly in the model, we assume that responses to these attitudinal questions are direct measures of attitudes, e.g. pro-environmental attitudes, and that they are exogenous, i.e. that the responses are uncorrelated with the error terms. If either assumption is violated, our model is misspecified and our parameters may be biased. To avoid some of the issues associated with measurement error and endogeneity bias, we can, for example, use a hybrid choice model. In this model, we assume that the responses to the attitudinal questions are mapped to a latent variable that is included directly in the marginal utility expression just like we would for any other individual characteristic. In our case, the latent variable is given by the following structural equation:

$$\hat{\mathcal{L}}_n = \widehat{\omega}_n \text{SPACING}\widehat{\boldsymbol{\omega}} \sim N(0, \hat{\sigma}_{\mathcal{L}}^2), \tag{3}$$

where $\widehat{\omega}_n$ is a normally distributed random disturbance with zero mean and standard deviation $\hat{\sigma}_{\mathcal{L}}$ to be estimated. Responses to our pro-environmental behaviour question are given on a three-point Likert scale, as explained below. Since the response is on an ordered scale, we need to use an ordered model for the measurement equations (Daly et al., 2012). Let us create an underlying continuous variable, i^*, that determines the observed response to the indicator question. For individual n, we assume the following relationship with the latent variable:

$$i_n^* = \hat{\zeta} + \hat{\psi}\hat{\mathcal{L}}_n + \varepsilon_n, \tag{4}$$

where $\hat{\zeta}$ is a constant to estimate, $\hat{\psi}$ represents the variation of the underlying continuous variable for a unitary variation in the latent variable and ε_n is an idiosyncratic random disturbance term assumed to be a deviate from an identically and independently standard logistic distribution. Now, we can map the value of i_n^* to the observed cardinal response to the three-point indicator question. Specifically, with l denoting the index for the indicator response (i.e. $l \in \{1,2,3\}$), we have:

$$\hat{\mathrm{i}}_n = \begin{cases} l = 1 \ \text{ if } -\infty < \hat{\zeta} + \hat{\psi}\hat{\mathcal{L}}_n \leq \hat{\tau}_1 \\ l = 2 \ \text{ if } \hat{\tau}_1 < \hat{\zeta} + \hat{\psi}\hat{\mathcal{L}}_n \leq \hat{\tau}_2 \\ l = 3 \ \text{ if } \hat{\tau}_2 < \hat{\zeta} + \hat{\psi}\hat{\mathcal{L}}_n \leq \infty \end{cases}, \tag{5}$$

where $\hat{\tau}_1$ and $\hat{\tau}_2$ are threshold parameters to be estimated. In order to preserve the positive signs of all of the probabilities and ensure that the support is over the entire real line, there is a strict ordering of threshold values that demarcate the observed ordinal levels of the indicator question, specifically $-\infty < \hat{\tau}_1 < \hat{\tau}_2 < \infty$, with $\tau_0 = -\infty$ and $\tau_L = \infty$. With this in place, the probability for the response to the indicator question for individual n can be represented by the ordered logit model:

$$\Pr\left(i_n \mid \hat{\mathcal{L}}_n, \hat{\zeta}, \hat{\psi}, \hat{\boldsymbol{\tau}}\right) = \prod_{l=1}^{L}\left[\Lambda\left(\hat{\tau}_{l-1} + \hat{\tau}_l - \hat{\zeta} + \hat{\psi}\hat{\mathcal{L}}_n\right) - \hat{\zeta} + \hat{\psi}\Lambda\left(\hat{\tau}_{l-2} + \hat{\tau}_{l-1} - \hat{\mathcal{L}}_n\right)\right]^{\mathbb{I}_{n_l}}, \tag{6}$$

where $\Lambda(.)$ represents the standard logistic cumulative distribution function and $\mathbb{I}_{n_l}$ is a variable equal to one when the indicator level l is responded by individual n and zero otherwise.

To estimate the ICLV model, we need to maximize the joint likelihood of the observed sequence of choices and the observed responses to the Likert scale questions gauging pro-environmental behaviour. We can write the overall likelihood function as follows:

$$\Pr\left(y_n, i_n \mid X_n, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\gamma}}, \hat{\boldsymbol{\Sigma}}, \hat{\sigma}_{\mathcal{L}}, \hat{\zeta}, \hat{\psi}, \hat{\boldsymbol{\tau}}\right) = \int \Pr\left(y_n \mid X_n, \hat{\boldsymbol{\mu}}, \hat{\boldsymbol{\gamma}}, z_n, \hat{\boldsymbol{\Sigma}}\right) \Pr\left(i_n \mid \hat{\mathcal{L}}_n, \hat{\zeta}, \hat{\psi}, \hat{\boldsymbol{\tau}}\right) \phi\left(\hat{\mathcal{L}}_n \mid 0, \hat{\sigma}_{\mathcal{L}}^2\right) d\left(\hat{\mathcal{L}}_n\right), \tag{7}$$

where $\phi\left(\hat{\mathcal{L}}_n \mid 0, \hat{\sigma}_{\mathcal{L}}^2\right)$ denotes the normal density with mean zero and variance $\hat{\sigma}_{\mathcal{L}}^2$. Note the probability now involves a K+1 dimensional integral.

3. Synthetic data generating process and approach

3.1 Data

We use Monte-Carlo experiments to generate synthetic datasets. This is particularly useful because we know the true parameters underlying the data generating process (DGP) and will enable us to judge model performance in terms of how close the model estimates are to the true values. For this demonstration, we construct a stated choice experiment characterized by three environmental attributes: "area" represents the protected area (in 1,000 km^2) with levels 2, 4, 6, 8, 10 and 12; "broadleaf" denotes the fraction of newly planted trees that are broad-leafed with levels 0.0, 0.2, 0.4, 0.6, 0.8, and 1.0; and, "recreation" is a zero-one indicator variable signifying if recreation opportunities are available. The "cost" attribute is specified as having six levels: €5, €10, €15, €20, €25 and €30. Next we generate a random experimental design consisting of 500 synthetic individuals completing six choice tasks comprising two alternatives.[3] For the indicator question we make use of a three-point Likert-scale indicating environmental tendency: anti-environmental tendencies, neutral-environmental tendencies and pro-environmental tendencies.

Our Monte-Carlo strategy involves 25 data generation processes. In all settings, the model specification used in the DGP is based on the ICLV model with a random parameters mixed logit kernel described above. Specifically, we assume:

$$\beta_{nk}=\mu_k+\gamma_k L_n+\sigma_k \upsilon_{nk}, \tag{8}$$

where υ_{nk} is an independent standard Normal deviate, meaning that σ_k can be interpreted as the standard deviation of the (underlying) Normal distribution.[4] The parameter vector $\boldsymbol{\gamma}$ determines the direction and strength of the relationship between the latent variable and the marginal utilities. To asses how findings are sensitive to different values of $\boldsymbol{\gamma}$ we consider different vectors. This goes from the case where the latent variable has no bearing on any of the marginal utility distributions (i.e. where $\gamma_k=0\forall K$) to one in which it plays a

---

[3] While this design ensures that all attribute levels can be estimated independently of each other, we recognise that a more efficient experimental design could have been used to minimise the variance of the parameters. However, in a Monte Carlo experiment with specified parameters it may be more appropriate to show that the results stand up in cases where the experimental design is not tailored too closely to the data-generating parameters. Indeed, this would be the case in a real-life empirical application.

[4] For the cost attribute, we specify $\beta_{nk}=-\exp(\mu_k+\gamma_k L_n+\sigma_k \upsilon_{nk})$ to ensure strictly negative values.

large role. Given our DGP of a positive correlation between the latent variable and environmental tendency, we achieve this by increasing the γ_k values for the non-cost attributes and decreasing it for the cost attribute. Furthermore, we consider different values of ψ to contrast the suitability of the indicator question as a manifestation of the underlying latent environmental tendency, respectively, from the case where the Likert responses are independent of environmental tendencies to one in which they are, for all intents and purposes, direct measures of environmental tendency. We make use of an orthogonal setup with five sets of parameters to control for the strength of relationship between the latent variable and the marginal utility parameters and five parameters to control the strength of relationship between the latent variable and the indicator response, thus producing 25 different DGPs enabling independent evaluation. The respective γ_k and ψ for each DGP is reported in Table 1. The other parameters remain constant across all DGPs. Respectively, for the cost, area, broadleaf and recreation attributes the values of μ_k are -1.0, 0.6, 2.5 and 1.4, and the values of σ_k are 0.4, 0.1, 0.8 and 0.4. For σ_L, ζ, τ_1 and τ_1 we use 1.2, -0.6, -1.0 and 0.1, respectively.

In practice, we generated a deviate for each synthetic individual from $N\sim(0,\sigma_L^2)$ to represent their specific latent variable, and independent deviates from $N\sim(0,\sigma_k^2)$ to obtain their specific marginal utility. Additionally, for each utility function and their underlying continuous variable relating to the indicator we retrieved deviates from independently and identically distributed type I extreme value distributions with variance $\pi^2/6$. The choices are produced by identifying the alternative associated with the largest utility value. The individual counterfactual response to the three-point indicator question is established by comparing the simulated indicator distribution against the demarcation thresholds. Since idiosyncratic results can arise from a single sample of individuals, we generate 100 replications for every simulation setting.

To determine how well the simulated data reflect the DGP, we report a number of Pearson correlation coefficients for each data generation setting in Table 1. Specifically, $\rho_{W_{k},L}$ and ρ_{W_{k},i^*} denote the correlation between the MWTP for attribute k and the latent variable and the underlying continuous variable relating to the indicator, respectively. The correlation between the latent variable and the underlying continuous variable relating to the indicator is signified by ρ_{L,i^*}. We can see that the correlations reflect the DGP and, most importantly, that we separately control for differences in the influence of the latent variable on preferences and the indicator. It is also noticeable that the latent variable has a relatively stronger influence on the area attribute, followed by broadleaf and, lastly, recreation. This is a deliberate artefact of the parameters we used in DGP, since it allows us to compare the implications under a wider range of settings.

3.2 Analysis

For each dataset generated, we estimate six candidate models. This includes a random parameters mixed logit model (MXL), a random parameters mixed logit model with the indicators mapping directly to the marginal utilities (MXLIND) and a hybrid random parameters mixed logit model (LVMXL), that matches the DGP, where the latent variable enters both the marginal utility expressions and measurement equation relating to environmental tendency. It is widely acknowledged that models relying on the strict notion of

Table 1. Parameters used in the DGPs and mean correlations across the 100 Monte-Carlo simulations.

DGP	γ_{cost}	γ_{area}	$\gamma_{broadleaf}$	$\gamma_{recreation}$	ψ	$\rho_{W_{area},\mathscr{L}}$	ρ_{W_{area},i^*}	$\rho_{W_{broadleaf},\mathscr{L}}$	$\rho_{W_{broadleaf},i^*}$	$\rho_{W_{recreation},\mathscr{L}}$	$\rho_{W_{recreation},i^*}$	$\rho_{\mathscr{L},i^*}$
1	0.000	0.000	0.000	0.000	0.000	-0.01 (-0.09 – 0.09)	0.00 (-0.07 – 0.09)	0.00 (-0.08 – 0.08)	0.01 (-0.09 – 0.10)	0.00 (-0.08 – 0.08)	0.01 (-0.07 – 0.09)	0.00 (-0.08 – 0.07)
2	0.000	0.000	0.000	0.000	0.700	0.01 (-0.09 – 0.10)	0.01 (-0.09 – 0.09)	0.01 (-0.07 – 0.09)	0.00 (-0.10 – 0.09)	0.01 (-0.08 – 0.09)	0.01 (-0.12 – 0.11)	0.54 (0.50 – 0.60)
3	0.000	0.000	0.000	0.000	1.400	0.00 (-0.08 – 0.08)	0.00 (-0.09 – 0.08)	0.00 (-0.08 – 0.08)	0.00 (-0.09 – 0.08)	0.00 (-0.08 – 0.09)	0.00 (-0.07 – 0.09)	0.79 (0.77 – 0.81)
4	0.000	0.000	0.000	0.000	2.100	0.00 (-0.09 – 0.08)	0.00 (-0.07 – 0.08)	0.00 (-0.08 – 0.08)	0.00 (-0.08 – 0.08)	-0.01 (-0.08 – 0.07)	-0.01 (-0.09 – 0.06)	0.89 (0.88 – 0.90)
5	0.000	0.000	0.000	0.000	2.800	0.00 (-0.08 – 0.09)	0.00 (-0.07 – 0.09)	0.00 (-0.08 – 0.09)	0.00 (-0.07 – 0.08)	0.00 (-0.09 – 0.10)	0.01 (-0.10 – 0.09)	0.93 (0.93 – 0.94)
6	-0.050	0.075	0.125	0.025	0.000	0.42 (0.36 – 0.47)	0.00 (-0.08 – 0.09)	0.22 (0.14 – 0.28)	0.00 (-0.08 – 0.07)	0.16 (0.08 – 0.24)	0.00 (-0.08 – 0.10)	0.00 (-0.08 – 0.09)
7	-0.050	0.075	0.125	0.025	0.700	0.41 (0.33 – 0.48)	0.23 (0.14 – 0.31)	0.22 (0.11 – 0.30)	0.12 (0.03 – 0.19)	0.16 (0.05 – 0.23)	0.09 (0.00 – 0.17)	0.54 (0.50 – 0.59)
8	-0.050	0.075	0.125	0.025	1.400	0.41 (0.35 – 0.47)	0.33 (0.27 – 0.39)	0.21 (0.15 – 0.28)	0.17 (0.09 – 0.24)	0.15 (0.07 – 0.23)	0.12 (0.05 – 0.19)	0.79 (0.77 – 0.81)
9	-0.050	0.075	0.125	0.025	2.100	0.41 (0.36 – 0.47)	0.37 (0.30 – 0.43)	0.21 (0.15 – 0.28)	0.19 (0.12 – 0.28)	0.15 (0.08 – 0.23)	0.13 (0.06 – 0.21)	0.89 (0.88 – 0.90)
10	-0.050	0.075	0.125	0.025	2.800	0.41 (0.36 – 0.48)	0.39 (0.33 – 0.45)	0.21 (0.13 – 0.30)	0.20 (0.12 – 0.29)	0.15 (0.07 – 0.25)	0.14 (0.06 – 0.24)	0.93 (0.93 – 0.94)
11	-0.100	0.150	0.250	0.050	0.000	0.64 (0.61 – 0.68)	-0.01 (-0.11 – 0.08)	0.40 (0.34 – 0.46)	0.00 (-0.09 – 0.10)	0.30 (0.23 – 0.36)	0.00 (-0.08 – 0.09)	-0.01 (-0.09 – 0.07)
12	-0.100	0.150	0.250	0.050	0.700	0.65 (0.61 – 0.68)	0.36 (0.30 – 0.41)	0.40 (0.33 – 0.46)	0.22 (0.16 – 0.29)	0.30 (0.22 – 0.37)	0.17 (0.09 – 0.26)	0.55 (0.50 – 0.61)
13	-0.100	0.150	0.250	0.050	1.400	0.64 (0.61 – 0.68)	0.51 (0.46 – 0.56)	0.40 (0.34 – 0.46)	0.32 (0.24 – 0.39)	0.30 (0.23 – 0.35)	0.23 (0.15 – 0.31)	0.80 (0.78 – 0.82)
14	-0.100	0.150	0.250	0.050	2.100	0.65 (0.60 – 0.69)	0.58 (0.52 – 0.62)	0.40 (0.33 – 0.46)	0.36 (0.29 – 0.42)	0.30 (0.22 – 0.37)	0.27 (0.18 – 0.33)	0.89 (0.88 – 0.90)
15	-0.100	0.150	0.250	0.050	2.800	0.64 (0.61 – 0.68)	0.60 (0.56 – 0.64)	0.40 (0.33 – 0.46)	0.38 (0.31 – 0.44)	0.29 (0.23 – 0.36)	0.28 (0.20 – 0.35)	0.93 (0.93 – 0.94)
16	-0.150	0.225	0.375	0.075	0.000	0.75 (0.71 – 0.78)	0.00 (-0.08 – 0.08)	0.53 (0.48 – 0.59)	0.00 (-0.07 – 0.09)	0.43 (0.37 – 0.49)	0.00 (-0.08 – 0.09)	0.00 (-0.08 – 0.08)
17	-0.150	0.225	0.375	0.075	0.700	0.75 (0.70 – 0.77)	0.41 (0.36 – 0.48)	0.53 (0.48 – 0.59)	0.29 (0.22 – 0.38)	0.42 (0.36 – 0.48)	0.23 (0.15 – 0.32)	0.55 (0.50 – 0.60)
18	-0.150	0.225	0.375	0.075	1.400	0.75 (0.72 – 0.77)	0.59 (0.55 – 0.64)	0.53 (0.48 – 0.57)	0.42 (0.35 – 0.50)	0.42 (0.35 – 0.47)	0.33 (0.27 – 0.39)	0.79 (0.77 – 0.82)
19	-0.150	0.225	0.375	0.075	2.100	0.75 (0.70 – 0.78)	0.67 (0.62 – 0.71)	0.54 (0.48 – 0.59)	0.48 (0.43 – 0.54)	0.42 (0.35 – 0.49)	0.38 (0.31 – 0.44)	0.89 (0.88 – 0.90)
20	-0.150	0.225	0.375	0.075	2.800	0.75 (0.71 – 0.77)	0.70 (0.65 – 0.73)	0.54 (0.48 – 0.58)	0.50 (0.44 – 0.55)	0.42 (0.34 – 0.47)	0.39 (0.32 – 0.46)	0.93 (0.93 – 0.94)
21	-0.200	0.300	0.500	0.100	0.000	0.79 (0.76 – 0.82)	0.00 (-0.07 – 0.10)	0.63 (0.59 – 0.66)	0.00 (-0.08 – 0.11)	0.51 (0.45 – 0.57)	0.00 (-0.07 – 0.08)	0.00 (-0.07 – 0.08)
22	-0.200	0.300	0.500	0.100	0.700	0.79 (0.75 – 0.82)	0.43 (0.37 – 0.49)	0.63 (0.57 – 0.67)	0.34 (0.26 – 0.41)	0.52 (0.47 – 0.56)	0.28 (0.20 – 0.35)	0.55 (0.50 – 0.60)
23	-0.200	0.300	0.500	0.100	1.400	0.79 (0.75 – 0.82)	0.63 (0.58 – 0.66)	0.62 (0.58 – 0.66)	0.49 (0.44 – 0.55)	0.51 (0.47 – 0.55)	0.41 (0.35 – 0.46)	0.79 (0.77 – 0.81)
24	-0.200	0.300	0.500	0.100	2.100	0.79 (0.75 – 0.82)	0.70 (0.66 – 0.74)	0.62 (0.58 – 0.66)	0.56 (0.51 – 0.60)	0.51 (0.47 – 0.55)	0.46 (0.39 – 0.51)	0.89 (0.88 – 0.90)
25	-0.200	0.300	0.500	0.100	2.800	0.79 (0.76 – 0.82)	0.74 (0.71 – 0.78)	0.62 (0.58 – 0.66)	0.58 (0.54 – 0.62)	0.51 (0.45 – 0.56)	0.48 (0.42 – 0.53)	0.93 (0.93 – 0.94)

Corresponding 2.5[th] and 97.5[th] percentiles are reported in the parentheses. Colour coding: -0.05 – 0.95 (mean) Pearson correlation coefficient.

independent random parameters can be inferior to those that accommodate correlation (Mariel and Artabe, 2020; Mariel and Meyerhoff, 2018). While this correlation can stem from observable characteristics (e.g. gender, age and income), it may also be an artefact of unobserved latent variables. The importance of this latter point is often not fully appreciated. Indeed, a pertinent question is whether or not—and in what settings—allowing for correlation is an acceptable substitute for hybrid latent variable models and, conversely, if it is possible to say anything about the potential aptness of considering a hybrid latent variable model based on an inspection of the correlation structure of random parameters. To explore these issues we also estimate the corresponding models that allow for correlated random parameters (MXL-CORR, MXLIND-CORR and LVMXL-CORR, respectively). Estimating all six candidate models allows us to compare the effects under correctly specified and misspecified cases and to make inferences regarding the consequences of the naïve assumption(s). Combined, this leads to a total of 15,000 mixed logit models to estimate (i.e. 25 simulation treatments times 100 replications times six model specifications).

All models are coded and estimated using the maxLik library in R (see Henningsen and Toomet (2011) and R Core Team (2020) for further details). We used maximum simulated likelihood estimation using 500 quasi-random scrambled Sobol sequences for the simulation of the random parameters and latent variable. For all models, we started the estimation iterations using the parameters that were specified as part of the DGP.

4. Results

In Table 2, we show the mean difference in log-likelihood for all 25 DGPs over the 100 simulated Monte-Carlo datasets, and the corresponding 2.5th and 97.5th percentiles. Note that for the latent variable models we focus only on the fit of the choice model component, which we denote using LL*. First, and unsurprisingly, in accordance with Mariel and Meyerhoff (2018), we see that the models allowing for correlations between the random parameters fits the data better, i.e. produces higher log-likelihood values. This result holds for all three model specifications. However, we do note that the improvements in log-likelihood reported here do not penalise for the increased number of parameters. Second, including the indicator directly in the utility expression leads to better choice predictions. Referring back to the correlations between the indicator and latent variable for each DGP in Table 1, we see that this improvement in model fit is increasing in the degree of correlation between the two (i.e. ρ_{W_{k},i^*}). The most important take-away from this is that we find that the reduced form model without the latent variable fits the data equally well, which is consistent with Vij and Walker (2016). This fact really brings the question of what the additional benefit of a hybrid choice model is in many contexts to the forefront.

Moving beyond model fit is necessary to fully understand what is going on. While the reduced form models do "just as well" at predicting the chosen alternative, do they also retrieve unbiased and consistent estimates of the parameters and welfare measures? To explore this, in Table 3 and Table 4 we show the degree of bias in the parameters associated with the latent variable: specifically, with Table 3 and Table 4 comparing the mean error (i.e. the mean of all differences between the estimated values and the true value for each data generation setting) and the corresponding 2.5th and 97.5th percentiles for the models without and with correlations, respectively. We report the absolute bias, but rela-

Table 2. Mean improvement in log-likelihood (choice) over respective DGP baseline MXL model across the 100 Monte-Carlo simulations.

DGP	$LL_{MXL\text{-}corr} - LL_{MXL}$	$LL_{MXLIND} - LL_{MXL}$	$LL_{MXLIND\text{-}corr} - LL_{MXL}$	$LL^*_{LVMXL} - LL_{MXL}$	$LL^*_{LVMXL\text{-}corr} - LL_{MXL}$
1	2.61 (0.72 – 5.80)	3.94 (1.09 – 8.67)	6.49 (3.21 – 12.21)	0.80 (-0.48 – 3.18)	2.63 (0.40 – 5.61)
2	2.80 (0.83 – 6.01)	3.84 (1.19 – 8.16)	6.64 (2.61 – 11.60)	0.56 (-0.62 – 3.36)	2.80 (0.48 – 6.56)
3	3.17 (0.63 – 8.25)	3.86 (0.96 – 7.64)	6.96 (2.01 – 13.27)	1.23 (-0.51 – 6.15)	3.30 (0.27 – 8.91)
4	2.81 (0.73 – 6.60)	4.01 (0.81 – 8.27)	6.78 (2.71 – 12.01)	0.84 (-0.56 – 4.86)	2.96 (0.43 – 7.96)
5	2.70 (0.92 – 5.29)	4.22 (1.28 – 9.37)	6.93 (3.22 – 12.04)	0.71 (-0.49 – 3.82)	2.61 (0.48 – 5.66)
6	2.89 (0.82 – 6.70)	4.66 (1.71 – 9.66)	7.55 (3.52 – 13.17)	0.97 (-0.59 – 4.83)	2.92 (0.52 – 7.27)
7	3.07 (0.56 – 7.59)	7.76 (2.83 – 15.80)	10.75 (5.28 – 18.33)	0.68 (-1.07 – 5.66)	2.87 (0.07 – 6.94)
8	2.72 (0.86 – 5.88)	11.83 (4.58 – 21.67)	14.48 (6.66 – 24.49)	0.41 (-1.28 – 3.24)	2.59 (0.07 – 6.46)
9	3.05 (0.83 – 6.21)	13.24 (5.62 – 22.97)	16.23 (7.57 – 26.06)	0.75 (-1.35 – 4.56)	2.84 (0.38 – 6.20)
10	2.81 (0.65 – 7.12)	12.90 (6.00 – 20.92)	15.65 (8.04 – 23.95)	0.68 (-1.22 – 4.91)	2.55 (0.39 – 6.38)
11	3.66 (0.82 – 7.42)	3.93 (1.31 – 7.87)	7.61 (2.83 – 13.01)	1.63 (-0.64 – 5.93)	4.12 (0.55 – 8.51)
12	3.71 (0.84 – 7.82)	18.59 (8.54 – 32.76)	21.87 (11.68 – 35.90)	1.58 (-0.97 – 7.00)	3.61 (-0.10 – 8.02)
13	3.70 (1.21 – 8.33)	30.00 (18.04 – 42.02)	33.13 (21.65 – 45.63)	1.17 (-1.49 – 5.31)	3.31 (0.56 – 7.62)
14	3.40 (0.62 – 6.72)	34.15 (22.01 – 45.37)	36.93 (24.01 – 47.72)	1.04 (-1.60 – 3.83)	2.88 (-0.20 – 7.35)
15	3.78 (0.77 – 9.93)	35.66 (22.41 – 48.34)	38.77 (25.67 – 52.04)	1.23 (-1.80 – 5.07)	3.27 (-0.10 – 8.61)
16	5.85 (1.61 – 11.58)	4.27 (1.36 – 9.39)	10.13 (3.86 – 16.62)	2.73 (-0.48 – 8.77)	6.36 (1.70 – 13.09)
17	5.85 (2.69 – 12.59)	29.71 (16.42 – 43.96)	34.34 (22.39 – 49.87)	3.76 (-0.05 – 9.34)	5.91 (1.19 – 12.17)
18	6.27 (1.37 – 12.99)	51.78 (36.64 – 72.98)	55.72 (39.63 – 76.49)	3.84 (-1.36 – 9.92)	5.85 (-0.16 – 12.08)
19	6.00 (2.31 – 12.71)	60.54 (44.45 – 75.73)	63.91 (46.58 – 78.43)	3.77 (-0.62 – 9.16)	5.64 (1.16 – 12.07)
20	6.24 (1.97 – 13.13)	64.96 (44.41 – 84.90)	68.66 (46.50 – 89.99)	4.06 (-0.33 – 10.46)	5.91 (0.67 – 12.63)
21	10.28 (3.78 – 19.30)	3.74 (1.08 – 7.40)	14.08 (7.11 – 23.76)	6.53 (-0.51 – 16.35)	10.76 (3.75 – 17.82)
22	9.13 (3.71 – 15.91)	39.60 (24.32 – 55.60)	46.02 (30.55 – 63.27)	7.40 (0.71 – 14.11)	9.31 (3.69 – 16.53)
23	9.70 (3.56 – 17.87)	72.77 (51.73 – 94.77)	78.00 (56.74 – 100.87)	7.60 (1.09 – 15.96)	9.37 (2.90 – 17.13)
24	9.52 (3.49 – 16.60)	87.04 (68.13 – 107.79)	91.25 (72.57 – 111.79)	7.38 (1.42 – 14.40)	9.21 (2.54 – 16.81)
25	9.55 (3.25 – 18.01)	93.25 (71.70 – 116.24)	97.53 (76.31 – 119.89)	7.13 (-0.84 – 14.97)	8.92 (2.58 – 16.78)

Corresponding 2.5^{th} and 97.5^{th} percentiles are reported in the parentheses. Colour coding: 0 – 100 (mean) improvement.

tive bias can be assessed by referring back to the true parameters associated with any given DGP in Table 1. Nonetheless, the tables are useful to compare the different DGPs and signing the bias.

First, looking at Table 3, the most striking result is that, in general, the standard deviation of the latent variable is underestimated (first column), the parameters for cost and recreation are overestimated while those for area and broadleaf are (for the most part) underestimated. We also remark that the latent variable interaction with the indicator shows a high degree of bias for all simulation settings. Indeed, in situations where ψ=0 we find that the interaction is underestimated, whereas for settings where ψ>0 we find that they are overestimated. Furthermore, while the extent of the bias is increasing in ψ, we see no such pattern for the standard deviation of the latent variable or the other estimated parameters. Recall that the DGP was based on the LVMXL model. While we can normally expect to see idiosyncratic bias because the integrals are simulated and the data randomly generated, the fact that we observe systematic bias is a cause for concern and should make any practitioner think twice about using hybrid choice models. The inablity to recover the true parameters—even when the DGP is known and we start the estimation at the true parameters—is disconcerting and underlines the point that these models are difficult to estimate even under "perfect" conditions.[5] So what does this say about our ability to

[5] We recognise that 500 quasi-random draws may not have been sufficient and that increasing the number of simulation draws may have led to a more stable set of parameter estimates. We justify this on the grounds that,

Table 3. Bias for parameters connected with the latent variable in the LVMXL model.

DGP	$\hat{\sigma}_{\mathcal{L}} - \sigma_{\mathcal{L}}$	$\hat{\gamma}_{\text{cost}} - \gamma_{\text{cost}}$	$\hat{\gamma}_{\text{area}} - \gamma_{\text{area}}$	$\hat{\gamma}_{\text{broadleaf}} - \gamma_{\text{broadleaf}}$	$\hat{\gamma}_{\text{recreation}} - \gamma_{\text{recreation}}$	$\hat{\psi} - \psi$
1	-0.79 (-1.20 – 1.10)	1.72 (0.09 – 3.88)	-0.17 (-0.63 – 0.44)	0.00 (-0.35 – 0.27)	-0.44 (-9.47 – 2.06)	-1.33 (-15.89 – 8.56)
2	-0.94 (-1.20 – 1.35)	2.45 (0.23 – 10.96)	0.02 (-0.32 – 0.58)	0.02 (-0.14 – 0.25)	-0.11 (-0.91 – 0.62)	0.96 (-0.64 – 4.79)
3	-0.93 (-1.20 – 0.06)	1.81 (0.32 – 6.72)	-0.02 (-1.01 – 0.52)	0.00 (-0.19 – 0.18)	0.00 (-1.14 – 1.39)	0.17 (-1.39 – 5.40)
4	-0.85 (-1.20 – 1.20)	1.42 (-0.02 – 7.48)	0.05 (-0.44 – 0.75)	0.04 (-0.18 – 0.63)	0.39 (-1.52 – 7.10)	0.16 (-2.18 – 10.25)
5	-0.89 (-1.20 – 0.23)	1.70 (0.41 – 8.64)	-0.03 (-0.46 – 0.46)	-0.01 (-0.25 – 0.18)	-0.09 (-1.99 – 1.50)	1.81 (-2.69 – 17.37)
6	-0.48 (-1.20 – 3.45)	1.27 (-0.16 – 2.97)	-0.24 (-1.85 – 1.37)	-0.07 (-0.49 – 0.90)	0.86 (-1.80 – 6.15)	-0.69 (-17.27 – 14.39)
7	-1.08 (-1.20 – -0.78)	2.04 (0.89 – 4.89)	-0.14 (-0.28 – 0.01)	-0.08 (-0.15 – 0.01)	0.08 (-0.30 – 0.66)	1.61 (-0.49 – 3.11)
8	-1.07 (-1.19 – -0.79)	1.81 (0.90 – 4.61)	-0.14 (-0.38 – -0.01)	-0.05 (-0.13 – 0.05)	0.11 (-0.33 – 0.75)	1.72 (-1.01 – 11.07)
9	-1.06 (-1.19 – -0.71)	1.79 (0.78 – 9.70)	-0.18 (-0.59 – -0.05)	-0.05 (-0.12 – 0.04)	0.10 (-0.46 – 0.65)	1.62 (-1.48 – 13.00)
10	-1.02 (-1.20 – -0.56)	1.71 (0.76 – 9.38)	-0.19 (-0.61 – 0.03)	-0.04 (-0.14 – 0.09)	0.06 (-0.72 – 0.85)	0.93 (-2.22 – 12.37)
11	-0.03 (-1.19 – 8.87)	1.31 (-0.07 – 4.69)	-0.36 (-6.72 – 1.43)	-0.01 (-1.71 – 5.00)	0.43 (-8.26 – 10.50)	-1.11 (-9.53 – 6.77)
12	-1.10 (-1.19 – -0.78)	2.21 (1.07 – 7.98)	-0.23 (-0.46 – -0.14)	-0.16 (-0.24 – -0.03)	0.09 (-0.28 – 0.55)	2.72 (-0.31 – 12.13)
13	-1.11 (-1.20 – -0.93)	1.84 (1.06 – 3.92)	-0.24 (-0.41 – -0.14)	-0.13 (-0.20 – 0.00)	0.15 (-0.20 – 0.65)	3.54 (-0.55 – 32.94)
14	-1.09 (-1.19 – -0.93)	1.60 (1.00 – 3.59)	-0.25 (-0.36 – -0.16)	-0.11 (-0.20 – 0.00)	0.19 (-0.17 – 0.50)	3.18 (-0.93 – 24.82)
15	-1.07 (-1.20 – -0.81)	1.85 (0.86 – 7.87)	-0.26 (-0.45 – -0.15)	-0.10 (-0.23 – 0.01)	0.22 (-0.12 – 0.71)	4.93 (-1.37 – 40.58)
16	0.37 (-1.19 – 6.28)	1.12 (-0.39 – 4.17)	-0.39 (-7.00 – 4.43)	-0.09 (-7.46 – 6.36)	0.62 (-17.72 – 14.88)	0.13 (-7.86 – 11.38)
17	-1.11 (-1.20 – -0.96)	2.13 (1.44 – 3.77)	-0.33 (-0.44 – -0.23)	-0.24 (-0.33 – -0.15)	0.14 (-0.25 – 0.50)	1.66 (-0.26 – 11.99)
18	-1.09 (-1.20 – -0.86)	1.60 (1.04 – 2.59)	-0.36 (-0.52 – -0.27)	-0.18 (-0.29 – -0.06)	0.26 (-0.10 – 0.80)	1.72 (-0.32 – 12.87)
19	-1.08 (-1.19 – -0.91)	1.53 (1.09 – 2.40)	-0.38 (-0.50 – -0.27)	-0.17 (-0.26 – -0.06)	0.27 (-0.10 – 0.64)	4.19 (-0.54 – 32.10)
20	-1.09 (-1.20 – -0.89)	1.62 (0.99 – 5.40)	-0.39 (-0.59 – -0.26)	-0.14 (-0.32 – -0.02)	0.30 (-0.02 – 0.70)	3.78 (-0.88 – 21.76)
21	0.82 (-1.20 – 8.75)	1.27 (-0.01 – 4.31)	-1.32 (-9.88 – 4.62)	0.90 (-8.21 – 8.95)	2.87 (-15.59 – 16.90)	-0.81 (-9.48 – 4.24)
22	-1.11 (-1.20 – -0.95)	2.04 (1.56 – 2.49)	-0.44 (-0.52 – -0.37)	-0.31 (-0.38 – -0.20)	0.23 (-0.06 – 0.45)	0.19 (-0.25 – 1.59)
23	-1.08 (-1.19 – -0.87)	1.80 (1.08 – 3.42)	-0.48 (-0.63 – -0.36)	-0.25 (-0.40 – -0.08)	0.32 (-0.04 – 0.73)	2.88 (-0.17 – 15.75)
24	-1.07 (-1.20 – -0.90)	1.58 (1.18 – 3.59)	-0.51 (-0.62 – -0.36)	-0.21 (-0.40 – -0.07)	0.36 (-0.09 – 0.83)	3.43 (-0.14 – 18.75)
25	-1.07 (-1.19 – -0.89)	1.72 (1.14 – 4.66)	-0.49 (-0.63 – -0.34)	-0.21 (-0.42 – -0.10)	0.33 (-0.09 – 0.73)	6.15 (-0.25 – 42.25)

Corresponding 2.5th and 97.5th percentiles are reported in the parentheses. Colour coding: -9.3 8.8 (mean) bias.

Table 4. Bias for parameters connected with the latent variable in the LVMXL-CORR model.

DGP	$\hat{\sigma}_{\mathcal{L}} - \sigma_{\mathcal{L}}$	$\hat{\gamma}_{\text{cost}} - \gamma_{\text{cost}}$	$\hat{\gamma}_{\text{area}} - \gamma_{\text{area}}$	$\hat{\gamma}_{\text{broadleaf}} - \gamma_{\text{broadleaf}}$	$\hat{\gamma}_{\text{recreation}} - \gamma_{\text{recreation}}$	$\hat{\psi} - \psi$
1	5.28 (-1.17 – 53.81)	-0.03 (-0.79 – 1.07)	-0.04 (-0.31 – 0.25)	-0.61 (-7.76 – 0.71)	-0.08 (-1.45 – 0.51)	-7.97 (-56.61 – -1.42)
2	3.62 (-1.18 – 40.18)	0.05 (-0.27 – 0.68)	-0.01 (-0.20 – 0.15)	-0.08 (-0.88 – 1.05)	-0.02 (-0.83 – 0.59)	-6.92 (-43.68 – -1.96)
3	1.60 (-1.16 – 23.06)	-0.02 (-1.03 – 0.50)	0.00 (-0.32 – 0.38)	0.05 (-1.66 – 2.00)	0.04 (-1.00 – 1.02)	-5.37 (-27.26 – -2.52)
4	1.51 (-1.16 – 22.59)	-0.13 (-1.79 – 0.61)	-0.03 (-0.57 – 0.22)	0.20 (-2.75 – 6.69)	0.02 (-1.40 – 1.74)	-5.82 (-27.49 – -3.13)
5	2.27 (-1.13 – 20.64)	-0.03 (-0.85 – 0.28)	-0.01 (-0.11 – 0.14)	-0.06 (-1.81 – 1.07)	-0.04 (-0.81 – 0.51)	-7.34 (-26.24 – -3.82)
6	3.76 (-1.18 – 50.47)	-0.21 (-1.12 – 1.04)	-0.10 (-0.96 – 0.35)	0.46 (-2.49 – 5.91)	-0.29 (-5.01 – 2.38)	-6.44 (-53.27 – -1.42)
7	5.15 (-1.16 – 44.43)	0.03 (-0.11 – 0.14)	-0.03 (-0.08 – 0.07)	-0.05 (-0.38 – 0.35)	0.00 (-0.23 – 0.24)	-8.53 (-47.93 – -2.02)
8	2.16 (-1.18 – 28.30)	-0.02 (-0.30 – 0.10)	-0.01 (-0.08 – 0.12)	-0.01 (-0.49 – 0.50)	-0.02 (-0.47 – 0.24)	-6.08 (-32.50 – -2.56)
9	2.55 (-1.17 – 25.35)	-0.04 (-0.32 – 0.08)	0.00 (-0.08 – 0.11)	-0.02 (-0.57 – 0.55)	0.00 (-0.33 – 0.40)	-7.06 (-30.25 – -3.18)
10	2.11 (-1.12 – 21.01)	-0.04 (-0.31 – 0.11)	0.00 (-0.07 – 0.15)	-0.05 (-0.63 – 0.53)	-0.02 (-0.40 – 0.41)	-7.16 (-26.61 – -3.83)
11	1.85 (-1.16 – 42.68)	-0.14 (-6.11 – 3.48)	-0.04 (-1.84 – 4.14)	0.18 (-10.73 – 13.76)	0.16 (-2.60 – 6.40)	-4.49 (-45.48 – -1.42)
12	6.38 (-1.19 – 45.15)	0.03 (-0.19 – 0.11)	-0.07 (-0.14 – 0.05)	-0.14 (-0.50 – 0.22)	-0.02 (-0.27 – 0.21)	-9.78 (-48.65 – -2.05)
13	3.58 (-1.17 – 32.89)	0.01 (-0.14 – 0.10)	-0.04 (-0.13 – 0.10)	-0.06 (-0.43 – 0.49)	-0.04 (-0.22 – 0.16)	-7.63 (-37.09 – -2.70)
14	2.99 (-1.18 – 23.36)	0.01 (-0.14 – 0.09)	-0.03 (-0.13 – 0.09)	-0.06 (-0.39 – 0.29)	-0.01 (-0.26 – 0.25)	-7.72 (-28.26 – -3.28)
15	3.47 (-1.14 – 20.67)	0.00 (-0.16 – 0.09)	-0.02 (-0.13 – 0.12)	-0.02 (-0.37 – 0.45)	-0.03 (-0.32 – 0.31)	-8.79 (-26.27 – -3.92)
16	3.30 (-1.17 – 50.50)	0.21 (-12.23 – 4.72)	0.08 (-3.78 – 5.22)	0.22 (-17.39 – 13.41)	-0.21 (-4.24 – 4.12)	-5.91 (-53.30 – -1.42)
17	4.30 (-1.18 – 44.74)	0.06 (-0.17 – 0.14)	-0.11 (-0.20 – -0.01)	-0.22 (-0.55 – 0.11)	-0.04 (-0.21 – 0.16)	-7.70 (-48.24 – -2.07)
18	2.30 (-1.17 – 27.76)	0.03 (-0.15 – 0.13)	-0.06 (-0.18 – 0.08)	-0.11 (-0.45 – 0.32)	-0.03 (-0.24 – 0.24)	-6.40 (-31.96 – -2.72)
19	2.58 (-1.19 – 23.51)	0.02 (-0.15 – 0.12)	-0.04 (-0.18 – 0.13)	-0.07 (-0.37 – 0.40)	-0.01 (-0.20 – 0.22)	-7.35 (-28.41 – -3.35)
20	2.67 (-1.18 – 20.77)	0.00 (-0.19 – 0.13)	-0.02 (-0.18 – 0.14)	-0.05 (-0.35 – 0.45)	-0.02 (-0.21 – 0.23)	-8.06 (-26.37 – -3.95)
21	2.38 (-1.17 – 42.52)	-1.40 (-9.36 – 8.25)	0.98 (-2.60 – 12.66)	1.56 (-8.79 – 16.31)	0.87 (-4.04 – 13.92)	-4.98 (-45.32 – -1.42)
22	2.29 (-1.19 – 41.54)	0.07 (-0.05 – 0.18)	-0.13 (-0.21 – 0.00)	-0.23 (-0.53 – 0.14)	-0.07 (-0.31 – 0.16)	-5.65 (-45.04 – -2.06)
23	4.80 (-1.19 – 34.98)	0.06 (-0.10 – 0.17)	-0.09 (-0.26 – 0.11)	-0.15 (-0.47 – 0.26)	-0.04 (-0.29 – 0.22)	-8.91 (-39.18 – -2.70)
24	2.28 (-1.17 – 22.37)	0.00 (-0.21 – 0.16)	-0.03 (-0.26 – 0.15)	-0.10 (-0.51 – 0.37)	-0.02 (-0.28 – 0.19)	-7.01 (-27.27 – -3.35)
25	3.67 (-1.19 – 22.58)	0.03 (-0.16 – 0.16)	-0.05 (-0.23 – 0.12)	-0.14 (-0.50 – 0.41)	-0.04 (-0.30 – 0.19)	-9.12 (-28.18 – -4.03)

Corresponding 2.5th and 97.5th percentiles are reported in the parentheses. Colour coding: -9.3 8.8 (mean) bias.

retrieve unbiased parameters in empirical settings when the DGP and its parameters are unknown?

---

in total, we estimated 15,000 mixed logit models. Increasing the number of draws would have entailed considerably more estimation time.

Turning our attention to the model with correlation reported in Table 4, we see some stark differences compared to the models without correlation. The most notable change is the switching signs and larger magnitude of the bias in the standard deviation of the latent variable and ψ. This shows overwhelming evidence that allowing for correlation in the random parameters when this was not part of the DGP leads to severe bias in the parameters associated with the latent variable when the latent variable is the only source of correlation in the data. Intuitively, this makes sense. We now have a whole correlation structure, in addition to the latent variable, trying to describe the influence of the latent variable. Crucially, the magnitude of the bias of ψ is important because it can lead to an entirely misleading interpretation of the latent variable. Note that given the true parameters of ψ in Table 1, the magnitude of the bias implies that the estimated value of the impact of the latent variable will be negative. Consequently, *ceteris paribus*, we would wrongly conclude that an increase in the latent variable is associated with an increase in the MWTP for the environmental attributes and a decrease in the tendency to report pro-environmental attitudes on our three-point Likert scale question. If we look at the bias in the γ parameters, we see that this is much smaller compared to the LVMXL model. While the bias for the standard deviation of the latent variable and the latent variable indicator interactions switch signs and are considerably larger, the bias for γ is much smaller, which makes it difficult to ascertain the net effect on welfare estimates. What it does highlight, and we cannot stress this enough, is that there appears to be a dilemma and a set of unforeseen trade-offs when it comes to hybrid choice model selection. The extent to which this is just an artefact of our DGP parameters and assumptions remain unclear, as this would require further simulation work under a broader range of settings. Nonetheless, it does show that model selection comes down to the analyst's belief about correlations and that model selection and the use of these models truly is "a road fraught with peril".

To determine how the above results affect MWTP, we compare the overlapped estimated area of the actual MWTP kernel density estimates to that of the kernel density of the distribution of the means of the individual-specific posterior MWTP. This is an easy way to quantify the similarities or differences between the actual and predicted MWTP distributions. To make the comparison more intuitive, we consider the difference in the percentage overlap of each model to the basic MXL model, which represents the most naïve assumptions about the DGP. To illustrate how this difference is sensitive to the correlation between the DGP MWTP and the latent variable, as well as the indicator, we plot the differences against $\rho_{W_{k},L}$ and $\rho_{W_{k},i^{*}}$, respectively. We sort the corresponding points by the correlation measure and graph this using a technique known as locally estimated scatterplot smoothing (LOESS).[6] We show these in Figure 1, Figure 2 and Figure 3 (and their associated 95 percent confidence level) for the area, broadleaf and recreation attributes, respectively. Specifically, the locally regressed and smoothed percentage point differences in overlap of each candidate model relative to the MXL model are plotted against: (i) the correlation between the actual MWTP and the latent variable in the left panel; and, (ii) the correlation between the actual MWTP and the underlying continuous variable relat-

[6] The LOESS method is a non-parametric approach where fitting is done locally (in our case with a neighbourhood proportion of 0.4). The result is a smooth curve, which makes it easier to detect trends. This was achieved using the stats library in R .

Figure 1. Percentage point difference in overlap of MWTP distributions relative to the true MWTP distribution for area.

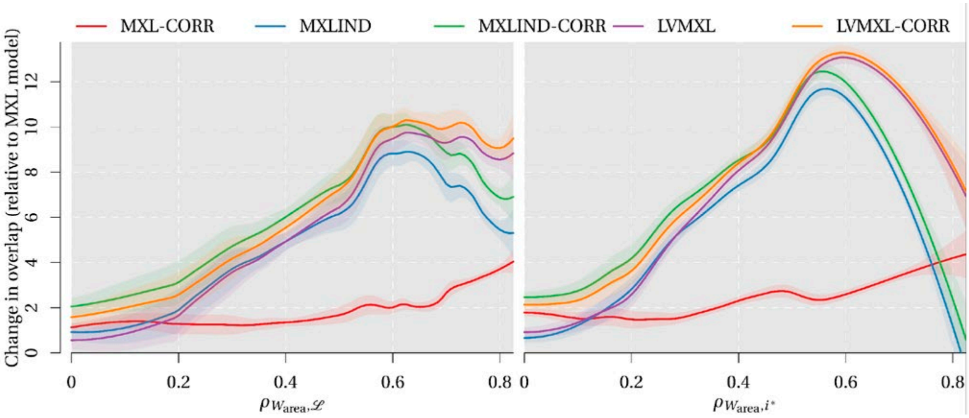
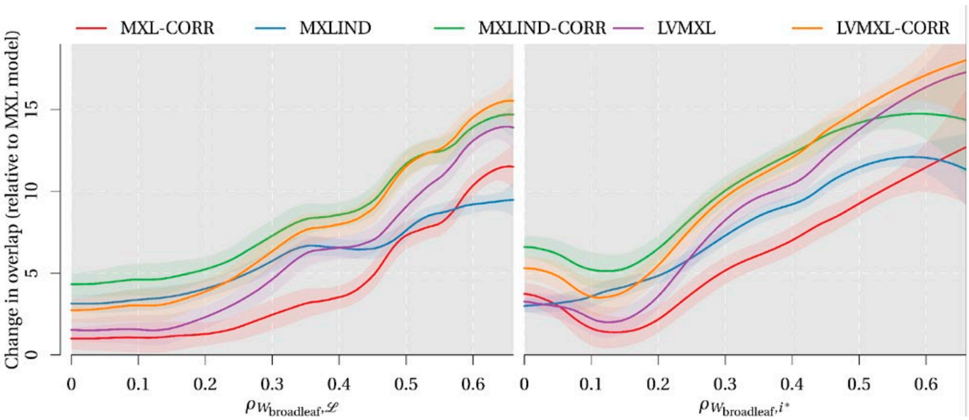


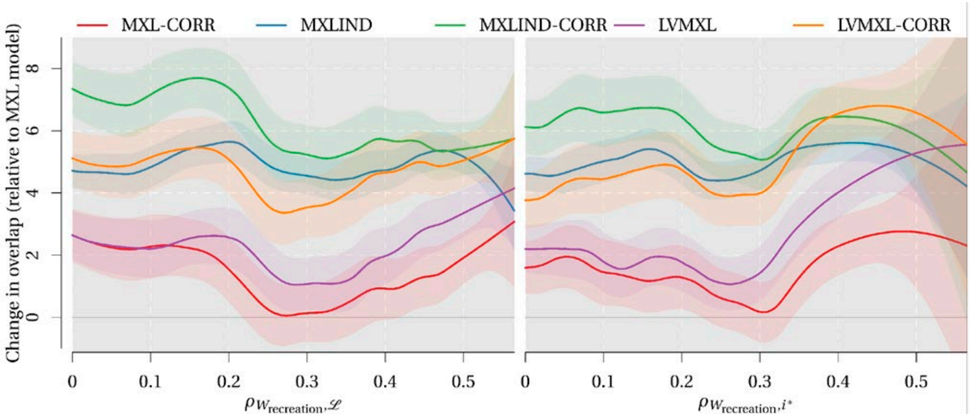
Figure 2. Percentage point difference in overlap of MWTP distributions relative to the true MWTP distribution for broadleaf.



ing to the indicator in the right panel. As we move from the origin to the right, the degree of correlation increases. The vertical axis shows the percentage point difference in overlap relative to the MXL model, meaning that a move up this axis signifies that the candidate model does better at predicting the true MWTP distribution relative to the MXL model.

As might be expected, a visual inspection of Figures 1-3 reveals that all models generally retrieve the same MWTP distributions when the correlation between MWTP and either the latent variable or indicator is low. But, as the degree of correlation increases, we can see that the models that accommodate correlated random parameters and/or environmental tendency (either directly or indirectly) are better at explaining the true MWTP distributions. Recall the discussion relating to the switching signs for the bias in the standard deviation of the latent variable and the interaction of the latent variable;

Figure 3. Percentage point difference in overlap of MWTP distributions relative to the true MWTP distribution for recreation.



this switch does not appear to affect the estimation of MWTP. Relatively speaking, for the most part, the LVMXL and LVMXL-CORR curves are closely aligned.

Focusing on Figure 1, we see that as the correlation with the latent variable (left panel) increases beyond 0.2, the models that directly or indirectly include the indicator outperform the MXL and MXL-CORR models. This is an important finding, since it suggests that simply allowing for correlation does not, in itself, allow us to recover the correct MWTP distribution. However, it must be noted that this result is strongest in cases where the correlation with the latent variable is moderate. As the strength of the relationship gets very high (ρ>0.6) there is a clear turning point, indicating that the relative importance of directly or indirectly including the indicator lessens. But this same finding is not observed for the MXL-CORR model, to the extent that just allowing for correlated random parameters does all most just as well at retrieving the correct MWTP to pay distribution. Importantly, this suggests that if the analyst believes that most, if not all, of the correlation between the random parameters are caused by a single unobserved latent variable, and that the effect of this variable is sufficiently strong, then simply estimating a standard mixed logit model with a full correlation structure may be sufficient if MWTP is the key measure of interest. Though, of course, this comes at the expense of not knowing the underlying source of heterogeneity, which may, or may not, be of interest. We also observe that the MXLIND and MXLIND-CORR are better able to uncover the true MWTP distribution compared to the LVMXL and LVMXL-CORR, respectively, when the strength of relationship between the latent variable and MWTP is weak or moderate. As the strength of relationship increases, however, we remark that this no longer holds. This additional insight implies that the relative advantage of ICLV models over simpler models to retrieve the correct MWTP distribution is dependent on the strength of the role that the latent variable plays on the distribution. While not a surprising finding, it reinforces the need to think twice about using hybrid choice models in situations where it is believed that the latent variable is weakly related. These findings are perhaps better illustrated when the change in overlap is plotted against the correlation with the indicator. The downward turn towards the MXL baseline is even more pronounced at higher levels of correlation for all but the MXL-CORR model and especially so for the models that directly include the indicator responses in the utility function. At this point, recall that the "area" attribute is also the one that is linked the strongest to the latent variable and the indicator. This explains why the difference between the models are so stark and why we see that this result is mitigated as the relationship between MWTP and the latent variable and indicator becomes weaker. For example, looking at the Figures 2 and 3, where the strengths of association are lower, we see that the predicted curves are more closely aligned and do not exhibit an inverted U-shape. This implies that, for these attributes, models that include the indicator (either directly or indirectly) do not produce markedly better predictions of the MWTP distribution compared to the MXL-CORR model, and this holds irrespective of the correlation between MWTP and either the latent variable or indicator. For the recreation attribute (Figure 3), which had the lowest association with the latent variable and indicator response, the predicted curves are relatively flat, suggesting that the prediction of the MWTP distribution is less sensitive to which of candidate models is used. While these are also obvious findings, the fact that we are able to retrieve, show and prove them through our simulation is reassuring.

In generating the results illustrated in Figures 1-3 we took account of all synthetic individuals per DGP. However, this may mask the relative performance of each candidate model to correctly predict MWTP for individuals who hold a particular environmental tendency. Indeed, one of the often-purported advantages of ICLV models is their ability to provide additional insight on preference heterogeneity, particular among those with different latent attitudes. While, as stated earlier, we should be prudent about making policy recommendations on the basis of a latent variable – as well as the, obvious, impracticality, and futility, of targeting policy on the basis of an indicator response – policy makers may still be interested in knowing how members of society with different environmental tendencies judge their policies. For this reason, in Figures 4-5, we plot the locally regressed and smoothed mean bias in MWTP for each attribute against the correlation between the attribute and the latent variable broken down by whether an individual holds anti-, neutral or pro-environmental tendencies, depicted on the first, second and third panel, respectively.[7] For comparison, in the fourth panel, we also present this for all individuals. Looking firstly at this fourth panel, we see that the curves essentially overlap and are not significantly different from zero when the correlation between MWTP and the latent variable is weak or moderate. In these cases, the ability to retrieve the mean MWTP (across all individuals) does not appear to be affected by the degree of correlation with the latent variable nor by which of the candidate models we use. As can be seen in Figures 4-5, however, these curves begin to diverge as the degree of correlation increases (ρ>0.5) to the extent that some are significantly different from zero. This insight suggests that if the main interest is on describing the means of the posterior MWTP distributions at the sample level, model choice is perhaps only consequential when the MWTP distribution

[7] For this, we subtract the actual individual-specific MWTP from the mean of the predicted individual-specific posterior MWTP and take the arithmetic mean for each data generation setting and model and, again, apply the LOESS method with a smoothing parameter of 0.4. The results are qualitatively similar for correlations between the attribute and the indicator and is omitted from the paper for brevity, but are available from the corresponding author upon request.

Figure 4. Mean bias in MWTP broken down by anti-, neutral and pro-environmental tendencies for area.

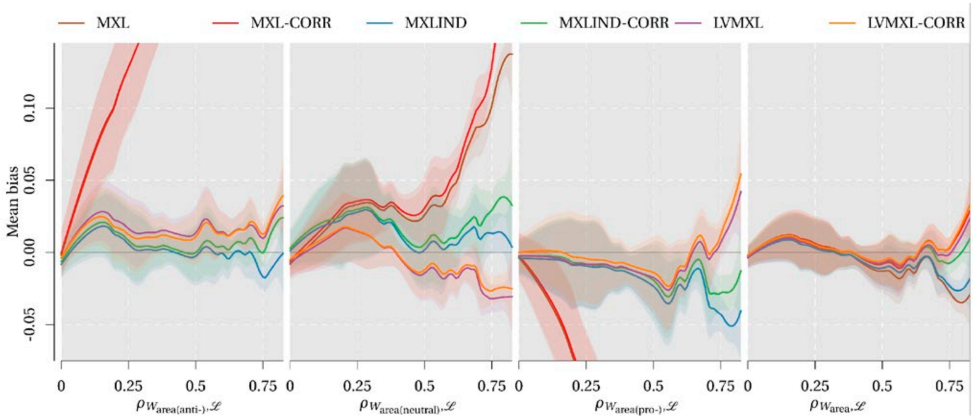
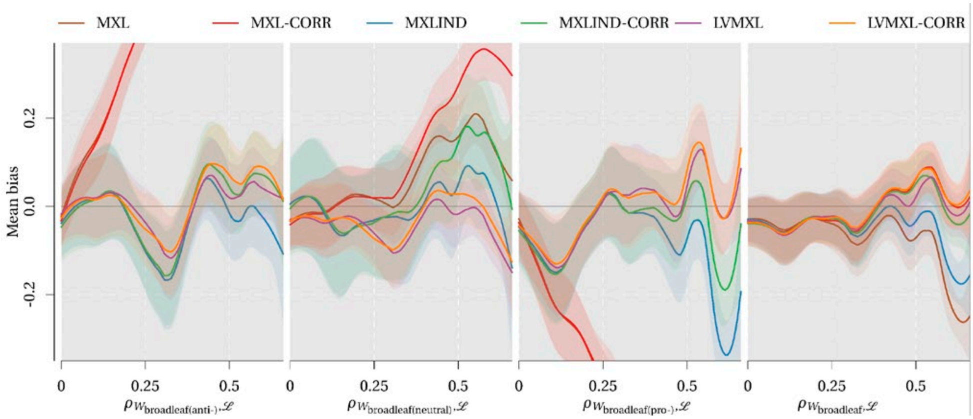


Figure 5. Mean bias in MWTP broken down by anti-, neutral and pro-environmental tendencies for broadleaf.



is believed to be strongly correlated with the latent variable. This is expected given the results above and that more flexible models are preferred if you suspect high degrees of correlation between MWTP and the latent variable. However, the corresponding curves produced for individuals who hold anti-, neutral or pro-environmental tendencies tell a somewhat different story. Only when the MWTP and latent variable distributions are uncorrelated do we find that all models produce relatively unbiased estimates of MWTP irrespective of environmental tendency. However, with any degree of correlation we see that the MXL and MXL-CORR models produce biased MWTP estimates for each subgroup. Specifically, these models overestimate individual-specific MWTP for individuals who hold anti- and (albeit to a lesser extent) neutral environmental tendencies, whereas they underestimate for the subgroup with pro-environmental tendencies. An important finding for analysts who make use of individual-specific posterior MWTP estimates is that the extent of these biases increase with the degree of correlation. While this trend of overestimating MWTP for anti- as well as neutral environmental tendencies and underestimating for pro-environmental tendencies still largely holds for the other candidate models, it is less evident and we observe it to be less sensitive to the degree of correlation. Nonetheless, there appears to be systematic differences between the models where we have included the indicators directly and their analogous latent variable models. For example, in Figures 4-5, relative to the MXLIND and MXLIND-CORR models, the LVMXL and LVMXL-CORR models, respectively, produce higher MWTP estimates for the anti- and pro-environmental tendency subgroups, but lower estimates for the neutral subgroup. Furthermore, these differences become more apparent as the degree of correlation between the MWTP and latent variable distributions increase. While the extent to which this finding can be generalized beyond our data generation settings is unclear, it does, nonetheless, further emphasise the difficultly associated with model selection when latent attitudes are believed to play an important role on MWTP.

5. Discussion and concluding remarks

In this paper, we generate a series of Monte-Carlo simulations that separately control for the strength of relationship between the latent variable and preferences and the strength of relationship between the latent variable and the indicator. In the real world, structural equations usually comprise standard socio-demographic characteristics and are often weak. To mimic this in the present paper, without complicating the DGP more than necessary, we treat the latent variable as normally distributed with zero mean and estimated standard deviation. This is exactly identical to a structural equation containing only an error-term. This also means that our reduced form model is a mixed logit model with an additional random error component (Vij and Walker, 2016). In the present paper, we used a simple three-point Likert scale question as an indicator of environmental tendency. This indicator was included in an ordered logit measurement equation. For each dataset generated, we estimate a random parameters mixed logit model, a random parameters mixed logit model with the indicators mapping directly to the marginal utilities and an ICLV random parameters mixed logit model, each with and without allowing for correlation among the random parameters.

From our results, it is clear that if you are only interested in choice prediction, then a mixed logit model with correlation may perform equally well to a hybrid choice model. While this is consistent with the general result of Vij and Walker (2016), who suggest that a reduced form model will fit the data at least as well, this is not in and of itself a reason to not use ICLV models. As we, and others, have shown, such models can offer greater insight into underling behavioural phenomena and contribute to decomposing marginal effects of the latent attitude on welfare estimates. But whether or not these additional behavioural insights outweigh the costs of estimating them remains an empirical question and will be entirely context dependent. In our simulations, we show that if the structural and measurement equations are weak (i.e. if observable characteristics are poor predictors of the latent variables, if appropriate indicators are not available and/or if the correlation between preferences, the latent construct and the indicator is weak), then the model's ability to separately identify the marginal effects are likely limited and the benefits of developing and using an ICLV model are less clear-cut. In the cases where we do have weak structural equations, the use of measurement equations can help explain the latent variable and improve the fit of our choice model. Unfortunately, in real world applications we do not know *a priori* whether an indicator is good or bad, nor is there much guidance on the strengths of correlation. But there are ways to identify *better* indicators using, for example, exploratory factor analysis (Hoyos et al., 2015; Mariel et al., 2018; Mariel and Meyerhoff, 2016). Ultimately, however, we show that model selection should be driven by the analyst's belief about the strength of correlations between preferences, the latent variable and indicator. In any case, we need to be careful and mindful of the criticisms of Chorus and Kroesen (2014) and Kroesen and Chorus (2018): given the potential endogenous relationship between the latent variable and choice and the cross-sectional nature of the data, it is impossible to ascertain a causal relationship between attitudes and behaviour meaning that we should be very careful recommending policies that target the latent variable itself.

While we have not spent significant time talking about prediction in the present paper, we do feel it is prudent to reiterate that hybrid choice models can lead to improved predictions, but that any improvements are only likely in the case where it would be possible to predict the future state of the latent variable itself (Vij and Walker, 2016; Yáñez et al., 2010). More likely than not, this type of data will not be available. This is also possibly why we see that our models fit the data equally well, i.e. in terms of explaining the sequence of choices made by individuals. That said, the conclusions in this paper echo those of many others (Chorus and Kroesen, 2014; Kroesen and Chorus, 2018; Vij and Walker, 2016), that we need to take better heed of the quality of our data and recognize the limitations of it. The usefulness of ICLV models hinge on the quality of the data, and an ICLV model applied to poor data may add nothing to explanatory power and even less to policy. Furthermore, it is clear from our simulation work that even under "perfect" conditions, we struggled to retrieve the true parameters of the model, and the appropriateness of the model itself came down to the degree of correlation between our attributes, latent variable and indicator. Taken together, this makes the use of latent variable models, perhaps especially to inform policy, a road fraught with peril.

Acknowledgements

Erlend Dancke Sandorf acknowledges funding from the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie grant agreement No 793163. Any remaining errors are the sole responsibility of the authors.

References

Alemu, M.H., Olsen, S.B., 2019. Linking Consumers' Food Choice Motives to their Preferences for Insect-based Food Products: An Application of Integrated Choice and Latent Variable Model in an African Context. *Journal of Agricultural Economics* 70, 241–258. https://doi.org/10.1111/1477-9552.12285

Bahamonde-Birke, F.J., Ortúzar, J. de D., 2017. Analyzing the continuity of attitudinal and perceptual indicators in hybrid choice models. Journal of Choice Modelling, IATBR 2015 - 14th International Conference on Travel Behaviour Research (IATBR) 25, 28–39. https://doi.org/10.1016/j.jocm.2017.01.003

Ben-Akiva, M., Mcfadden, D., Train, K., Walker, J., Bhat, C., Bierlaire, M., Bolduc, D., Boersch-Supan, A., Brownstone, D., Bunch, D.S., Daly, A., De Palma, A., Gopinath, D., Karlstrom, A., Munizaga, M.A., 2002. Hybrid Choice Models: Progress and Challenges. Marketing Letters 13, 163–175. https://doi.org/10.1023/A:1020254301302

Bhat, C.R., Dubey, S.K., Nagel, K., 2015. Introducing non-normality of latent psychological constructs in choice modeling with an application to bicyclist route choice. *Transportation Research Part B: Methodological* 78, 341–363. https://doi.org/10.1016/j.trb.2015.04.005

Boyce, C., Czajkowski, M., Hanley, N., 2019. Personality and economic choices. *Journal of Environmental Economics and Management* 94, 82–100. https://doi.org/10.1016/j.jeem.2018.12.004

Chorus, C.G., Kroesen, M., 2014. On the (im-)possibility of deriving transport policy implications from hybrid choice models. *Transport Policy* 36, 217–222. https://doi.org/10.1016/j.tranpol.2014.09.001

Daly, A., Hess, S., Patruni, B., Potoglou, D., Rohr, C., 2012. Using ordered attitudinal indicators in a latent variable choice model: a study of the impact of security on rail travel behaviour. *Transportation* 39, 267–297. https://doi.org/10.1007/s11116-011-9351-z

Guevara, C.A., Ben-Akiva, M., 2010. Addressing Endogeneity in Discrete Choice Models: Assessing Control-Function and Latent-Variable Methods, in: Hess, S., Daly, A. (Eds.), Choice Modelling: The State-of-the-Art and The State-of-Practice. Emerald Group Publishing Limited, pp. 353–370. https://doi.org/10.1108/9781849507738-016

Henningsen, A., Toomet, O., 2011. maxLik: A package for maximum likelihood estimation in R. *Computational Statistics, Computational* S 26, 443–458. https://doi.org/10.1007/s00180-010-0217-1

Hess, S., 2012. Rethinking heterogeneity: the role of attitudes, decision rules and information processing strategies. *Transportation Letters* 4, 105–113. https://doi.org/10.3328/TL.2012.04.02.105-113

Hess, S., Stathopoulos, A., 2013. Linking response quality to survey engagement: A combined random scale and latent variable approach. *Journal of Choice Modelling* 7, 1–12. https://doi.org/10.1016/j.jocm.2013.03.005

Hoyos, D., Mariel, P., Hess, S., 2015. Incorporating environmental attitudes in discrete choice models: An exploration of the utility of the awareness of consequences scale. *Science of The Total Environment* 505, 1100–1111. https://doi.org/10.1016/j.scitotenv.2014.10.066

Kassahun, H.T., Nicholson, C.F., Jacobsen, J.B., Steenhuis, T.S., 2016. Accounting for user expectations in the valuation of reliable irrigation water access in the Ethiopian highlands. *Agricultural Water Management* 168, 45–55. https://doi.org/10.1016/j.agwat.2016.01.017

Kroesen, M., Chorus, C., 2018. The role of general and specific attitudes in predicting travel behavior – A fatal dilemma? *Travel Behaviour and Society* 10, 33–41. https://doi.org/10.1016/j.tbs.2017.09.004

Kroesen, M., Handy, S., Chorus, C., 2017. Do attitudes cause behavior or vice versa? An alternative conceptualization of the attitude-behavior relationship in travel behavior modeling. *Transportation Research Part A: Policy and Practice* 101, 190–202. https://doi.org/10.1016/j.tra.2017.05.013

Lundhede, T., Jacobsen, J.B., Hanley, N., Strange, N., Thorsen, B.J., 2015. Incorporating Outcome Uncertainty and Prior Outcome Beliefs in Stated Preferences. *Land Economics* 91, 296–316. https://doi.org/10.3368/le.91.2.296

Mariel, P., Artabe, A., 2020. Interpreting correlated random parameters in choice experiments. *Journal of Environmental Economics and Management* 103, 102363. https://doi.org/10.1016/j.jeem.2020.102363

Mariel, P., Hoyos, D., Artabe, A., Guevara, C.A., 2018. A multiple indicator solution approach to endogeneity in discrete-choice models for environmental valuation. *Science of The Total Environment* 633, 967–980. https://doi.org/10.1016/j.scitotenv.2018.03.254

Mariel, P., Meyerhoff, J., 2018. A More Flexible Model or Simply More Effort? On the Use of Correlated Random Parameters in Applied Choice Studies. *Ecological Economics* 154, 419–429. https://doi.org/10.1016/j.ecolecon.2018.08.020

Mariel, P., Meyerhoff, J., 2016. Hybrid discrete choice models: Gained insights versus increasing effort. *Science of The Total Environment* 568, 433–443. https://doi.org/10.1016/j.scitotenv.2016.06.019

McFadden, D., 1986. The Choice Theory Approach to Market Research. *Marketing Science* 5, 275–297. https://doi.org/10.1287/mksc.5.4.275

McFadden, D., Train, K., 2000. Mixed MNL models for discrete response. *Journal of Applied Econometrics* 15, 447–470. https://doi.org/10.1002/1099-1255(200009/10)15:5<447::AID-JAE570>3.0.CO;2-1

R Core Team, 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Taye, F.A., Vedel, S.E., Jacobsen, J.B., 2018. Accounting for environmental attitude to explain variations in willingness to pay for forest ecosystem services using the new environmental paradigm. *Journal of Environmental Economics and Policy* 7, 420–440. https://doi.org/10.1080/21606544.2018.1467346

Vij, A., Walker, J.L., 2016. How, when and why integrated choice and latent variable models are latently useful. *Transportation Research Part B: Methodological* 90, 192–217. https://doi.org/10.1016/j.trb.2016.04.021

Yáñez, M.F., Raveau, S., Ortúzar, J. de D., 2010. Inclusion of latent variables in Mixed Logit models: Modelling and forecasting. *Transportation Research Part A: Policy and Practice* 44, 744–753. https://doi.org/10.1016/j.tra.2010.07.007

Zawojska, E., Bartczak, A., Czajkowski, M., 2019. Disentangling the effects of policy and payment consequentiality and risk attitudes on stated preferences. *Journal of Environmental Economics and Management* 93, 63–84. https://doi.org/10.1016/j.jeem.2018.11.007

Stampato da Logo s.r.l. Borgoricco (PD)

Bio-based and Applied Economics Focus and Scope

The journal *Bio-based and Applied Economics* (*BAE*) provides a forum for presentation and discussion of applied research in the field of bio-based sectors and related policies, informing evidence-based decision-making and policy-making. It intends to provide a scholarly source of theoretical and applied studies while remaining widely accessible for non-researchers.

BAE seeks applied contributions on the economics of bio-based industries, such as agriculture, forestry, fishery and food, dealing with any related disciplines, such as resource and environmental economics, consumer studies, regional economics, innovation and development economics. Beside well-established fields of research related to these sectors, *BAE* aims in particular to explore cross-sectoral, recent and emerging themes characterizing the integrated management of biological resources, bio-based industries and sustainable development of rural areas. A special attention is also paid to the linkages between local and international dimensions. *BAE*'s objectives are:

- to stimulate cross-fertilization between the above mentioned research fields;
- to synthesize and integrate lessons learned from current strands of literature in economics;
- to provide a forum for well-established scholars as well as promising young researchers;
- to increase the knowledge about assessment, design and evaluation of public policies;
- to promote the debate on issues relating to the economics profession and its consultancy activities;
- to discuss future research pathways on the above issues.

***BAE* publishes high quality research and review papers, after a timely and rigorous double blind peer review process. *BAE* also publishes book reviews.**

€ 16,00