Opt-Out Forced Choice Effect in Combined Revealed and Stated Preference Discrete Choice Models: A Gender Perspective

Talpur, Musharaf and Brouwer, Roy and Koetse, Mark

Abida Taherani Sindh Development Studies Centre (ATSDSC), University of Sindh, Jamshoro, Pakistan., Institute for Environmental Studies (IVM), Vrije Universiteit (VU University), Amsterdam, The Netherlands, the Water Institute, The University of Waterloo, Canada.

10 October 2019

Opt-Out Forced Choice Effect in Combined Revealed and Stated Preference Discrete Choice Models: A Gender Perspective

Musharaf A. Talpur[1,2], Roy Brouwer[2,3] and Mark J. Koetse[2]

Abstract

In this study, we assess the convergent validity of preferences and willingness-to-pay (WTP) values for beach quality improvements from a gender perspective by isolating opt-out forced-choice effect from the SP1 DCE data (that is a forced-choice situation when a respondent was asked to select among the competing labelled alternatives if they chose an opt-out). Following this approach, we combine the RP discrete choice model and SP1 DCE datasets by splitting them into female and male sub-samples and then investigate whether estimated preferences and WTP values are susceptible to this effect from a gender perspective. Using the multinomial logit (MNL) models, we find that female visitors' preferences are compatible across RP and SP1 data if the forced-choice effect is isolated from SP1 data, whereas this is not true for the male visitors. However, WTP values appear similar for both the female and male RP and SP1 sub-samples. Also, the sources of opt-out forced choices appear more promising for females than those of male counterparts in the estimated binary logit models. Our results, therefore, suggest that preferences' similarity is a gender-specific if the opt-out forced-choice effect is isolated, but WTP similarity is not.

Keywords: Revealed preference; Discrete choice model, Discrete choice experiments; Opt-out forced-choice effect; Gender perspective

1. Introduction

Non-market valuation or more specifically environmental valuation involves both revealed preferences (RP) and stated preferences (SP) models. However, there are relatively a very limited number of studies on the valuation of coastal recreation that has combined RP and SP models (e.g. Cheng and Lupi, 2016; Cameron, 1992). As compared to applying either revealed or stated preference discrete choice models for valuing coastal recreation, combining the RP and SP choice models have several advantages, which include an estimation of robust parameters, a decrease in the collinearity resulted from SP designs, an increase in information availability, the possibility of identifying potential 'environmental goods and services', and the creation of 'environmentally sustainable projects' (Birol, et al., 2006; Adamowicz et al., 1997).

[1]Corresponding author: Musharaf A. Talpur is an assistant professor at Abida Taherani Sindh Development Studies Centre (ATSDSC), University of Sindh, Jamshoro, Pakistan. Email: musharaf.talpur@usindh.edu.pk
[2] Institute for Environmental Studies (IVM), Vrije Universiteit (VU University), Amsterdam, The Netherlands
[3] Prof. Roy Brouwer is an Executive Director at the Water Institute, The University of Waterloo, Canada.

It has been witnessed that RP and SP discrete choice models are successfully combined to assess preference equality using the same random utility modelling (RUM) framework (e.g. Birol, et al., 2006; Adamowicz, et al., 1994; 1997). In this context, the RP discrete choice models have been pooled with the SP discrete choice experiments (DCE), which include choice situations with multiple alternatives including an 'opt-out', which depicts either 'status-quo' or 'neither' constant alternative, phrased as neither A or B home garden (e.g. Birol, et al., 2006) or 'none of these' option framed as 'stay at home or do other recreational activity (e.g. Talpur et al., 2018; Adamowicz et al., 1994). When combining RP and SP DCE data collected from the same respondents and analysed using the same RUM framework, one of the main issues (i.e. the inclusion of an opt-out alternative in DCE that makes the number of alternatives in both the RP and SP choice sets unequal) has always been overlooked (e.g. see Birol, et al., 2006; Adamowicz et al., 1994)[1]. Although non-inclusion of an 'opt-out' would generate biased welfare estimates because of principally two reasons: first, it is not being consistent with the demand theory if an 'opt-out is not included (Hanley, Mourato and Wright, 2001), and second, in real life there is always an 'opt-out' situation (Veldwijk et al., 2014), it is nevertheless essential to have an equal number of alternatives, i.e. possibly excluding an 'opt-out' or creating a forced choice, so that on one hand the respondents have the similar RP and SP utility circumstances, and on the other hand, this can furthermore minimize the choice task complexity and/or cognitive burden for the respondents (Veldwijk et al., 2014). Our research adopts this novel approach and combines the RP discrete choice model and SP1 DCE data while excluding an opt-out to have the equal number of alternatives and to possibly avoid any opt-out bias with this type of data enrichment.

In the prior literature, substantial research work has emphasized the impact of alternative 'opt-out' and forced-choice effects on preferences and welfare estimates (Pederson et al., 2011; Kontoleon and Yabe, 2003; Banzhaf, Johnson, and Mathews, 2001), however, limited research exists on the impact of the inclusion or exclusion of an 'opt-out' (Pederson et al., 2011). According to Boxall et al (2009), the number of respondents choosing 'opt-out' or 'status-quo' increases with more complex choices and consequently, this affects welfare measures. Dhar and Simonson (2003) observed that 'attraction effect', i.e. the likelihood of choosing an opt-out as an easy exit from the complex choice situation, becomes stronger if an 'opt-out' alternative is available, whereas simultaneously 'compromise effect', i.e. the likelihood of choosing between competing alternatives other than an opt-out, becomes weaker. This means the compromise effect tends to

[1] For instance, under the RUM modelling framework, say there are two alternatives in the RP choice set, however, there is a common practice to include the third alternative as an 'opt-out' or 'status-quo' in a series of choice situations designed to create an unforced SP DCE because of several advantages as mentioned above. If the RP and SP data is not combined, then adding an 'opt-out' has those advantages. Hence, combining RP and SP data under this situation with unequal number of alternatives in both data sets, and knowing that the respondents also have the least or most preferred utility for all alternatives, including an 'opt-out, alternative, the objective to test preference equality across both data sets becomes a questionable.

become stronger in a forced-choice situation; however, this results in the systematic violation of the independence of irrelevant alternative (IIV) property (Dhar and Simonson, 2003). In contrast, following the similar experiments but using more than two attributes as compared to Dhar and Simonson (2003), Brazell et al. (2006) found that IIV property is not violated and concluded that when alternatives have more attributes there is no clear compromise alternative. This finding raises the question: Does a compromise or forced-choice effect still exist if there are more alternatives with more than two attributes and simultaneously an 'opt-out' is excluded as we adopted this novel approach our case study?

In this study, we combine revealed preference (RP) discrete choice model and stated preference (SP1) DCE based on the non-market valuation of multiple site attributes of various beaches located along the coastal area of Sindh province in Karachi city, Pakistan. Using sub-samples of females and males in a forced-choice situation of SP1 DCE (i.e. when a respondent was asked to select among the competing labelled alternatives if they chose an opt-out), and simultaneously including and excluding respondents who selected these 'opt-out' forced choices, this study investigates 'convergent validity' from a gender perspective (i.e. isolating or opt-out forced-choice effect on public preferences and their welfare measures to assess whether convergent validity is a gender-specific or not). Putting differently, we included and excluded the female and male respondents from SP1 DCE sub-sample, who selected forced choices embodying their compromise effect, hereby called a forced-choice effect, because of having no 'opt-out' alternative, and combined female and male RP and SP1 sub-samples, respectively, and then separately applied multinomial logit (MNL) models assess the convergent validity of the estimated preferences and WTP values from a gender perspective. This means we included and excluded these forced choices of the uncertain respondents who made less consistent, monotonic and random choices and chose frequently 'opt-out' choice (Brouwer et al., 2017).

In the subsequent sections, we review the literature and discuss how our study is different from previous studies by formulating and presenting hypotheses in section 3. The rest of the paper is structured as follows. Section 4 reports the case study area, whereas it is followed by Section 5 that elaborates a combined RP discrete choice and the SP1 DCE survey design approach. Section 6 represents a choice modeling framework followed by section 7 that demonstrates the model estimation results. Hypotheses testing results are briefly elaborated in Section 8. The paper is concluded with a brief discussion in Section 9.

2. Earlier literature

Earlier literature comprises a considerable number of studies on using various 'opt-out' or 'status-quo'[2] formats in DCE and analysing their impact on preferences and/or welfare estimates (e.g. Campbell and

[2] In different DCE studies, various formats, including 'opt-out', status-quo', 'none of these', 'no purchase' or 'no choice' have been interchangeably used. In this paper, we also used these terms interchangeably where it was necessary.

Erdem, 2018; Boxall et al., 2009; Pederson et al., 2011; Kontoleon and Yabe, 2003; Banzhaf, Johnson, and Mathews, 2001; Dhar, 1997), however, differentiating specifically opt-out forced-choice effects from a gender perspective and analyzing their impact on preferences and welfare measures using female and male sub-samples derived from RP and SP discrete choice models make our study unique and, to our knowledge, has never been conducted in either the DCE or the combined RP-SP literature. Apart from knowing various advantages of including an 'opt-out', it is nevertheless a fact that inclusion of an 'opt-out' is not a realistic approach in some choice situations, for instance, preferences for alternative transport modes (e.g. Rose and Bliemer, 2009; Hensher and Rose, 2007), not adding an 'opt-out' alternative to other hypothetical alternatives is not consistent with the theory of demand (Hanley, Mourato and Wright, 2001). Dhar (1997) concluded that the preferences for 'no-choice' or an 'opt-out' alternative can be decreased if an inferior (or non-dominant) alternative is added to a choice situation, whereas preferences for an 'opt-out' can be increased if one more attractive (or dominant) alternative is added, that eventually put the consumer in a complex choice situation. According to Dhar and Simonson (2003), this complex choice situation makes the respondent indifferent to choose between two dominant alternatives and the availability of an 'opt-out' alternative attracts him/her to decisively choose that alternative.

Banzhaf, Johnson, and Mathews (2001) in their DCE study conducted on anglers' preferences for fishing sites used a split-sample approach with one sample included an 'opt-out' alternative framed as 'not to go fishing', whereas the other included 'status-quo' alternative framed as 'anglers' routine site'. In their study, they concluded that some respondents because of being not usual or new anglers found even a 'status-quo' situation as a forced-choice situation and recommended including both 'status-quo' and 'none of these' situations to avoid a forced-choice effect on preferences in such circumstances. Likewise, Kontoleon and Yabe (2003) assessed consumer preferences for genetically modified contents in food using two 'opt-out' alternatives two identical DCE using a split-sample approach with one sample of the respondents received 'no purchase' alternative and the other received 'purchase my own brand'. Using the mixed logit (MIXL) models, the authors demonstrated that with a comparison to 'no purchase' choice situation, preferences heterogeneity, relative choice share and response consistency was higher with a lower respondents' fatigue in 'purchase my own brand' opt-out situation and concluded that all model results varied because 'no purchase' alternative was found to be a forced choice situation by most of the respondents. Authors thus recommended that a similar approach would have been adopted as implemented beforehand by Banzhaf, Johnson, and Mathews (2001).

In the recent past, Pedersen and Gyrd-Hansen (2013) profoundly criticised the use of 'status-quo' and 'opt-out' alternatives in the SP DCE with their applications in health economics followed by some recommendations. In this respect, they criticise a study conducted by Kiiskenen et al (2010) on patients' preferences for the private versus public dental services in Finland and recommended otherwise the use of ‘status-quo’ instead of using ‘no choice’ alternative. They also criticise another similar DCE study conducted by Burge et al. (2004) on patients’ choices for quicker treatment given under the London Patient Choice Project (LPCP) and once again recommended adding ‘status-quo’ alternative instead of incorporating ‘neither’ as an opt-out alternative. In both studies, according to Pedersen and Gyrd-Hansen (2013), giving ‘no choice and ‘neither’ alternatives respectively means that patients are forced to choose ‘not to undergo an operation’, which has ‘zero utility’ and may bias the preference estimates.

To further the discussion of an ‘opt-out’ dilemma, as thoroughly reviewed above, there exist very few studies on the other hand that include both ‘opt-out’ and ‘status-quo’ alternatives. Lancsar et al. (2007) investigated preferences of asthma patients using DCE entailing hypothetical medication alternative, and current medication as a ‘status-quo’ and no asthma medication as a ‘no choice’ alternatives, respectively. Another DCE study in sports economics conducted by Pedersen, Kiil and Kjær (2011) assessed preferences of soccer attendees in Fiona Park Stadium in Denmark for soccer club managers. The authors in their study used two hypothetical stadiums in addition to the current stadium as a ‘status-quo’ and alternative activity as ‘opt-out’ options, respectively, and demonstrated that using both ‘status-quo’ and ‘opt-out’ alternatives together capture preferences of both the current users and the future potential users of soccer. To this end, we may, therefore, conclude that inclusion and exclusion of an opt-out, status-quo and/or their types, such as ‘none of these’, ‘neither’, ‘no choice’ and so forth, largely depend on varying choice situations which suit the applicability of these different types of opt-outs.

All the previous studies in the DCE literature used different types of ‘opt-out’ formats employed generic (or unlabelled) DCE designs, however, not even a single DCE study exists which entailed labelled DCE design. By saying so, we mean that DCE if includes all hypothetical possible alternatives which exist in the reality, adding a ‘status-quo’ and ‘opt-out’ alternative, that no matters carry a zero-utility, does not portray a realistic scenario. In an SP DCE study, Talpur, et al. (2018) included only ‘none of these’ alternative while simultaneously entailed a full set of eight beach alternatives. In their situation, including a ‘status-quo’ was an impractical approach; specifically when their study also included a full set of eight alternatives to capture RP data, which elicited beach visitors’ actual choices representing their perceived individual-specific ‘status-quo’ choices. Although Adamowicz et al. (1994; 1997) in environmental economics and Birol et al. (2006) used the same DCE designs, however, these authors used two unlabeled SP hypothetical scenarios in addition to ‘none of these ’ opt-out situations. In the above studies, it was possible to include instead both ‘status-quo’ and ‘none of these’ alternatives, respectively, as both studies used unlabelled DCE designs.

Apart from studies addressing the ‘opt-out’ forced-choice effect, there are a few DCE studies that addressed gender issues from different perspectives. Using a split-sample approach, Ladenburg and Olsen (2008) studied the influence of price starting point bias (SPB) on both female and male preferences as well as welfare estimates across both samples and found that price SPB affected only female preferences and their WTP values. Differently, another study conducted by Keane et al (2016) estimated and compared female and male cross preferences for community-based conservation (CBC) initiative in Kenya. The researchers found that females' preferences differed from those revealed by males in that females placed a higher value on conservancy membership attribute, but less value on cultivation, access to conservancy land for grazing, and wage-income for herding cattle, whereas the opposite was found in case of males.

In comparison to all studies discussed above, our study adopts a disaggregated novel approach that assesses the influence of including and excluding 'out-out' forced-choice effect on preferences and WTP values from a gender perspective and finds out whether the 'convergent validity' is a gender-specific or not. Putting differently, we included and excluded the female and male respondents and compared their preferences and WTP values influenced by their random choice behaviour increasingly resulting from the choice of an 'opt-out' alternative (Brouwer et al., 2017). Besides, we analysed sources of forced-choice opt-out effect including various respondent and design characteristics that influence the selection of an opt-out forced-choice using separate binary logit models for female and male sub-samples, respectively.

3. Hypotheses formulation

To assess the convergent validity by isolating (or disentangling) opt-out forced-choice effects from a gender perspective, we test a series of the following hypotheses;

3.1 *Preference equality*

This first series of hypotheses $H^{\mathbf{1}}$ tests the convergent validity in terms of the equivalence of the estimated preference β and their scales μ parameters across RP and SP1 DCE sub-samples of females and males by disentangling opt-out forced-choice affects. These tests involve a two-step test procedure assessing the equality of preference and scale parameters across RP and SP1 models (Swait and Louviere, 1993). The first step starts testing whether the preference parameters for the two samples are equal and can be combined while allowing the scale parameters to vary across both the samples:

$$H^{STEP-1}\text{: } \beta_1^{RP} = \beta_i^{SP1} = \beta_i^{RP+SP1\ Combined} \tag{3}$$

A grid-search procedure for the best-combined model fit is performed over a range of scale parameter values of one sample while keeping the value of the scale parameter of another sample constant (i.e. equal to 1). To test H^{STEP-1} hypothesis, a chi-square test using the log-likelihoods for the separate and combined best fit model is implemented and then, a standard Log-Likelihood Ratio (LR) test is performed to assess whether the preference parameters across the two samples are equal:

$$\lambda^1: -2[\,LL^{RP} + LL^{SP1} - (LL^{RP+SP1\,Combined})] \text{ with d.f.} = |\beta_i| - 1$$

where $|\beta_i|$ refers to the restriction regarding the number of imposed parameters.

If the LR test confirms that preference parameters are equal between the two samples, then we implement the second step involving equality imposed on both preference and scale parameters.

$$H^{STEP-2}: \mu_i^{RP} = \mu_i^{SP1} = \mu_i^{RP+SP1\,Combined} \quad (4)$$

To test H^{STEP-2} hypothesis, a chi-square test using the log-likelihoods for both the combined models (i.e. the one with the equal scale and the other with the varying scale) is implemented and then, as usual, a standard Log-Likelihood Ratio (LR) test statistic is performed to assess whether the scale parameters across the two samples are equal:

$$\lambda^2: -2[LL_{RP-SP\,Combined}^{\mu^{RP}=\mu^{SP1}} - LL_{RP-SP\,Combined}^{\mu^{RP}\neq\mu^{SP1}}] \text{ with d.f.} = 1$$

If both H^{STEP-1} and H^{STEP-2} hypotheses are accepted, then we conclude that preference parameters across RP and SP1 models are equal. Using the above grid-search procedure, we test the following series of hypotheses from a gender perspective:

$\boldsymbol{H^1}$: $\beta_{i\,females}^{RP} = \beta_{i\,females}^{SP1}$ and $\beta_{i\,males}^{RP} = \beta_{i\,males}^{SP1}$ *if isolating 'opt-out' forced-choice effects from a gender perspective*

H^{1A}: $\beta_{i\,females}^{RP} = \beta_{i\,females}^{SP1}$ *if we do not isolate 'opt-out' forced-choice effects* (Table 3)

H^{1B}: $\beta_{i\,females}^{RP} = \beta_{i\,females}^{SP1}$ *if we isolate 'opt-out' forced-choice effects* (Table 4)

H^{1C}: $\beta_{i\,males}^{RP} = \beta_{i\,males}^{SP1}$ *if we do not isolate 'opt-out' forced-choice effects* (Table 5)

H^{1D}: $\beta_{i\,males}^{RP} = \beta_{i\,males}^{SP1}$ *if we isolate opt-out' forced-choice effects* (Table 6)

3.2 *WTP equality*

This second series of hypotheses H^2 tests the convergent validity in terms of the equivalence of WTP values across RP and SP1 DCE sub-samples of females and males by isolating (or disentangling) opt-out forced-choice effects. These tests require the combinatorial test process assessing differences in mean WTP values (Poe, Giraud and Loomis, 2005).

$\boldsymbol{H^2}$: $WTP_{i\,females}^{RP} = WTP_{i\,females}^{SP1}$, and $WTP_{i\,males}^{RP} = WTP_{i\,males}^{SP1}$ *if isolating 'opt-out' forced-choice effects from a gender perspective*

H^{2A}: $WTP_{i\,females}^{RP} = WTP_{i\,females}^{SP1}$ *if we do not isolate 'opt-out' forced-choice effects* (Table 3)

H^{2B}: $WTP_{i\,females}^{RP} = WTP_{i\,females}^{SP1}$ *if we isolate 'opt-out' forced-choice effects* (Table 4)

H^{2C}: $WTP_{i\,males}^{RP} = WTP_{i\,males}^{SP1}$ *if we do not isolate 'opt-out' forced-choice effects* (Table 5)

H^{2D}: $WTP_{i\,males}^{RP} = WTP_{i\,males}^{SP1}$ *if we isolate 'opt-out' forced-choice effects* (Table 6)

3.3 *Sources of opt-out forced-choice effect*

This third series of hypotheses H^2 tests the sources of opt-out forced-choice effect such as respondents' and design characteristics influencing opt-out forced choices selected by the females and males. To assess these hypotheses, we first applied binary logit models separately estimated from female and male SP1 sub-samples and then implemented both the Wald and Log-likelihood Ratio (LR) tests.

$\boldsymbol{H^3}$: $\beta_i^{SP1}\ design\ characteristics = 0$ (Table 8)

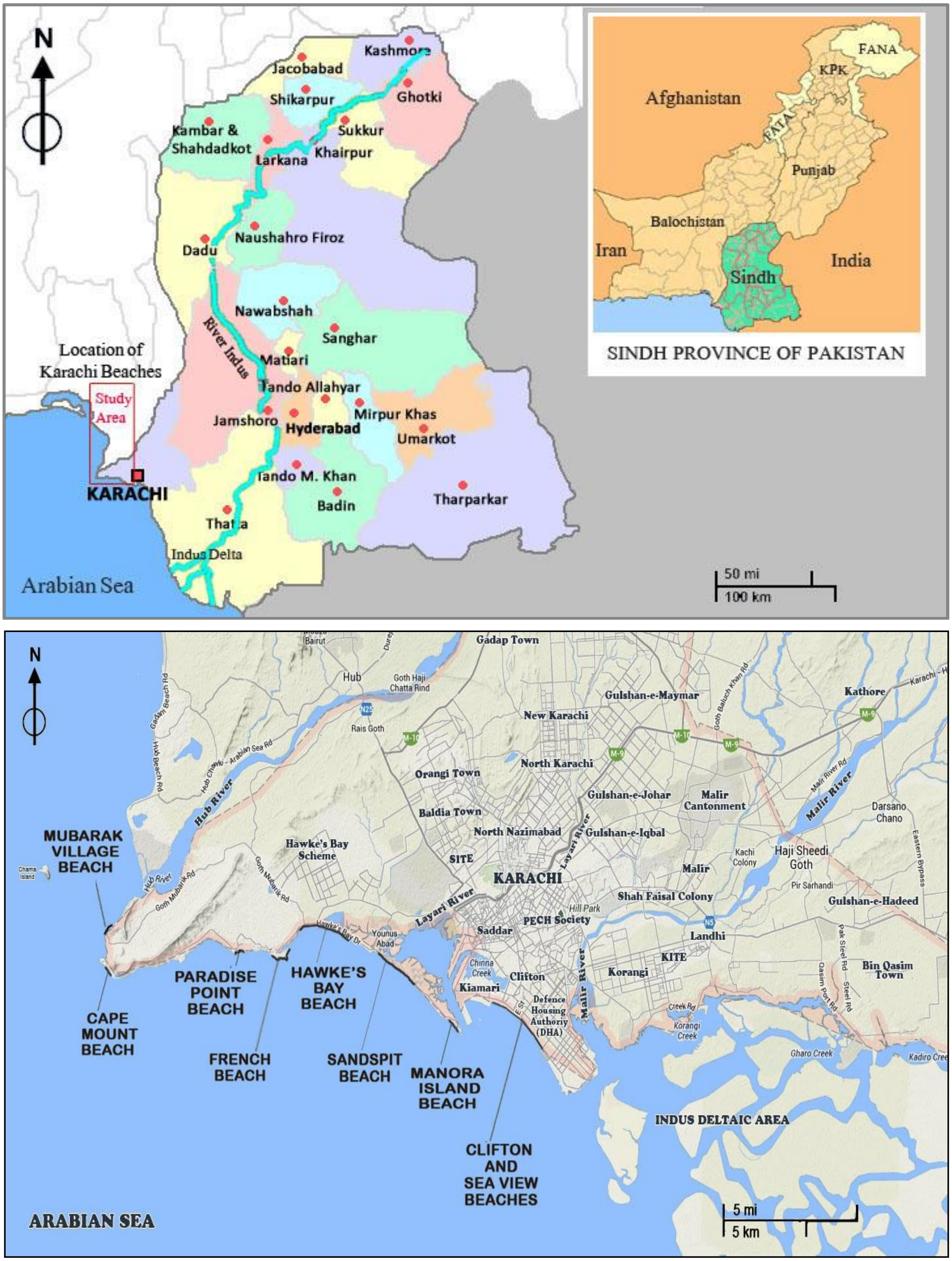
H^{3A}: $\beta_{females}^{SP1}\ design\ characteristics = 0$

H^{3B}: $\beta_{males}^{SP1}\ design\ characteristics = 0$

4. Study description

With a population of over 16 million and its economy's capacity to account for 20 per cent of country's GDP, Karachi city, which is the capital city of Sindh province in Pakistan, is the major port, the most urbanised and industrialised city (KSDP, 2007), City's 70 km long coastline, which starts from the Mubarak Village beach in the west and ends at Korangi creek neighbouring the Indus Delta in the east (see Figure 1), provides a wealth of recreational and other opportunities to the residents of the city, including fisheries, mangrove forest products, seaside agriculture, coastal wildlife, marine resources, and shoreline stabilisation services (Khalil, 1999). However, the monetary importance of many of these environmental services, such as beach recreation, coastal water quality and its depending aquatic life, is not fairly recognised at policy and decision-making levels, due mainly to public good nature of these resources. As a result, coastal resource managers and policy-makers lack the information required at the stages of designing and implementing coastal zone development on a sustainable basis, particularly when the future welfare losses are associated with the degradation of recreational use values of beaches due to rapid population and industrial growth since last few decades.

Figure 1: Names, locations and directions main beaches in Karachi city, Sindh province of Pakistan



The residents of Karachi enjoy a variety of water and beach-related recreational opportunities (e.g. swimming, diving, and walking) provided by the city's several beaches. Stretching from east to west, we included eight of these beaches in our study (see Figure 1). These beaches include; Clifton/Sea View, Manora Island, Sandspit, Hawke's Bay, French, Paradise Point, Cape Mount, and Mubarak Village beach. We excluded three beaches: Russian beach, located within Korangi creek near the Indus Delta, is a distant beach, so it is only infrequently visited by some visitors, Nathiagali beach is managed by the Pakistan Naval forces, so only these militaries have access to it, and Sunehra beach is used for angling only.

Over many decades, socio-economic activities in the city, such as housing and construction, fisheries, ports and shipping, power plants and industry, recreation and tourism, have increased massively with its resulting burden on coastal resources, particularly on the quality of shoreline, including beaches and its recreational waters. In the metropolitan area of Karachi, there are around 6,000 industries that release effluents directly or indirectly through the *Layari* and *Malir* rivers into the Arabian Sea. Likewise, some seaside power plants and industrial units are the principal sources of thermal pollution, jointly discharging around 1,500 million cubic meters of warm water yearly into the coastal waters, which consequently raise the temperature levels in the coastal waters that endanger aquatic life. Approximately 330 million gallons of urban sewage water flows to the Arabian Sea daily through these rivers, drainage and other waterways. According to KSDP (2007), only around 30 per cent of the metropolitan sewage water is treated and the remainder is disposed into the coastal waters as raw sewage and untreated manufacturing rubbish. Besides, coastal agricultural contaminants, such as fertilizers, pesticides and herbicides, as well as scrape from construction sites and waste from poultry farms, make their way into the coastal waters.

Apart from declining coastal water quality, beach littering is another critical aspect adversely affecting resident's use benefits of their beach recreation in the city. According to SACEP (2007) report, the Karachi Municipal Corporation (KMC) regularly gathers and discards on average 25 per cent of the 8,000 –10,000 tons of litter generated day after day, whereas the remaining litter is either left in the city or dumped near the river banks and waterways, where it is eventually blown away by the wind or washed down to the coastal waters. Regularly, only two urban beaches, Clifton/Sea View and Manora Island get cleaned by the authorities in Karachi, whereas the remaining beaches are cleaned only rarely, either by village organisations or local non-governmental organisations. Overcrowding is another crucial factor that causes beach littering as visitors usually leave behind their trash. Overcrowding is very difficult to control because the public has free access to the beaches.

Until recently, there is no specific legislation to control coastal water pollution, beach littering and seaside crowding in Pakistan, even though the country is a signatory to various international conventions, protocols and treaties on coastal environment-related issues, including MARPOL 73/78, the Convention on Marine Biological Diversity, and several others, and the Environmental Protection Act from 1997, that vaguely aims to protect the coastal and marine environment altogether (Talpur and Jariko, 2001). To deal with environmental problems, such as coastal water quality, beach litter and overcrowding, only a Marine Pollution Control Board exists (SACEP, 2007). Thus, on the whole, there is no clear policy planning, coordination and implementation among the concerned authorities.

5. Combined RP-SP survey design and implementation

There are some advantages and disadvantages associated with both RP and SP designs. SP designs are usually criticised because they do not capture the observed behaviour (Birol et al., 2006; Carson and Mitchell, 1993), hence, they fail to obtain information on real market situations (Louviere et al., 2000). However, these designs help to estimate the value of non-marketed environmental goods and services without substitute market values. The SP methods, specifically DCE, depict a broad range of attribute levels associated with the varying quantity and quality improvements pertaining to the future benefits of public goods. In contrast, RP designs have an advantage of obtaining information on real or actual choice behaviour, however, attributes and their levels of non-marketed goods and services in these designs do not vary over time if we use a single cross-section survey, hence, the data on the varying quantity and quality improvements in the proposed future projects due to changing policies cannot be acquired (Louviere et al., 2000). Besides, RP methods sometimes generate coefficients with unexpected signs or erroneous magnitudes due to collinearity (Hensher and Rose, 2007; Louviere et al., 2006).

In the recent times, the number of research studies on combining both RP and SP methods, which is also known as 'data enrichment paradigm', has considerably grown intending to flourish the strengths and lessen the weaknesses of each method (Birol et al., 2006; Adamowicz et al., 1994, 1997; Swait and Louviere, 1993). This data fusion or 'data enrichment', that combines both RP and SP discrete choice data using the same RUM framework, reduces the problem of collinearity because of combining RP method with the experimentally designed SP method, generates more robust and efficient estimates, improves the significance levels of the parameters, and produces more information (Louviere et al., 2000; Birol et al., 2006; Adamowicz et al., 1994, 1997). In this study, we followed the similar approach previously implemented by Swait and Louviere (1993) and Adamowicz et al. (1994, 1997), however, our approach is unique in terms of disentangling opt-out forced choice and estimating the determinants of this forced choice effect from a gender perspective using the MNL models.

Our combined RP-SP survey design has four sections. The first section collected revealed preferences (RP) data from the respondents relating to their past and current site choice and travel behaviour, such as distance from their residence to the different beaches they travelled during the current year, transportation costs, travelling time, on-site costs and time, trip type and purpose, and beach activities they undertake when visiting these beaches. In the second section, we recorded visitors' opinions regarding beach attributes, such as water quality, beach cleanliness, crowding (and noise), facilities and distance from their residence using a scale between 1 and 10, where 1 refers to 'not important at all' and 10 'most important'. In the third section, we recorded perceptions of the respondents and asked them to rate the conditions of all the site quality attributes, including water quality, beach cleanliness (or littering), crowding and noise, and facilities across all 8 beaches. Afterwards, respondents were asked to reveal their actual preferred beach out of a choice set of eight alternative beaches that they have visited during the current year (see Figure 2). Thus, we collected actual or RP discrete choice data that respondents revealed in terms of their perceived baseline quality levels for each attribute at each beach. For respondent's convenience, we used a city map illustrating the exact locations of the eight beaches in the Karachi metropolitan area to help respondents to assess the perceived distance as a proxy of travel cost from their residence to each of the eight beaches. Subsequently, we collected SP data concerning respondent's stated choices using discrete choice experiments (DCE). Finally, the fourth section collected information regarding the socio-economic characteristics of the respondents.

We collected SP data using two separate versions of the DCE: SP1 and SP2. In SP1 version visitors are faced with a future increase in travel cost as an implicit payment vehicle, i.e. a future increase in fuel expenditures and the opportunity cost of travel time. In SP2 version visitors face the same increase in future travel costs, but this version also adds an entrance fee as an explicit payment vehicle to the implicit travel cost. For this study purpose, however, we used only the SP1 version. Like the RP version, the SP1 version includes the eight beaches located along the Karachi coast as labelled alternatives (see Figure 3). The labelled DCEs are designed to capture site selection behaviour of visitors, who were asked to choose their preferred beach out of a choice set of eight alternative beaches and an opt-out alternative (see Annex 1 for the full details of the design). Since RP choice set does not have an opt-out alternative, we designed a forced SP1 DCE that helped respondents to choose their preferred beach out of a choice set of eight alternative beaches, so that zero-utility effect could be avoided and the respondents can face an equal number of alternatives in both RP and SP1 datasets. (e.g. see Figures 2 and 3). In both RP and SP1 data sets, choice situations are identical in their non-monetary site attributes, i.e., coastal water quality, beach cleanliness, crowding and site facilities, and one monetary attribute, i.e. travel cost (see Table 1).

Table 1: Discrete choice data sets: RP and SP1 versions, attributes and levels

Attributes / variables	Levels	RP	SP1
Water quality	Poor, Moderate, Good		
Beach cleanliness	Very littered, Moderately littered, No litter / Clean		
Crowding	Sparsely crowded, Moderately crowded, Very crowded		
Facilities	Low, Medium, High		
Travel cost	Travel cost (i.e. trip expenditures and the opportunity cost of time)		

For each site attribute, three levels were selected, which reflect and depict low, moderate and high site quality improvements. For the reason that coastal water regulation in Pakistan does not exist, we comprehensively outlined poor, moderate and good levels for coastal water quality along the beaches in Karachi, similar to the classification of water bodies in the European Water Framework Directive (WFD), and illustrated using the US-EPA water quality ladder (Carson and Mitchell, 1993). These water quality levels describe whether water is suitable for swimming, catching fish that are safe to eat and supporting plants, fish and other aquatic life. Throughout the focus groups and pre-tests, we observed that characterising water quality levels objectively or physically (e.g. turbidity, dissolved oxygen, etc.) was not easily recognised by the respondents. Following Schaafsma and Brouwer (2013), we, therefore, developed a colour-coded water quality ladder, representing water clarity, colour, contamination, odour, chemical status as well as its suitability for recreational uses such as swimming, bathing, and wading in the water, and illustrated these levels using easily understandable pictograms (full descriptions of the attribute levels as shown to the survey respondents are exhibited in Annex 1). To avoid statistical insignificance of site attributes due to lack of public understanding, we applied the colour-coding and pictograms, and at the same time facilitate comparison of elicited WTP values in this study with WTP values found elsewhere in the literature using similar water quality ladders (e.g. Bateman et al., 2011).

Likewise, we followed an approach first implemented by Smith et al. (1997) to estimate the use benefits of decreased beach littering, so we described three levels for the attribute beach cleanliness (very littered, moderately littered and a clean beach with no litter). Since there is no legislation addressing coastal litter control in Pakistan, we adopted the definition mentioned in the EU Marine Strategy Framework Directive (MSFD) for reductions in coastal or marine litter, specifically in terms of its visibility (e.g. 25 units of litter per 50 meters of beach length). We adopted this definition because it was observed during the focus groups and pre-tests that the general public found it easier to assess the impacts of beach cleanliness regarding human health hazards and beach suitability for recreational uses such as walking, playing beach sports, playing in the sand and other recreational activities.

For the recreational valuation of site quality improvement, crowding is usually selected as an essential attribute because an increase in the number of visitors may negatively affect the recreational experience of visitors regarding reduced movement, privacy, and noise (Talpur et al., 2018; Hanley et al., 2002; Brouwer, 1999). In contrast, crowding may affect users' experiences also positively (Taylor and Longo, 2010). Crowding is mostly incorporated in the recreational site choice models as a dummy variable (e.g. Hanley et al., 2002). On the contrary, our study incorporates it as a categorical variable for the sake of testing likely non-linear effects (e.g. Taylor and Longo, 2010), and also combined an element of noise to it (Brouwer, 1999). Along these lines, three levels were selected (sparsely crowded, moderately crowded and very crowded) and pictograms were also shown to the visitors imagine crowding, described as beach visitors' density. Subsequently, site facilities were chosen as the fourth non-monetary attribute based on feedback from focus groups and pre-tests. Once again, three levels were selected for this attribute in a cumulative manner (parking, parking and food stores, and parking, food stores, and washrooms).

Finally, we need a monetary attribute that computes marginal WTP estimates, which demonstrates the rates at which beach visitors are willing to trade off-site attribute levels with income (Louviere et al., 2000). Both the RP version and SP1 DCE version includes only travel costs variable as a common implicit payment vehicle (SP1). For outdoor recreation, visitors mostly travel from their homes to the chosen beach by implicitly bearing travel costs. So during surveys, respondents were asked about their perceived distances and travel times for visiting the 8 selected beaches along the Karachi coast, which enabled us to calculate these travel costs. The travel cost calculation approach, where distance acts as a proxy for travel costs for each respondent (Hanley et al., 2002), is presented in Annex 2 of this paper. An example of an SP1 choice card is demonstrated in Figure 3. Each respondent faced six choice situations and asked to choose his or her preferred beach to visit in the next 12 months under different site quality improvements from the set of eight alternative beaches, or to choose an 'opt-out' alternative, framed as 'None of these, I prefer to stay at home or do other non-beach activities' (e.g. Talpur, et al., 2018). Those respondents who chose 'none of these' alternatives, we further asked them about their forced-choice by asking them about the choice situation in which this 'opt-out' alternative was not available, of course bearing in mind their trip expenditures and travel times to each of the beaches.

Figure 2. Choice card example (RP version)

Water quality levels

Good 

Moderate 

Poor 

What is your perception about the overall water quality at the following beaches during the last 12 months?
*(Put **0 = Poor water quality**, **1 = Moderate water quality**, **2 = Good water quality**, if you do not know the answer or even if you have never been there, then provide with a better guess).*

Clifton & Sea View beaches	**Manora Island beach**	**Sandspit Beach**	**Hawke's Bay beach**	**French beach**	**Paradise Point beach**	**Cape Mount beach**	**Mubarak Village beach**

Cleanliness Levels

No Litter / Clean 

Moderately Littered 

Very Littered 

What is your perception about the overall cleanliness at the following beaches during the last 12 months?
*(Put **0 = Very littered**, **1 = Moderately littered**, **2 = No litter/Clean**, if you do not know the answer or even if you have never been there, then provide with a better guess).*

Clifton & Sea View beaches	**Manora Island beach**	**Sandspit Beach**	**Hawke's Bay beach**	**French beach**	**Paradise Point beach**	**Cape Mount beach**	**Mubarak Village beach**

Crowding levels

Sparsely Crowded 

Moderately Crowded 

Very Crowded 

What is your perception about the overall crowding at the following beaches during the last 12 months?
*(Put **0 = Sparsely Crowded & Quiet**, **1 = Moderately crowded**, **2 = Very crowded & Noisy**, if you do not know the answer or even if you have never been there, then provide with a better guess).*

Clifton & Sea View beaches	**Manora Island beach**	**Sandspit Beach**	**Hawke's Bay beach**	**French beach**	**Paradise Point beach**	**Cape Mount beach**	**Mubarak Village beach**

Facilities Levels

Low 

Medium 

High 

What is your perception about the overall facilities at the following beaches during the last 12 months?
*(Put **0 = Low**, **1 = Medium, 2 = High**, if you do not know the answer or even if you have never been there, then provide with a better guess).*

Clifton & Sea View beaches	**Manora Island beach**	**Sandspit Beach**	**Hawke's Bay beach**	**French beach**	**Paradise Point beach**	**Cape Mount beach**	**Mubarak Village beach**

Your Choice

☐	☐	☐	☐	☐	☐	☐	☐

Figure 3. Unforced and Forced choice card example (SP1) capturing the out-out forced-choice effect

Given the beach attributes described below, and bearing in mind **future increase in your personal trip expenditures and travel times** as well as **entrance fee** to each of the beaches in addition to the proposed entrance fees, which beach would you prefer to visit?

Beach / **Attributes**	**Clifton / Sea View**	**Manora Island**	**Sandspit / Turtle**	**Hawke's Bay**	**French**	**Paradise Point**	**Cape Mount**	**Mubarak Village**	**None of these**
Water quality	Poor	Good	Poor	Good	Moderate	Moderate	Good	Moderate	**I prefer to stay at home or Do other non-beach activities**
Cleanliness	No litter / Clean	Moderately littered	No litter / Clean	Moderately littered	Very littered	Moderately littered	Very littered	No litter / Clean	
Crowding	Very crowded	Moderately crowded	Sparsely crowded	Very crowded	Sparsely crowded	Sparsely crowded	Very crowded	Moderately crowded	
Facilities	Low	Medium	High	Medium	Low	Low	Low	High	
Your Choice	☐	☐	☐	☐	☐	☐	☐	☐	☐

OK, you have selected 'none of these beaches' opt-out. Imagine this opt-out alternative was not available in this choice set. Given this situation, and bearing in mind future increase in your personal trip expenditures and travel times to each of the beaches, which beach would you still prefer to visit?

Your Choice	☐	☐	☐	☐	☐	☐	☐	☐

We used only one choice card for our RP version, whereas we applied orthogonal designs for the SP1 version of the DCE to generate prior parameter values using a small sample (N = 48). For SP1 pilot datasets, we estimated separate multinomial logit (MNL) models. In the MNL model, all parameters had the expected signs (e.g., a negative sign for distance from a visitors' residence) and orders of magnitudes (e.g., higher levels being valued more than lower levels), and were statistically significant at the 1% and 5% levels, except for the moderately crowded coefficient, which was insignificant in the SP1 pilot version. Based on these prior values, two separate D-efficient designs of 36 choice tasks were generated, using the Ngene software (version 1.1.1; Choice Metrics, 2014). To reduce respondents' cognitive burden during the interview, these 36 choice tasks were blocked into six versions of six choice tasks. Hence, each respondent was randomly assigned one of the six versions.

Using an off-site random cluster sampling approach, we conducted in-person interviews with the residents of Karachi. We preferred the off-site residential survey to an on-site survey for ensuring high response rates from both (i) visitors and non-visitors to avoid endogenous stratification (Talpur et al., 2018), and (ii) men and women. Thus, only residents of the city were interviewed, not non-residents, who might visit the beaches only occasionally from nearby cities. The survey was implemented by selecting different clusters of municipal neighbourhoods in Karachi city. Within each cluster, residential homes were selected for in-person interviews through a simple random sampling approach. Eventually, we conducted surveys to obtain the required information from 294 respondents for RP and 152 respondents for SP1, respectively. In this way, we finally obtained two datasets, including RP and SP1, respectively. Table 2 exhibits descriptive statistics for the RP and SP1 versions.

Table 2: Descriptive statistics of beach visitors

Data set / Visitors' characteristics[3]	RP version	SP1 version	Pakistan population	Test applied	Test-statistic
Gender (% Female =1)	35.1	35.5	48.0	Chi-squared	0.188
Age (%)				Mann-Whitney	- 1.436
18 – 39 years	66.9	76.3	31.1	Mann-Whitney	- 2.806***
40 – 59 years	31.4	21.7	13.5	Mann-Whitney	2.332**
≥ 60 years	1.7	2.00	5.5	Mann-Whitney	0.000
Marital status (% Married =1)	53.4	55.9	63.0	Chi-squared	4.568**
Average household size (people)	6.4	6.7	6.4	Mann-Whitney	- 4.740***
Average number of children/household	1.8	1.8	2.9	Mann-Whitney	- 1.163***
Education distribution (%)				Mann-Whitney	2.261**
Primary and secondary education	26.2	22.4	74.9	Mann-Whitney	- 3.725***
Higher (college / university) education	72.8	77.0	6.4	Mann-Whitney	2.004**
Informal education	1.0	0.6	0.4		+
Employed (% Employed =1)	80.6	76.3	79.0	Chi-squared	18.748***
Average one-way distance travelled (km)	38.1	36.5		Mann-Whitney	- 2.969***
Average household income (PKR/month)	61760.20	59491.07	209.1	Mann-Whitney	4.564***
Average travel cost (PKR per person/day) without opportunity cost of time	1234.02	1112.81		Mann-Whitney	- 3.224***
Average travel cost (PKR per person/day) with the opportunity cost of time	1231.27	1348.34	----	Mann-Whitney	- 2.507**
Sample respondents (N)	294	152	----	----	----
Number of choice sets/respondent	1	6	----	----	----
Observations	294	912	----	----	----

6. Modelling framework

Following the random utility theory (RUT) and its basic modelling framework (McFadden, 1974), we applied the MNL (or conditional) logit model to estimate preferences elicited by the beach visitors using separate revealed preferences (RP) and stated preferences (SP) discrete choice models. Let's assume that a beach visitor i $(i = 1, 2, \ . \ . \ . \ , n)$ makes a single-day trip to beach alternative j out of $k = 1 \ldots K$ beach alternatives along the Karachi coast on a given day during the current and next years, his/her indirect

[3] We applied Mann-Whitney (MW) test for contineous variables (e.g. age, education, etc.) and Chi-Squared (CS) test for dummy variables (e.g. female =1, married = 1, etc.) across the RP and SP1 datasets. Aestriks **,*** indicate statisitical significant differences at 5% and 1% levels. + Education groups do not match across both versions.

revealed and stated utility *U* for both the current and next years can be expressed as a linear function of the vector of beach attributes (or characteristics) *X*:

$$U_{ijc}^{RP} = V_{ijc} + \varepsilon_{ijc} = \alpha_j + \beta_i X_{ijc} + \varepsilon_{ijc} \qquad \forall\, j \in K,\ c^{RP} = 1 \tag{1}$$

$$U_{ijc}^{SP1} = V_{ijc} + \varepsilon_{ijc} = \alpha_j + \beta_i X_{ijc} + \varepsilon_{ijc} \qquad \forall\, j \in K,\ c^{SP1} = 1, 2, .., 6$$

where V_{ijc} is the deterministic component of the linear additive utility function, and where ε_{ijc} is the error term which is assumed to follow independently and identically distributed (IID) Gumbel extreme value type 1 distribution. Given RP and SP1 utility functions, the observable deterministic component comprises a set of alternative-specific constants (ASC) α_j, the vector of site attributes X_{ijc}, and the vector of associated preference parameters or the marginal utilities β_i. In RP utility function, a visitor faces only one real or actual discrete choice set *c = 1*, whereas the same visitor faces six experimentally designed hypothetical discrete choice sets *c = 1, 2, . . , 6* in the SP1 utility function. Now, the probability that a respondent *i* chooses to visit a beach alternative *j* out of *k = 1, . ., K* alternatives given RP or SP1 choice situation can be expressed as:

$$P_{ij} = \frac{e^{(\alpha_j + \beta_i X_{ijc})}}{\sum_{k=1}^{K} e^{(\alpha_j + \beta_i X_{ijc})}} \qquad \forall\, j \in K,\ c^{RP} = 1,\ c^{SP1} = 1, 2, .., 6 \tag{2}$$

The above expression is known as the MNL (or conditional) model specification. According to Lancaster (1966), the observable deterministic component of the indirect utility function V_{ijc} can be decomposed into the vector of site attributes ($X_{1jc}, X_{2jc}, \ldots, X_{njc}$) for each alternative. So, we estimated preferences and WTP estimates using the MNL models.

Using expression (2), we estimate preference parameters in both the RP and SP2 versions. Using travel cost as a monetary parameter, we estimate marginal WTP values applying the Krinsky and Robb (1986) bootstrapping method. For modelling framework, we accommodate possible non-linear threshold effects and avoid confounding effects (Louviere et al., 2000) by applying effects coding instead of dummy coding.[4] Following the above model specification, we, therefore, test a series of hypotheses in terms of assessing the convergent validity from a gender perspective, and further estimate the determinants behind opt-out forced-

[4] Because the sum of the effects codes is equal to 0, the sum of the attribute with three levels is also equal to zero, that is $\beta_0 + \beta_1 + \beta_2 = 0$, which can also be expressed as $\beta_0 = -\beta_1 - \beta_2$. For calculating WTP for the medium and the higher site quality improvements, we applied the following formulas using the above definition of effects coding: $WTP_{medium} = (\beta_1 - \beta_0)/\beta_{payment\ vehicle} = (\beta_1 - (-\beta_1 - \beta_2))/\beta_{payment\ vehicle} = (2*\beta_1 + \beta_2))/\beta_{payment\ vehicle}$ and $WTP_{high} = (\beta_2 - \beta_0)/\beta_{payment\ vehicle} = (\beta_2 - (-\beta_1 - \beta_2))/\beta_{payment\ vehicle} = (2*\beta_2 + \beta_1))/\beta_{payment\ vehicle}$.

choice effects from a gender perspective using binary logit models (see Figure 1 in Annex 3). This helped us to test hypotheses regarding the influence of both respondents' and design characteristics on opt-out forced choices made by both females and males, respectively.

7. Estimation results

7.1 *Descriptive statistics*

In total, 312 respondents were interviewed during the survey. After excluding incomplete questionnaires, 294 respondents for the RP, 152 respondents for the SP1 and 156 respondents for the SP2 versions were ready for the analysis. The survey interviews on average took 20 minutes to obtain information from the respondents. To determine the similarities (or the differences) between RP and SP1 sample characteristics, we applied semi-parametric Mann-Whitney (MW) test to the continuous variables, like age, education, household size, household income, distance travelled and travel costs, whereas we applied Chi-Squared test to the dummy variables, like gender, employment and marital status. We found no significant differences in terms of gender and overall age groups at 1% significance level; however, there are significant differences among age-groups. In both the samples, around 35% were female and 64% were male and, 80.6% and 76.3% of respondents were employed, whereas about 53% and 55% of the respondents were married. The average household size was 6.7 and 6.2 persons per household, which is nearly closed to the national average household size. The average monthly household income was PKR. 61760.20, and PKR. 59491.07 per month across both versions, respectively. Besides, there exist significant differences for marital status, household size, children per household, education, employment, average household income, the average distance travelled and average travel cost (both with and without the opportunity cost of time) across both samples at 5% and 1% levels, respectively. However, some socio-economics characteristics of visitors, including gender, marital status, and average household size, are much comparable to the country's population, whereas some are not, indicating that the most of the beach visitors belong to the young generation with higher education. The average distance travelled by the visitors interviewed is 38.1 kilometres and 36.5 kilometres in RP and SP1, respectively (see Table 2).

7.2 *Model results*

Multinomial logit (MNL) choice models are estimated using NLOGIT version 5.0. Both RP and SP1 MNL models for female and male sub-samples are presented in Tables 3, 4, 5 and 6. These models are identical in terms of their non-monetary and monetary parameters (i.e. travel cost). The RP discrete choice models for female sub-sample are the same as their single choice situation don not include an 'opt-out', however, SP1 DCE choice models are different since they include an 'opt-out' alternative using six different choice situations, framed as 'none of these: stay at home or do another non-beach activity' (see Tables 3 and 4). For male sub-sample, we also adopted this similar approach combining both RP and SP1 versions (see Tables 5 and 6).

Overall, both RP and SP1 discrete choice models for female sub-sample perform better in terms of LR test and Pseudo R^2, respectively, despite the similarities in statistical significances of parameters and, specifically alternative-specific constants (ASCs) in SP1 models estimated for male sub-sample. The magnitude of most of the coefficients in RP version of female sub-sample is higher than the same version of male sub-sample, indicating that female beach visitors have relatively higher preferences for site quality improvements than their male counterparts. Using Swait and Louviere (1993) approach, we estimated combined RP-SP1 MNL models and found that these joint models perform better than the single RP and SP1 models for both the female and male sub-samples, regardless of including and excluding female and male individuals from those sub-samples in SP1 versions who selected a forced choice in case of not having an 'opt-out' alternative (see Figure 3). Although the combined MNL models for females perform better than that of male visitors, it is nonetheless that the significances of parameters have largely improved for all the combined models (see Tables 3, 4, 5 and 6), demonstrating that the joint estimation hence increases the precision and robustness of model coefficients (Swait and Louviere, 1993).

In all the single and combined models, travel cost as a monetary parameter has the expected negative sign and is relatively highly significant in SP1 and combined models estimated from both female and male sub-samples. This demonstrates that both female and male visitors would less likely to visit beaches in the present and future if their travel cost increases by living farther away from the beaches of their choice. Across both single and combined models for female and male beach visitors, all the coefficients have the expected positive signs, except sparsely crowded (or quiet) parameters with an unexpectedly negative sign in SP1 model and a joint RP-SP1 model for male counterparts, when 'opt-out' forced-choice effects are excluded from SP1 data. This validates the results that the uncertain male beach visitors because of mostly choosing random choices, including 'opt-out' choice, influence the stability of choice parameters (e.g. Brouwer et al., 2017). Besides, female visitors have higher WTP than male visitors across both RP and SP1 versions regardless of opt-out forced-choice effects are taken into account.

We also separately estimated binary logit models for females and males sub-samples drawn from SP1 data set to detect the sources, such as respondents' and design characteristics, behind the selection of opt-out forced choices. The respondents' characteristics, such as age, education, income, marital status, household size, travelling group size, family and weekdays visitors are all statistically significant at 1 % and 5% levels, respectively. Although some characteristics affect opt-out choice behaviour of female and male visitors in a similar way (e.g. age, income, beach activities affect negatively, and education, income-squared, household size, weekdays visitor affect positively), however, signs of some characteristics, such as marital status, travelling group size and visiting with family members, appear differently indicating that these variables influence female and males visitors' opt-out forced-choice behaviour in the different ways (see Table 7). For instance, married females are more likely to choose an opt-out forced choice, whereas their male counterparts do the other way around. This reveals that married female visitors are more selective when deciding to choose from alternative beaches as compared to 'none of the beach' opt-out alternative in SP1 choice situations. The same is the case of other variables, including the travelling group and visiting with family members, for female visitors. Besides, travelling group size for female and beach activities for male visitors are insignificant indicating that travelling group size does not influence female visitors' opt-out choice behaviour, where the same is true about beach activities for male visitors.

As discussed above, we further extended our binary logit model by incorporating design characteristics, including choice card number, chosen alternative beach, travel cost (a monetary attribute), self-reported certainty, choice consistency, choice monotonicity, and randomness in choice behaviour (see Table 7). The design variables, such as choice card number and chosen alternative beach, are statistically significant and have the expected negative signs. For instance, the more interested to choose a beach and the higher the travel cost, the less likely it becomes for both the female and male visitors to choose opt-out forced choice. Self-reported certainty, though insignificant for female visitors' opts-out forced-choice behaviour, have a counter-intuitive sign, however, choice consistency is highly significant at 1% level for both the females and males and have the expected negative sign. Choice monotonicity is highly significant across both sub-samples but has an unexpected positive counter-intuitive sign for male visitors. The similar is the case of randomness in a choice variable which is highly significant at 1% level but has unexpectedly a negative for female visitors' opt-out forced-choice behaviour. Although binary logit model for males capturing the influence of the sources of opt-out forced-choice behaviour performs better than the same model for females, binary logit model for female visitors overall looks relatively more convincing in terms of both variable magnitudes and their signs. For instance, the magnitudes of variables, including marital status, visiting with family members, are relatively higher, and signs of variables, such as choice card number, choice consistency and choice monotonicity are expectedly valid for a binary logit model estimated from female SP1 sub-sample.

Table 3: Estimated MNL models including respondents who selected forced choices (RP and SP1 sub-samples of **female visitors**)

Data sets / ***Beach Attributes***	**RP version**	**SP1 version**	**RP-SP1 Combined**[2] ($\mu_{RP} \neq \mu_{SP1}$)	**RP-SP1 Combined** ($\mu_{RP} = \mu_{SP1}$)	**RP version**	**SP1 version**	**Poe et al test**
	Coefficient (SE)	***Coefficient (SE)***	***Coefficient (SE)***	***Coefficient (SE)***	***WTP***	***WTP***	***p-values***
Moderate water quality	0.694 (0.267)***	0.731 (0.165)***	0.749 (0.131)***	0.769 (0.135)***	5260.68	4111.16	0.360
Good water quality	1.528 (0.324)***	1.633 (0.160)***	1.548 (0.134)***	1.597 (0.138)***	6759.06	5309.57	0.358
Moderately littered	0.217 (0.238)	0.379 (0.126)***	0.434 (0.107)***	0.441 (0.110)***	2333.71	2463.02	0.556
No litter / Clean	0.861 (0.253)***	1.102 (0.133)***	1.076 (0.108)***	1.110 (0.113)***	3493.24	3428.45	0.511
Moderately crowded	0.175 (0.178)	0.278 (0.122)**	0.198 (0.093)**	0.187 (0.098)*	1732.57	1094.04	0.278
Sparsely crowded / Quiet	0.612 (0.227)***	0.267 (0.176)	0.391 (0.125)***	0.428 (0.132)***	2519.33	1079.53	0.135
Medium facilities	0.325 (0.188)*	0.140 (0.135)	0.286 (0.095)***	0.297 (0.099)***	3475.91	1211.96	**0.058**
High facilities	1.279 (0.292)***	0.632 (0.124)***	0.754 (0.104)***	0.768 (0.108)***	5193.12	1865.14	**0.056**
Travel cost	- 0.0005 (0.0003)*	- 0.0007 (0.0002)***	- 0.0007 (0.0001)***	- 0.0007 (0.0001)***	-----	-----	-----
Alternative-Specific Constants							
Sea View / Clifton beaches	1.601 (0.897)*	0.490 (0.376)	0.709 (0.320)**	0.706 (0.321)**			
Manora Island beach	1.413 (0.672)***	0.146 (0.367)	0.501 (0.306)	0.504 (0.306)*			
Sandspit / Turtle beach	0.930 (0.554)*	0.413 (0.341)	0.558 (0.279)*	0.556 (0.279)**			
Hawke's Bay beach	0.350 (0.686)	0.557 (0.342)	0.524 (0.285)*	0.517 (0.285)*			
French beach	0.631 (0.465)	0.515 (0.337)	0.494 (0.263)*	0.501 (0.263)*			
Paradise Point beach	0.772 (0.521)	- 0.047 (0.367)	0.219 (0.289)	0.221 (0.289)			
Cape Mount beach	0.478 (0.504)	0.488 (0.314)	0.453 (0.259)*	0.446 (0.259)*			
Log-Likelihood	-149.200	-386.320	- 548.806	- 548.602			
Pseudo R^2	0.25	0.40	0.36	0.36			
AIC	3.208	2.483	2.645	2.644			
No. of choice sets (Individuals)	01 (103)	06 (54)	(157 = 103 + 54)	(157 = 103 + 54)			
No. of observations	103	324	427 (= 103 + 324)	427 (= 103 + 324)			
Hypothesis-1 Swait & Louviere LR test	$H^{1A}: \beta^{RP} = \beta^{SP1}$	Reject preferences equality?	$H^{2A}: \mu^{RP} = \mu^{SP1}$	Reject scale equality?	**Hypothesis-2** $H^{2A}: WTP^{RP} = WTP^{SP1}$		
χ^2 (RP-SP1 MNL models)[1]	26.15	Yes	---	---	Do not reject equality for most of the WTP values		

[1] Critical values of χ^2 test with 10 and 1 degree(s) of freedom at 1% level of significance are 23.21 and 6.63, respectively. Asterisks *,**,*** indicate statistical significance at 10%, 5% and 1% levels.

[2] LR (μRP = μSP1) = – 2 [– 548.602 – (– 149.200 + – 386.320)] = 26.15

Table 4: Estimated MNL models excluding respondents who selected forced choices (RP and SP1 sub-samples of **female visitors**)

Data sets / ***Beach Attributes***	**RP version**	**SP1 version**	**RP-SP1 Combined ($\mu_{RP} \neq \mu_{SP1}$)**	**RP-SP1 Combined ($\mu_{RP} = \mu_{SP1}$)**	**RP version**	**SP1 version**	**Poe et al test**
	Coefficient (SE)	***Coefficient (SE)***	***Coefficient (SE)***	***Coefficient (SE)***	***WTP***	***WTP***	***p-values***
Moderate water quality	0.694 (0.267)***	0.720 (0.199)***	0.752 (0.146)***	0.778 (0.152)***	5260.68	3897.44	0.334
Good water quality	1.528 (0.324)***	1.795 (0.199)***	1.642 (0.152)***	1.706 (0.159)***	6759.06	5192.25	0.350
Moderately littered	0.217 (0.238)	0.198 (0.142)	0.308 (0.115)***	0.311 (0.119)***	2333.71	1805.80	0.408
No litter / Clean	0.861 (0.253)***	1.103 (0.155)***	1.025 (0.117)***	1.067 (0.122)***	3493.24	2894.84	0.416
Moderately crowded	0.175 (0.178)	0.206 (0.139)	0.159 (0.100)**	0.144 (0.105)	1732.57	704.93	0.163
Sparsely crowded / Quiet	0.612 (0.227)***	0.172 (0.194)	0.340 (0.133)***	0.378 (0.141)***	2519.33	664.33	**0.080**
Medium facilities	0.325 (0.188)*	0.049 (0.153)	0.267 (0.101)***	0.276 (0.107)***	3475.91	1081.55	**0.053**
High facilities	1.279 (0.292)***	0.801 (0.147)***	0.897 (0.117)***	0.919 (0.121)***	5193.12	1968.87	**0.070**
Travel cost	- 0.0005 (0.0003)*	- 0.0008 (0.0002)***	- 0.0007 (0.0002)***	- 0.0007 (0.0002)***	-----	-----	-----
Alternative-Specific Constants							
Sea View / Clifton beaches	1.601 (0.897)*	0.568 (0.429)	0.782 (0.353)**	0.781 (0.354)**			
Manora Island beach	1.413 (0.672)***	0.214 (0.419)	0.596 (0.337)*	0.602 (0.337)*			
Sandspit / Turtle beach	0.930 (0.554)*	0.523 (0.392)	0.654 (0.308)**	0.655 (0.308)**			
Hawke's Bay beach	0.350 (0.686)	0.627 (0.391)	0.563 (0.315)*	0.557 (0.315)*			
French beach	0.631 (0.465)	0.503 (0.389)	0.486 (0.289)*	0.496 (0.289)*			
Paradise Point beach	0.772 (0.521)	0.184 (0.409)	0.391 (0.311)	0.395 (0.311)			
Cape Mount beach	0.478 (0.504)	0.787 (0.349)	0.652 (0.279)**	0.645 (0.279)**			
Log-Likelihood	-149.200	- 305.100	- 465.666	- 465.641			
Pseudo R^2	0.25	0.41	0.36	0.36			
AIC	3.208	2.488	2.668	2.668			
No. of choice sets (Individuals)	01 (103)	06 (43)	(146 = 103 + 43)	(146 = 103 + 43)			
No. of observations	103	258	361(= 103 + 258)	361(= 103 + 258)			
Individuals excluded who selected opt-outs	---	11 (= 54 – 43)	11 (= 157 – 146)	11 (= 157 – 146)			
Hypothesis-1 Swait & Louviere LR test	$H^{1B}: \beta^{RP} = \beta^{SP1}$	Reject preferences equality?	$H_0^{2B}: \mu^{RP} = \mu^{SP1}$	Reject scale equality?	**Hypothesis-2** $H^{2B}: WTP^{RP} = WTP^{SP1}$		
χ^2 (RP-SP1 MNL models)[1]	22.68	**No**	0.05	**No**	Do not reject equality for most of the WTP values		

[1] Critical values of χ^2 test with 10 and 1 degree(s) of freedom at 1% level of significance are 23.21 and 6.63, respectively.

[2] LR (μRP = μSP1) = – 2 [– 465.641 – (– 149.200 + – 305.100)] = 22.68, whereas LR (μRP ≠ μSP1) = – 2 [– 465.666 – (– 465.641)] = 0.05

Table 5: Estimated MNL models including respondents who selected forced choices (RP and SP1 sub-samples of **male visitors**)

Data sets / ***Beach Attributes***	**RP version**	**SP1 version**	**RP-SP1 Combined** $(\mu_{D1} \neq \mu_{D2})$	**RP-SP1 Combined** $(\mu_{RP} = \mu_{SP1})$	**RP version**	**SP1 version**	**Poe et al test**
	Coefficient (SE)	***Coefficient (SE)***	***Coefficient (SE)***	***Coefficient (SE)***	***WTP***	***WTP***	***p-values***
Moderate water quality	0.286 (0.164)*	0.607 (0.115)***	0.456 (0.085)***	0.475 (0.088)***	2868.85	2610.83	0.461
Good water quality	1.135 (0.212)***	1.578 (0.113)***	1.381 (0.086)***	1.430 (0.089)***	4295.44	3518.69	0.357
Moderately littered	0.464 (0.163)***	0.209 (0.089)**	0.256 (0.073)***	0.253 (0.075)***	2633.59	1280.07	**0.089**
No litter / Clean	0.638 (0.197)***	0.951 (0.094)***	0.881 (0.074)***	0.912 (0.077)***	2925.33	1974.24	0.227
Moderately crowded	0.039 (0.120)	0.385 (0.088)***	0.260 (0.065)***	0.262 (0.068)***	710.18	683.03	0.512
Sparsely crowded / Quiet	0.343 (0.164)**	- 0.039 (0.112)	0.011 (0.083)**	0.024 (0.087)	1219.51	285.60	**0.071**
Medium facilities	0.138 (0.123)	0.118 (0.093)	0.163 (0.067)***	0.165 (0.074)**	2289.28	676.34	**0.012**
High facilities	1.087 (0.182)***	0.488 (0.089)***	0.634 (0.071)***	0.639 (0.071)***	3884.52	1022.29	**0.007**
Travel cost	- 0.0006 (0.0002)**	- 0.0011 (0.0002)***	- 0.0009 (0.0001)***	- 0.0009 (0.0001)***	-----	-----	-----
Alternative-Specific Constants							
Sea View / Clifton beaches	1.530 (0.619)**	0.652 (0.279)**	0.884 (0.230)***	0.886 (0.231)***			
Manora Island beach	0.486 (0.465)	0.114 (0.268)	0.274 (0.221)	0.275 (0.222)			
Sandspit / Turtle beach	1.078 (0.357)***	0.493 (0.245)**	0.642 (0.193)***	0.642 (0.193)***			
Hawke's Bay beach	0.430 (0.412)	0.496 (0.250)**	0.448 (0.200)**	0.446 (0.201)**			
French beach	0.234 (0.232)	0.308 (0.247)	0.211 (0.189)	0.221 (0.190)			
Paradise Point beach	0.027 (0.388)	0.229 (0.248)	0.172 (0.199)	0.175 (0.199)			
Cape Mount beach	0.117 (0.375)	- 0.185 (0.253)	- 0.106 (0.202)	- 0.118 (0.203)			
Log-Likelihood	- 312.779	- 709.778	- 1050.601	- 1047.556			
Pseudo R^2	0.17	0.39	0.33	0.33			
AIC	3.443	2.468	2.738	2.730			
No. of choice sets (Individuals)	01 (191)	06 (98)	(289 = 191 + 98)	(289 = 191 + 98)			
No. of observations	191	588	779 (= 191 + 588)	779 (= 191 + 588)			
Hypothesis-1 Swait & Louviere LR test	$H_0^{1C}: \beta^{RP} = \beta^{SP1}$	Reject preferences equality?	$H_0^{3C}: \mu^{RP} = \mu^{SP1}$	Reject scale equality?	**Hypothesis-2** $H^{2C}: WTP^{RP} = WTP^{SP1}$		
χ^2 (RP-SP1 MNL models)[1]	49.99	Yes	---	---	Do not reject equality for half of the WTP values		

[1] Critical values of χ^2 test with 10 and 1 degree(s) of freedom at 1% level of significance are 23.21 and 6.63, respectively.

[2] LR (μRP = μSP1) = – 2 [– 1047.556 – (– 312.779 + – 709.778)] = 49.99

Table 6: Estimated MNL models excluding respondents who selected forced choices (RP and SP1 sub-samples of **male visitors**)

Data sets / Beach Attributes	RP version	SP1 version	RP-SP1 Combined ($\mu_{RP} \neq \mu_{SP1}$)	RP-SP1 Combined ($\mu_{RP} = \mu_{SP1}$)	RP version	SP1 version	Poe et al test
	Coefficient (SE)	***Coefficient (SE)***	***Coefficient (SE)***	***Coefficient (SE)***	***WTP***	***WTP***	***p-values***
Moderate water quality	0.286 (0.164)*	0.603 (0.125)***	0.456 (0.085)***	0.484 (0.092)***	2868.85	2671.42	0.485
Good water quality	1.135 (0.212)***	1.627 (0.126)***	1.381 (0.086)***	1.465 (0.094)***	4295.44	3636.56	0.394
Moderately littered	0.464 (0.163)***	0.096 (0.095)	0.256 (0.073)***	0.184 (0.078)**	2633.59	1105.26	**0.062**
No litter / Clean	0.638 (0.197)***	0.979 (0.102)***	0.881 (0.074)***	0.913 (0.080)***	2925.33	1936.63	0.226
Moderately crowded	0.039 (0.120)	0.373 (0.094)***	0.260 (0.065)***	0.255 (0.071)***	710.18	570.79	0.430
Sparsely crowded (and Quiet)	0.343 (0.164)**	- 0.140 (0.118)	0.011 (0.083)**	- 0.031 (0.091)	1219.51	87.32	**0.034**
Medium facilities	0.138 (0.123)	0.069 (0.094)	0.163 (0.067)***	0.158 (0.073)**	2289.28	690.66	**0.015**
High facilities	1.087 (0.182)***	0.593 (0.096)***	0.634 (0.071)***	0.719 (0.077)***	3884.52	1184.74	**0.013**
Travel cost	- 0.0006 (0.0002)**	- 0.0011 (0.0002)***	- 0.0009 (0.0001)***	- 0.0009 (0.0001)***	-----	-----	-----
Alternative-Specific Constants							
Sea View / Clifton beaches	1.530 (0.619)**	0.848 (0.302)** *	0.884 (0.230)***	1.034 (0.243)***			
Manora Island beach	0.486 (0.465)	0.227 (0.288)	0.274 (0.221)	0.359 (0.233)			
Sandspit / Turtle beach	1.078 (0.357)***	0.562 (0.265)**	0.642 (0.193)***	0.689 (0.203)***			
Hawke's Bay beach	0.430 (0.412)	0.589 (0.269)**	0.448 (0.200)**	0.494 (0.211)**			
French beach	0.234 (0.232)	0.416 (0.263)	0.211 (0.189)	0.272 (0.198)			
Paradise Point beach	0.027 (0.388)	0.343 (0.262)	0.172 (0.199)	0.233 (0.207)			
Cape Mount beach	0.117 (0.375)	- 0.151 (0.268)	- 0.106 (0.202)	- 0.095 (0.211)			
Log-Likelihood	- 312.779	- 632.328	- 970.492	- 967.691			
Pseudo R^2	0.17	0.39	0.33	0.33			
AIC	3.443	2.456	2.744	2.736			
No. of choice sets (Individuals)	01 (191)	06 (88)	(279 = 191 + 88)	(279 = 191 + 88)			
No. of observations	191	528	719 (= 191 + 528)	719 (= 191 + 528)			
Individuals excluded who selected opt-outs	---	10 (= 98 – 88)	10 (= 289 – 279)	10 (= 289 – 279)			
Hypothesis-1 Swait & Louviere LR test	$H^{1D}: \beta^{RP} = \beta^{SP1}$	Reject preferences equality?	$H^{2D}: \mu^{RP} = \mu^{SP1}$	Reject scale equality?	**Hypothesis-2** $H^{2D}: WTP^{RP} = WTP^{SP1}$		
χ^2 (RP-SP1 MNL models)[1]	45.17	Yes	---	---	Do not reject equality for half of the WTP values		

[1] Critical values of χ^2 test with 10 and 1 degree(s) of freedom at 1% level of significance are 23.21 and 6.63, respectively.

[2] LR (μRP = μSP1) = – 2 [– 967.691 – (– 312.779 + – 632.328)] = 45.17

8. Hypotheses testing results

8.1 *Preferences equality: Isolating opt-out forced-choice effects from a gender perspective*

To test our first series of hypotheses, we apply Swait and Louviere (1993) LR test procedure using MNL models to assess the equivalence of the estimated preference β and their scales μ parameters across RP and SP1 DCE estimated from sub-samples of females and males by isolating (or disentangling) opt-out forced-choice affects. We reject the null hypothesis of equality of preference and scale parameters at 1% significance level if forced-choice effects are not excluded from the female sub-sample (see Tables 3), however, we cannot reject the same hypothesis if forced-choice effects are excluded from female sub-sample at 1% level of significance (see Table 4). This validates that the respondent's choice uncertainty that leads to their random choice behaviour and thus their frequent selection of 'opt-out' alternative (Brouwer et al., 2017) in a forced-choice situation. No matter with a little evidence this study provides, we can still overcome this problem by asking the respondents about their forced-choice by not giving them 'opt-out' choice. On the contrary, we reject the hypotheses of equality of preference and scale parameters at 1% levels of significance if forced-choice effects are first included and then excluded from the male sub-sample (see Tables 5 and 6). This once again validates that male visitor' preferences are more stable than their female counterparts. By isolating this forced choice effect, we only witnessed the equality of preference and scale parameters for female sub-sample (see Table 4).

8.2 *WTP equality: Isolating opt-out forced-choice effect from a gender perspective*

To test our second series of hypotheses, we apply Poe, Giraud and Loomis (2005) combinatorial test to assess the differences in average WTP estimates across RP and SP1 MNL discrete choice models estimated from sub-samples of females and males by disentangling opt-out forced-choice effects. If forced-choice opt-out effects are taken into account, we cannot reject the hypothesis of equality of WTP values for site quality improvements in the most of the attributes for female sub-sample (see Tables 3 and 4), however, the same is not true for male sub-sample, i.e. we reject WTP equality for half of the beach attributes (see Tables 5 and 6).

8.3 *Sources of opt-out forced choices influencing preference equality of female and male visitors*

Although numerous studies have addressed the impact of including (and excluding) an opt-out alternative on preferences (e.g. Pederson et al., 2011; Kontoleon and Yabe, 2003; Banzhaf, Johnson, and Mathews, 2001; Veldwijk et al., 2014), however, according to our knowledge, modelling the sources (or determinants) of opt-outs has never been studied before. We selected each respondent from a gender perspective selecting either eight beaches or opt-out as a binary random variable when facing six different choice situations for SP1 version and then applied a binary logit [5] with explanatory variables, including the visitor and design characteristics. The rationale behind this novel approach is to determine the impact of visitor and design characteristics influencing the selection of opt-out choices, which eventually influences preferences equality across RP and SP1 versions. The results of binary logit models derived from both female and male SP1 sub-samples are demonstrated in Table 7.

A binary logit model for female sub-sample exhibits that the female visitors' selection of an opt-out alternative followed by forced-choice is significantly affected by gender (female visitors tend to choose more opt-outs than their male counterparts), age (the older visitors tend to choose fewer opt-outs), education (educated visitors are likely more opt-outs, income (visitors with the higher income tend to choose more opt-outs than beach alternatives), and marital status (married couples are likely to choose more opt-outs). Both household size and travelling group size significantly but negatively affect neither beach (or opt-out) choice (respondents with an increase in the number of persons per households and the number of persons in their travelling group tend to choose less number of opt-outs), whereas respondents with the increased number of children in the travelling group are likely to select more number opt-outs (Table 7). The visitors travelling to beaches either with their family or friends tend to choose less number of opt-outs, which indicates their liking for coastal recreation. Besides, the visitors who engage in water activities, such as swimming and bathing, are less likely to choose an opt-out alternative, however, visitors who visit beaches during weekdays are more likely to choose opt-outs as compared to weekend visitors.

Also, our analysis has included design characteristics which significantly but differently influence the opt-out choice behaviour of both female and male visitors. As expected, the female respondent facing varying choice cards and varying beach alternative sites as compared to neither beach (or opt-out) alternative tend to choose the less number of opt-outs as compared to their male counterparts, who choose the more number of opt-outs. Travel cost (monetary attribute) influences negatively visitor's likelihood of choosing neither beach alternative, which indicates that the higher travel cost is borne by the female and male visitors the higher the likelihood not to choose an opt-out alternative. Self-reported certainty of the respondents does not adversely affect their choice of an opt-out alternative. Furthermore, we included three unique variables, including consistency, monotonicity and randomness, in our analysis. Choice consistency is usually defined as choosing an identical choice repeatedly when facing the varying choice situations with the similar alternative, choice monotonicity is described as the selection of the dominant alternative by the same respondent (Mattmann et al., 2019; Determann et al., 2011), and choice randomness can be defined as the

[5] We also applied Poison model because the choice of opt-out as a count variable was highly skewed. Besides, we also applied Zero-Inflated Poisson model because of observing too many zeros for the count variable, but it did not perform better than Poisson model. Also, we tested for over-dispersion to better apply the Negative Binomial model; however, we found that there was no such problem. The results are available with the author upon request. Overall, Binary Logit model performed better.

non-existence of neither choice consistency nor choice monotonicity (i.e. all choice alternatives are differently chosen when going from the first to the last choice situation). Whether it is consistent, monotonic or even random choices made by the visitors, these all variables significantly and negatively influence the selection of neither beach alternatives for the female visitors only. Putting differently, a respondent, who makes consistent, monotonic and random choices, tends to choose the less number of opt-out alternatives. However, this is only true for male respondents in terms of their consistent choice. Finally, Table 8 demonstrates that the design characteristics in addition to visitors' (or respondents') characteristics are much more important. Both the Wald and Log-Likelihood Ratio (LR) tests confirm that the null hypothesis of indicating that the design characteristics are not significant is rejected at 5% level of significance.

Table 7: Sources of opt-out forced choices from a gender perspective in the SP1 version (Binary logit model results)

Visitors characteristics	*Variable coding / range*	**Expected sign**	**SP1 version females** *Coefficient (SE)*		**SP1 version males** *Coefficient (SE)*	
Constant	Intercept	Minus /Plus	0.397	(1.015)	- 8.919	(1.168)***
Age	18 – 63 years	Minus	- 0.145	(0.015)***	- 0.039	(0.009)***
Education	0 – 18 years	Plus	0.177	(0.060)***	0.285	(0.057)***
Income	5 – 175 x 10^3 PKR / month	Minus	- 00017	(0.00002)***	- 0.00004	(0.130)***
$Income^2$	Income-squared	Plus	$1.35x10^{-09}$	$(1.37x10^{-10})$***	$3.39x10^{-10}$	$(3.46x10^{-10})$***
Marital status	1 = Married	Minus /Plus	3.805	(0.256)***	- 0.796	(0.205)***
Household size	1 – 22 persons/household	Plus	0.084	(0.026)***	0.075	(0.020)***
Travelling group size	1 – 48 persons / visit	Plus	0.109	(0.017)***	- 0.253	(0.018)***
Children in the travelling group	0 – 25 children / group	Plus	0.016	(0.030)	0.639	(0.037)***
Family members	1 = Visiting with family members	Plus	1.439	(0.252)***	- 0.574	(0.153)***
Beach activities	1 = Yes	Minus	- 0.386	(0.160)**	- 0.189	(0.131)
Weekdays visitor	1 = Yes	Plus	2.322	(0.213)***	2.716	(0.178)***
Design characteristics						
Choice card number	36 choice cards in total = 6 sets/ respondent x 6 blocks	Minus	- 0.039	(0.007)***	0.076	(0.008)***
Chosen alternative beach	0 = Opt-out (base), 1 – 8 alternative beaches	Minus	- 0.271	(0.033)***	- 0.147	(0.031)***
Travel cost (monetary attribute)	0 – 1854.95 PKR / visitor	Minus	- 0.0004	(0.0003)*	- 0.0037	(0.0005)***
Self-reported certainty	0 = Not important at all, . . . , 10 = Most important	Minus	0.076	(0.061)	0.417	(0.066)***
Choice consistency	1 = Chosen the same alternative twice or more times	Minus	- 0.912	(0.213)***	- 2.515	(0.189)***
Choice monotonicity	1 = Dominant alternative chosen in the choice set	Minus	- 0.540	(0.208)***	1.420	(0.324)***
Randomness in choice	1 = Neither consistent nor monotonic choice	Minus	- 1.629	(0.335)***	1.515	(0.372)***
Log-Likelihood (df = 10)	Model with respondents' characteristics only		- 894.613		- 1333.764	
Log-Likelihood (df = 17)	Model with respondents' and design characteristics		- 820.485		- 1058.446	
LR test (χ^2) (df =10)			1005.27***		1049.390***	
LR test (χ^2) (df =17)			1153.52***		1600.030***	
Pseudo R^2			0.41		0.43	
AIC			1676.97		2152.89	
No. of observations (N)	912 (8,208) (152 respondents)		2916 (324)		5292 (588)	

Asterisks *,**,*** indicate statistical significance at 10%, 5% and 1% levels.

Table 8: Hypothesis testing results (Visitor and design characteristics form a gender perspective)

Hypothesis-3	**Wald test (p-value)** $\beta_i = 0 \quad \forall\ i = 1, 2, 3, \ldots, n$	**LR test (p-value)** LR = – 2 (LL Model d. f = 10) – (– LL Model d. f = 17)
H^{3A}: $\beta^{SP1}_{females} = 0$	Wald = 127.88 (0.000)[1]	LR = – 2 [– 894.613 – (– 820.485)] LR = 148.26 (0.0000)[2]
H^{3B}: $\beta^{SP1}_{males} = 0$	Wald = 346.34 (0.000)[1]	LR = – 2 [–1333.764 – (–1058.446)] LR = 550.64 (0.0000)[2]

[1] Chi-square critical value with degrees of freedom = 7 at 5% significance level is 14.067

[2] As compared to Wald test, which is one model test, LR test is based on two models, so we used the restricted and unrestricted log-likelihood ratios to test our third series of hypotheses (see Table 7).

9. Conclusions and discussion

By taking into account of opt-out forced-choice effects from a gender perspective, this paper implemented a novel approach based on assessing the convergent validity of preferences and WTP values of beach visitors estimated from female and male sub-samples of revealed preference (RP) discrete choice model and stated preference (SP1) DCE. We carried out this study for non-market valuation of beaches in Karachi, highly populated and industrialised city in Sindh province of Pakistan, along the Arabian Sea. We isolated (or disentangled) forced-choice effects using first an unforced choice situation followed by a forced-choice situation in SP1 DCE by excluding (and including) the female and male respondents. Later, we combined SP1 DCE data with RP data by creating sub-samples of female and male visitors and finally assessed the convergent validity of preferences and WTP values from gender perspective using the MNL models. The results demonstrate that when opt-out forced choices are not excluded, preferences of both the female and male visitors do not remain equivalent, however, preferences of the female visitors remain equivalent or similar once opt-out forced-choice effects are excluded from SP1 DCE version. This indicates that preferences estimates could be biased if forced-choice effects are not taken into account.

Empirically, our results are different from previous studies conducted by Birol, et al. (2006) and Adamowicz, et al (1994; 1997) which found that combined RP and SP data are overall compatible. To this end, one can reasonably agree to a larger extent, but knowing that preference parameters and WTP measures are susceptible to various design and individual-specific factors, for instance, choice task complexity (Hanley et al., 2002), labelled versus unlabeled choice sets and spatial heterogeneity (e.g. Logar and Brouwer et al., 2018), framing and substitution effects (Schaafsma and Brouwer, 2013), single and multiple monetary parameters (e.g. Talpur et al., 2018), starting point bias (Ladenburg and Olsen, 2008), and so on so forth, our study suggests that preferences stability if affected by the different factors as some mentioned above can be further investigated by disaggregating and isolating various design and individual-specific characteristics. For example, Landenburg and Olsen (2013) found in their study that that preferences and welfare measures are influenced by starting point bias that is a gender-specific. By adopting such an approach, we envisage that it is by and large possible that preference similarity can be achieved. Parallel to this, our study suggests that preference equality is a gender-specific if out-out forced-choice effects are isolated and demonstrate that various individual and design characteristics as further discussed below are the factors influencing the selection of opt-out alternative that eventually effect preferences equality of at least female respondents.

Estimating two separate binary logit models from the female and male SP1 sub-samples, our paper further analysed the influence of respondent and design characteristics of female and male visitors on their choice behaviour of selecting more or less number of opt-out alternatives. Although the binary logit model for the males performed relatively better, however, the same model for the female respondents makes much more sense so far as the design characteristics are concerned. For instance, choice consistency, monotonicity and randomness in choosing beach site alternatives have a negative impact on selecting a forced-choice opt-out alternative in case of female visitors as compared to male visitors, which means the higher is a consistent, monotonic and random choice, the less will be the likelihood to select an opt-out forced choice that eventually deviates preference equality, in our case of female visitors. Besides, respondent (or visitor) characteristics, including age, education, income, household size, children in the travelling group, beach activities and weekdays visitor, affect both female and male visitors' selection of opt-out forced-choices similarly, whereas marital status, travelling group size, and travelling with family members affect both types of visitors differently. These results, therefore, conclude that an opt-out forced-choice behaviour of female and male visitors is partially analogous (and partially different), and so there is a need to further investigate this research by using an approach based on varying alternatives in SP1 choice set with and without opt-outs and then combine these samples with the same RP version to finally assess the convergent validity from a gender perspective.

References

Adamowicz, W., Louviere, J. and Williams, M. (1994), Combining stated and revealed preference methods for valuing environmental amenitie*s, Environmental Economics and Management* 26, 271-292

Adamowicz, W., Swait, J., Boxall, P., Louviere, J. and Williams, M. (1997), Perceptions versus objective measures in combined revealed and stated preference models of environmental valuation, *Environmental Economics and Management* 32, 65–84

Arnberger, A. and Eder, R. (2011), The influence of age on recreational trail preferences of urban green-space visitors: a discrete choice experiment with digitally calibrated images, *Environmental Planning and Management* 54 (7), 891 – 908

Bakhtiari, F., Jacobsen, J., Thorsen, B. J., Lundhede, T., Strange, N, and Boman, M. (2018), Disentangling distance and country effects on the value of conservation across national borders, *Ecological Economics* 147, 11 – 20

Banzhaf, M. R., Johnson, F. R., and Mathews, K. (2001), Opt-out alternatives and anglers' stated preferences, *The choice modelling approach to environmental valuation*, 157-177

Bateman, I. J., Brouwer, R., Ferrini, S., Schaafsma, M., Barton, D., Dubgaard, A., Hasler, B., Hime, S., Liekens, I, Navrud, S., De Nocker, L., Šceponaviciute, R., and Semeniene. (2011), Making benefit transfers work: Deriving and testing principles for value transfers for similar and dissimilar sites using a case study of the non-market benefits of water quality improvements across Europe, *Environmental and Resource Economics* 50, 365–387

Birol, E., Kontoleon, A. and Smale, M. (2006), Combining revealed and stated preference methods to assess the private value of agro-biodiversity in Hungarian home gardens, *EPT Discussion Paper 156*, International Food Policy and Research Centre (IFPRI), Washington DC, USA

Bjerke, T., Østdahl, T., Thrane, C., and Strumse, E.. (2006), Vegetation density of urban parks and perceived appropriateness for recreation, *Urban Forestry and Urban Greening*, 5, 35–44

Brouwer, R., Logar, I and Sheremet, O. (2017), Choice consistency and preference stability in test-retest s of discrete choice experiments and open-ended willingness to pay elicitation formats, *Environmental and Resource Economics* 68 (3), 729 – 751

Brouwer, R. (1999), Public right of access, overcrowding and the value of peace and quiet: the validity of contingent valuation as an information tool, *CSERGE Working Paper* GEC 99-05, University of East Anglia, UK

Cameron, T. (1992), Combining contingent valuation and travel cost data for the valuation of non-market goods, *Land Economics* 68 (3), 302 – 317

Campbell, D. and Erdem, S. (2018), Including opt-out options in discrete choice experiments: Issues to consider, *The Patient – Patient Centered Outcome Research* 12 (1), 1 – 14, Springer International

Carson, R. T. and Mitchell, R. C. (1993), The value of clean water: The Public's willingness to pay for boatable, fishable, and swimmable quality water, *Water Resources Research* 29, 2445–2454

Cheng, L. and Lupi, F. (2016), Combining revealed and stated preference methods for valuing water quality changes to Great Lake beaches, a selected paper for presentation for th2 2016 Agriculture and Applied Economics Association, Boston, M. A., July 31 – August 2, 2016

Choice Metrics, C. (2014), Ngene Software 1.1 , 1 User manual and reference guide, Sydney, Australia

Dhar, R. (1997), Consumer preference for no-choice option, *Consumer Research* 24 (2), 215 – 231

Dhar, R. and Simonson, I. (2003), The effect of forced-choice on choice, *Journal of Marketing Research*, 40(2), 146-160

Determann et. al. (2011), Impact of survey administration mode on the results of a health-related discrete choice experiment: Online and paper comparison, Value in Health 20 (7), 953 – 960

Hanley, N., Wright, R. and Koop, G. (2002), Modelling recreation demand using choice experiments: Climbing in Scotland, *Environmental and Resource Economics* 22 (3), 449-466

Hanley, N., Mourato, S., and Wright, R. E. (2001), Choice modelling approaches: A superior alternative for environmental valuation?, *Journal of Economic Surveys,* 15(3), 435-462.

Hensher, D. and Rose, J. (2007), Development of commuter and non-commuter mode choice models for the assessment of new public transport infrastructure projects: A case study, *Transportation Research Part-A: Policy and Practice* 41 (5), 428 – 443

Khalil, S. (1999), Economic valuation of the mangrove ecosystem along the Karachi coastal areas, in Joy E. Hecht (Ed.)., *The Economic Value of the Environment: Cases from South Asia*, IUCN

Keane et al (2016), Gender-differentiated preferences for community-based conservancy (CBC) initiative, *PLoS ONE Journal* 11 (3), doi:10.1371/journal.pone.0152432

Kemperman, A.D.A.M. and Timmermans, H.J.P., 2006. Heterogeneity in urban park use of ageing visitors: A latent class analysis, *Leisure Sciences*, 28, 57–71

Krinsky, I. and Robb, A.L. (1986), On approximating the statistical properties of elasticities, *Review of Economics and Statistics* 68, 715–719

Kontoleon, A., and Yabe, M. (2003), Assessing the impacts of alternative 'opt-out' formats in choice experiment studies: consumer preferences for genetically modified content and production information in food, *Journal of Agricultural Policy and Resources, 5*(1), 1-43

KSDP: Karachi Strategic Development Plan (2007), *Karachi Strategic Development Plan-2020*, Master Plan Group of Offices, City District Government, Karachi, Pakistan

Ladenburg, J. and Olsen, S. B. (2008), Gender-specific starting point bias in choice experiments: Evidence from an empirical study, *Environmental Economics and Management* 56, 275 – 285

Lancaster, K. (1966), A new approach to consumer theory, *The Journal of Political Economy*, 74, 132 – 157

Lancsar, J., Hall, J., Madeleine, K., Kenny, P., Louviere, J., Fiebig, D., Hossain, I., Thien, F., Reddel, K., and Jenkins, C. (2007), Using discrete choice experiments to investigate subject preferences for preventive asthma medication, *Respirology* 12, (1) 127-136

Logar, I and Brouwer, R. (2018), Substitution and spatial preference heterogeneity in single – and multiple – site choice experiments, *Land Economics* 94 (2), 320-322

Louviere, J., Hensher, D. and Swait, J. (2000), *Stated Choice Methods: Analysis and Applications*, Cambridge University Press

Mattmann, M., Logar, I. and Brouwer, R. (2019), Choice certainty, consistency, and monotonicity in discrete choice experiments, *Journal of Environmental Economics and Policy*, 8 (2), 109 – 127

McFadden, D. (1974), Conditional logit analysis of qualitative choice behaviour, In Zarembka, P., editor, *Frontiers in econometrics*, pages 105–142, Academic Press

Nlogit 5.0 / Limdep 10, (1986 – 2002), Econometric Software, Inc., New York, The USA

Pedersen, L. B. and Gyrd-Hansen, D. (2013), Implications of researchers dubious use of the 'neither' option, and recommendations on the future use of 'status quo' and 'opt-out' options in choice experiments, a presented at the International Choice Modelling Conference–2013, Sydney, Australia, July 3–5, 2013

Pedersen, L. B., Kjær, T., Kragstrup, J., and Gyrd-Hansen, D. (2011), Does the inclusion of a cost attribute in forced and unforced choices matter?: Results from a web survey applying the discrete choice experiment, *Journal of Choice Modelling,* 4(3), 88-109

Pedersen, L.B., Kiil, A., and Kjær, T. (2011), Soccer attendees' preferences for facilities at the Fionia Park Stadium: An application of the discrete choice experiment, *Journal of Sports Economics*, 12, (2) 179 – 199

Poe, G. L., Giraud, K. L., and Loomis, J. B. (2005), Computational methods for measuring the difference of empirical distributions, *American Journal of Agricultural Economics* 87 (2): 353 – 65

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

SACEP (2007), Marine Litter in the South Asian Seas Region, *A report*, September-2007, South Asia Cooperative Environment Programme (SACEP), Colombo, Sri Lanka

Schaafsma, M. and Brouwer, R. (2013), Testing geographical framing and substitution effects in spatial choice experiments. *Journal of Choice Modelling* 8, 32 – 48

Smith, V. K., Zhang X., and Palmquist, R.B. (1997), Marine debris, beach quality and non-market values. *Environmental and Resource Economics* 10, 223–247

Swait, J. D. and Louviere, J. (1993), The role of scale parameter estimation in the estimation and comparison of Multinomial Logit models, *Journal of Marketing Research* 30, 304–314

Tahvanainen, L., Tyrväinen, L, Ihalainen, M., Vuorela, N. and Kolehmainen, O. (2001), Forest management and public perceptions – visual versus verbal information, *Landscape and Urban Planning* 53 (1/4), 53–70

Talpur, M. A., Koetse, M. J., and Brouwer, R. (2018), Accounting for implicit and explicit payment vehicles in a discrete choice experiment, *Journal of Environmental Economics and Policy*, 7 (4), 363 – 385

Talpur, M.A., and Jariko, G.A. (2001), Role of environmental legislation and administration in protecting the environment: the experience of Pakistan, *Biannual Research Journal of Grassroots* 23, 75–84

Taylor, T., and Longo, A. (2010), Valuing algal bloom in the Black Sea Coast of Bulgaria: A choice experiments approach, *Journal of Environmental Management* 91, 1963–1971

Veldwijk, J., Lambooij, M., de Bekker-Grob, E., Smit, H. A. and de Wit, G. (2014), The effect of including an opt-out option in discrete choice experiments, *PLoS ONE Journal* 9(11), doi:10.1371/journal.pone.0111805

Thompson, W. C., Aspinall, P., Bell, S. and Finlay, C. (2005), "It gets you away from everyday life": Local woodlands and community use – what makes a difference?, *Landscape Research* 30 (1), 109–146

ANNEXURES

Annex 1. Detailed definitions and illustrations of the attribute levels used in the survey

Water quality levels

Good

Water has high clarity, so it is clean and clear, looks blue. Suitable for swimming, bathing, wading and playing in the water. Water does not smell, so it does not affect walking and other beach-related activities. (Almost) No visible signs of pollution and a large number of diverse finfish and shellfish species. Fish is definitely safe to eat.

Moderate

Water is slightly clear, so it is less dirty and looks moderately blue. Only suitable for wading & playing in the water. Water has a slightly bad odour (it smells now and again), so it moderately adversely affects walking and other beach-related activities. Few visible signs of pollution and reasonable numbers of finfish and shellfish species. Fish is probably safe to eat.

Poor

Water has low clarity, so it is dirty and cloudy, looks dark, not blue. Not suitable for swimming, bathing, wading and playing in the water. Water has a bad odour (it is smelly), so it highly adversely affects walking and other beach-related activities. Visible signs of pollution and little finfish & little variety of shellfish. Fish is not definitely safe to eat.

Cleanliness levels

No Liter

The beach is clean, (almost) without litter. At first glance no visible litter, only sometimes if you look closer, which means 0 – 5 units of litter per 50 meters of beach length. There are almost no chances of getting injured, so the beach is highly suitable for walking, and other beach activities.

Moderately Littered

The beach is moderately littered. A significant part of the beach contains litter, which means 5 – 25 units of litter per 50 meters of beach length. The chances of getting injured are moderate so the beach is moderately suitable for walking and beach other activities.

Very Littered

The beach is very littered. Litter is visible nearly everywhere in all shapes and sizes, which means more than 25 units of litter per 50 meters of beach length. The chances of getting injured are high, so the beach is not suitable for walking and other beach activities.

Crowding levels		
Sparsely Crowded 	**Moderately Crowded** 	**Very Crowded** 
Very few visitors, so the beach is generally quiet. Suitable for walking, playing sports, getting relaxed and having privacy. Crowd occupancy is less than 10 people per 50 meters of beach length.	A substantial number of visitors, leading most of the times to some crowding. Moderately suitable for walking, playing sports, relaxing and having some privacy. Crowd occupancy is between 10 – 20 people per 50 meters of beach length.	A lot of visitors, leading most of the times to overcrowding and noise. Not suitable for walking, playing sports, relaxing and having some privacy. Crowd occupancy is more than 25 people per 50 meters of beach length.
Facilities levels		
Low 	**Medium** 	**High** 
The beach has parking facilities.	The beach has parking facilities and food stores/restaurants	The beach has parking facilities, food stores/restaurants and toilets

Annex 2. Travel cost calculation approach

Since distance acts as a proxy for travel cost (Hanley et al., 2002; Adamowicz et al., 1997), we converted the distance into travel cost (TC) using the standard formula:

$$TC = 2 * (Distance\ cost + Opportunity\ cost\ of\ travel\ time) \quad \text{(A2.1)}$$

Travel costs are measured as the sum of a visitor's round distance (fuel) cost and the opportunity cost of travelling time to a beach. As people in Pakistan customarily bear the fuel cost or transport fares for their family members and relatives when travelling with them, we measured distance cost as the two-way fuel costs borne by the head of the household and, where applicable, shared by the total number of adults in the travelling group.

The opportunity cost of time (OCT) is calculated as the benefits forgone of a visitor's time and fixed at 30 per cent of his annual income, as is common practice in the travel cost literature (e.g. Talpur et al., 2018). Since working hours differ widely in both the private and public sector in Pakistan due to the lack of labour law enforcement, we included an estimate of each visitor's varying working hours (if employed) instead of a fixed number of working hours (e.g. 40 hours per week) and multiplied these varying working hours with 52 weeks to arrive at a visitor's total annual working hours. Thus, the opportunity cost of travelling time is calculated as:

$$OCT = (0.3 * MI * 12 * EM) * (RTT/WHrs) \quad \text{(A2.2)}$$

where MI is a visitor's monthly income, EM indicates whether a visitor is employed or not, RTT is the round-trip travel time and WHrs is the varying annual number of working hours. EM equals 1 if a visitor is employed and 0 if he is unemployed. In the latter case, OCT is zero (Adamowicz et al., 1997; Talpur et al., 2018) and the travel cost only depends on distance-based transportation costs. RTT is calculated by dividing each respondent's perceived travel distance by an average travelling speed of 60 km/hour based on own observations whilst travelling to the different beaches during the survey. For SP1 data, travel cost was not converted into the future increase in travel cost, so, therefore, travel cost definition remained the same in both RP and SP1 versions. The purpose behind this approach was based on comparing WTP values more rationally across RP and SP1 choice models using the same travel cost formula with an assumption that travel cost does no increase at least in very near future, i.e. the next year in case of SP1 data.

Annex 3.

Figure 1: Sources of opt-out forced-choice effects influencing SP1 female and male preferences