Non-market Valuation of Urban Ecosystem Services in Penang, Malaysia

A thesis

submitted in fulfilment

of the requirements for the degree

of

Doctor of Philosophy in Economics

at

The University of Waikato

by

CHIA CHIA KOAY

THE UNIVERSITY OF

WAIKATO

Te Whare Wānanga o Waikato

2024

Abstract

This thesis undertakes an empirical examination of the non-market values attributed to ecosystem services provided by urban green spaces in Penang, Malaysia. Recognizing the central role of urban green spaces in enhancing urban dwellers' well-being through ecosystem services, the study addresses the growing threats posed by urbanization to these areas. This research employs non-market valuation techniques, aiming to provide a comprehensive understanding of the contributions of urban green spaces to the local economy.

This thesis comprises five topics. The first topic employs the best-worst scaling method to explore urban residents' preferences for green space characteristics. The study answers the research questions on attribute importance and preference consistency across different survey designs. I find the air quality emerges as the most important attribute, emphasizing the high demand for improved air quality among urban residents. Moreover, I find inconsistencies in respondent preferences, indicating challenges in maintaining stable attribute hierarchies with additional elements in survey designs.

The second topic extends the analysis to individuals' visit preferences and behaviors for actual urban green spaces in Penang using travel cost method. It addresses three research questions: the contribution of travel costs to visit patterns, the impact of preferences for ecosystem assets and services on visit choices, and the reliability of estimates for travel and recreation time values. The study reveals variations in travel cost estimates and emphasizes the importance of incorporating time-related costs for a better understanding of individuals' travel behavior and preferences.

The third topic employs a discrete choice experiment to understand the trade-offs between characteristics of urban green spaces and travel distance. The research uses travel distance, which is then converted into travel costs as a payment vehicle. I find that residents highly value improved air quality, reflected in the highest WTP. However, despite incorporating individual travel time in travel cost estimations, the study reveals no enhancement in model fit, indicating challenges in assessing the value of travel time.

The fourth topic explores the impact of individual spatial data on willingness to pay for urban green space attributes, utilizing a Seemingly Unrelated Regression model. I find that while the study has limited explanatory power, it provides valuable insights into how spatial characteristics influence preferences. Specifically, residents in areas abundant with open and recreational lands demonstrate lower willingness to pay for facility improvements, while those in commercial zones express lower WTP for tree species and ecosystems improvements.

The fifth topic uses a simulation framework to investigate the relationship between attribute changes in Penang Island's urban green spaces and their impact on welfare values. The study focuses on five selected UGS on Penang Island, employing a simulation analysis to assess changes in individual-level consumer surplus. The results reveal that improvements in UGS attributes lead to higher consumer surplus estimates, emphasizing the economic benefits beyond direct travel costs.

These findings contribute insights into resident preferences, spatial influences, and economic implications, providing guidance for local policymakers on sustainable urban development.

Acknowledgements

I would like to express my sincere gratitude to the Waikato Doctoral Scholarship and the Department of Economics for their generous funding, which made this research possible. I extend my deepest gratitude to my supervisors Professor Riccardo Scarpa and Professor Frank Scrimgeour for their support, guidance, and supervision throughout the entire research process. I express my appreciation to Dr. Andrea Pellegrini for guiding the design of the experiment. I am also grateful to Professor Andrew Tan Khee Guan and Dr. Abdul Rais Abdul Latiff for their guidance in conducting surveys in Penang, Malaysia. Lastly, I want to acknowledge the support of my family and friends throughout this journey. Your support has played an important role in my success.

Contents

Chapter 1

Introduction

Urban green spaces refer to publicly accessible green areas such as parks, gardens, playgrounds, and neighbourhood fields in urban areas that are administered by the local government (Tapsuwan et al., 2021). These green spaces provide multiple benefits in socioeconomic, cultural, and environmental aspects, which are critical for urban dwellers' well-being. Specifically, the ecosystem services provided by these spaces can help mitigate the negative effects of rapid urbanisation while improving the quality of life for urban dwellers (Xu et al., 2018). These green spaces provide important ecosystem services such as climate regulation, air filtration, carbon storage, noise regulation, recreation, and habitat maintenance (United Nations et al., 2021). However, the growing pressures from human activities due to urban expansion have threatened the capacity of these green areas. This gradually reduces the provision of ecosystem services by these green areas, deteriorates people's health, and increases climate risks.

Ecosystems in urban areas are often neglected and undervalued because they are not directly consumed by individuals (Tapsuwan et al., 2021; Villamagna et al., 2013). Recognizing the need to quantify the value of these ecosystems, the United Nations' System of Environmental-Economic Accounts (SEEA) plays a crucial role in measuring and valuing changes in urban ecosystems. The development of the SEEA has attracted significant interest from researchers and policymakers. The SEEA, as outlined by United Nations et al. (2014b) and United Nations et al. (2021), collects data on ecosystem assets and the flow of ecosystem services in physical units, subsequently translating this information into monetary values. This framework aims to establish connections between the environment and the economy, offering insights into the adverse effects of environmental degradation on human well-being.

The primary objective of this thesis is to assess the economic value of ecosystem assets and services in urban settings, employing non-market valuation techniques. While the physical presence of ecosystem assets and the tangible flow of ecosystem services can be directly recorded, their transformation into monetary value requires specific non-market valuation techniques. In this study, I propose using the travel cost method (TCM) and discrete choice experiment (DCE) to estimate the economic values associated with urban green spaces. The TCM is chosen because it can capture individuals' actual visit behaviour and associated travel expenses in the market, while the DCE allows for the estimation of changes in welfare resulting from variations in the quantity or quality of urban green spaces.

Given the limited discussion and implementation of the non-market valuation of ecosystem assets and services in urban contexts within most developing countries, including Malaysia, I have chosen Malaysia as my study area to address this research gap. Specifically, this research focuses on the island of Penang in Malaysia (Penang Island). This choice is motivated by the current conditions in the city centre of George Town, where urban green spaces are lacking despite a high population density of 4624 individuals per square kilometre (more details are discussed in Section 1.3) (Penang Institute, 2023; Penang Town and Country Planning Department, 2018). These factors highlight the importance of prioritising and urgently addressing the values of urban green spaces in the urban context of Penang Island.

Following the estimation of the economic value of these ecosystems, the primary objective of this research is to integrate these non-market values as a complementary addition to the existing SEEA framework. The SEEA currently records only exchange values, excluding these significant non-market values. This integration aims to provide a more comprehensive understanding of the economic contributions of ecosystems, going beyond the traditional focus on exchange values (Caparrós et al., 2017).

Furthermore, these elicited non-market values will provide valuable insights to authorities and other agencies regarding the benefits derived from environmental expenditures when compared to the returns on other forms of public expenditure. The use of more comprehensive benefit estimations in ecosystem accounts can enhance the understanding of the relationship between ecosystem assets, services, and human activities within the urban context. The research outcomes could enhance our understanding of these benefits, mitigate the risks of ecosystem deterioration, and strengthen efforts towards environmental sustainability.

The study begins with a focus group discussion and best-worst scaling analysis to identify key attributes of urban green spaces in the context of Penang Island. It will then use the TCM, designed to capture individuals' travel expenses to reach green areas from places of residences. Subsequently, the DCE will be employed to estimate the welfare values of urban green spaces, assuming that no market data is available.

Finally, the study will discuss how these derived values can be simulated, assuming a specific market scenario, to provide insights and inform further analyses.

The following sections discuss the study's background, introduce the study area, outline the research questions, underscore the research's significance, provide a preview of the thesis structure, and address research ethics.

1.1 Background

1.1.1 Research methodology in Economics

Research methodology in economics has long been influenced by various philosophical positions, which play a significant role in shaping economic ideas and guiding economic research. Among the many philosophical positions, including positivism, normativism, and pragmatism (Blaug, 1992), this research aligns with the neopositivist approach.

According to (Fischer, 1998), the neopositivist approach encompasses two main viewpoints. First, it suggests that knowledge can only be acquired through conventional natural or valid scientific methods, primarily through empirical falsification or verification via hypothesis testing. Second, it mandates a strict separation of observational facts from theory. Empirically derived facts are established independently of normative context or consequences.

Suppose that establishing a causal relationship for a phenomenon relies on the consistent observation of a relationship between a potential cause and its effect. However, drawing an absolute or universally valid conclusion remains elusive, even if numerous individual observations seem to support it, because the consistent relationship may fail on occasion. In fact, any universal statement can be refuted by just one singular counterexample (Blaug, 1992; Hume, 1751; Popper, 1959). In essence, a hypothesis or theory is subject to potential falsification through empirical observations. Even if it manages to escape from falsification, it can only be considered provisionally true, never conclusively true (Popper, 1959).

In this research, the chosen research methodologies are deeply influenced by neopositivism, which emphasizes a quantitative approach and statistical analysis (Blaug, 1992). Specifically, this study focuses on quantifying environmental value. To achieve this, data were collected through surveys, and various non-market valuation techniques were employed to derive multiple dimensions of value and corresponding indicators, which serve as the foundation for policy recommendations.

The non-market valuation techniques used in this study include the TCM, DCE, and a combination of both approaches. The objective is to investigate whether the values obtained through these valuation methods align with the theoretical implications that reflect societal welfare changes. Additionally, this research examines whether combining travel costs with the DCE offers a more accurate representation of social welfare changes.

Should the empirical data contradict the implications of the theory, it would falsify the existing theory, prompting the construction of a new theory that can account for these contradictions. Conversely, if the empirical data does not falsify the existing theory, it is considered provisionally valid (Blaug, 1992; Hands, 1993).

In an economic context, unlike many goods and services that can be easily measured in monetary terms using market prices, numerous values or benefits associated with natural resources, such as urban green spaces and forests, cannot be readily quantified in monetary terms. These values are referred to as non-market values. Non-market valuation techniques have been developed and are thus useful tools for assessing the value of environmental resources. These techniques enable authorities to evaluate the costs and benefits of policy changes by incorporating non-market values.

In the pursuit of efficiently allocating resources from the perspective of increasing social welfare, it is crucial not only to consider the costs associated with preserving or conserving environmental resources but also to account for the benefits these resources offer to the public. Local governments and authorities must take into account various factors that influence both the costs and benefits of environmental resources to maximize overall social welfare. However, the current SEEA's ecosystem accounts primarily emphasize exchange values, aligning with the System of National Accounts (SNA). This focus on exchange values implies that the monetary values presented in the accounts may not provide a reliable economic measure of welfare (Caparrós and Schweppe-Kraft, 2020). Economists typically estimate changes in welfare by estimating changes in consumer and producer surpluses (Harberger, 1971). In particular, exchange values, which are based on market prices, exclude consumer surplus but include producer surplus and production costs. Consequently, the emphasis on exchange values suggests that the final estimate of economic flows' value will likely be lower than values obtained using non-market valuation techniques. Therefore, this study undertakes a comparison of values derived from various non-market valuation methods to offer a more precise reflection of social welfare.

1.1.2 Economic valuation of ecosystem services

Ecosystem services involve the benefits produced in economic and other activities, spanning various forms including direct consumption, passive enjoyment, and indirect receipt. These services can be categorized into two key groups: intermediate and final ecosystem services. Within the context of ecosystem services, provisioning services involve the supply of essentials like food, fibre, fuel, and water. Regulating and maintenance services include activities such as climate regulation, air filtration, water purification, and noise attenuation. Lastly, cultural services focus on the experiential and non-material services related to the perceived or realized qualities of ecosystems, facilitating diverse cultural benefits for individuals (United Nations et al., 2021). The framework used to conceptualize these ecosystem services is the SEEA Ecosystem Accounting, which aligns with the Common International Classification of Ecosystem Services (CICES) 5.1 framework (Haines-Young and Potschin, 2018). This framework aligns with international standards, ensuring reliable and comparable assessment of ecosystem services. By adopting this framework, this study aims to provide a comprehensive understanding of ecosystem services and their contributions to urban dwellers' well-being.

Most of the ecosystem services have characteristics of 'public goods', and face the challenge of excluding individuals from receiving these benefits, which are non-rival in nature. Many people can enjoy the same scenic views at the same time without excluding others from enjoying it. Given these characteristics, the private sector faces difficulties in marketing or selling these ecosystem services. Although these services do not have market prices, they contribute utility to individuals, thus having inherent value. This value is quantified as an individual's net willingness to pay (WTP) or consumer surplus (Loomis et al., 2018; Patterson et al., 2013).

Many valuation techniques, whether used directly or indirectly in previous studies, have aimed to estimate individuals' WTP to secure benefits, or willingness to accept (WTA) to forgo environmental goods or services. In a graph featuring traditional supply (marginal cost) and demand (marginal benefit) curves typical of marketed goods or services, the value included in the SNA is the product of market price p and quantity q. The consumer surplus, representing welfare value, refers to the difference between individuals' WTP for an environmental resource and the amount that individuals should pay for it (Roger, 2019), is represented by the area between the demand curve and the market price. The producer surplus or net rent is depicted by the area between the market price and the supply curve. The total economic value of a good or service is the sum of producer and consumer surplus. Regarding the supply curve, if economic system actions cannot influence the quantity of supply of ecosystem services, their supply curves tend to be more vertical (Costanza et al., 1997).

1.1.3 Economic valuation techniques

Environmental resource valuation techniques can be categorized into two groups based on the types of data employed: revealed preference (RP) and stated preference (SP) techniques.

Stated preference techniques, involving questionnaire-based approaches, aim to uncover individuals' preferences when market-derived WTP information is not readily available. These techniques elicit individuals' monetary valuations of costs and benefits, referred to as WTP or WTA. To elicit WTP, individuals are asked questions about the amount they are willing to pay or whether they would pay a specified amount for environmental benefits, with the intention of simulating individuals' behaviour in the market (Bateman et al., 2002). Two main stated preference techniques are the contingent valuation method (CVM) and the choice experiments (CE). The CVM was developed by Mitchell and Carson (1989), which elicits values that respondents are willing to pay or accept for the changes in the quantity or quality of a non-market good or service. Conversely, the CE requires respondents to make choices from a range of alternatives comprising non-market goods, each defined by specific attributes, accompanied by a hypothetical price (Hanley, 1989; Adamowicz et al., 1989; Twerefou and Ababio, 2012).

Revealed preference techniques use information from markets associated with the goods or services under evaluation and are applicable when the relevant WTP information can be collected from individuals' real-life choices (Bateman et al., 2002). Two main revealed preference techniques are the hedonic pricing method (HPM) and the TCM. The HPM assumes that the market value of property or labour wages are linked to the benefits such as the environmental facilities derived from it. This method employs two techniques: the property value approach and the wage differential approach. Based on this assumption, the value of the environmental asset/good can be derived from the property or labour market. The TCM is commonly employed to estimate the use values of visiting recreational sites. This method assumes that the travel cost that individuals incur to visit a site serves as the price for accessing that site. Therefore, the WTP can be elicited based on the number of trips that they undertake (Twerefou and Ababio, 2012).

In this study, two techniques were used to elicit individuals' WTP for urban green space and its characteristics, including the TCM and the discrete choice method.

Travel cost method

The TCM is used to estimate the costs incurred by individuals when travelling to a recreational site to use environmental resources. It is assumed to represent the value individuals place on the recreational site (Faccioli et al., 2016). This method takes into account both monetary transactions already recorded in national accounts and non-monetary transactions that are not included in those accounts.

The fundamental concept behind this method is that the number of visits to a site can be explained by information on individual travel costs, such as distance and time spent traveling. This information is then used to derive a demand curve and estimate the recreational use value of the site (Gillespie, 1997). This method also relies on the premise that individual preferences are reflected in their actual visit behaviours and the expenses they incur to access a recreational site or enjoy environmental goods and services (Ward and Loomis, 1986; Phaneuf and Smith, 2005; Parsons, 2003; Adamowicz et al., 1994; Kerry and Desvousges, 1986). Based on this notion, this method integrates the concept of weak complementarity. Specifically, it models recreational trips as the private good and an environmental attribute (e.g., quality of the site) as a weak complement. In this framework, the travel cost, representing the full price of reaching the site, is not systematically related to the level of the environmental attribute. By operationalizing the model through observed trips and estimating the demand equation considering travel costs and other factors, the TCM allows for the assessment of individuals' willingness to pay for changes in the environmental attribute, incorporating both the intrinsic value and the role of weak complementarity in shaping preferences (Phaneuf and Requate, 2017).

There are various types of travel cost models. The zonal travel cost model is employed to assess how changes in the proportion of visitors from the same area are influenced by distance and travel costs (Gillespie, 1997). The individual travel cost model, on the other hand, illustrates how the number of visits an individual makes to a site is influenced by travel costs, as well as other factors like socio-demographic characteristics and site characteristics (Champ et al., 2003). The random utility travel cost model relies on random utility theory, in which the probability of choosing one option over another is determined by the utility associated with each choice. This utility can be divided into explained and unexplained components (Champ et al., 2003), making the model particularly relevant when environmental quality plays a role in attracting visitors.

Discrete choice experiment

The Discrete Choice Experiment (DCE) is widely used to estimate both use and non-use values of environmental resources. Its advantage over the Travel Cost Model (TCM) is its capability to include non-use values, especially for ecosystems valued for cultural, aesthetic, or intrinsic reasons. Using hypothetical scenarios, DCE captures individuals' preferences for diverse attributes and the estimation of related values (Johnston et al., 2015), offering versatility in assessing ecosystems beyond direct use.

In ecosystem valuation studies, DCE provides multiple advantages by enabling the calculation of the value associated with each ecosystem attribute and elucidating trade-offs among them. The derived values hold significance for ecosystem management and carry significant policy implications (Rakotonarivo et al., 2016).

1.2 Research questions

This thesis has five topics to be discussed. The research questions are summarized as follows:

Topic 1: Analysis of Best-Worst Scaling Data

This study employs the best-worst scaling method to explore urban residents' preferences for urban green space characteristics in Penang Island, Malaysia. The main research question is: How can the utilization of best-worst scaling studies contribute to the exploration of individuals' preferences for urban green space characteristics? This research question is segmented into two sub-questions:

1. What are the most and least important attributes influencing individuals' decision to visit a green space?
2. How consistent and reliable are preference representations for urban green space characteristics across different designs of best-worst scaling question sets when new attributes are included?

The first research question explores the most and least important attributes that individuals perceive through responses to best-worst scaling questions. An attribute that is perceived as important affects individuals' probability of visit to green sites (everything else equal).

The second research question assesses the consistency and reliability of individuals' preferences for urban green space attributes. It investigates whether individuals' preferences for a particular attribute are influenced by the inclusion of additional attributes are to be considered by respondents.

Topic 2: Analysis of Travel Cost Data

Following the best-worst scaling analysis, this study further investigates individuals' visit preferences and behaviours related to two actual urban green spaces within the same study area. It tests the model using various travel costs derived from different methods and includes variables concerning preferences for ecosystem assets and services, obtained in the best-worst scaling study. Three research questions have been identified:

1. Do travel costs to an urban green site contribute to explaining an individual's visit to the site?

2. Do the preferences for ecosystem assets and services at green sites affect visitors' preferences?

3. How reliable are the estimates of travel and recreation time values derived from travel cost models?

The first research question is raised to explore the relationship between travel costs and individuals' visit patterns for visiting urban green sites, given that the WTP for travel expenses is expected to be significant in determining demand of visits. Travel cost models were developed to examine the validity of travel costs in a real situation.

The second research question explores the connections between individuals' perceptions of ecosystem assets and services and visit preferences through travel cost models. The individual difference scores for perceptions of ecosystem assets and services estimated in Chapter 4 were used as independent variables the value of which is to be estimated in travel cost models.

The third research question examines the reliability of time values of travel time and recreational time when they are added to the travel costs. The individual-specific values of time were estimated in travel cost models to investigate their validity in real and hypothetical situations.

Topic 3:Analysis of Time Value Effects in DCE

This research focuses on the decision-making process involved in selecting a recreational site in the context of ecosystem services valuation. It explores the trade-offs between site characteristics and travel distance made by individuals to understand residents' WTP for ecosystem assets and services at the urban green spaces. Two research questions have been identified:

1. Can a distance-based discrete choice model be employed as a valuable tool to comprehend individual sensitivities towards travel costs and preference heterogeneity for urban ecosystem services on the Island of Penang, Malaysia?

2. Should the time spent travelling to a recreational site be perceived as a cost of visitation?

The first research question is raised concerning the validity of distance-based discrete choice models in investigating the preference of individuals for site characteristics. The second research question is motivated by a number of published studies that explore the validity of travel time in travel cost models.

Topic 4: Assessing Spatial Heterogeneity in WTP for Urban Ecosystem Services

Following the DCE analysis, this study investigates the relationship between individual spatial data and the WTP for urban green space attributes. In this context, two research questions have been identified:

1. How is the spatial distribution of individual-level mWTP for a specific attribute level across Penang Island characterized?

2. How do land use categories, such as forests, open and recreational lands, commercial areas, agricultural lands, and industrial areas in individual neighbourhoods of residences, impact their mWTP for attributes associated with urban green spaces?

The first research question aims to visually represent the distribution of individual-level mWTP on a map, enabling an exploration of areas with relatively higher or lower mWTP values. The second research question seeks to investigate the relationship between land use patterns within a 2-kilometre radius (in the neighbourhood) of the respondent's residence and their individual mWTP for attributes related to hypothetical urban green sites. To examine the validity of the effects of neighbourhood land use patterns on individual mWTP, a Seemingly Unrelated Regression (SUR) model was developed.

Topic 5: Analyzing the Impact of Improved Urban Green Space Attributes: Simulated Changes in Consumer Surplus

This study employs a simulation analysis to estimate the welfare changes resulting from various scenarios or policy interventions related to urban green spaces on Penang Island. Within this context, a key research question arises:

1. How do changes in the provision of ecosystem assets and services across different UGS impact welfare values under different scenarios?

This research question provides a basic understanding of the effects of changes in the flow of ecosystem services and the stock of ecosystem assets. This lays the foundation for understanding the relationship between the quality and extent of ecosystem services and assets and the welfare values attributed to these spaces.

1.3 Study area

The study area is the Island of Penang in the State of Penang in Malaysia. The population in Penang State reached 1.738 million in 2022, with a land area of 1031 km^2. Penang State has the highest population density in Malaysia and has a high urbanization rate, the average population density in the year 2022 was 1686 km^2 (Penang Institute, 2023). The Penang State is made up of five districts, two of which (Northeast Penang Island and Southwest Penang Island) belong to the Island which is located on the northwest coast of Peninsular Malaysia, while the other three (North Seberang Perai, Central Seberang Perai, and South Seberang Perai) are located on the Penang mainland.

The study area includes both Northeast Penang Island and Southwest Penang Island. The land areas of these two districts are 119.2 km^2 and 173.5 km^2 respectively, accounting for a total of 28% of the State's total land area. As of 2022, the total population on the island had reached 790,200, while the population on Northeast Penang Island and Southwest Penang Island had reached 551.1 thousand and 239.1 thousand, accounting for approximately 31.7% and 13.7% of the total population of Penang State respectively (Department of Statistics Malaysia, 2023; Penang Institute, 2023). The population density at Northeast Penang Island had a peak of 4624 per km^2, while Southwest Penang Island had a population density of 1378 per km^2, showing an uneven distribution across the land of the state (Penang Institute, 2023). Because of this high density of population, and the presence of a diverse range of landscapes, including mountains, forests, agricultural lands, heathlands, wetlands, etc., Penang Island is an ideal area to study the recreation preferences of urban residents. Moreover, the state's environment has experienced significant changes in recent decades, due to the rapid growth of urbanization in Penang Island. As a consequence, the future development of initiatives directed to nature conservation will require guidance about the preferences of the local population, which has not been investigated to date.

1.3.1 Background of Penang Island's Urban Green Spaces

In the context of Penang Island, Malaysia, a pressing issue has arisen regarding the insufficient availability of green spaces, despite the island's rich biodiversity and ecosystems. As of 2022, the island comprised a total area of 29,300 hectares, with approximately 14,009 hectares of forests (48%), 4,326 hectares of agricultural lands, and 530 hectares of open and recreational lands (Peninsular Malaysia Town and Country Planning Department, 2023). In terms of overall green space, the island still possesses a relatively abundant supply. However, due to the absence of proper infrastructure and amenities, many of these forested areas are either inaccessible or difficult to reach. This presents significant challenges for potential visitors, restricting urban residents to just 530.42 hectares of accessible open and recreational land for their recreational activities.

The deficiency of green spaces became evident when examining the Penang Green and Sustainable Index. The indicator related to green and water spaces, which assesses the proportion of green and water spaces relative to the total land area in Penang state, recorded a mere 0% in the 2009 final report (Penang Green Council, 2020a). Figures on the open space and recreational land area for the two districts in Penang Island from 2017 to 2022 are presented in Figure 1. This figure illustrates a significant reduction in the land area classified as open and recreational between 2017 and 2018, followed by a slight increase between 2020 and 2022. Although there was a modest improvement, with the proportion rising to 1.8% in 2022, the overall performance remained unsatisfactory (Peninsular Malaysia Town and Country Planning Department, 2023). A visual representation of Penang Island's land use distribution in 2022 is shown in Figure 2. The light green regions on the map represent open and recreational lands.

There are three major urban green spaces located in Northeast Penang Island, namely Penang Botanic Gardens (latitude: 5.43786; longitude: 100.29101, with a land area of approximately 2.41 square kilometres), Penang City Park (latitude: 5.43189; longitude: 100.29732, covering approximately 0.16 square kilometres), and Bukit Dumbar Park (latitude: 5.38295; longitude: 100.31249, spanning about 0.07 square kilometres). These parks are situated within highly urbanized areas and experience high visitation rates. They offer a wide range of urban ecosystem services, including climate regulation, water regulation, air filtration, noise regulation, habitat maintenance, visual amenities, and recreational opportunities. Regarding site characteristics, these parks possess a variety of facilities, including playgrounds, athletic fields, sports equipment, picnic shelters, and trails. Additionally, they support a rich diversity of plant and animal species, contributing to the generation of ecosystem services.

The limited availability of green spaces on Penang Island can be attributed to several factors, including land use pressure, urban sprawl, and cost considerations. Existing urban trees have been cut down for activities such as road widening, and urban green spaces have been converted into more financially lucrative land uses. Furthermore, government departments often fail to fully recognize the importance of urban ecosystems and lack the necessary knowledge to effectively manage existing urban green spaces. This has led to a prioritization of maintenance costs over the benefits of enhancing urban ecosystems during green space design (Penang Green Council, 2020b). Additionally, there is a lack of understanding and appreciation among the public regarding Penang's urban ecosystems and biodiversity. Part of this issue is caused by a lack of awareness of the monetary and cultural values that urban ecosystems provide. The absence of comprehensive information contributes to the public's indifference towards these ecosystems (Penang Green Council, 2020b).

By 2030, it is expected that the urban population in Penang will constitute over 95% of the total population and nearly 50% of Penang's land area will be transformed into built-up areas (Penang Town and Country Planning Department, 2018). Consequently, there is an urgent need to address issues related to urban ecosystems, particularly the availability of urban green spaces.

Both the federal government of Malaysia and the state government of Penang recognize the significance of urban green spaces and have taken proactive measures to address the problems of reducing green spaces. The Department of Landscaping within the state government has made substantial efforts to promote and protect urban ecosystems and biodiversity within public urban green spaces. Furthermore, the Federal Government of Malaysia introduced a National Policy on Biological Diversity (2016-2025) in 2015, offering a comprehensive framework for targets, actions, and policy implementation regarding biodiversity protection. One of the focal areas of this framework is urban biodiversity in Penang (Penang Green Council, 2020b).

In 2017, the State Government of Penang launched the Penang Green Agenda (PGA) 2030 to chart a sustainable development path for the state. The PGA 2030 Land Use Planning Report underscores the importance of resilience-focused land use planning for sustainable development, aligning with the UN Sustainable Development Goals (SDGs). To achieve its goal of becoming a Climate-Proof and Biodiversity-Sensitive state by 2025, the report recommends optimizing the utilization of urban green spaces (Penang Green Council, 2020d). Additionally, the Penang Structure Plan 2030 mandates that new development plans include green spaces amounting to at least 10% of the total development areas, while the National Landscape Policy aims to transform at least 30% of urban development areas into green spaces (Department of Landscape Malaysia, 2019; Peninsular Malaysia Town and Country Planning Department, 2018; Penang Green Council, 2020c).

Moreover, the Penang City Council is working towards establishing 10 acres of interconnected green spaces by linking existing public green spaces in the urban area of Penang, specifically George Town. This proposal is anticipated to deliver significant benefits to urban residents in George Town, as studies indicate that each acre of green space, national park, or botanical garden contributes to an annual reduction of 2600 kg of carbon dioxide (City Council of Penang Island, 2022).

The government of Penang allocates a substantial budget for environmental protection annually, as indicated in the Penang Economic and Development Report 2019/2020 (Penang Institute, 2020). According to the Annual Economic Survey conducted by the Department of Statistics, Malaysia, in 2022, the total environmental protection expenditure in Penang for the year 2021 amounted to nearly RM199 million (approximately 72 million NZD). Despite these significant investments in environmental protection, such as the establishment of public green spaces, the design of these spaces is heavily influenced by cost considerations. To minimize expenses, government agencies like the Department of Landscaping often prioritize single-species tree planting and artificial landscapes. Moreover, there is a lack of knowledge regarding planting designs that enhance urban ecosystems, largely due to the limited communication and formal relationships between government agencies and professionals. However, well-designed green spaces have the potential to enhance urban ecosystems and biodiversity while simultaneously reducing expenditure costs (Penang Green Council, 2020b).

1.4 Significance of research

The significance of this research includes the exploration of the non-market valuation of urban green spaces through different methods, which has implications for urban planning, policy development, and the preservation of green spaces in an urban context.

The non-market valuation of urban green spaces provides stakeholders with critical information, including the economic value of green spaces and their attributes, enabling informed decisions regarding land use, urban development, and green space preservation. The introduction aligns the thesis with the SEEA framework, aiming to integrate non-market values as a complementary addition. This integration highlights the tension between traditional economic welfare measurements and the SEEA's methodological preference for exchange value, emphasizing the importance of capturing the full range of benefits provided by urban green spaces for comprehensive environmental-economic accounting. Government agencies, particularly the Department of Statistics Malaysia, involved in the development of the Malaysian SEEA

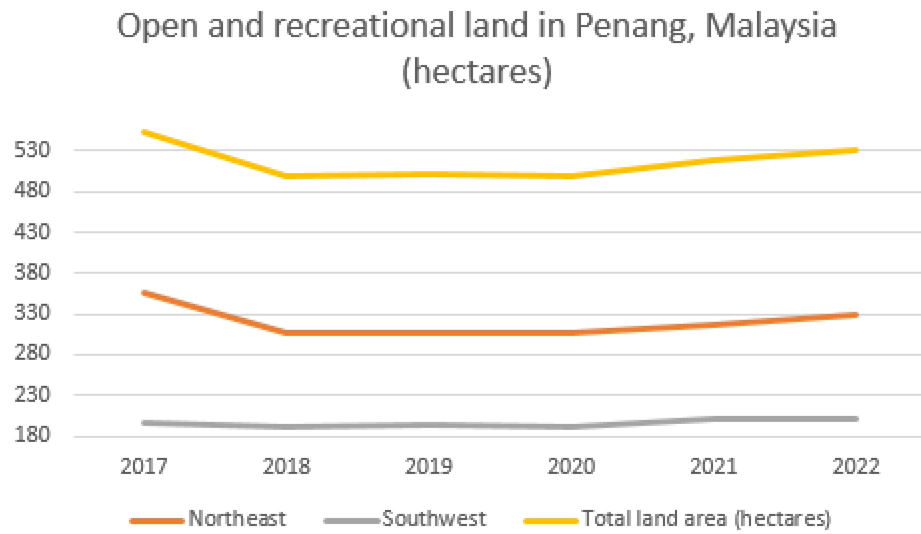


Figure 1: Statistics of open and recreational land area in Penang Island (Peninsular Malaysia Town and Country Planning Department, 2023)

framework, can use this data to enhance ecosystem accounts for policy formulation.

Understanding the economic benefits of urban green spaces offers stakeholders, such as park managers, owners, and investors, reasonable justifications for allocating limited resources to the development and maintenance of these spaces. This research aims to identify and recognize the preferred characteristics of these green spaces, guiding site management and financial planning. Additionally, it informs public investors about the most beneficial investment alternatives to maximize social benefits.

By facilitating better-informed decisions by government agencies and internal stakeholders, this study contributes to improving urban green space attributes and enhancing the provision of ecosystem services. As a result, it improves the quality of life for external stakeholders, including visitors and local communities, while also advancing urban sustainability through pollution reduction and the enhancement of overall ecosystem services.

1.5 Thesis outline

This chapter provides an overview of the background, study area, research questions, and contributions of the study. Chapter 2 introduces the research methods used in the study. Chapter 3 introduces the survey design and data collection process.

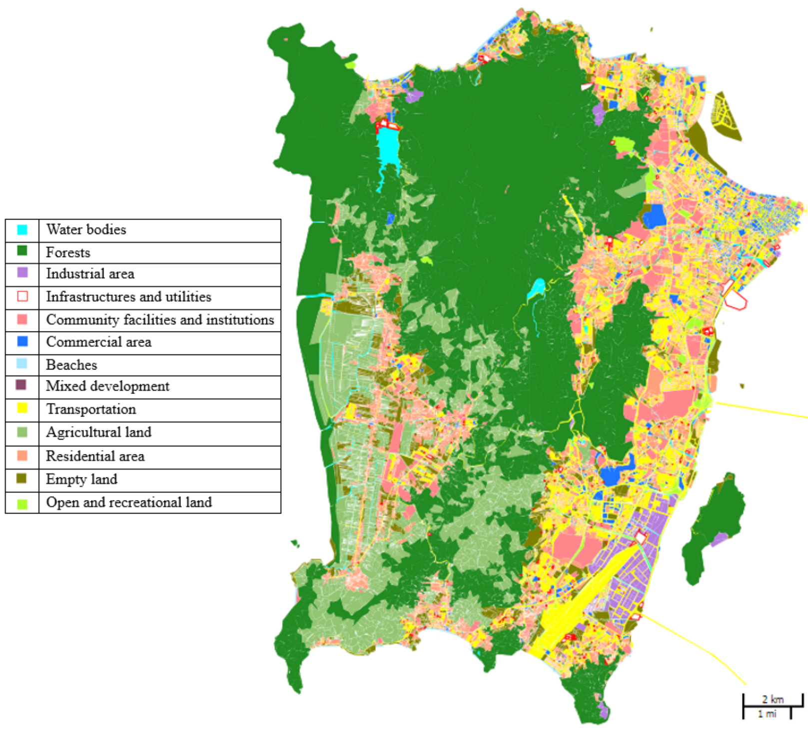


Figure 2: Land use map of Penang Island (Peninsular Malaysia Town and Country Planning Department, 2023)

Chapters 4 to 8 cover the five topics included (refer to Section 1.2). Finally, Chapter 9 offers concluding remarks on the research.

1.6 Research ethics

During the course of this study, primary data were collected from individuals through an online survey. To ensure ethical research conduct, applications for preliminary and final ethics approvals were submitted to the WMS Ethics Committee at the University of Waikato. A complete copy of the preliminary ethical approval application and ethical consent documentation is available upon request. Furthermore, this study adheres to the University of Waikato's regulations and guidelines on Human Research Ethics Regulations 2008.

In the survey design, proactive measures were taken to safeguard the rights and well-being of participants. Participation in the survey was entirely voluntary. Before participating, individuals were fully informed about the researchers' identities and the research's objectives. Participants were under no obligation to take part and retained the freedom to withdraw from the study at any point.

All information collected during the survey is held in strict confidence and anonymity. To formalize this commitment, participants were provided with a written consent form that clearly outlined the nature of the research. This consent form had to be signed by participants before they could start the survey.

Lastly, this study has no conflicts of interest that could compromise the research objectives or data treatment.

1.7 Conclusion

In conclusion, this chapter serves as an introduction to the study of non-market valuation of urban ecosystem services in the context of Urban Penang Island in Malaysia. The subsequent chapter will provide an overview of the research methodologies employed in this study.

Chapter 2

Research Methods

2.1 Introduction

This chapter reveals the research methods and techniques employed to address the research objectives, providing insight into the different methods employed to derive non-market values of urban green spaces from survey data.

Before designing the questionnaire, I engaged with residents in the study area through focus group discussions. Section 2.2 outlines the methodology employed for conducting these focus group discussions. The chapter then commences with an exploration of Best-Worst Scaling (BWS) in Section 2.3, while Chapter 4 elaborates its application within the study's context. Section 2.4 introduces the steps involved in implementing the Travel Cost Method (TCM), and Chapter 5 describes its application to the major urban green spaces within the study area. Additionally, Section 2.5 explains the methods used for analysing the choice data obtained from the Discrete Choice Experiment (DCE). The case study of DCE is discussed in Chapter 6. Furthermore, this study includes a spatial analysis of individual preferences, and the Seemingly Unrelated Regression Model (SUR) serves as a statistical framework for investigating the relationships among variables, including spatial information and individual preferences. The application of SUR is discussed in Chapter 7.

This chapter provides a structured discussion of the research methods employed throughout this study, providing a foundation for the presentation and the data analysis, the results of which are detailed in the subsequent chapters.

2.2 Focus group discussion to explore key urban green space attributes

In the course of this study, primary data were collected from individuals through an online survey. The objective was to investigate individuals' perceptions of urban green spaces and identify key urban green space attributes. To achieve this, I initially conducted a focus group study followed by an online follow-up survey.

The focus group study involved three groups, each consisting of ten participants, conducted in December 2021. Participants were recruited from NGOs, government-linked organizations, and park user groups. However, this selection may not adequately represent the intended target population as it excludes broader public perspectives. The purpose of these focus groups was to explore urban green space characteristics not identified by me and to gain an in-depth understanding of participants' prioritization of urban green space attributes.

During the focus group interviews, participants were asked to share their expectations of urban green spaces and identify important characteristics specific to George Town. These discussions encompassed several key points, including participants' experiences visiting urban green spaces in George Town, their preferences for certain urban green space characteristics, evaluations of different attributes, WTP for improvements, and considerations regarding the cost and time associated with travel. The discussions were structured around these topics but allowed participants to raise additional points as well. Participants were provided with paper for note-taking.

Participants were recruited through various channels, including non-governmental organizations, universities, social media, and email invitations. Although I aimed for a balanced ethnic composition reflective of Penang Island's inhabitants (comprising Malay, Chinese, and Indian ethnicities), I encountered a lower participation rate from Malay and Indian individuals, resulting in an imbalanced representation in the focus groups. Figure 46 illustrates the ethnic composition of the focus groups.

Regarding gender composition, Groups 1 and 2 maintained an equal male-to-female ratio, while Group 3 had a male majority. Additionally, participants in Groups 2 and 3 generally held at least a bachelor's degree, while those in Group 3 had a lower average education level.

The age composition was diverse, with two focus groups being age-homogeneous and one being age-heterogeneous. This diversity allowed me to explore preferences across different age groups, recognizing that access to urban green spaces can vary by age. Younger individuals typically have greater access to urban green spaces and are more comfortable speaking up during focus group discussions. Conversely, older individuals often have less access due to transportation limitations or health issues.

Groups 1 and 2 represent specific age groups, while Group 3 includes participants spanning a wider age range.

The follow-up survey comprised four sections: (1) Socio-demographic questions; (2) Visit patterns; (3) Perceptions of urban green space characteristics; and (4) Perceptions of urban green space ecosystem services. Part (1) collected basic demographic information, including gender, age, education level, employment status, and residency status. Part (2) focused on respondents' frequency and days of the week of visits to urban green spaces. The most critical sections were parts (3) and (4). In part (3), respondents ranked the importance of various urban green space characteristics based on a provided list. In part (4), respondents selected the five most important ecosystem services provided by urban green spaces, rated their importance using a 5-point scale, and indicated the reasons for their selections.

In response to the initial focus group discussion dynamics, where some participants dominated the conversation, a follow-up survey was introduced to collect anonymous views from every participant. An invitation link was sent to participants in Group 1, and follow-up phone calls were made to ensure completion. For three participants unfamiliar with online surveys, data were collected through Computer-Assisted Personal Interviews (CAPI) during face-to-face meetings. Participants in Groups 2 and 3 were requested to complete the online survey before the group discussions concluded, using a QR code provided by the moderator. They were given 10 minutes to complete the survey.

2.2.1 Methods

The focus group study employed a qualitative research design to explore perceptions and attitudes toward the characteristics and ecosystem services of urban green spaces among participants in three focus groups. The data gathered from these sessions underwent analysis through a combination of techniques, including word cloud analysis, text network analysis, and thematic analysis.

During the focus group discussions, which were conducted with participants' consent, both audio recordings and notes taken by the moderator were utilized. Subsequently, the audio recordings were transcribed and subjected to analysis. A word cloud analysis was applied to identify the words and themes most frequently mentioned in the conversations. This text-mining technique generates a visual representation where word size corresponds to frequency, providing insight into the significance of specific terms. Word cloud analysis is a widely accepted approach in social science studies for the qualitative analysis and visualization of data (Cameron et al., 2012; Lacey and Luff, 2009; McLeod, 2012).

The word cloud analysis was conducted using the statistical software R. Processing steps included rearranging the transcribed audio recordings, which involved removing stop words, punctuation, numbers, and unrelated terms, as well as applying text stemming. To be included in the word cloud, a word needed to appear at least four times in the discussions, with a maximum of 200 words featured in the visualization. Moreover, a specific analysis was performed to find words strongly associated with the term *trees*, only including words with a correlation coefficient of at least 0.3. The processed data were then employed to generate the word clouds using the R function *wordcloud*. The use of word cloud analysis for qualitative data is consistent with previous research across various fields (Barnett et al., 2016; Ecker et al., 2015; Wilkinson et al., 2015).

Text network analysis was another tool used in this study to explore relationships among important words in the focus group discussions. This method, also performed in R, enabled visualization of textual relationships. Similar to the word cloud analysis, various data processing techniques were applied to rearrange the focus group data within R. The resulting text network was visualized using the *textplot_network* function, incorporating only the top 30 most frequently mentioned words.

Thematic analysis was employed to identify recurring themes and patterns within the data. Keywords were identified and categorized based on the themes discussed in the focus groups. This approach facilitated the organization of a substantial amount of data into meaningful categories. Given that the primary focus of the groups was to identify essential attributes of urban green spaces for discrete choice questions, the analysis specifically honed in on data relevant to this objective (Braun and Clarke, 2006; Clarke and Braun, 2017; Fereday and Muir-Cochrane, 2006; Guest et al., 2012; Macintyre et al., 2019).

Lastly, quantitative research methods were employed to analyze data collected from the online follow-up survey. This analysis aimed to assess the importance levels and rankings of urban green space characteristics and their associated ecosystem services. Mean rankings and importance scores were computed to facilitate further analysis.

The results of this focus group study are discussed in Appendix B.

2.3 Best-worst scaling analysis

Best–Worst Scaling (BWS) is a method where participants select the best and worst options from a given set, revealing valuable insights into their preferences. Three BWS cases have been identified in previous research. Case 1, introduced by Finn and Louviere (1992), involves individuals directly choosing the best and worst options from a set of objects or attributes. In Case 2, individuals evaluate a set of attributes with varying levels, choosing the best and worst attribute levels within that set. Case 3 presents individuals with various choice sets dictated by an underlying design, where they choose the best and worst profiles from these sets (Marley and Pihlens, 2012). Aligned with the research objective, this study uses Case 1 of BWS to examine preferences regarding UGS characteristics.

After conducting focus group discussions, 13 attributes were identified as important by the participants and aligned with my research objectives. These attributes were subsequently included in the Best-Worst Scaling (BWS) questionnaire. The BWS data collection process consisted of two phases. In the first phase, all 13 attributes from the focus groups were included, treating it as a pilot survey. In the second phase, only 6 attributes were included.

In each BWS question, respondents were presented with 4 attributes (Phase 1) or 3 attributes (Phase 2) selected from the list of 13 (Phase 1) or 6 (Phase 2) attributes. They were then asked to identify the most and least important attributes influencing their decision to visit urban green sites in each question.

2.3.1 Population-level Normalized Difference Scores

The normalized difference scores for each attribute are calculated based on the frequency of selection as the most important (N_i^{best}), least important (N_i^{worst}), and the total number of appearances in the entire question set (N_i^{total}). It is calculated as

$$NDiff_i = \frac{N_i^{best} - N_i^{worst}}{N_i^{total}} \,, \tag{2.1}$$

where i is a particular attribute i.When the attribute i is selected as the best (most important) attribute, the value is 1; when it is selected as the worst (least important) attribute, the value is -1. Therefore, according to Equation 2.1, the difference scores are normalized and are bounded between 1 and -1 (Marley et al., 2016).

2.3.2 Multinomial Logistic Regression

In addition to that, BWS data can be analyzed by random utility theory. Random utility theory relies on the assumption that a person's relative preference for attribute i over attribute j is a function of the relative number of times attribute i preferred to attribute j. Therefore, the theory allows individuals to make stochastic decisions with errors (Louviere et al., 2013). Random utility theory was proposed by Thurstone in 1927 to inspire the development of the pairwise comparisons method to measure perceived preferences, in which individuals select the best item from two items. This method allows for errors to be made by individuals during the decision-making process and enables the model parameters to be produced (Thurstone, 1927).

This random utility theory model was generalized by McFadden in 1974, the generalization process enables the model to be estimated when the choices from the sets are more than two items. In the case that item i is chosen as the best item, McFadden considered that the model parameter, which is also called the scale value (S_i) of item i, consists of explainable observed utility value (V_i) and unexplainable random errors (ε_i). Random errors imply that a person's choice cannot be exactly predicted but can be explained up to a probability (McFadden, 1974). The choice probability of item i over item j is expressed as

$$P_i = P[(\mathbf{a_i x_i} + \varepsilon_i) > (\mathbf{a_j x_j} + \varepsilon_j)], \tag{2.2}$$

where $V_i = \mathbf{a_i x_i}$, $\mathbf{a_i}$ is a vector of coefficient estimation and $\mathbf{x_i}$ is a vector of variables, and the probability of choice for item i with closed-form solution is defined as

$$P_i = \frac{\exp(\mathbf{a_i x_i})}{\sum_{i=1}^{n} \exp(\mathbf{a_j x_j})}. \tag{2.3}$$

Although the best-worst scaling data can be estimated by multinomial logistic regression (MNL), it does not need the logit regressions modelling to calculate the utility coefficients (Lipovetsky and Conklin, 2014). The probability P_i of any binary event can be calculated as

$$P_i = \frac{N_i^{total} - N_i^{worst} + N_i^{best}}{2N_i^{total}}\ , \tag{2.4}$$

where N_i^{total} is the number of times item i is shown across the question sets, N_i^{worst} is the number of times item i is selected as the worst item, and N_i^{best} is the number of times item i is selected as the best item. The P_i obtained from equation 2.4 equals the P_i obtained from the pairwise logit regression (see Lipovetsky and Conklin (2014) for explanation):

$$P_i = \frac{\exp{(\mathbf{a_i x_i})}}{1 + \exp{(\mathbf{a_i x_i})}}\,, i = 1, 2, ..., n. \tag{2.5}$$

The value P_i can be used to calculate utility coefficient β_i by simply estimate the logarithm of odds of P_i:

$$\beta_i = \ln\left(\frac{P_i}{1 - P_i}\right)\ . \tag{2.6}$$

The coefficient β_i can be taken for a_i in equation 2.2. Therefore, the choice probabilities for item i can be represented as

$$P_i = \frac{\exp(\mathbf{a_i x_i})}{\sum_{j=1}^{n} \exp(\mathbf{a_j x_j})} = \frac{\exp(\mathbf{x_i} \ln \frac{P_i}{1-P_i})}{\sum_{j=1}^{n} \exp(\mathbf{x_j} \ln \frac{P_j}{1-P_j})} = \frac{(\frac{P_i}{1-P_i})^{\mathbf{x_i}}}{\sum_{j=1}^{n} (\frac{P_j}{1-P_j})^{\mathbf{x_j}}}. \tag{2.7}$$

The standard error can be calculated by taking the differential, which is represented as

$$d\beta_i = d\left(\ln \frac{P_i}{1-P_i}\right) . \tag{2.8}$$

The proportion's standard error is represented as

$$\Delta P_i = \sqrt{\frac{P_i(1-P_i)}{2N_i}}. \tag{2.9}$$

The parameter's standard error in the logit model is represented as

$$\Delta\beta_i = \frac{\Delta P_i}{P_i(1-P_i)} = \frac{1}{\sqrt{2N_i P_i(1-P_i)}}. \tag{2.10}$$

This approach is recommended by (Lipovetsky and Conklin, 2014), where the utility coefficients, standard errors, and choice probabilities can be derived.

The MNL model was estimated using the R statistical software to derive the utility estimations that respondents attach to different attributes of urban green spaces. An R package 'bwsTools,' introduced by White II (2021), was employed. It enables the estimation of utility coefficients and standard errors from the MNL model described in the previous section. The required data include N_{ni}^{total}, N_{ni}^{worst}, and N_{ni}^{best} to derive utility coefficients, standard errors, 95% confidence intervals, and choice probabilities.

A positive and statistically significant coefficient in the MNL model indicates that the attribute contributes to the overall utility and is preferred by respondents. Attributes with higher utility and stronger preference typically have larger coefficients. MNL estimation allows for a direct comparison of the relative utility of different attributes.

2.3.3 Individual-level normalized difference scores

The individual-level scoring has been used for different purposes (White II, 2021). For instance, Lee et al. (2019) used the scores to predict psychological values, Erdem and

Rigby (2013) estimated individual perceptions of risk, and de Magistris et al. (2014) used the scores to study individual preferences for wine. In this study, the individual scores were used to predict respondents' perceptions of ecosystem assets and ecosystem services by examining the importance level of each attribute which was perceived by respondents. The estimated results were used as indicators of independent variables which were then used in the Poisson regression models explaining number of visits to urban green sites in the travel costs analysis detailed in Chapter 5.

Similar to the population normalized difference scores, the individual-level normalized difference scores were derived from the number of times item i was selected as the most important attribute by individual n (N_{ni}^{best}), the number of times attribute i was selected as the least important by individual n (N_{ni}^{worst}), and the number of times attributes i was presented in the question set for individual n (N_{ni}^{total}). It is represented as

$$NDiff_{ni} = \frac{N_{ni}^{best} - N_{ni}^{worst}}{N_{ni}^{total}} \; . \tag{2.11}$$

The difference scores fall between the values 1 and -1 due to the normalization when the difference between the best count and worst count is divided by the total times the attribute appeared for every individual.

2.4 Travel cost method analysis

2.4.1 Poisson regression models

In the analysis of travel cost data, Poisson regression models were employed, as the dependent variable represents the number of visits to a site in a year, which is discrete, and its frequency can be accommodated by a Poisson distribution (Smith, 1988; Soe Zin et al., 2019). The Poisson regression model was first introduced by Shaw (1988) and has been used in numerous travel cost demand analyses (Hellerstein and Mendelsohn, 1993). This model accounts for the non-negative nature of the observed number of visits and the small number of visit counts empirically observed (Hellerstein and Mendelsohn, 1993; Chae et al., 2012; Englin et al., 2003).

Suppose the random variable denoting the number of recreational visits to urban parks by respondent n is Y_n, I intend to describe the distribution of travel visits as distributed Poisson with parameter $\lambda > 0$,

$$Pr(Y_n = y_n) = \frac{\lambda^y}{y_n!} e^{-\lambda}, \tag{2.12}$$

and the parameter λ is also the mean and variance of random variable Y_n, therefore,

$$E(Y_n) = Var(Y) = \lambda. \tag{2.13}$$

As the expected value of the random variable is non-negative, the model specification has a log-linear form, therefore, it can be expressed as

$$E(Y_n) = \exp(\mathbf{X_n}\beta), \tag{2.14}$$

where n=1,...,N represents the individuals who are included in the data, $E(Y_n)$ represents the expected number of visits, $\mathbf{X_n}$ represents a 1 by $\mathbf{h}$ vector of independent variables and β is an $\mathbf{h}$ by 1 vector of parameters of independent variables. The exponential criterion is required in the model to limit the $E(Y_n)$ to be positive (Phaneuf and Requate, 2017; Soe Zin et al., 2019; Creel and Loomis, 1990).

The log-likelihood for an ordinary Poisson model is represented as

$$\ln\ L = \sum_{n=1}^{N} -\exp(\mathbf{X_n}\beta) + Y_n\mathbf{X_n}\beta - \ln(Y_n!). \tag{2.15}$$

The demand model for our two main visitation sites in George Town are Penang Botanic Gardens and Penang Youth Park can be formulated as

$$Y_n = f(\mathbf{X_n}, \beta) + \mu_n, \tag{2.16}$$

where Y_n is the number of visits of individual n and μ_n is the error term.

2.4.2 Equidispersion test

An equidispersion test examines whether the count data are equidispersed, overdispersed or underdispersed. The equidispersion assumes that $Var(y|x) = E(y|x)$, while the overdispersion assumes that $Var(y|x) = E(y|x) + \alpha^2 E(y|x)$. In rare cases, the observed count data can be underdispersed. According to Cameron and Trivedi (1990), the overdispersion test can be carried out by running a dependent variable's auxiliary regression, which is calculated as $\left\{(y-\hat{\mu})^2 - y\right\}/\hat{\mu}$, with no intercept term. Next, a t test is performed to test the null hypothesis on the coefficient $\hat{\mu}$ to be equal to zero. If the null of zero is not rejected, equidispersion cannot be rejected. If the null of zero is rejected, overdispersion or underdispersion exists. There are different ways to model the data if the overdispersion or underdispersion exists. For the common case of overdispersed data, the most common approach is to use the negative binomial model. Moreover, the generalized Poisson model and quasi-Poisson model can be used to model both overdispersed and underdispersed data.

2.4.3 Robust estimate of the variance-covariance matrix of the estimator

When the count is underdispersed or overdispersed, the assumption of a Poisson distribution is inappropriate. If the conditional mean function is still defined by equation 2.17, the robust version of the variance-covariance matrix estimator (VCE) can be employed. It is called the sandwich estimator of the variance-covariance matrix (White, 1980), which provides robust standard errors (Freedman, 2006). When heteroskedasticity is present, the VCE with this assumption is consistent and does not depend on the formal model of the heteroskedasticity structure. The employment of robust VCE cannot completely remove heteroskedasticity, but proper assumptions can be drawn on the conditional mean parameters (White, 1980). Although the Poisson maximum likelihood estimator (MLE) is maximized, this is not a real likelihood but is a pseudo-likelihood. Therefore, the Poisson model with the use of robust VCE is also called the pseudo-Poisson or quasi-Poisson model Cameron and Trivedi (1990). It is represented as:

$$\hat{V}_{rob}(\hat{\beta}_P) = \left(\sum_{n=1}^{N} \exp\left(x_n'\hat{\beta}_P\right) x_n x_n'\right)^{-1} \left\{\sum_{n=1}^{N} \left(y - \exp\left(x_n'\hat{\beta}_P\right)\right)^2 x_n x_n'\right\} \left(\sum_{n=1}^{N} \exp\left(x_n'\hat{\beta}_P\right) x_n x_n'\right)^{-1}. \tag{2.17}$$

The point estimates are derived by maximizing the Poisson sample likelihood by MLE, but the robust VCE is employed.

For the ordinary Poisson MLE, the estimator $\hat{\beta}_P$ is consistent for the population parameter β, the covariance matrix is represented as

$$\hat{V}(\hat{\beta}_P) = \left(\sum_{n=1}^{N} \hat{\mu}_n x_n x_n'\right)^{-1}. \tag{2.18}$$

In the most common of overdispersion, the variance is larger when using the equation 2.17 compared to Equation 2.18, and $(y - \hat{\mu})^2 > \hat{\mu}_n$. It is the opposite when the observed count is underdispersed.

2.4.4 Willingness to pay and consumer surplus estimations

The WTP is the maximum amount that a visitor is willing to pay for a single visit. It is not equivalent to the travel costs incurred by the visitor for a single visit. Travel cost is less than the WTP: if this were not the case the trip would not be taken. The difference between WTP and travel costs is known as a CS. The expected CS a visitor derives from visiting a green space once can be approximated by the formula (Cameron and Trivedi, 1990):

$$CS_{trip} = -\frac{1}{\hat{\beta}_{TC}}, \tag{2.19}$$

where $\hat{\beta}_{TC}$ is the ML estimate for the coefficient of travel cost variable in the Poisson model. On the other hand, the CS can be derived by estimating the area under the trip demand curve and above the actual travel cost. The demand function is expressed as $Y = e^{\beta_0 + \beta_{TC} X_{TC}}$, where Y is the expected count of trips, β_0 is a constant term, X_{TC} is the total travel cost for a single visit.

2.5 Stated choice analysis and inference

2.5.1 Multinomial logit model

A discrete choice model assumes utility maximization by the decision-maker when making choices (Train, 2003). The models that have a derivation from utility maximization are called random utility models (Marschak, 1960; Train, 2003). The random utility framework proposes that each decision maker n maximize his utility by choosing alternative i over several mutually exclusive alternatives from the choice set J (McFadden, 1974). If decision maker n chooses alternative i over alternative j, the utility of choosing alternative i is greater than that of choosing any other alternative $j \in J$. It can be expressed as

$$(U_{ni} > U_{nj})\, \forall j \neq i, j, i \in J, \tag{2.20}$$

where U_{nit} and U_{njt} are the utilities that individual n obtains from choosing alternative i and j respectively at choice situation t, j=1,…,J. The utility U is made up of V and ε. Where V is an observed component of utility, in other words, it contains some elements that are known by the researcher and observable from the data. While ε is an unobserved component of utility, which is not known by the researcher but is known by the decision maker. It can be expressed as

$$U_{njt} = V_{njt} + \varepsilon_{njt}. \tag{2.21}$$

The error terms, ε are random and are assumed to be independently and identically distributed, which follow Gumbel and type I extreme value distributions. Random utility theory suggests that the choice probability that decision maker n chooses alternative i over alternative j at choice situation t is then represented as

$$\begin{aligned} P_{nit} &= P\left[(U_{nit} > U_{njt})\ \ \forall j \neq i\right] \\ &= P\left[(V_{nit} + \varepsilon_{nit} > V_{njt} + \varepsilon_{njt})\ \ \forall j \neq i\right] \\ &= P\left[(V_{nit} - V_{njt} > \varepsilon_{nit} - \varepsilon_{njt})\ \forall j \neq i\right]. \end{aligned} \tag{2.22}$$

With this assumption, the MNL model can be derived from a random utility framework (McFadden, 1974; Ben-Akiva and Lerman, 1985; Swait and Louviere, 1993). According to McFadden (1974), the logit choice probability that individual i chooses alternative j at choice situation t, assuming the error terms follow independently a Gumbel distribution and are identically distributed, can be expressed as

$$\begin{aligned} P_{nit} &= P\left[(V_{nit} + \varepsilon_{nit} > V_{njt} + \varepsilon_{njt})\,\forall j \neq i\right] \\ &= P\left[(\varepsilon_{njt} < V_{nit} - V_{njt} + \varepsilon_{nit})\,\forall j \neq i\right] \\ &= e^{-e^{-(V_{nit} - V_{njt} + \varepsilon_{nit})}}. \end{aligned} \tag{2.23}$$

If the assumption of independence of errors is made, the probability of choosing item i at choice situation t is

$$P_{nit}|\varepsilon_{nit} = \prod_{j \neq i} e^{-e^{-(V_{nit} - V_{njt} + \varepsilon_{nit})}}. \tag{2.24}$$

As the ε_{nit} is not provided, the unconditional probability is the integral of the conditional probability above over all ε_{nit}, weighted by density $f\left(\varepsilon_{nit}\right) = e^{-\varepsilon_{nit}} e^{-e^{-\varepsilon_{nit}}}$. Therefore, it can be expressed as

$$P_{nit} = \int \left(\prod_{j \neq i} e^{-e^{-\left(V_{nit} - V_{njt} + \varepsilon_{nit}\right)}} \right) e^{-\varepsilon_{nit}} e^{-e^{-\varepsilon_{nit}}} d\varepsilon_{nit}. \tag{2.25}$$

The unconditional probability has a closed form, which can be rewritten as

$$P_{nit} = \frac{e^{V_{nit}}}{\sum_m e^{V_{njt}}}. \tag{2.26}$$

As $V_{nit} = \beta' \mathbf{x_{nit}}$, the MNL choice probability is written as

$$P_{nit} = \frac{e^{\beta' \mathbf{x_{nit}}}}{\sum_j e^{\beta' \mathbf{x_{njt}}}}, \tag{2.27}$$

where β' is a vector of fixed coefficients, $\mathbf{x_{nit}}$ is a vector of attributes associated with alternatives i.

2.5.2 Mixed logit model

The MXL model, which is also called the random parameter logit model, is developed from a random utility framework, it can approximate any model derived from random utility theory (McFadden and Train, 2000). It solves three limitations of the MNL model by allowing for random taste variation, unrestricted substitution patterns, and correlation in unobserved factors (Train, 2003). In MXL models with normally distributed parameters, if individual n chooses alternative j among J alternatives at choice situation t, it can be represented as

$$U_{njt} = \beta'_{\mathbf{n}} \mathbf{x_{njt}} + \varepsilon_{njt}, \tag{2.28}$$

where $\mathbf{x_{njt}}$ is a vector of observed variables associated with individual n, alternatives j and choice situation t. $\beta'_{\mathbf{n}}$ is a vector of random coefficients of observed variables that varies over individual n, each element of the coefficient vector can be expanded into $\bar{\beta}^k + \sigma^k \eta_{njt}$, where $\eta_{njt} \sim N(0, 1)$ denotes the estimation of the distributional moments of random parameter β^k for k attribute. The $\bar{\beta}^k$ denotes the mean value of β_n^k, σ^k is the estimated standard deviation, and η_{njt} is a draw from a normal distribution with zero mean (Scarpa et al., 2012). ε_{njt} is a random error term that represents unobserved utility components (e.g. attributes) and is assumed to be independent and individually distributed (iid) extreme value Gumbel distribution. The variance of ε_{njt} is different among individuals, which is represented as

$$Var\left(\varepsilon_{njt}\right) = s_n^2 \frac{\pi^2}{6}, \tag{2.29}$$

where s_n is the scale parameter for individual n. More generally, the coefficient is a random term which is distributed among individuals in the population with density $f(\beta)$. $f(\beta)$ is a function of parameters θ which represents the mean and covariance of the coefficient. The difference between MXL and MNL is that coefficients vary across individuals in the MXL model instead of being fixed (Train, 2003; Hess and Train, 2017).

The logit choice probability conditional on coefficient β_n is written as

$$L_{nit}(\beta_n) = \frac{e^{\beta'_{\mathbf{n}}\mathbf{x_{nit}}}}{\sum_j e^{\beta'_{\mathbf{n}}\mathbf{x_{njt}}}}. \tag{2.30}$$

However, in MXL, the coefficient is random, so the choice probability is unconditional, and therefore, the unconditional probability, which is also called the MXL choice probability is expressed as the integral of the conditional logit probability over all values of β_n:

$$P_{nit} = \int \prod_{t=1}^{T} L_{nit}\left(\beta_n\right) f\left(\beta|\theta\right) d\beta = \int \prod_{t=1}^{T} \frac{e^{\beta'_{\mathbf{n}}\mathbf{x_{nit}}}}{\sum_j e^{\beta'_{\mathbf{n}}\mathbf{x_{njt}}}} f\left(\beta|\theta\right) d\beta, \tag{2.31}$$

where $f\left(\beta|\theta\right)$ is the density function of β associated with θ parameter. The density depends on the parameters to be estimated. When the coefficients are allowed to be random, it indicates that taste heterogeneity among individuals is allowed to vary randomly.

If F is discrete, the MXL is expressed as a weighted sum:

$$P_{nit} = \sum_{r \in S} L_{nit}\left(\beta_r\right) \pi_r\left(\beta_r|\theta\right), \tag{2.32}$$

where π_r is the probability mass function associated with F, S is a finite set which includes r elements of probability mass (Hess and Train, 2017).

2.5.3 Correlation among coefficients

Correlation analysis among coefficients in MXL has been applied in several studies (Revelt and Train, 1998; Train, 1998; Scarpa and Rose, 2008). This research intends to find correlation patterns between the attributes used in the choice experiment. For instance, an individual who prefers having better air quality may also tend to like having a lot of tree species in the ecosystem; and an individual who prefers having more tree species in the ecosystem may tend to enjoy fewer facilities provided in the green area. Correlations between attributes may arise and they can be either positively or negatively correlated, instead of independent of each other. If an individual's preference for a higher level of one attribute is related to her preference for a higher level of another attribute, both attributes are positively correlated. If an individual's preference for a higher level of one attribute is related to her preference for a lower level of another attribute, these attributes are negatively correlated (Hess and Train, 2017).

Although two possibilities exist, the researcher will only get one estimation result when the correlation between these two attributes is examined. A positive correlation does not mean the preferences for attributes are only positively correlated, but it means they are not independent. It does not mean the negative correlation is impossible, just less likely. In short, a positive correlation between random coefficients simply means the probability of jointly observing same sign in the two coefficient is higher than observing different signs.

2.5.4 WTP-space specifications

It is often useful to interpret utility coefficients as marginal WTP (mWTP) values so as to obtain more reliable estimates of the spread for mWTP. In these cases, the MXL model with utility in WTP-space can be estimated. The WTP-space concept was first introduced by Cameron and James (1987); Cameron (1988) in a fixed-coefficient binary choice for contingent valuation method referendum data, it was extended to models with random coefficients by (Train and Weeks, 2005). The model is different from the MXL model in preference space, in terms of the coefficients' interpretation. Although a substantial body of evidence suggests that the MXL model in preference space fits the data better when compared to the MXL model in WTP-space, the distribution of WTP, which is derived from the MXL model in preference space, has a substantially large variance. A large variance in WTP indicates that the individuals may be willing to pay either a considerably high or exceptionally low amount of money. This level of variability can lead to unreasonable interpretation. In contrast, when using WTP-space specifications in the model, the WTP distributions can be directly controlled. One of the key advantages of employing WTP-space specifications is that it allows analysts to have more direct control over the characteristics of WTP distributions (Thiene and Scarpa, 2009). For instance, it becomes possible to set and test for specific restrictions on WTP distributions directly (Thiene and Scarpa, 2009). Consequently, the MXL model implemented in WTP-space can result in a reduced variance in WTP, making its interpretation more reasonable (Train and Weeks, 2005). Therefore, the MXL model in WTP-space specification is also suggested in the estimation of the welfare values in the context of this thesis.

As stated above, the utility of the conventional MXL model includes a vector of observed variables $\mathbf{x_{njt}}$, a vector of coefficients of observed variables $\beta_\mathbf{n}$ which can be fixed or random, and a random error term ε_{njt}. As described in Train and Weeks (2005) and Scarpa and Rose (2008), the combined component $\beta'_\mathbf{n}\mathbf{x_{njt}}$ can be separated into two different components, which is written as

$$U_{njt} = -\alpha_n \rho_{njt} + \beta'_\mathbf{n}\mathbf{x_{njt}} + \varepsilon_{njt}, \tag{2.33}$$

where α_n is the price coefficient, ρ_{njt} is the price attribute, $\beta'_\mathbf{n}$ is now a vector of non-price coefficients and $\mathbf{x_{njt}}$ is the vector of non-price attributes. Similarly, ε_{njt} is a random term which is iid extreme value. The variance of ε_{njt} varies among individuals: $Var\,(\varepsilon_{njt}) = s_n^2(\frac{\pi^2}{6})$, where s_n is the scale parameter for individual n. This formula shows that the scale parameter S_n is different across individuals. The scale parameter represents the standard deviation of utility over individuals which estimates the variance in unobserved factors. If the scale parameter is assumed to be random, the error term ε_{njt} has a variance's variance, which is called the "scale parameter" (Train and Weeks, 2005).

The recent utility is then divided by the scale parameter s_n to generate a scale-free formula. By doing this, the behaviour is unaffected but a new error term in which the variance is similar over individuals is generated. It is written as

$$U_{njt} = -(\alpha_n/s_n)\rho_{njt} + (\beta_\mathbf{n}/s_n)'\mathbf{x_{njt}} + \varepsilon_{njt}, \tag{2.34}$$

where the error term ε_{njt} is iid type one extreme value, dividing out the scale parameter, it now has a fixed variance $\frac{\pi^2}{6}$. To simplify the formula, the utility is written as

$$U_{njt} = -(\lambda_n)\rho_{njt} + (\eta'_\mathbf{n})\mathbf{x_{njt}} + \varepsilon_{njt}, \tag{2.35}$$

where $\lambda_n = \alpha_n/s_n$ and $\eta_\mathbf{n} = \beta_\mathbf{n}/s_n$. As the scale parameter s_n is the denominator of all coefficients, the correlation among coefficients exists if s_n is proved to vary randomly. For example, a varying s_n with fixed α_n and $\beta_\mathbf{n}$ in the same formula indicates that the coefficients are perfectly correlated. While a varying s_n with varying α_n and $\beta_\mathbf{n}$ will cause the correlation of coefficients to be less than one.

To compute WTP for attributes, the non-price utility coefficient must be divided by the price coefficient, such as $\omega_\mathbf{n} = \eta_\mathbf{n}/\lambda_n$. Therefore, the model in WTP-space is represented by the following reparameterization of U_{njt}:

$$U_{njt} = -(\lambda_n)\rho_{njt} + (\lambda_n\omega_\mathbf{n})'\mathbf{x_{njt}} + \varepsilon_{njt}. \tag{2.36}$$

The variations in price coefficient λ_n and wtp $\omega_\mathbf{n}$ are separated, while λ_n includes scale as the scale parameter is the denominator of the price coefficient, and $\omega_\mathbf{n}$ is free from scale as the scale parameter is divided out from $\omega_\mathbf{n} = \eta_\mathbf{n}/\lambda_n = (\beta_\mathbf{n}/s_n)/(\alpha_n/s_n) = \beta_\mathbf{n}/\alpha_n$. The formula has equivalent utility expressions. The distribution specification of coefficients in equation 2.35 is equivalent to that of equation 2.36. If the non-price coefficient is normally distributed and the price coefficient is log-normally distributed, the mWTP is the ratio of normal to log-normal term. If the mWTP is normally distributed and the price coefficient is log-normally distributed, the non-price coefficient has a combination of normal and log-normal terms.

2.5.5 Individual coefficients from mixed logit model

The individual-level coefficients and mWTP values for each individual can be estimated following the estimation of the MXL model in WTP-space.

Train (2003) and Scarpa et al. (2005) discussed the derivation of the individual-level conditional estimator of an attribute coefficient for each individual by applying the Bayes rule. The mean of the individual distribution of $\tilde{\beta}^k$ can be obtained by calculating the expectation of the parameter, considering the parameter distribution conditional on observed choices and parameter estimates:

$$\bar{\beta_n} = \int \beta \cdot h(\beta|y_n, x_n, \mu, \Omega) d\beta, \tag{2.37}$$

where $h(\beta|y_n, x_n, \mu, \Omega) = \frac{P(y_n|x_n,\beta)\ g(\beta|\mu,\Omega)}{P(y_n|x_n,\mu,\Omega)} = \frac{P(y_n|x_n,\beta)\ g(\beta|\mu,\Omega)}{\int P(y_n|x_n,\beta)\ g(\beta|\mu,\Omega)}$, $g(\beta|\mu, \Omega)$ represents the assumed distribution of the parameter within the entire population. Therefore, Equation 2.37 can be rewritten as

$$\bar{\beta_n} = \frac{\int \beta\ P(y_n|x_n, \beta) g(\beta|\mu, \Omega) d\beta}{\int P(y_n|x_n, \beta) g(\beta|\mu, \Omega) d\beta}, \tag{2.38}$$

where $P(y_n|x_n, \beta)$ represents the conditional density of the observed data for the panel of choices made by individual n, given the value of x_n and β.

The integrals in Equation 2.38 do not have a closed form, therefore, they cannot be computed directly but they can be computed by simulation methods. Draw samples of β from the probability density function $g(\beta|\mu, \Omega)$. Compute the weighted average of these draws, where the weight assigned to each draw β^r is proportional to

$P(y_n|x_n, \beta^r)$. The simulated sub-population mean is expressed as

$$\check{\beta}_n = \sum_r w^r \ \beta^r, \tag{2.39}$$

where the weights are

$$w^r = \frac{P(y_n|x_n, \beta^r)}{\sum_r \ P(y_n|x_n, \beta^r)}. \tag{2.40}$$

2.5.6 mWTP estimation for models in preference space

The mWTP for the attribute level can be calculated by dividing the negative coefficient of that attribute level by the coefficient of the cost variable, which represents the distance between the individual's house and the hypothetical green site. This ratio provides the marginal rate of substitution between the attribute level and distance, indicating the maximum distance that an individual is willing to travel for an additional unit of the attribute level. However, the estimated value is not expressed in monetary terms, as the base of the calculation is measured in distance (kilometres). To express the WTP value in monetary terms, the calculated WTP value is multiplied by a conversion factor, representing the amount of money an individual is willing to pay per kilometre of distance. The conversion factor is RM0.205 per kilometre, and the calculated WTP is multiplied by 2 (for a return trip) and is expressed in Malaysian Ringgit (RM).

In this context, the distribution of WTP values for an attribute level can be derived by dividing the distribution of attribute level coefficients by the distribution of the negative distance coefficients. To simulate the distribution of coefficients, the *drawnorm* command in Stata 14.2 is used to draw 1000 random samples from a distribution with the mean and SD of the coefficients in the model. Random samples are used to represent a range of possible coefficient values for attribute levels. These samples are used to estimate the distribution of WTP values and derive the mean and SD of WTP for an attribute level. Random samples are only generated if the coefficient is random. If the coefficient is fixed or if the standard deviation of the coefficient is not statistically significant at the 10% level, random draws are not performed. In such cases, the coefficient estimation from the model is used directly.

2.5.7 Welfare measures

Welfare measures consistent with random utility theory are well established (Domencich and McFadden, 1975; Hanemann, 1982; Small and Rosen, 1981; McConnell, 1995).

The welfare measures obtained from this theory can be derived by measuring the Compensating Variation (CV) from the changes in one or multiple non-price attributes k, from x^0 to x^1, where $\mathbf{x}$ is the vector of non-price attributes, and conditional on individual parameter estimates $\beta_{\mathbf{n}}$ are logit (Hynes et al., 2008), are expressed as

$$CV_n = \frac{[\ln[\sum \exp(\beta_{\mathbf{n}} \mathbf{x}_{\mathbf{n}}^{\mathbf{1}})] - \ln[\sum \exp(\beta'_{\mathbf{n}} \mathbf{x}_{\mathbf{n}}^{\mathbf{0}})]]}{-\lambda_n}, \tag{2.41}$$

where λ_n is the price coefficient. To account for preference heterogeneity, the expected CV estimate needs the integration of the population's taste distribution:

$$\widehat{CV_n} = \int CV_n \; f(\hat{\beta}|\hat{\mu}, \hat{\Omega}) d\beta = \int \frac{[\ln[\sum \exp(\hat{\beta}'_{\mathbf{n}} \mathbf{x}_{\mathbf{n}}^{\mathbf{1}})] - \ln[\sum \exp(\hat{\beta}_{\mathbf{n}} \mathbf{x}_{\mathbf{n}}^{\mathbf{0}})]]}{-\hat{\lambda}_n} f(\hat{\beta}|\hat{\mu}, \hat{\Omega}) d\beta, \tag{2.42}$$

where the integral can be estimated through the simulation from draws of the estimated distributions for the random parameters.

2.6 Seemingly unrelated regression model analysis

This study also developed a SUR model to investigate the impacts of spatial and socio-demographic factors on individual mWTP for attribute levels. There are 10 mWTP values corresponding to 10 attribute levels, and it is important to recognize that not all variables impact these mWTP values in the same manner. In reality, different variables bring varying influences on mWTP values. Consequently, it is not feasible to incorporate all mWTP values into a single model. Campbell (2007) proposed a solution involving the use of dummy variables for each mWTP value to address this issue. However, this approach may lead to a large number of coefficients being estimated, potentially resulting in statistical issues in practical applications. An alternative approach, suggested by Czajkowski et al. (2017), involves estimating the model separately for each mWTP value. The Czajkowski's approach allows for capturing the unique relationships between independent variables and each individual mWTP value.

To explain the individual-level conditional mWTP values obtained from the DCE while considering individual spatial data and socio-demographic information, one could initially employ a simple ordinary least squares (OLS) model. However, it is important to note that all conditional mWTP values are derived from the same MXL model, leading to correlations among these values. To address this correlation, a SUR model is recommended.

The SUR model, initially discussed by Zellner (1962), accommodates dependent variables with correlated disturbances that are influenced by similar factors, allowing for their estimation within a system of interrelated equations (Wang and Kockelman, 2007). It comprises several linear regression models, each with its own dependent variable and the same or different sets of independent variables. When employing the same set of observations, the SUR model accommodates correlated error terms across equations (Sheng and Sharp, 2019; Rentziou et al., 2012). Therefore, compared to the Ordinary Least Squares (OLS) method, which estimates these interrelated models separately, the SUR model efficiently estimates coefficients by explaining all mWTP values through the independent variables collectively.

According to Zellner (1962), suppose y_{it} is a dependent variable, $\mathbf{x_{it}} = (1, x_{it,1}, x_{it,2}, ..., x_{it,Ki-1})'$ is a K_i-vector of explanatory variables for equation i, and μ_{it} is an unobservable error term, where the double index it denotes the t^{th} observation of the i^{th} equation in the system. A classical linear SUR model consists of N linear regression equations,

$$y_{it} = \mathbf{x_{it}} + \mu_{it},\ i = 1, ..., N\, and\, t = 1, ..., T. \tag{2.43}$$

Denote $L = K_1 + ... + K_N$. Assume that all equations N have the same K_i-vector of explanatory variables of $\mathbf{x_{it}}$, the error terms U_t in each equation i are i.i.d. over observation t with mean $E[\mu_t|X] = 0$ and homoskedastic variance $\sum = E[\mu_t\mu_t'|X]$. Moreover, the $\sum$ is also assumed to be positive definite and denote by σ_{ij} the $(i,j)^{th}$ element of $\sum$, that is, $\sigma_{ij} = E[\mu_{it}\mu_{jt}|X]$. Under this assumption, the covariance matrix of the entire vector of disturbances $\mathbf{U} = [U_1, ..., U_T]$ is given by $E[vec(\mathbf{U})(vec(\mathbf{U}))'] = \sum \bigotimes I_T$ (Moon and Perron, 2008).

2.7 Conclusions

In conclusion, this chapter serves as an introduction to the methods employed in this study to elicit individual preferences and WTP for various attributes, particularly those associated with ecosystem assets and services within urban green spaces in Penang Island, Malaysia. Environmental economics offers numerous valuation techniques, and this research employs a combination of BWS, TCM, and DCE to achieve its research objectives. Furthermore, this research discusses on the DCE design, post-estimation processes to understand the relationship between spatial information and WTP for these attributes, and the simulation of attribute values under different scenarios.

The following chapter (Chapter 4) provides comprehensive details of the examination of the BWS study.

Chapter 3

Data and Survey Design

This chapter introduces the design of the study's survey instrument (Section 3.1) and the process of data collection (Section 3.2). These sections highlight the strategies adopted to ensure the collection of high-quality data.

3.1 Survey development and design

After conducting the focus group sessions, I created an online survey using the Qualtrics platform. This platform allows question blocks to appear randomly to participants, which is particularly useful for the discrete choice model questions. These questions were divided into three blocks, each containing 12 questions. Furthermore, the survey was translated into three languages: English, Mandarin, and Malay. Respondents were asked to select the language they were most comfortable with for completing the survey. Once they made their selection, they were directed to the corresponding version of the survey. To ensure the questionnaire's clarity, I conducted a pre-test by sending it to a select group of individuals, including my professors, NGO's members, and friends. This pre-testing step helped improve the survey design.

The survey is organized into four sections:

1. The first section collects responses for a DCE.

2. The second section involves a best-worst scaling exercise.

3. The third section gathers socio-demographic information.

4. The final section collects data regarding past visits to selected urban recreational sites.

3.1.1 First section: the discrete choice experiment

Development of attributes and levels

Before developing the attributes, three focus group discussions were conducted with Penang residents to gain a comprehensive understanding of the characteristics of urban green sites that are important to the residents of Penang. Detailed information about these focus group discussions can be found in Appendix B.

Six attributes were selected to describe the characteristics of urban green sites in the DCE. These attributes were determined through a follow-up survey conducted after discussions in focus groups, as depicted in Figure 52 in Appendix B. During this survey, participants were asked to assess the importance of various urban green space characteristics, such as *distance*, *facilities*, *trees and plants*, *abundance of animals*, *landscape naturalness*, *noise level*, and *air quality*. Notably, the attribute *abundance of animals* ranked the lowest, leading to its removal from consideration. Similarly, *landscape naturalness* was deemed unsuitable due to the predominantly man-made nature of urban green spaces and was thus excluded.

In addition to green space characteristics, participants were also asked to evaluate the importance of urban ecosystem services. Ecosystem services refer to the benefits humans derive from ecosystems, such as clean air, water purification, and recreational opportunities. This understanding provides a contextual framework for participants to assess urban green spaces and their contributions to human well-being in urban areas. As shown in Figure 53 in Appendix B, *ecosystem and species appreciation*, *air filtration*, and *nursery population and habitat maintenance* emerged as the three most crucial attributes. *Ecosystem and species appreciation* encompassed trees, plants, and animals, though *abundance of animals* ranked lowest. Consequently, the animal species aspect was removed, and this attribute was redefined to include trees and plant species, merging with the *trees and plants* attribute from green space characteristics.

Considering the association between *air filtration* and *air quality*, the latter was included as one of the six attributes. *Nursery population and habitat maintenance*, a less-explored aspect in existing literature, was included as one of the six attributes based on its perceived importance.

In summary, the DCE included six attributes: *air quality*, *distance*, *facilities*, *noise levels*, *nursery habitat maintenance*, and *tree species and ecosystems*.

Table 1 displays the definition of attribute levels. The *distance* attribute, which is a six-level cost attribute representing the travel distance between the respondent's residential address and the hypothetical urban green site. Given that the average distance from respondents' houses to the two major green spaces (Penang Botanic Gardens and Penang Youth Park) is 14.7 km, I have chosen 15 km as the central point. The first three levels proposed distances shorter than this, while the subsequent levels were longer than this. The progression starts with the shortest distance, 1 km, followed by 5 km, 10 km, 20 km, 30 km and 40 km.

The attributes and their respective levels were graphically represented using various forms, including numbers and symbols. This visual representation of attribute levels in DCE is believed to reduce bias in estimation results (Umberger and Mueller, 2010). Furthermore, (Breitmeyer et al., 2004) have suggested that the visual presentation of attributes can unconsciously influence individuals' choices. Consequently, a series of graphics representing the attribute levels were incorporated into the DCE questions. The visual representations of these attribute levels can be seen in Figure 3. Apart from the visual representation, the description of attribute levels is provided in a separate sheet, and a link to this sheet is included in the online questionnaire to help respondents better understand the attribute levels.

The *air quality* attribute of a site is quantified by the Air Quality Index (AQI) and is categorized into three levels. The AQI boundaries for each level follow the U.S. AQI, established by the U.S. Environmental Protection Agency (AirNow, 2023). This attribute relates to the CICES framework's Code 2.1.1.2, which involves the mediation of anthropogenic wastes by living processes, specifically classified as filtration, sequestration, storage, and accumulation by plants (Haines-Young and Potschin, 2018).

The *facilities* attribute describes the amenities provided on the site and varies across three levels: a few, normal, and a lot. The definitions for each level align with the current conditions of urban green spaces in the study area.

The *noise levels* attribute of a site is measured in decibels (dB) and is divided into three levels. The dB thresholds for each level adhere to guidelines provided by the U.S. Centers for Disease Control and Prevention (CDC), where the moderate level is set at 60 dB, the highest allowed. To maintain balanced three-level noise categories, each level has a range of 30 dB (Centers for Disease Control and Prevention, 2023). In CICES framework, this attribute relates to the mediation of anthropogenic nuisances through noise attenuation, classified as Code 2.1.2.2 (Haines-Young and Potschin, 2018).

The *nursery habitat maintenance* attribute indicates the level of maintenance of the nursery habitat, which supports the reproduction of trees and plants of a particular species. While *nursery habitat maintenance* is often considered an intermediate service, it serves as an important input for sustaining species populations that contribute to final ecosystem services (United Nations et al., 2021). Although this service is an intermediate service, its inclusion in the valuation framework is justified by the significant role it plays in supporting the provision of final ecosystem services. This attribute was described by using three levels, each representing different degrees of maintenance, and it attribute relates to the regulation of physical, chemical, and biological conditions, particularly focusing on lifecycle maintenance, habitat, and gene pool protection, as classified under Code 2.2.2.3 in CICES framwork (Haines-Young and Potschin, 2018).

The *tree species and ecosystems* attribute were described using three levels of diversity in tree species at the site. In the DCE questions, attribute levels for *nursery habitat maintenance* and *tree species and ecosystems* are visually represented, enabling respondents to compare and assess the perceived levels of diversity against recognizable benchmarks.

Attributes	Levels	Descriptions
Air quality	Poor	Air quality index: 101-150; Sensitive groups should reduce prolonged or heavy outdoor exertion.
	Normal	Air quality index: 51-100; Unusually sensitive people should consider limiting prolonged outdoor exertion.
	Good	Air quality index: 0-50; Air quality is satisfactory and poses little or no risk.
Distance	1KM 5KM 10KM 20KM 30KM 40KM	Distance from residential address to the hypothetical urban green site (KM).
Facilities	A few	Include only walking paths and benches.
	Normal	Include only walking paths, benches, cycling tracks and playground.
	A lot	Include walking paths, benches, cycling tracks, playground, football field, baseball court, hiking track and rock-climbing wall.
Noise levels	Loud	70-90dB; Noise from truck, car or hair dryer.
	Normal	40-60dB; Noise from conversation or refrigerator.
	Quiet	10-30dB; Noise from whispering or breathing.
Nursery habitat maintenance	Poor	Nurseries that make a little contribution to the reproduction of trees and plants from a particular species, where juveniles occur at lower densities.
	Normal	Nurseries that make an intermediate contribution to the reproduction of trees and plants from a particular species, where juveniles occur at normal densities.
	Good	Nurseries that make a large contribution to the reproduction of trees and plants from a particular species, where juveniles occur at higher densities.
Tree species and ecosystems	A few	Low diversity of tree species.
	Normal	Intermediate diversity of tree species.
	A lot	High diversity of tree species.

Table 1: Attributes and levels in DCE

Development of efficient designs

This section outlines the experimental design process of the case study, which involved two phases of survey collection. In the first phase, an orthogonal (*Orth*) design was employed, and in the subsequent phase, a Bayesian *D-efficient* (*Baye*) design was employed. The *Baye* design was constructed using the Bayesian *D-error* minimization criterion, applied to the parameter estimates of the MNL obtained from the pilot choices, using the *Orth* design. Both designs were generated using Ngene software (ChoiceMetrics, 2018).

An orthogonal (*Orth*) design must meet the condition that for each attribute column, its attribute levels are with minimal correlation. To test the validity of an *Orth* design, orthogonal vectors are commonly used. An *Orth* design is considered truly orthogonal when the sum of the multiplication of two attribute columns equals zero, as represented by (ChoiceMetrics, 2018):

$$\sum_{s=1}^{S} x_{j_1 k_1 s_1} x_{j_2 k_2 s_1} = 0 \tag{3.1}$$

where j_1 represents the alternative, k_1 represents the attribute, and s_1 denotes the choice situation. In this case, an attribute with three levels (0, 1, 2) was transformed into codes (-1, 0, 1). As illustrated in Table 2, the attribute b is multiplied by attribute c, resulting in bc. The sum of bc across 36 choice situations equals zero. The syntax for generating an *Orth* design in Ngene is:

```
Design
;alts = choice1, choice2
;rows = 36
;orth = seq
;block = 3
;model:
U(choice1) = b1
+ b2 * A{[}0,1,2{]}
+ b3 * B{[}1,5,10,20,30,40{]}
+ b4 * C{[}0,1,2{]}
+ b5 * D{[}0,1,2{]}
+ b6 * E{[}0,1,2{]}
+ b7 * F{[}0,1,2{]}
/
U(choice2) = b2 * A + b3 * B + b4 * C + b5 * D + b6 * E + b7 * F
```

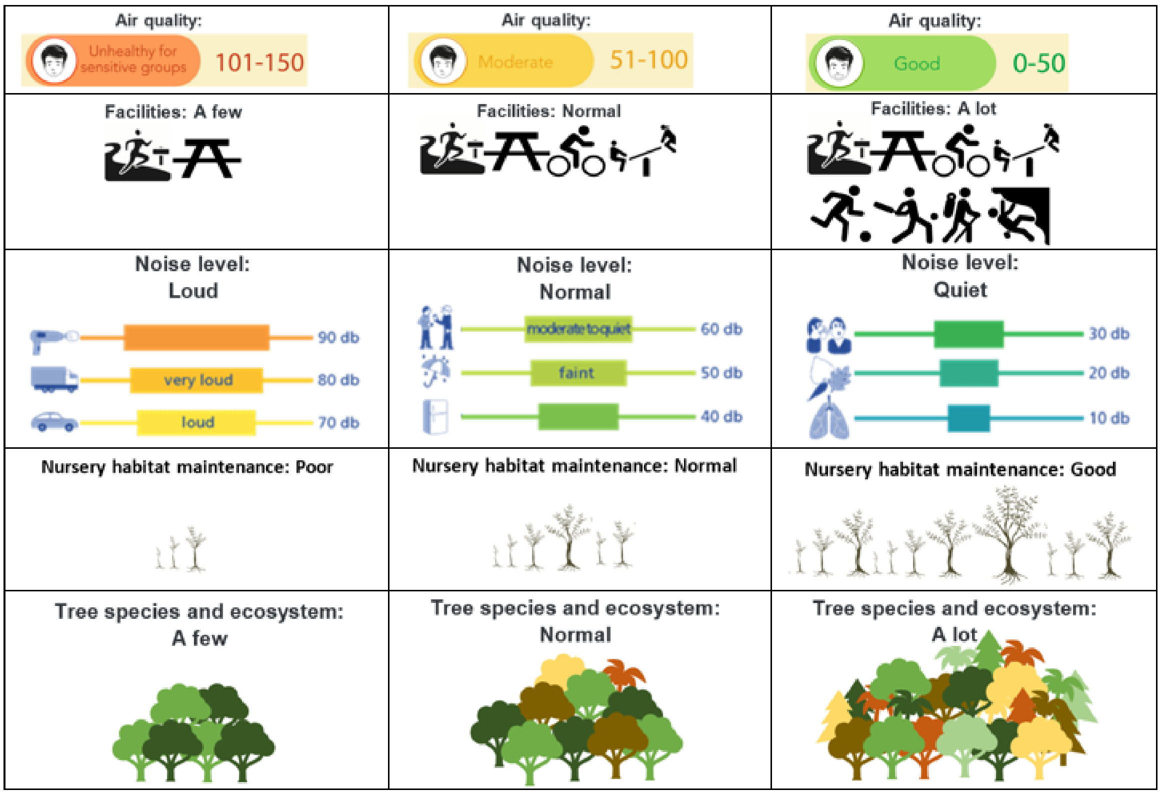


Figure 3: Visual representation for attribute levels

Choice situation	a	b	c	d	e	f	i. e.: $b \times c$
1	5	1	1	0	0	1	1
2	30	0	0	1	1	0	0
3	40	1	1	1	1	0	1
4	10	0	0	0	0	1	0
5	30	1	-1	-1	0	0	-1
...	...	...	...	...	...	...	...
36	5	1	-1	0	1	0	-1
Sum							0

Table 2: Validity test for orthogonal design

s	1.a	1.b	1.c	1.d	1.e	1.f	2.a	2.b	2.c	2.d	2.e	2.f	Block
1	5	2	2	1	1	2	5	1	2	0	2	2	1
2	30	1	1	2	2	1	40	1	0	2	1	1	2
3	40	2	2	2	2	1	40	2	1	0	1	0	3
4	10	1	1	1	1	2	40	2	2	0	0	2	1
5	30	2	0	0	1	1	10	1	1	1	1	2	1

Table 3: Output of orthogonal design for the first five choice situations (s)

$

This design comprises two alternatives and 36 choice situations, which are further divided into 3 blocks, with 12 choice situations in each block. Ngene's *orth* function provides the capability to generate either a sequential *Orth* design or a simultaneous *Orth* design. A sequential *Orth* design, denoted by the code *seq*, ensures orthogonality within a specific alternative, while a simultaneous *Orth* design, denoted by the code *sim*, extends orthogonality across alternatives. In this study, a sequential *Orth* design was employed to generate the DCE questions. Table 3 displays the output of the *Orth* design for the first five rows.

To create the Bayesian *D-efficient* (*Baye*) design for choice tasks, a multinomial logit model was estimated using 117 choices gathered from the orthogonal design choice tasks. The aim was to obtain coefficient estimates and standard errors, which were subsequently used to generate *a priori* multivariate distribution of the model parameters for the *Baye* design, ultimately improving the model's accuracy while maintaining a degree of uncertainty around the point estimates. The estimation results for the MNL model from the pilot survey responses and *Orth* design are shown in Appendix I. With the exception of *fac1*, all coefficient estimates were significant at the 5% level. Notably, the coefficient estimate for the distance attribute (*dis*) was negative, while the coefficients for the other attributes were positive. The syntax used to generate a *Baye* design in Ngene is:

```
Design
;alts = choice1, choice2
;rows = 36
;eff = (mnl,d,mean)
;bdraws = halton(500)
;block = 3
;model:
U(choice1) = b0[(n,.1220257,.0735307)]
+ b1.dummy[(n,1.343686, 0.1134044)|(n,2.481235, 0.1388363)]*A[1,2,0]
+ b2[(n, -0.0252043, 0.004418)]*B[1,5,10,20,30,40]
+ b3.dummy[(n,0.1043531, 0.1304324)|(n,0.2582007,0.1349382)]*C[1,2,0]
+ b4.dummy[(n,0.6102,0.111)|(n,0.4191,0.115)]*D[1,2,0]
+ b5.dummy[(n,0.2520511,0.142872)|(n,0.4927782,0.1327828)]*E[1,2,0]
+ b6.dummy[(n,0.295656,0.1174269 )|(n,0.4901873,0.1259081)]*F[1,2,0]
/
U(choice2) = b1 * A + b2 * B + b3 * C + b4 * D + b5 * E + b6 * F
;formatTitle = '<blocknumber>'
;formatTableDimensions = 3, 7
;formatTable:
1,1 = '' /
1,2 = 'Air quality' /
1,3 = 'Distance (KM)' /
1,4 = 'Facilities' /
1,5 = 'Noise level' /
1,6 = 'Nursery habitat maintenance' /
1,7 = 'Tree species and ecosystem'
$
```

In this case, an efficient design was created for the MNL model with the goal of minimizing the *D-error*. The *mean* parameter specified in the third argument of the efficiency method *eff* indicates that the mean value of the efficiency measure is computed when calculating the Bayesian *D-error* over a set of random draws (ChoiceMetrics, 2018). To achieve this, the design requires multiple draws from the Bayesian distributions. In this instance, 500 Halton random draws were used to average the design D-errors. The Bayesian prior parameters were set to follow normal

s	1.a	1.b	1.c	1.d	1.e	1.f	2.a	2.b	2.c	2.d	2.e	2.f	Block
1	1	30	2	0	0	1	0	5	1	2	1	0	3
2	1	1	1	2	0	2	2	40	0	1	2	0	1
3	0	30	1	1	1	1	0	5	2	2	2	0	2
4	1	30	0	0	2	0	0	5	2	2	0	1	1
5	2	10	2	2	1	2	2	20	1	0	2	1	1

Table 4: Output of Bayesian *D-efficient* design

distributions with specified means and standard deviations from the MNL model of the pilot.

Similar to the *Orth* design, the *Baye* design consisted of two alternatives and generated 36 choice situations, organized into three blocks, each containing 12 choice situations. During the survey, participants were randomly presented with one of the three blocks. Table 4 presents the output of the *Baye* design for the first five rows.

Designs of DCE questions

The DCE aims to collect empirical data on respondents' preferences for urban green sites' characteristics with different levels. The section begins with introducing a hypothetical situation:

'Suppose that you have decided to visit a green space in George Town, for example, a public park or a garden. Suppose you are going there alone by car (the distance from your house to the destination will affect your travel cost). Finally, suppose there are only the following two green spaces available, each with different characteristics. Which of the two green spaces do you prefer to visit?'

The DCE questions were organized into three blocks, with each block containing 12 choice tasks. To minimize order bias, each respondent was randomly assigned to receive only one block of 12 choice tasks. The random assignment was conducted using the Qualtrics platform to ensure that each respondent received only one randomly selected block of questions. Each task included two alternative sites labelled as 'Green Space A' and 'Green Space B.' Each alternative site was described by six different site attributes with varying levels: distance, facilities, tree species and ecosystems, noise levels, air quality, and nursery habitat maintenance. Respondents were required to select their preferred alternative site for each task.

Once respondents made their choice, they were asked to indicate how many times they would choose to visit their selected site if they were to repeat this decision 20 times:

'Imagine you are to repeat this decision 20 times. How many out of these 20 times would you choose to visit your selected destination?'

Figure 4 presents an example of a choice task.

The DCE questions underwent a redesign after an initial round of data collection in the first phase. In this first phase, an orthogonal design was employed in developing the DCE questions. Subsequently, in the second phase, the DCE questions were designed using a Bayesian efficient design with prior values of the parameters based on the estimates of the first phase.

In DCE applications, the status quo (SQ) alternative often serves as a benchmark. However, the decision to use a paired alternative choice set design without an SQ alternative was made to focus respondents on trade-offs between new options, reducing the anchoring effect of the current state and revealing preferences for new options. This approach also mitigates status quo bias and simplifies choice tasks, improving response quality.

Likewise, the design has no 'opt-out' response for participants. The assumption was made that respondents were already intending to visit the UGS, thereby the opt-out option is excluded. Moreover, this decision was made to ensure that respondents provide active choices rather than opting out by default. By requiring participants to make a choice between alternatives, this design encourages respondents to consider the trade-offs between different options, leading to better insights into their decision-making processes. Additionally, excluding the 'opt-out' option could weaken the attraction effect and strengthen the compromise effect (Dhar and Simonson, 2003). Moreover, this design simplifies the choice data analysis and allows for a more straightforward interpretation of results.

This research applied fixed conversion factor for changing distance to cost in the DCE specifications. The assumption describes respondents driving alone to a UGS, yet some might walk or drive with others. To address these situations, I conducted a Mixed Logit model in WTP space specification, assuming that the return-trip travel costs for individuals are derived based on the transport mode used by the individuals. The results are shown in Appendix A. The results indicate that when travel costs are adjusted for the actual transport mode, the WTP for attribute levels is generally lower. This suggests that individuals optimize their transport choices to minimize costs, reflecting more realistic behaviour. Despite these findings, I maintain the fixed conversion factor for the primary analysis. This choice was made to maintain

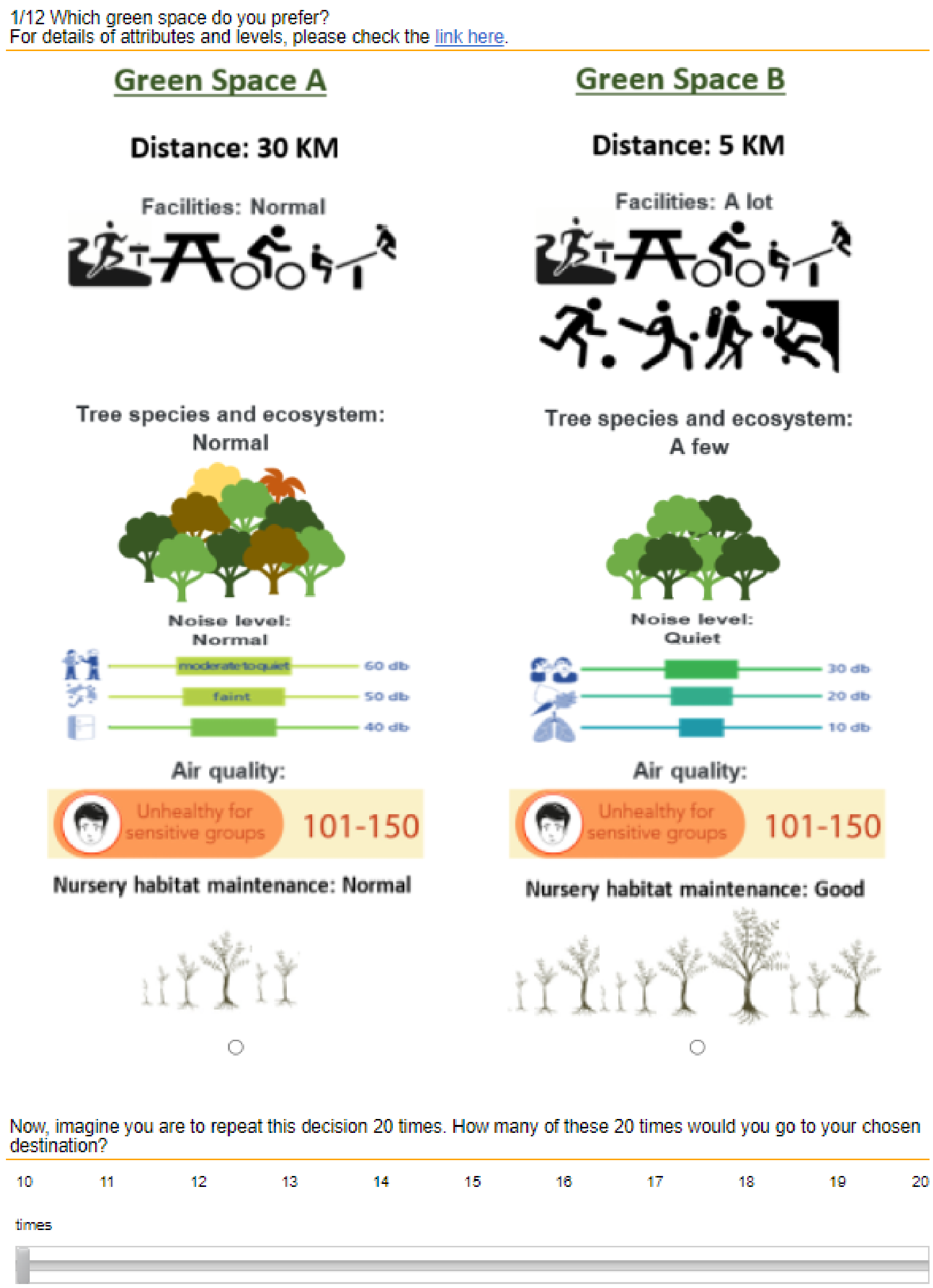


Figure 4: Example of DCE question

simplicity and to provide respondents with a clearer picture of the hypothetical scenarios in the DCE. The assumption of a fixed conversion factor ensures consistency and ease of understanding for respondents. Nonetheless, this analysis demonstrates the robustness of the welfare estimates and highlights the importance of considering diverse transport modes for a comprehensive understanding of WTP.

3.1.2 Second section: best-worst scaling questions

Generation of experimental designs

The series of questions was developed using balanced incomplete block designs (BIBD), as recommended by Louviere et al. (2013). These designs ensure that the included attributes are presented an equal number of times, and that pairwise comparisons of attributes are also shown an equal number of times (White II, 2021). I employed the *bwsTools* R package to generate these BIBDs. To be considered balanced, an incomplete block design must meet the requirements outlined in the following equation.

$$\lambda = r(k-1)/(t-1), \tag{3.2}$$

where λ is the number of times each pairwise comparison occurs, r is the number of times each item is repeated, k is the number of items in each block, t is the number of items. Both λ and r are integers.

The initial version of the survey included 13 attributes. To comply with the requirements outlined in Equation 3.2, I generated 13 choice tasks, each containing 4 items. In the revised version of the survey, I narrowed down the attributes to 6, aligning them with those used in the DCE, which were presented in 10 choice tasks, each featuring 3 attributes. Participants were instructed to identify the most and least important attributes influencing their decisions to visit urban green spaces in each choice task. The attributes that participants selected as the most and least important in each choice task had the greatest difference in terms of the utility scale (Lancsar et al., 2013; Franco et al., 2015).

BWS questions

Respondents were posed with the following question:
"Please select one most and one least important factors that affect your decisions to visit urban green spaces."
An example of a choice task is presented in Figure 5.

Please select one most and one least important factors that affect your decisions of visiting urban green spaces.

Most important		Least important
○	Noise level	○
○	Landscape naturalness	○
○	Animal species and ecosystem	○
○	Size	○

Figure 5: Example of best-worst scaling choice task

3.1.3 Third section: socio-demographic questions

This section comprises nine socio-demographic questions, which cover respondents' year of birth, gender, first language, employment status, highest level of education, marital status, household size, the number of children under 18 residing in the same household, and the duration of residence in Penang.

3.1.4 Fourth section: past visit patterns

This section consists of two parts. The first part gathers general information about respondents' visit patterns to urban parks. This includes their visiting days of the week and the frequency of their visits in the past year.

The second part collects information about respondents' visit patterns to specific locations, namely Penang Botanic Gardens and Penang Youth Park. This information covers several aspects, such as the number of visits, total time spent during visits, mode of transportation used, the number of individuals accompanying them during the visit, the purpose of their visit, and their home address. In cases where respondents were unwilling to provide their home address, they were requested to provide either the nearest road to their house or the general area of their place of residence. This section is designed for travel cost analysis and aims to collect empirical data regarding respondents' visit behaviour and travel expenses.

3.2 Data collection

The online survey was conducted between February and June 2022. To reach a diverse group of respondents, three enumerators were hired to distribute leaflets inviting the public to participate in the survey. These enumerators were stationed at various green spaces, including Penang Botanic Gardens, Penang Youth Park, Bukit Dumbar Park, and several smaller neighbourhood parks. They conducted surveys during both morning and evening shifts, seven days a week, with each shift lasting for two hours. Potential respondents who received the leaflet were invited to complete the survey on their own smart phones. A QR code linked to the survey web page was provided in the leaflet, allowing respondents to complete the survey at their convenience. Before answering the questions, respondents received a brief introduction to the study and an explanation of information confidentiality.

The decision to use intercept surveying with an online mode was made after considering various sampling strategy alternatives. This method was used because it can efficiently reach individuals with diverse backgrounds and ensure geographical representation. The online mode enhances accessibility and reduces logistical constraints, offering a cost-effective method to gather data for the study.

After excluding the outliers, respondents took an average of 11.23 minutes to complete the survey, and the standard deviation was 5.2 minutes. Respondents who completed the survey received RM5.00 (NZD1.80) as a token of appreciation, the equivalent of 2 cups of iced coffee or a bowl of noodles in Penang at that time.

At the conclusion of the data collection process, a total of 417 complete surveys were gathered. Of these, 228 were collected during the first phase, and 189 were collected during the second phase. Within the entire set of 417, 69 respondents were interviewed twice, completing both versions of the questionnaires at both phases 1 and 2.

3.2.1 The discrete choice experiment

The DCE data comprise a complete set of 404 respondents. During phase 1, in which a total of 228 complete surveys were collected, 13 were excluded due to incomplete responses from participants. This led to a total of 215 valid DCE responses. During phase 2, 189 valid DCE questionnaires were obtained. Out of the total of 404, 69 respondents completed both Phase 1 and Phase 2 questionnaires, contributing to a cumulative total of 138 responses.

3.2.2 Best-worst scaling analysis

Out of the initially collected 417 questionnaires, the final sample size for the best-worst scaling data was adjusted to 416. During phase 1, a total of 228 complete surveys were collected. During phase 2, in which a total of 189 complete surveys were collected, one was excluded due to incomplete response from participant.

3.2.3 Socio-demographic data

In the socio-demographic data section, 69 were excluded because the same respondents had completed the survey twice, resulting in 348 attempted responses by different respondents. Out of the 348, 28 respondents did not complete the socio-demographic questions, leaving a total of 320 completed socio-economic sets of values.

3.2.4 Travel cost study

Out of the 320 mentioned earlier, one did not complete the fourth section, which inquired about past visits to selected urban green spaces, reducing the valid TCM data responses to 319. Additionally, 73 participants reported zero visits to Penang Botanic Gardens in the past year, while 84 participants stated that they had zero visits to Penang Youth Park during the same period. Furthermore, 54 respondents either did not provide their residential address or resided outside the study area. As a result, the travel cost study ultimately utilized a final sample that ranged between 220 and 258 observations.

3.2.5 Descriptive statistics

From Table 5, it can be observed that the gender ratio of the respondents was approximately 2:3 (male to female), and the average age of respondents was 37 years, with an age range spanning from 18 to 75 years. Around 62.5% of respondents were below the age of forty, while the remaining participants were over forty. Among the respondents, only 41% were employed either full-time or part-time, while the remaining respondents identified as unemployed, homemakers, students, or retired individuals. Approximately 43% of respondents held at least a bachelor's degree, while roughly 42% held an O-level certificate or had educational qualifications below that level. The average household size was 4 people, with a standard deviation of 1.6, and household sizes ranged from a minimum of 1 member to a maximum of 10 members. Additionally, in terms of the number of people younger than 18 living in the same household, the mean was 0.6, indicating that most households did not have any members younger than 18. However, there was a maximum of 9 household members who were younger than 18 and living in the same house. On average, respondents had been residing in Penang for 26.8 years, with a minimum duration of less than one year and a maximum of 80 years.

Table 6 displays the visitation patterns of respondents to urban green spaces within the study area over the preceding year, reflecting data from 2021. The majority of respondents (52.81%) reported exclusively visiting green spaces on weekends, while

Variable	Observations	Mean	St. Dev.	Min	Max
gender (male)	314	0.414	0.493	0	1
age	320	36.850	17.026	18	75
employed	320	0.409	0.492	0	1
educ	320	3.738	1.586	1	7
household	320	4.097	1.576	1	10
under 18	320	0.600	1.012	0	9
years in Penang	320	26.847	20.486	0	80

Table 5: Summary of Respondent Characteristics

	Categories	Freq.	%	Cum.
Days of visit	Weekdays only	29	9.06	9.06
	Weekends only	169	52.81	61.88
	Weekdays and weekends	122	38.13	100.00
	Total	320	100.00	
Visit freq	Never	21	6.56	6.56
	A few times a year	127	39.69	46.25
	A few times a month	81	25.31	71.56
	Once a week	39	12.19	83.75
	More than once a week	52	16.25	100.00
	Total	320	100.00	

Table 6: Summary of visitation patterns in urban green spaces

38.13% visited them on both weekdays and weekends. Furthermore, nearly 40% of respondents indicated that they visited green spaces a few times a year, which is the highest percentage among the five categories of visit frequency. About one-fourth (25.31%) of respondents visited green spaces a few times a month. Those who reported visiting green spaces 'more than once a week' and 'once a week' accounted for 16.25% and 12.19%, respectively. Approximately 6.56% of respondents stated that they had never visited a green space.

3.3 Conclusions

This chapter covers the design of the survey and the process of data collection, along with the presentation of descriptive statistics concerning the survey participants.

Chapter 4

Analysis of Best-Worst Scaling Data

4.1 Introduction

Urban green spaces are increasingly recognized as one of the important green infrastructures in urban areas, providing substantial benefits to the environment and urban dwellers (Bolund and Hunhammar, 1999; Antrop, 2004; Lafortezza et al., 2013). They play a comprehensive role in urban areas by enhancing people's well-being, promoting mental and physical health (Kaplan and Kaplan, 1989), improving residents' quality of life through recreation and relaxation (Madureira et al., 2011), and supporting social interaction and integration (James et al., 2009; Kabisch et al., 2015; Tzoulas et al., 2007; van den Berg et al., 2015; Tyrväinen et al., 2005; Lafortezza et al., 2009).

Urban green spaces can be presented in different ways and have different characteristics, such as the levels of naturalness (Cilliers et al., 2013), the types of facilities provided, and the species of vegetation (Aronson et al., 2014; Arnberger and Eder, 2015). Due to the variety of green space characteristics, it is important to understand residents' preferences and perceptions of these spaces' characteristics. This information is crucial for policymakers to manage and finance green spaces in urban areas (Lindholst et al., 2016; Schipperijn et al., 2010; Wan and Shen, 2015).

However, it is important to acknowledge that individuals have heterogeneous preferences for the characteristics of green spaces. Therefore, this study aims to establish a hierarchy of attribute importance using the Best-Worst Scaling (BWS) method, and correlate it with ecosystem services. The ultimate goal is to reduce the gap between individuals' preferences and perceptions and the future provision of urban green spaces.

There is a substantial body of literature that seeks to understand individual preferences for and/or the perception of urban green spaces in various aspects. One group of studies focuses on preferences for and/or perceptions of urban green spaces (Lo and Jim, 2012; Jim and Shan, 2013; Éva Terézia Vesely, 2007; Özgüner, 2011).

Another group of studies analyzes preferences for and/or perceptions of urban green space characteristics. Adinolfi et al. (2014) reported that safety, accessibility, facilities, the degree of naturalness, and flora and fauna species richness affect individuals' preferences to visit a green space. Bertram and Rehdanz (2015a) found that animal and plant species richness, biological diversity, cleanliness, and the existence of a playground received higher scores in relation to the importance of park characteristics.

Bullock (2008) studied the values of parks with different attributes such as size, maintenance, the abundance of trees, the presence of water elements, distance from home, and the types of facilities provided, including playgrounds, paths, benches, trails, and cycle paths. Giles-Corti et al. (2005) discovered that accessibility, distance, attractiveness, and size affect the likelihood of using open space in Perth, Western Australia. Voigt et al. (2014) found that the park's naturalness, tranquillity, accessibility, the beauty of the landscape, the provision of shaded areas, facilities, and proximity to water received higher scores in relation to the importance of park characteristics.

Zhang et al. (2013) investigated the landscape characteristics in urban green spaces that are preferred for recreational activities. Additionally, some studies investigate preferences for and/or the perception of the provision of ecosystem services in urban green spaces or neighbourhoods. Bonaiuto et al. (2003) measured the perceived environmental qualities of neighbourhoods, such as air and noise conditions, and the provision of recreational services on the city of Rome.

Jim and Chen (2006a) studied the perception of 25 ecosystem services provided by urban green spaces in Guangzhou, China, and found that oxygen release, aesthetic enhancement, noise abatement, carbon sequestration, and air pollution absorption are among the five most important ecosystem services.

While these research studies aim to identify individual preferences for urban green space attributes, they have yielded diverse results. The variations in observations can be attributed to differences in research methodologies, including the use of focus groups, face-to-face interviews, or questionnaire surveys (Madureira et al., 2018; Kabisch et al., 2015). Moreover, research studies reveal diverse findings, and it is crucial to recognize that these outcomes may be influenced by socioeconomic factors. This implies that the composition and dynamics of the population should guide future planning for green areas. For example, an aging population may necessitate different park attributes. Similarly, a more educated population might have distinct preferences. Therefore, it is essential to take these factors into account when interpreting observations related to preferences for urban green spaces.

Following the focus group discussions and their analysis, this study developed a best-worst scaling method to explore urban residents' preferences for selected urban green space characteristics in Penang, Malaysia. The main research question is: How can the utilization of best-worst scaling studies contribute to the exploration of individuals' preferences for urban green space characteristics? This research question is segmented into two sub-questions: (1) What are the most and least important attributes influencing individuals' decision to visit a green space? (2) How consistent and reliable are preference representations for urban green space characteristics across different designs of best-worst scaling question sets when new attributes are included? The first research question explores the most and least important attributes that individuals perceive through responses to best-worst scaling questions. An attribute that is perceived as important affects individuals' probability of visit to green sites (everything else equal). The second research question assesses the consistency and reliability of individuals' preferences for urban green space attributes. It explores fragility of the hierarchy of attribute importance when additional attributes are included.

The chapter is structured as follows. Section 4.2 describes the best-worst scaling study design. The survey design for the best-worst scaling questions is addressed in Section 3.1.2, and the data collection process is covered in both Section 3.2 and Section 3.2.2. The methods of study are discussed in Section 2.3. Section 4.3 reports the results of the best-worst scaling empirical analysis. Section 4.4 discusses the key findings of this study. Section 4.5 concludes the findings of the research. The results have implications for the planning, funding, design, and management of future urban green spaces in Penang, by investigating the importance level of green space characteristics and hence focusing on the most important characteristics to improve benefits to the wider community.

4.2 Study design

After the focus group study had been conducted, the contribution of participants in the focus groups led to the short-listing of thirteen attributes: *distance; facilities; cleanliness; size; animal species and ecosystems; tree species and ecosystems; noise levels; air quality; landscape naturalness; landscape variety;* and *security*. These characteristics and ecosystem services were used to adopt the best-worst scaling case 1 (the object case) study to measure an individual's preference for urban green space characteristics in Penang, Malaysia. In addition to the mentioned attributes, two additional attributes, namely *nursery habitat maintenance* and *education and research*, were included in the study. These attributes were recommended by the SEEA EA framework and have received limited attention in previous literature regarding the valuation of urban green spaces.

The details of the study area are discussed in Section 1.3. Individuals were presented with a series of hypothetical scenarios with different urban green space characteristics. They were asked to choose the best and the worst attributes in each scenario.

The survey was conducted in two phases. In the first phase, thirteen attributes were included in the best-worst scaling questions: *distance; facilities; cleanliness; size; animal species and ecosystems; tree species and ecosystems; noise levels; air quality; landscape naturalness; landscape variety; security; nursery habitat maintenance;* and *education and research*. In the second phase, seven attributes were excluded, leaving six attributes to be evaluated: *distance; facilities; tree species and ecosystems; noise levels; air quality;* and *nursery habitat maintenance*. These attributes are in line with the attributes used in the DCE in the same survey (see Section 3.1.1 for details), which aims to estimate the economic values of ecosystems and ecosystem services.

In designing the DCE, I included facilities as an attribute and excluded cleanliness and safety. The primary objective of this research is to examine the ecosystem services provided by urban green spaces, such as recreational opportunities and aesthetic value. Facilities are policy-relevant because they can be directly influenced and improved through urban planning and investment. Understanding the preferences for different types of facilities can inform policymakers on how to allocate resources and design green spaces. In Penang, green spaces are differentiated by the facilities they offer, attracting different user groups. For instance, spaces with sports facilities draw children and young adults, while areas with more natural features appeal to those seeking tranquility. Including facilities helps us understand these interactions and preferences. While cleanliness and safety are important, they are often baseline expectations for all green spaces. We assumed these factors to be consistent across all options, allowing us to focus on more variable attributes like facilities. Additionally, cleanliness and security are difficult to parameterize in levels, making them challenging to include in the analysis.

In addition to the primary BWS analysis for all respondents, a secondary analysis was conducted to investigate possible differences in utility among respondents who had completed both phases of the survey. In phase 1, the survey included 13 attributes, while in phase 2, these were reduced to 6 to produce more attentive responses. This analysis examined whether the number of attributes used in the choice tasks affects the hierarchy of ranks of individual's preferences. In this case, all attributes were ranked based on their utility coefficients, and an investigation was carried out to detect changes in the relative ranking of specific attributes in both versions of the survey.

4.3 Results

4.3.1 Normalized Difference Scores

A total of 228 eligible respondents completed the 13-scenario choice set during the first phase of data collection, and a total of 188 eligible respondents answered the 10-scenario choice set during the second phase of data collection, resulting in a total of 2964 scenarios and 1880 scenarios, respectively.

Table 7 summarizes the total number of times each attribute was chosen as the best and worst choices, the total number of times each attribute was presented in the question set, the difference in the number of times each attribute was chosen as the best and worst, and the normalized difference scores during the first phase of data collection. Table 8 shows the choice data for the second phase of data collection. Table 7 shows that *air quality* has the highest score, followed by *cleanliness*, *security*, and *tree species and ecosystems*. The *size* of green spaces has a considerably low score, indicating that it is the least important characteristic of green spaces. The number of attributes was reduced to 6 during the second phase of data collection. As shown in Table 8, *air quality* has the highest score, followed by *tree species and ecosystems* and *noise levels*. *Facilities* receive the lowest score, indicating that it is the least important attribute for respondents.

4.3.2 Multinomial Logistic Regression

The difference scores do not account for the errors that inevitably occurs during the decision-making process. Therefore, a multinomial logistic regression (MNL) is used to account for this. Table 9 shows the utility coefficients, standard errors, 95% lower and upper bounds (confidence intervals), and choice probabilities of attributes. The results indicate that *air quality* is perceived as the most important attribute, followed by *cleanliness*, *security*, and *tree species and ecosystems*, while *size* is considered the least important attribute. These results align with the simple difference scores, which typically exhibit a high correlation with each other (Marley et al., 2016; Marley and Louviere, 2005). The sixth column of table 9 shows that *air quality* has the highest choice probability score, indicating that it is the most likely attribute to be chosen as the most important from the question set, while *size* has the lowest probability score,

Attribute	N_i^{total}	N_i^{best}	N_i^{worst}	$N_i^b - N_i^w$	$NDiff_i$	Rank
air quality	912	520	48	472	0.518	1
cleanliness	912	357	96	261	0.286	2
security	912	307	112	195	0.214	3
tree species and ecosystem	912	335	150	185	0.203	4
landscape naturalness	912	231	128	103	0.113	5
animal species and ecosystem	912	232	179	53	0.058	6
nursery habitat maintenance	912	180	145	35	0.038	7
noise level	912	229	200	29	0.032	8
facilities	912	152	264	-112	-0.123	9
landscape variety	912	152	270	-118	-0.129	10
distance	912	158	400	-242	-0.265	11
education and research	912	95	346	-251	-0.275	12
size	912	16	626	-610	-0.669	13

Table 7: Normalized difference scores of the attributes (first phase survey)

Attribute	N_i^{total}	N_i^{best}	N_i^{worst}	$N_i^b - N_i^w$	$NDiff_i$	Rank
air quality	940	717	20	697	0.741	1
tree species and ecosystem	940	310	210	100	0.106	2
noise level	940	289	277	12	0.013	3
distance	940	243	453	-210	-0.223	4
nursery habitat maintenance	940	183	440	-257	-0.273	5
facilities	940	138	480	-342	-0.364	6

Table 8: Normalized difference scores of the attributes (second phase survey)

Attribute	Coef.	Std. err.	L	U	Prob.	Rank
air quality	1.146	0.055	1.039	1.253	0.204	1
cleanliness	0.589	0.049	0.493	0.685	0.117	2
security	0.434	0.048	0.34	0.528	0.1	3
tree species and ecosystem	0.411	0.048	0.318	0.505	0.098	4
landscape naturalness	0.227	0.047	0.134	0.319	0.081	5
animal species and ecosystem	0.116	0.047	0.024	0.208	0.073	6
nursery habitat maintenance	0.077	0.047	-0.015	0.169	0.07	7
noise level	0.064	0.047	-0.028	0.155	0.069	8
facilities	-0.247	0.047	-0.339	-0.154	0.051	9
landscape variety	-0.26	0.047	-0.353	-0.168	0.05	10
distance	-0.544	0.049	-0.639	-0.449	0.038	11
education and research	-0.565	0.049	-0.66	-0.47	0.037	12
size	-1.617	0.063	-1.741	-1.494	0.013	13

Table 9: Multinomial logistic regression results (first phase survey)

suggesting that it is the least likely attribute to be chosen as the most important in the question set.

Table 10 presents the results of the MNL analysis using the best-worst scaling data from the second phase of data collection. Among the six attributes, *air quality* is identified as the most important, followed by *tree species and ecosystems* and *noise levels*. *Facilities* are perceived as the least important, followed by *nursery habitat maintenance* and *distance*. Comparing the results shown in tables 9 and 10, the rankings of the six utility coefficients have changed. *Air quality* and *tree species and ecosystems* remain the two most important attributes, while *nursery habitat maintenance*, which was ranked third in table 9, is ranked fifth in table 10. *Noise levels*, which was ranked fourth before, is ranked third in the latter table. *Facilities*, which was ranked fifth before, is ranked sixth in the later table. Lastly, *distance*, which was ranked sixth before, is ranked fourth later.

4.3.3 Secondary analysis: respondents with two attribute evaluations

The secondary analysis is based on data from 69 respondents who completed both phases of the best-worst scaling questions – the complete (first phase) and the reduced attribute set (second phase). The estimation results using Multinomial Logit (MNL) are presented in Tables 11 and 12.

Attribute	Coef.	Std. err.	L	U	Prob.	Rank
air quality	1.908	0.069	1.773	2.042	0.631	1
tree species and ecosystem	0.214	0.046	0.123	0.304	0.116	2
noise level	0.026	0.046	-0.065	0.116	0.096	3
distance	-0.454	0.047	-0.547	-0.362	0.059	4
nursery habitat maintenance	-0.561	0.048	-0.655	-0.467	0.053	5
facilities	-0.763	0.05	-0.86	-0.666	0.044	6

Table 10: Multinomial logistic regression results (second phase survey)

Attribute	Coef.	Std. err.	Lower	Upper	Prob.	Rank
air quality	1.496	0.11	1.281	1.712	0.267	1
distance	-0.661	0.09	-0.837	-0.485	0.031	6
facilities	-0.366	0.087	-0.536	-0.197	0.041	5
noise levels	0.094	0.085	-0.073	0.261	0.066	4
nursery habitat maintenance	0.218	0.086	0.05	0.386	0.074	3
tree species and ecosystems	0.366	0.087	0.197	0.536	0.086	2

Table 11: Secondary analysis MNL regression results (first version survey)

Table 11 does not include seven attributes that were not estimated in the second phase survey. Therefore, the six attributes presented in both tables are comparable. *Air quality* and *tree species and ecosystems* emerged as the two most important attributes in both phases of the survey. Nevertheless, the hierarchy of importance for *distance, facilities, noise levels*, and *nursery habitat maintenance* varied between the two phases, highlighting the instability of attribute importance hierarchy when additional attributes were introduced into the question sets.

Attribute	Coef.	Std. err.	Lower	Upper	Prob.	Rank
air quality	1.91	0.114	1.687	2.133	0.631	1
distance	-0.339	0.077	-0.491	-0.188	0.067	4
facilities	-0.766	0.082	-0.926	-0.605	0.043	6
noise levels	0.087	0.076	-0.062	0.236	0.102	3
nursery habitat maintenance	-0.693	0.081	-0.851	-0.535	0.047	5
tree species and ecosystems	0.163	0.076	0.013	0.312	0.11	2

Table 12: Secondary analysis MNL regression results (second version survey)

4.4 Discussions

In the best-worst scaling results of the first phase, the five most important attributes of urban green spaces included three from the category of ecosystems and ecosystem services: *air quality*, *tree species and ecosystems*, and *landscape naturalness*. Notably, *air quality* emerged as the top-ranked attribute, a result aligned with the views expressed by participants in focus group discussions. This ranking reflects the high demand for improved air quality, a particularly important factor among urban residents grappling with persistent air pollution issues. Urban green spaces serve as vital sources of fresh air in urban areas, which explains why people specifically seek them out to breathe in cleaner air, as emphasized in the focus group discussions.

It is unsurprising that *tree species* and *ecosystem and landscape naturalness* would receive high rankings, given their close relationship with *air quality*. The presence of trees and natural landscapes forms the very essence of green spaces. A green space would be incomplete without either of these two attributes.

Interestingly, two attributes, *cleanliness* and *security*, which are not typically categorized under ecosystems and ecosystem services, have emerged among the top three most important attributes overall. This suggests that park visitors prioritize attributes that offer direct and immediate benefits over those related to ecosystems and ecosystem services, which may not provide such immediate advantages. As a result, the cleanliness and security standards of green spaces play pivotal roles in influencing people's decisions to visit them.

During the focus group discussions, participants expressed disappointment with the current lower standards of cleanliness and security in Penang's urban green spaces. Consequently, these two attributes hold significant value for visitors, as they aspire to witness improvements in these areas. These findings hold considerable implications for the planning and development strategies of urban green spaces in Penang, and especially for George Town residents.

Among the 13 attributes in the best-worst scaling analysis, *animal species and ecosystems* in green spaces were ranked sixth, indicating a moderate level of importance to visitors. While animal species and their ecosystems play a significant role in sustainable development, their presence in urban green spaces is not always favoured due to potential safety concerns for visitors, as mentioned in the focus group discussions. In contrast, the *nursery habitats maintenance* attribute was ranked seventh, slightly less important than *animal species and ecosystems*. Surprisingly, *noise levels* has a lesser impact on people's preferences compared to other attributes in the ecosystems and ecosystem services category. Although green spaces generally have lower noise levels, it is worth noting that if a green space is located adjacent to or amidst busy roads, higher noise levels can be expected. During the focus group discussions, participants expressed complaints about the noise generated by cars at Bukit Dumbar Park, one of the prominent urban green spaces in George Town. However, it should be acknowledged that this noise issue cannot be easily mitigated due to the park's proximity to a major road. Therefore, based on people's preferences, the noise levels are considered less important, perhaps showing a form of adjustment to background noise by urban dwellers.

The rankings also indicate a relatively low level of importance placed on *facilities*, *landscape variety*, *distance*, *education and research*, and *size*. These attributes were more frequently chosen as the least important rather than the most important, as evidenced by the negative signs on their utility coefficients. *Facilities* and *landscape variety* of green spaces received nearly equal utility rankings. Results from a follow-up survey of focus groups reveal that while certain community facilities, like public toilets, parking lots, and footpaths are considered essential in recreational green spaces, other recreational facilities such as playgrounds, swimming pools, basketball courts, and tennis courts appear to be slightly less important. This suggests that visitors tend to prefer natural environments over artificial ones, as reflected in the high rankings of *tree species and ecosystems* and *landscape naturalness*. However, *landscape variety*, despite offering numerous ecosystem services, does not provide a high level of utility.

Two attributes with low utilities are *distance* and *education and research*, both of which have nearly the same utility level. The *distance* to a park is typically associated with travel costs, where a shorter distance implies lower costs. However, compared to most other attributes, *distance* appears to hold minor importance. This can be explained by the relatively low travel costs, primarily driven by low vehicle operation costs. For instance, as mentioned in Chapter 5, the average travel cost for visiting Penang Botanic Gardens is RM1.00, and RM0.98 for visiting Penang Youth Park, without factoring in the value of time. The *education and research* attribute falls within the category of ecosystem services. The low level of importance indicates that people do not expect to receive significant educational or research-related benefits when visiting urban green spaces. Among the 13 attributes, *size* is considered the least important, suggesting that the size of urban green spaces does not significantly influence people's preferences.

In the second phase of the best-worst scaling analysis, only six attributes were included. Four of these attributes belong to the group of ecosystems and ecosystem services, while the remaining two attributes are *distance* and *facilities*. The rankings reveal that the importance level of *air quality* and *tree species and ecosystems* remains relatively stable, while the rankings of other attributes change after first round of survey. In comparison to the first analysis, *noise levels* and *distance* now hold higher ranks, indicating an increased preference among people for lower noise levels and shorter distances. On the other hand, *nursery habitat maintenance* and *facilities* show lower ranks, suggesting a decreased emphasis on these attributes. This implies that people are giving higher priority to reducing noise and travel distances, as opposed to prioritizing better nursery habitat maintenance and improved facilities.

To examine changes in people's preferences over time and across different choice task designs, a secondary analysis was conducted, focusing on a specific group of respondents who completed both sets of surveys. The results demonstrate that the rankings of attributes from this group of respondents for both versions of the survey are similar to the rankings derived from the full sample. Therefore, it is observed that the rankings of attributes changed after two phases of the survey. These findings suggest that the preferences of this group of respondents for attributes are unstable when the number of attributes presented to them changes. Two possible interpretations for these findings can be considered. Firstly, it is possible that this group of respondents became confused when presented with many attributes simultaneously, leading to potentially inaccurate decisions. Secondly, this group of respondents became more familiar with the question sets and had more time to contemplate their answers between completing the first and second surveys. However, further studies are recommended to determine the underlying cause of this instability.

4.5 Conclusions

This study assesses the degree of importance associated with attributes of urban green spaces. Additionally, this study examines the stability of attribute rankings over different designs of question sets and over time by conducting two phases of BWS survey.

The findings of this study hold significant policy implications, offering valuable insights for urban planners, relevant authorities, and green space managers involved in decision-making regarding the development of urban green spaces in the specific context of Penang Island. However, it is essential to acknowledge several limitations in this study.

Firstly, this study relies on a single case study approach. Despite conducting a two-phase survey with different designs, the use of the same experimental design (BIBD) across both phases means that the results presented here are specific to the selected case study and experimental design for analysis. Consequently, this limits the ability to conduct further analysis of the potential impact of different experimental designs, which could play a crucial role in terms of experimental design efficiency. While this study does not address this issue, it is recommended that future research explore various experimental designs to analyze the data from different perspectives.

Moreover, the two-phase surveys, differing in the number of attributes and scenarios, were conducted at different points in time, despite a maximum time lapse of only four months (valid for only the 69 respondents who completed both surveys). This raises questions about whether the inconsistency in attribute rankings among this group of respondents is attributable to the design of the BWS questions or the time gap between the surveys. This group of respondents may have been less familiar with the question sets during the first survey and may not have had sufficient time to consider their responses. In contrast, during the second survey, they might have become more familiar with the questions and had more time to reflect, potentially resulting in more accurate responses. Consequently, attributing the inconsistency in attribute ranking solely to the design of the BWS questions may introduce bias and may not provide a complete understanding of the underlying factors. It is advisable for future research to investigate similar studies with minimal time gaps or even by fielding two surveys simultaneously to gain a clearer understanding of the factors contributing to ranking inconsistencies, and eliminating time effects.

Chapter 5

Analysis of Travel Cost Data

5.1 Introduction

Urban green spaces, such as parks, gardens, and recreational areas, have become increasingly valuable assets in rapidly urbanizing areas as they contribute to the well-being of city dwellers. To sustainably maximize the benefits of urban green spaces, effective planning, financing, and management are essential. Accurate valuation of these spaces is crucial for making informed decisions. Valuing urban green spaces requires a deep understanding of visitors' and residents' preferences and behaviours related to these areas. Various non-market valuation methods have been employed to assess the economic value of urban green spaces. These methods include the travel cost method (TCM) (Karunarathne and Gunawardena, 2020; Chintantya and Maryono, 2018), the contingent valuation method (Jim and Chen, 2006b; Chintantya and Maryono, 2018; Tyrväinen and Väänänen, 1998), choice experiments (Roberts et al., 2022), and the hedonic price method (Chang et al., 2017; Engström and Gren, 2017; Trojanek et al., 2018; Lategan et al., 2022; Panduro and Veie, 2013; Samad et al., 2020).

This chapter employs the TCM to estimate the values of two major urban green spaces in Penang, Malaysia: Penang Botanic Gardens (latitude: 5.43786; longitude: 100.29101, land area: approximately 2.41 square kilometres), Penang City Park (latitude: 5.43189; longitude: 100.29732, land area: approximately 0.16 square kilometres). The TCM has a long history of use for estimating the non-market value of environmental goods and services. It has been applied to estimate the non-market values of various resources, including rock climbing (Shaw and Jakus, 1996), forests (Willis and Garrod, 1991), lakes (Fleming and Cook, 2008), beaches (Zhang et al., 2015), urban open spaces (Hanauer and Reid, 2017), urban forests (Bertram and

Larondelle, 2017; Dwyer et al., 1983; Wanyu et al., 2014), and natural areas (Ezebilo, 2016). The TCM has also been used in numerous studies to value urban open spaces and recreational sites in different contexts, such as recreational value, public park services, and individual benefits (Hanauer and Reid, 2017; Iamtrakul et al., 2005; Liu et al., 2014).

In this context, the values of the two urban green spaces in Penang are linked to the costs associated with accessing these sites, including monetary costs, travel time, and recreation time components. By integrating the value of time into valuation models, the estimation results can offer better insights into the values that individuals assign to urban green spaces. Several studies have explored the validity of both time components within the framework of the TCM (Bockstael et al., 1987; Cesario, 1976; Lloyd-Smith et al., 2020a; McConnell, 1992; Smith et al., 1983). However, the methods used to quantify these components can vary significantly, depending on the specific approach and data availability. This chapter aims to investigate the impact of different time valuation methods on individuals' visit preferences and behaviours. Therefore, two assumptions were made when estimating the value of time: (1) All visitors have a non-zero value of time, and (2) Only employed visitors have a non-zero value of time. By incorporating the time component using different methods, the total value that individuals ascribe to these urban green spaces may vary.

In this context, three research questions have been identified: (1) Do travel costs to a green site contribute to explaining an individual's visit preferences and behaviours? (2) Do the preferences for ecosystem assets and services at green sites affect visitors' preferences? (3) How reliable are the estimates of travel and recreation time values derived from travel cost models? The first research question is raised to explore the relationship between travel costs and individuals' visit patterns for visiting green sites, given that the willingness to pay (WTP) for travel expenses is significant for valuing a green site. Travel cost models were developed to examine the validity of travel costs in real situations. The second research question explores the connections between individuals' perceptions of ecosystem assets and services and visit preferences through travel cost models. The perceptions of ecosystem assets and services were assessed using individual coefficients for *air quality* and *tree species and ecosystems*. These coefficients were estimated from the Discrete Choice Experiment (DCE). Since respondents answered both travel cost and DCE questions in the same survey, it is possible to identify individual coefficients corresponding to each respondent. The DCE data is represented in the Willingness-to-Pay (WTP) space, with all coefficients being randomly and normally distributed. The estimation was conducted in R using the *Apollo* package, and the code is provided in Appendix C. The third research question examines the reliability of time values of travel time and recreational time when they are added to the travel costs. It is expected that the out-of-pocket travel costs do not sufficiently express the whole picture of visit preference, therefore, the individual-specific values of time were estimated in travel cost models to investigate their validity in real and hypothetical situations.

This study is structured as follows. Section 5.2 introduces the study areas of the travel cost analysis case study, namely Penang Botanic Gardens and Penang Youth Park. The survey design for the travel cost study was outlined in Section 3.1.4, and the details of the data collection process were elaborated in both Section 3.2 and Section 3.2.4. In this study, sample sizes ranging from 220 to 258 observations are utilized. The methods of study were discussed in Section 2.4. Section 5.3 describes the travel cost study design. In this section, Section 5.3.1 introduces the travel cost model specification. Section 5.4 presents the findings of the study. Section 5.5 discusses the key findings of this study. Section 5.6 concludes the findings and discusses the limitations of the study. The findings from this study will provide insights into urban green space planning and management, with the aim of improving social benefits.

5.2 Case study areas

Penang Botanic Gardens

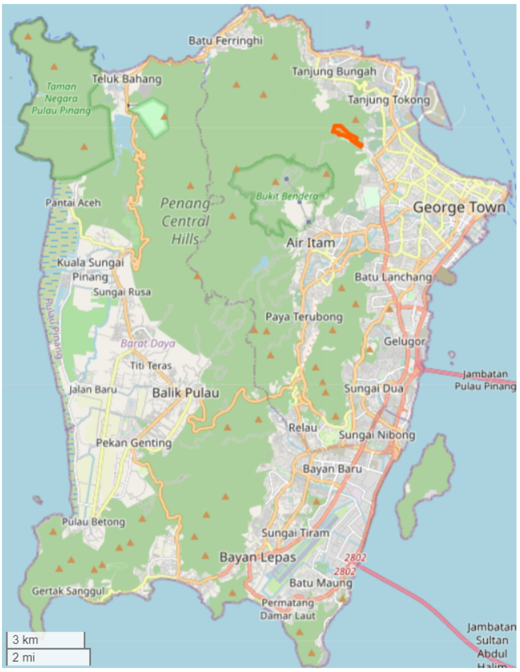
The study areas for the travel cost study are Penang Botanic Gardens and Penang Youth Park. These two parks are situated in the northeastern part of Penang Island, as depicted in Figure 6 (OpenStreetMaps, 2023a). Penang Botanic Gardens is also known as the 'Waterfall Gardens' due to its proximity to a waterfall. Established in 1884, it is one of the oldest green spaces in Penang.

The primary functions of Penang Botanic Gardens include supporting conservation and education programs, as well as providing a recreational environment. The Gardens collaborate with various research institutes in the country, engaging in botanical and ecological research. Additionally, Penang Botanic Gardens serves as a popular recreational site and tourist attraction, acting as a green oasis in the heart of urban Penang. It boasts a diverse array of flora and fauna and enjoys a high visitation rate, attracting city dwellers for activities such as jogging, walking, hiking, and pic-nicing. Figure 7a shows the structure of Penang Botanic Gardens (OpenStreetMaps, 2023b). The gardens cover a total area of 241.2 hectares, 30 hectares of which are hosting infrastructure, which are made up of numerous theme gardens which collect rare plant species, such as 'Secret Garden', 'Cactus House', 'Herbs Garden', 'Sunken Garden', 'Japanese Garden', 'Fern House (*Rumah Paku-Pakis* in Malay)', 'Economic Garden', 'Bambusetum' etc. In addition to these features, there is a section for plant nurseries situated in the heart of the Gardens, and a river known as the 'Waterfall River' flows within the premises (Penang Green Council, 2020b). Furthermore, the Gardens are interconnected with several hiking trails, including the 'Rifle Range Trail', 'Bukit Cendana Trail', and 'Moongate Trail'. These trails lead to Penang Hill and Mount Olivia, which are located in the northern part of Penang Island.

Penang Youth Park

Penang Youth Park, also referred to as the 'City Park', is located near Penang Botanic Gardens. The park is maintained by the Penang Island City Council and was established in 1972 with the primary purpose of serving as a recreational site and tourist destination. Like Penang Botanic Gardens, Penang Youth Park is nestled in the heart of Penang's urban area and serves as a green oasis for city residents. Consequently, it is a popular destination for Penang residents seeking recreational activities. While Penang Botanic Gardens is renowned for its collection of flora and fauna, Penang Youth Park is known for its wide range of public recreational amenities that cater to both parents and children. The 16-hectare park contains a skating rink, a bridal corner (*laman mempelai* in Malay), swimming pools (*kolam rekreasi* in Malay), jungle trails, giant chess boards (*laman catur* in Malay), street arts, playgrounds, graffiti corner, reflexology track, outdoor stage (*panggung terbuka* in Malay) and multi-purpose field. In addition to its public facilities, the park is also home to the Penang Graffiti Park, showcasing graffiti art to visitors. The park's layout and structure are depicted in Figure 7b (OpenStreetMaps, 2023c). Beyond its recreational amenities, the park is linked to a hiking trail known as the 'Point 3 Trail', which connects to the 'Moongate Trail'. Furthermore, within Penang Youth Park, you can find the offices of two Penang State City Council departments: the Community Service Department and the Landscape Department, both of which have made the park their official home. Moreover, the park serves as the headquarters for numerous sports organizations, including archery clubs, hiking clubs, Tai Chi clubs, and aerobic clubs.

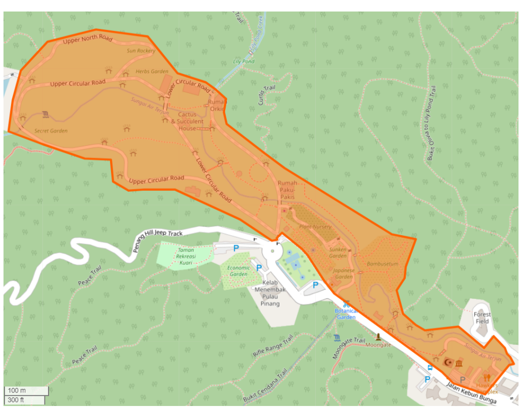
These green sites offer opportunities to host a substantial number of ecosystem assets, including plants, trees, and wildlife, which contribute to the provision of various ecosystem services, such as climate regulation, water regulation, air filtration, noise regulation, nursery habitat maintenance, visual amenities, and recreational opportunities.



(a) Penang Botanic Gardens

(b) Penang Youth Park

Figure 6: Location of two urban green spaces in George Town



(a) Penang Botanic Gardens

(b) Penang Youth Park

Figure 7: Structure map of urban green spaces

5.3 Study design

This study employs the TCM to develop a demand model for green sites and estimate the recreational value of both Penang Botanic Gardens and Penang Youth Park, which are the two largest urban green spaces in Penang. By examining these two sites simultaneously, this chapter aims to explore their potential as substitute sites for each other.

Additionally, this study investigates potential biases introduced by various methods of measuring the value of time in the models. Among the most significant opportunity costs considered are the time spent travelling to and from the site and the time allocated for recreational activities on-site. Therefore, the accuracy of measurements for the value of time becomes crucial for this investigation.

5.3.1 Model specification

Derivation of travel cost variables

In a travel cost model, the primary independent variable is travel costs. The computation of individual-level travel costs for recreational visits typically involves considering both the monetary costs of travelling to the green spaces and the non-monetary opportunity costs of the visit (Lloyd-Smith et al., 2019). These opportunity costs encompass the income forgone for the time spent travelling to and from the green areas and, to a lesser extent, the time spent actually visiting the recreational site (Brainard, 1999; Hanauer and Reid, 2017).

In this study, travel costs were divided into three components: 1) the monetary costs associated with vehicle operation (VC_n), 2) the non-monetary opportunity costs related to travel time (OTC_n), and 3) the non-monetary opportunity costs associated with recreational time (ORC_n). It is important to note that these travel costs do not include entrance fees and parking fees, as no such fees are charged when visiting urban parks or gardens in Penang.

To calculate vehicle operation costs (VC_n), information about respondents' home addresses and the modes of transport used to reach the destinations was collected through the survey. The home address served as the point of origin for each respondent's trip, and it was used to calculate the distance travelled to the destinations with the assistance of Google Maps. The round-trip vehicle operation costs are expressed as:

$$VC_n = \frac{2 \times voc_m}{nop_n}, \tag{5.1}$$

where nop_n is the number of people visiting together, and voc_m is the single-trip vehicle operating cost that varies depending on the mode of travel (m),

$$\text{voc}_m = \begin{cases} 0.205 \times \text{KM}_n, & \text{if } m = \text{car} \\ 0.8 \times \text{KM}_n + 1, & \text{if } m = \text{taxi/e-hailing} \\ 0.04 \times \text{KM}_n, & \text{if } m = \text{motorbike} \\ 0, & \text{if } m = \text{bicycle} \\ 1.40, & \text{if } m = \text{bus} \\ 0, & \text{if } m = \text{walking}, \end{cases} \tag{5.2}$$

where KM_n represents the travel distance from the origin to the destination. The vehicle operation costs were estimated based on average fuel consumption of 0.1 litre per kilometre for cars and 0.02 litre per kilometre for motorcycles. In Malaysia, unleaded 95 petrol is priced at RM2.05 per litre. Additionally, the average cost of taking a taxi is RM0.80 per kilometre, with an additional base fare of RM1.00. Lastly, the one-way bus fare for travel within Penang is RM1.40.

The opportunity costs of travel time were derived based on the travel time on the road and the per-hour value of time of the visitor. Travel time spent on the road was determined using spatial information collected from the survey, and two different per-hour value of time calculations were applied (OTC_n and $AOTC_n$).

The first per-hour value of time (wc_n) was computed by taking one-third of the expected hourly income and multiplying it by the total round-trip travel time (Champ et al., 2003; Hanauer and Reid, 2017; Kinghorn et al., 2014; Mendes and Proença, 2011). To calculate expected hourly income, median salaries and wages data for different age groups provided by the Department of Statistics Malaysia (Department of Statistics Malaysia, 2021) was used. In the calculation process of converting time into a monetary value, secondary data on income by age groups is applied to relevant respondents. The decision to not collect primary data on respondents' income was made to streamline data collection processes and minimize respondent burden. By using secondary data on income, the respondents' income can be efficiently estimated without imposing additional survey questions. However, this approach may introduce uncertainties and potential inaccuracies in income estimation, which could affect welfare estimates. This method was applied only to respondents who reported being employed, either full-time or part-time. For respondents who indicated being unemployed, homemakers, students, or retired, a zero value of time was assigned. This is obviously an underestimate as these people would have opportunity cost of time, but unpriced by the market.

Therefore, the second calculation method assumed that all respondents, regardless

Age	Median income per month (RM)	Median income per hour (RM)
15-24	1468	9.18
25-34	2001	12.51
35-44	3077	19.23
45-54	3061	19.13
55-64	2760	17.25

Table 13: Median income by age

of employment status, had a positive value of time. Similar to the first method, respondents who reported being unemployed, homemakers, students, or retired were assigned an expected hourly income based on the median income for different age groups (Table 13). The hourly value of time (awc_n) was then calculated as one-third of the expected hourly income multiplied by the total round-trip travel time.

The travel time opportunity costs (OTC_n and $AOTC_n$) are expressed as

$$OTC_n = \frac{wc_n}{60} \times ot_{nm} \times 2, \tag{5.3}$$

$$AOTC_n = \frac{awc_n}{60} \times ot_{nm} \times 2, \tag{5.4}$$

where wc_n and awc_n represent the per-hour value of time for individual n, and these values were divided by 60 to obtain the per-minute value of time. Meanwhile, ot_{nm} represents the single-trip travel time in minutes for individual n, which varies depending on the mode of travel (m).

The calculation of recreational time opportunity costs (ORC_n) was based on the time spent by respondents during their visits to the green spaces, as reported in the survey. Similar to the computation of travel time costs, two different approaches were used for recreational time costs. The first approach assumed that only respondents with full-time or part-time employment had a value of time, while the second approach considered that all respondents had a value of time. The formulas for recreational time opportunity costs (ORC_n and $AORC_n$) are as follows:

$$ORC_n = \frac{wc_n}{60} \times or_n, \tag{5.5}$$

$$AORC_n = \frac{awc_n}{60} \times or_n, \tag{5.6}$$

where or_n is the amount of time spent (in minutes) on the green spaces for each individual n.

Travel costs	Equations
TC1	$TC_n = VC_n$
TC2	$TC_n = VC_n + AOTC_n$
TC3	$TC_n = VC_n + AOTC_n + AORC_n$
TC4	$TC_n = VC_n + OTC_n$
TC5	$TC_n = VC_n + OTC_n + ORC_n$

Table 14: Equations of travel costs

There were five versions of return-trip travel costs *(TC1 – TC5)* calculated for each visitor (Table 14). *TC1* includes only the vehicle operation cost. *TC2* includes the vehicle operation cost and the opportunity costs for travel time, assuming that all respondents have a value of time ($AOTC_n$). *TC3* combines the costs of *TC2* with the opportunity costs for recreational time ($AORC_n$), again assuming that all respondents have a value of time. *TC4* considers both the vehicle operation cost and the opportunity costs for travel time (OTC_n). Lastly, *TC5* encompasses the costs of *TC4* along with the opportunity costs for recreational time (ORC_n), with the assumption that only individuals who are employed have a value of time.

Other variables in the model

In this study, two models were developed. The first model has the number of visits to Penang Botanic Gardens in the past year as the dependent variable, while the independent variables include Travel Cost, Age, Education Level, Employment Status, Household Size, Years in Penang, Number of Visits to Substitute Site, Travel costs for those who visited urban green space a few times in the past year, and Perceptions of Ecosystems and Ecosystem Services. The second model, on the other hand, has the number of visits to Penang Youth Park in the past year as the dependent variable, with independent variables similar to those in the first model. Information on these variables was collected through the survey questionnaire.

From Table 15, five measures of travel costs were generated for each individual (*TC1-TC5*). The respondent's *age* was calculated as the difference between the year of completing the survey and the year of birth as reported in the survey. The variable *educ* represents the respondent's highest education level, divided into seven categories based on their highest education levels. Each category was assigned a numerical value, reflecting a hierarchical order. The *employed* variable is a binary variable with a value of *1* for individuals who work either full-time or part-time, and *0* for those with other employment statuses. *Household* records the household size, indicating the number of people in the respondent's household. *Years in Penang* denotes the number of years the respondent has lived in Penang. Lastly, the variable *pbg visit/youth visit* represents the number of visits to the substitute site, which is Penang Youth Park for the first model and Penang Botanic Gardens for the second model.

Moving on to the *perceptions of ecosystems and ecosystem services*, this aspect of the analysis aimed to capture respondents' perceived importance levels of ecosystems and ecosystem services. Specifically, this study focused on the importance of two attributes: *tree species and ecosystems* and *air quality* in green spaces. These two attributes ranked as the most important among all others in BWS survey, as evident from their high utility coefficients and rankings (refer to Tables 7, 8, 9 and 10 in Chapter 4). To derive the importance levels, this study relied on respondents' choices in the DCE questions, which were discussed in Section 3.1.1. The individual-level coefficient estimates for these two attributes (Model 2 in Table 30 in Chapter 6) were used as an estimator of the independent variable *perceptions of ecosystems and ecosystem services* in this chapter. Both attributes, *air quality* (with levels *air1* and *air2*) and *tree species and ecosystems* (with levels *tree1* and *tree2*), each with two levels, were simplified for ease of interpretation. The coefficient estimates for *air1* and *air2* were aggregated into a single variable, as were those for *tree1* and *tree2*. This process resulted in two sets of sub-variables: *tree species and ecosystems* and *air quality*.

Variables	**Variables**	**Description**
Y_{pbg}	pbg visit	Number of visits to Penang Botanic Gardens in the past year
Y_{youth}	youth visit	Number of visits to Penang Youth Park in the past year
X_1	TC	Travel costs (TC1 – TC5)
X_2	age	Year of completing survey - Year of birth
X_3	educ	Education level: 1=primary school education; 2=secondary school education; 3=O-level or equivalent; 4=A-level/Diploma or equivalent; 5=Bachelor's degree or equivalent; 6=Master's degree or equivalent; 7=Doctoral degree or equivalent
X_4	employed	Employment status: 1=Employed either full-time or part-time; 0=Otherwise
X_5	household	Household size: number of people in the household
X_6	pbg visit/ youth visit	Number of visits to substitute site
	Perceptions of ecosystems and ecosystem services	Respondent's perceptions of the importance level of ecosystems and ecosystem services
X_7	tree	The importance level of tree species and ecosystems
X_8	air	The importance level of air quality
X_9	years in Penang	Number of years that the respondent has been living in Penang
X_{10}	visit freq	Respondent's visit frequency to urban green sites in the past year: 0=never; 1=a few times a year; 2=a few times a month; 3=once a week; 4=more than once a week

Table 15: Description of variables used in the model

Variable	Observations	Mean	St. Dev.	Min	Max
pbg visit	246	7.691	11.071	1	100
pbg time spent	277	87.942	44.552	15	210
pbg no. of people	272	3.438	2.293	1	20
youth visit	235	7.400	12.434	1	120
youth time spent	266	88.929	48.159	15	210
youth no. of people	259	3.498	2.569	1	20

Table 16: Summary of visitation patterns at Penang Botanic Gardens (*pbg*) and Penang Youth Park (*youth*) (exclude zero-visit participants)

Correlation analysis

Correlation analysis techniques were employed to select the most appropriate variables for the model. Since the demand model involved several categorical variables, it was essential to assess the relationships between these variables. The findings of the correlation analysis among independent variables were used to detect multicollinearity. When a strong correlation was observed between independent variables, one of them was removed from the travel demand models. Pearson's correlation coefficient and Spearman's rank correlation analysis were used for this purpose.

The correlation analysis results are reported in Appendix D. In both tests, a strong correlation was found between *age* and *years in Penang* ($r > 0.7$). Consequently, the variable *years in Penang* was excluded from the travel demand model analysis. Additionally, the variable *visit freq*, being closely related to the dependent variable, was also excluded.

Travel cost models

As a result, the travel cost model is expressed as:

$$\begin{aligned} Y_{pbg}/Y_{youth} = \exp\big(&\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 \\ &+ \beta_6 X_6 + \beta_7 X_7 + \beta_8 X_8\big) + \mu_n \end{aligned} \tag{5.7}$$

5.4 Results

5.4.1 Summary of visit behaviour

		Penang Botanic Gardens			Penang Youth Park		
	Categories	Freq.	%	Cum.	Freq.	%	Cum.
Mode of transport	Car	222	69.59	69.59	210	67.09	67.09
	Taxi/E-hailing	7	2.19	71.79	3	0.96	68.05
	Motorbike	41	12.85	84.64	40	12.78	80.83
	Bicycle	3	0.94	85.58	0	0.00	80.83
	Bus	3	0.94	86.52	5	1.60	82.43
	Walk	1	0.31	86.83	4	1.28	83.71
	Never visited	42	13.17	100.00	51	16.29	100.00
	Total	319	100.00		313	100.00	
Activities	Walking	137	43.08	43.08	127	39.94	39.94
	Jogging	45	14.15	57.23	37	11.64	51.57
	Hiking	23	7.23	64.47	24	7.55	59.12
	Cycling	4	1.26	65.72	2	0.63	59.75
	Play other sports or exercise	24	7.55	73.27	32	10.06	69.81
	Enjoy nature	40	12.58	85.85	38	11.95	81.76
	Meeting with friends	2	0.63	86.48	2	0.63	82.39
	Other	0	0.00	86.48	3	0.94	83.33
	Never visited	43	13.52	100.00	53	16.67	100.00
	Total	318	100.00		318	100.00	

Table 17: Summary of visitation patterns at Penang Botanic Gardens and Penang Youth Park (include zero-visit participants)

	Variable	Observations	Mean	St. Dev.	Min	Max
Penang Botanic Gardens	TC1	258	1.006	1.315	0.000	11.788
	TC2	258	4.303	3.476	0.033	28.079
	TC3	258	10.512	5.757	0.033	36.704
	TC4	258	2.643	3.240	0.000	28.079
	TC5	258	5.615	6.634	0.000	36.704
Penang Youth Park	TC1	247	0.978	1.940	0.000	23.370
	TC2	248	3.880	3.861	0.014	39.278
	TC3	247	10.209	5.799	0.014	40.716
	TC4	248	2.397	3.470	0.000	39.278
	TC5	247	5.478	6.664	0.000	40.716

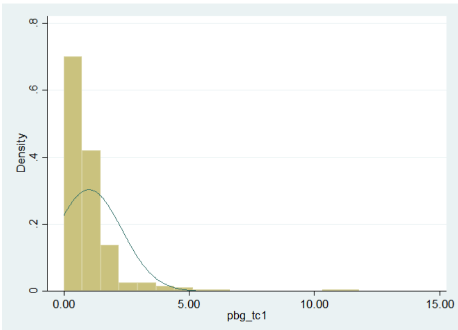
Table 18: Travel costs of visits

As shown in Table 16 and Table 17, in terms of visitation patterns, about 23% of respondents stated that they had not visited Penang Botanic Gardens in the previous year. Among those who had visited, the average number of visits was 7.7 (standard deviation: 11.1), with an average duration of 87.9 minutes per visit (standard deviation: 44.6). Typically, there were 3.4 individuals in the visiting party (standard deviation: 2.3). Modes of transportation included 70% by car, 13% by motorbike, and 0.3% on foot, while reasons for visiting comprised 43% for walks and 14% for jogging. Additionally, about 13.5% of respondents claimed they had never visited Penang Botanic Gardens before.

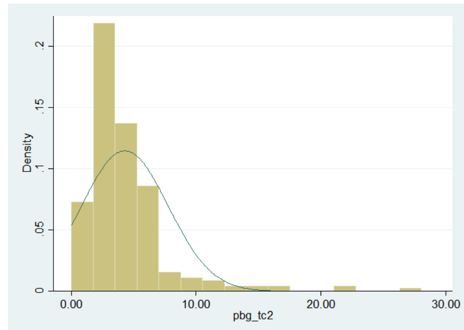
In the context of Penang Youth Park, approximately 26% of respondents reported no visits in the past year. For those who had visited, the average number of visits was 7.4 (standard deviation: 12.4), and the mean visit duration was 88 minutes (standard deviation: 48.2 minutes). On average, each visit involved 3.5 individuals (standard deviation: 2.6). Transportation methods were diverse, with 67% by car, 12.8% by motorbike, and 1.3% on foot. Reasons for visiting included 40% for walks, nearly 12% for enjoying nature, and 11.6% for jogging. Moreover, approximately 16% claimed they had never visited Penang Youth Park before.

5.4.2 Summary of travel costs

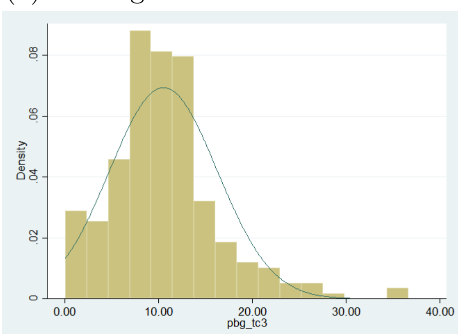
Table 18 shows the descriptive statistics of travel costs of visits to Penang Botanic Gardens and Penang Youth Park, and Figures 8 and 9 show the histograms of travel costs. As shown in the table, the travel costs for visiting Penang Botanic Gardens



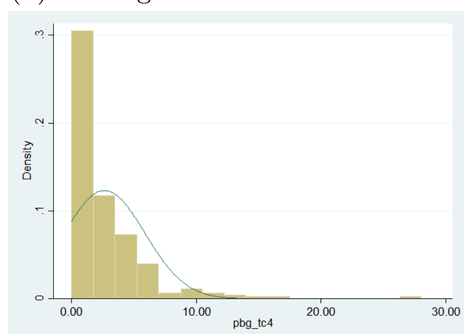
(a) Penang Botanic Gardens - TC1



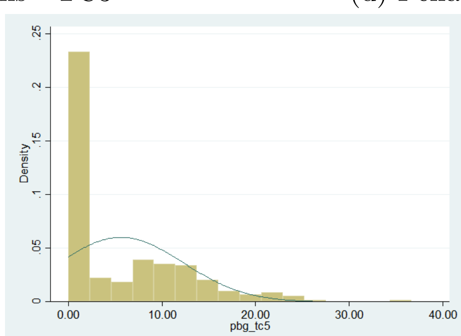
(b) Penang Botanic Gardens - TC2



(c) Penang Botanic Gardens - TC3

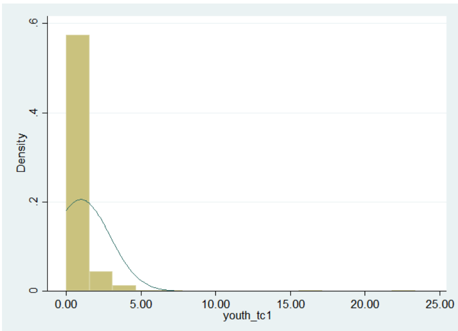


(d) Penang Botanic Gardens - TC4

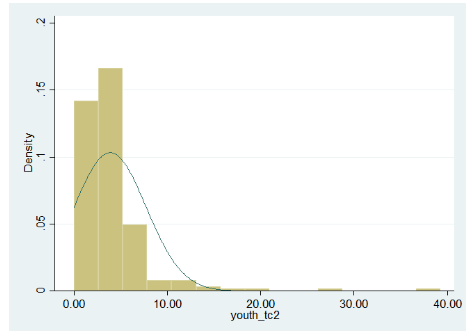


(e) Penang Botanic Gardens - TC5

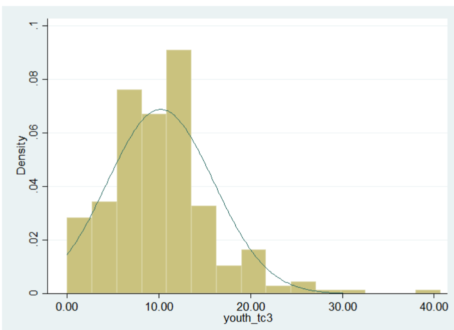
Figure 8: Histograms of travel costs of visit to Penang Botanic Gardens



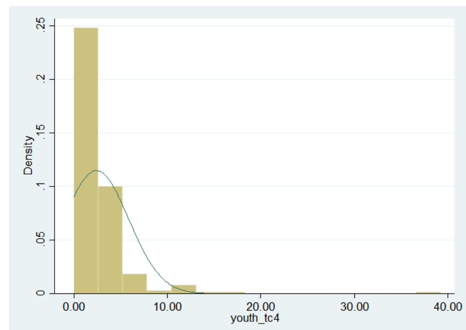
(a) Penang Youth Park - TC1



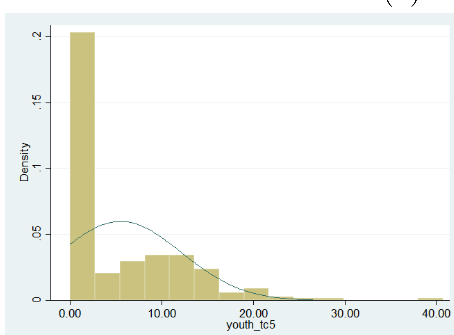
(b) Penang Youth Park - TC2



(c) Penang Youth Park - TC3



(d) Penang Youth Park - TC4



(e) Penang Youth Park - TC5

Figure 9: Histograms of travel costs of visit to Penang Youth Park

are slightly higher than those for visiting Penang Youth Park.

Without considering opportunity costs, the vehicle operation cost (TC1) averaged around RM1.00 (NZD0.34) for trips to Penang Botanic Gardens, with a maximum cost of RM11.78 (NZD4.02). When opportunity costs (time value) are factored in, the total travel costs (TC2-TC5) increase significantly for trips to Penang Botanic Gardens, ranging from RM4.30 (NZD1.47) to RM10.51 (NZD3.59), depending on the criteria used for cost calculation. The highest travel cost, including values for travel time and recreation time, reaches RM36.70 (NZD12.52).

In contrast, travel costs for trips to Penang Youth Park are generally slightly lower. TC1 averages only RM0.98 (NZD0.33), with a maximum cost of RM23.37 (NZD7.97). When considering opportunity costs (time value), TC2, TC3, TC4, and TC5 range from RM2.39 (NZD0.82) to RM10.20 (NZD3.48), with the highest value among these travel costs reaching RM40.72 (NZD13.89). [1]

5.4.3 Travel Cost Model regression results - Penang Botanic Gardens

Poisson regression analysis - Equidispersion tests

The Poisson regression models were estimated using Stata version 14.2. As there were five estimated travel costs, five Poisson regression models were employed. Each model corresponds to one travel cost estimation, as indicated in Columns 2 to 6 of Table 53 in Appendix E.

An equidispersion test was conducted to empirically investigate the presence of underdispersion or overdispersion in the Ordinary Poisson model. This study follows the method introduced by (Cameron and Trivedi, 1990) and employs Stata version 14.2 software for estimation. Equidispersion is violated when the conditional variance of the dependent variable differs from the conditional mean. Specifically, when the conditional variance is smaller or larger than the conditional mean, it indicates the presence of underdispersion or overdispersion, respectively.

As presented in Table 19, the coefficients of $\hat{\mu}$ for all models are negative, which signifies underdispersion in the Poisson model. To appropriately model underdispersed count data, this study employs the quasi-Poisson model. This model maximizes the Poisson maximum likelihood estimation but incorporates a robust estimate of the variance-covariance estimator (VCE), as proposed by Cameron and Trivedi (1990).

[1] In 2022, the exchange rate between Malaysian Ringgit (RM) and New Zealand Dollar (NZD) was approximately RM2.93 per NZD, with slight variations.

	(1)	(2)	(3)	(4)	(5)
	y*	y*	y*	y*	y*
$\hat{\mu}$	-0.0872***	-0.0901***	-0.0889***	-0.0873***	-0.0877***
	(13.21)	(12.65)	(13.46)	(13.14)	(13.35)
N	257	257	257	257	257

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 19: Equidispersion test

Robust estimate of VCE for Poisson model

The robust estimate of VCE for the Poisson model is also calculated using Stata version 14.2 software. In general, the robust standard errors are larger than those in the ordinary Poisson model. Additionally, several independent variables exhibit higher p-values, suggesting weaker evidence to reject the null hypothesis of no effect.

Table 20 displays the coefficient estimates, which remain consistent with the ordinary Poisson model. Among these models, Model 2, which includes travel costs accounting for the value of travel time, exhibits the highest log-likelihood. Conversely, Model 1, featuring only vehicle operation costs, records the second lowest log-likelihood. Among the five travel cost variables, only *TC2* and *TC4* are statistically significant at 10 % level.

Most of the coefficient estimates are statistically insignificant. However, the coefficient estimates of *age* and the number of visits to the substitute park, Penang Youth Park (*youth visit*), are positive and significant at 10 % level. Most of the *air* and *tree* variables are found to be insignificant, except the coefficient estimates of *air* in Model 4.

The marginal effects of these variables were estimated based on the model, and the results are summarized in Table 21. Specifically, every one-unit decrease in travel cost in Malaysian Ringgit increases the expected number of visits by a range between 0.09 and 0.51. Additionally, every one-unit increase in the number of visits to Penang Youth Park is associated with an approximate 0.2 increase in the expected number of visits to Penang Botanic Gardens, so they are complement in consumption. Lastly, each one-year increase in respondents' age is associated with approximately a 0.07 increase in the expected number of visits.

	(1) pbg_visit	(2) pbg_visit	(3) pbg_visit	(4) pbg_visit	(5) pbg_visit
pbg_visit					
pbg_tc1	-0.0780				
	(1.28)				
pbg_tc2		-0.0902***			
		(2.62)			
pbg_tc3			-0.0186		
			(1.57)		
pbg_tc4				-0.0565*	
				(1.78)	
pbg_tc5					-0.0156
					(0.93)
age	0.0121**	0.0120**	0.0119**	0.0124***	0.0120**
	(2.51)	(2.51)	(2.52)	(2.62)	(2.52)
educ	0.0325	0.0548	0.0319	0.0375	0.0284
	(0.50)	(0.85)	(0.49)	(0.59)	(0.45)
employed	0.0503	0.0844	0.0485	0.236	0.190
	(0.33)	(0.57)	(0.31)	(1.40)	(0.92)
household	-0.0225	-0.00544	-0.0189	-0.0188	-0.0229
	(0.39)	(0.10)	(0.33)	(0.33)	(0.39)
youth_visit	0.0374***	0.0398***	0.0385***	0.0374***	0.0374***
	(4.53)	(5.52)	(5.03)	(4.58)	(4.60)
air	0.00911	0.00909	0.00942	0.0101*	0.00962
	(1.51)	(1.62)	(1.54)	(1.67)	(1.56)
tree	0.0157	0.0154	0.0156	0.0156	0.0163
	(1.32)	(1.43)	(1.34)	(1.35)	(1.38)
_cons	0.440	0.537	0.526	0.321	0.375
	(0.66)	(0.81)	(0.79)	(0.48)	(0.56)
N	257	257	257	257	257
Log pseudolikelihood	-1087.6	-1045.7	-1084.4	-1081.4	-1090.5

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 20: Robust estimate of VCE for Poisson model (Penang Botanic Gardens)

	(1)	(2)	(3)	(4)	(5)
	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx
pbg_tc1	-0.46				
pbg_tc2		-0.51***			
pbg_tc3			-0.11		
pbg_tc4				-0.33*	
pbg_tc5					-0.09
age	0.07**	0.07**	0.07**	0.07***	0.07**
educ	0.19	0.31	0.19	0.22	0.17
employed	0.29	0.48	0.28	1.38	1.12
household	-0.13	-0.03	-0.11	-0.11	-0.13
youth_visit	0.22***	0.22***	0.23***	0.22***	0.22***
air	0.05	0.05	0.06	0.06	0.06
tree	0.09	0.09	0.09	0.09	0.1

t statistics in parentheses

* p <0.10, ** p <0.05, *** p <0.01

Table 21: Marginal effects of variables from Poisson model (Penang Botanic Gardens)

5.4.4 Travel Cost Model regression results - Penang Youth Park

Poisson regression analysis - Equidispersion tests

In addition, the Poisson regression models for visitation to Penang Youth Park were also estimated. The estimation results are reported in Table 58 in Appendix E.

An equidispersion test was conducted for the ordinary Poisson model. As shown in Table 22, the coefficients of $\hat{\mu}$ for all models are negative. This indicates that the dependent variable's conditional variance is smaller than the conditional mean, and the underdispersion in the Poisson model occurred. Therefore, this study uses the quasi-Poisson model to model underdispersed count data, which uses the robust estimate of the variance-covariance estimator (VCE) (Cameron and Trivedi, 1990).

Robust estimate of VCE for Poisson model

The results of the robust estimate of VCE for the Poisson model are presented in Table 23. In this robust estimation, most of the p-values for coefficient estimations have increased. Among the five travel cost variables, only the coefficient estimate for *TC3* remains statistically significant. However, it is important to highlight that the

	(1)	(2)	(3)	(4)	(5)
	y*	y*	y*	y*	y*
$\hat{\mu}$	-0.0902***	-0.0897***	-0.0873***	-0.0899***	-0.0894***
	(11.19)	(11.18)	(10.88)	(11.19)	(11.19)
N	245	246	245	246	245

t statistics in parentheses

* p <0.10, ** p <0.05, *** p <0.01

Table 22: Equidispersion test

positive coefficient estimate for *TC3* raises some questions and may require further investigation.

The number of visits to Penang Botanic Gardens (*pbg visit*) variable's coefficient estimates are positive and statistically significant, indicating that respondents who visit Penang Botanic Gardens more frequently also visit Penang Youth Park more frequently. The coefficients of the *educ* variable have significant negative signs, indicating that respondents with lower educational attainment are more likely to visit Penang Youth Park. The coefficients of the importance level of the air quality variable are statistically significant and negative. Surprisingly, these results diverge from my expectations, suggesting that a lower perceived importance of air quality is associated with an increased number of visits.

The estimation results of variables' marginal effects are displayed in Table 24. In Model 3, for instance, every one-unit increase in travel costs is associated with an expected increase in visit frequency of 0.12. On average, every categorical level increase in the education level reduces the expected number of visits by a range between 0.71 and 0.78. An increase in the number of visits to Penang Botanic Gardens is associated with an expected increase in the number of visits to Penang Youth Park by approximately 0.26 visits, and an increase in the importance level of the air quality is associated with an expected decrease in the number of visits by approximately 0.04 visits.

5.4.5 The validity of value of time in Travel Cost Model

In this study, the value of time was calculated in two ways. The first way assumes that only individuals who are employed have a positive value of time. For unemployed individuals, the value of time was set as zero. The value of travel time based on this assumption is identified as OTC_n, and the value of recreational time is identified as ORC_n.

	(1) youth_visit	(2) youth_visit	(3) youth_visit	(4) youth_visit	(5) youth_visit
youth_visit					
youth_tc1	-0.0477				
	(1.38)				
youth_tc2		0.00433			
		(0.28)			
youth_tc3			0.0247**		
			(2.30)		
youth_tc4				-0.00763	
				(0.63)	
youth_tc5					0.00985
					(1.09)
age	0.000119	0.00000795	-0.000746	0.000185	-0.000452
	(0.03)	(0.00)	(0.18)	(0.05)	(0.11)
educ	-0.145***	-0.152***	-0.161***	-0.148***	-0.157***
	(2.84)	(2.89)	(3.15)	(2.84)	(2.95)
employed	0.0860	0.0642	0.0245	0.0933	-0.0255
	(0.61)	(0.44)	(0.17)	(0.66)	(0.16)
household	0.00978	0.00801	-0.00312	0.0106	0.00524
	(0.29)	(0.23)	(0.09)	(0.31)	(0.15)
pbg_visit	0.0534***	0.0542***	0.0548***	0.0538***	0.0540***
	(15.60)	(16.18)	(16.20)	(15.43)	(15.66)
air	-0.00806**	-0.00843**	-0.00950**	-0.00815**	-0.00868**
	(2.06)	(2.09)	(2.32)	(2.06)	(2.21)
tree	0.000396	0.000967	0.00329	0.000667	0.000872
	(0.04)	(0.09)	(0.32)	(0.06)	(0.08)
_cons	2.210***	2.207***	2.130***	2.185***	2.275***
	(4.61)	(4.59)	(4.82)	(4.48)	(4.66)
N	245	246	245	246	245
Log pseudolikelihood	-845.6	-851.1	-832.4	-851.0	-847.1

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 23: Robust estimate of VCE for Poisson model (Penang Youth Park)

	(1) dy/dx	(2) dy/dx	(3) dy/dx	(4) dy/dx	(5) dy/dx
youth_tc1	-0.23				
youth_tc2		0.02			
youth_tc3			0.12**		
youth_tc4				-0.04	
youth_tc5					0.05
age	0	0	0	0	0
educ	-0.71***	-0.74***	-0.78***	-0.72***	-0.77***
employed	0.42	0.31	0.12	0.46	-0.13
household	0.05	0.04	-0.02	0.05	0.03
pbg_visit	0.26***	0.26***	0.27***	0.26***	0.27***
air	-0.04**	-0.04**	-0.05**	-0.04**	-0.04**
tree	0	0	0.02	0	0

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 24: Marginal effects of variables from Poisson model (Penang Youth Park)

The second way assumes everyone has a positive value of time regardless of employment status. The value of travel time and recreational time are identified as $AOTC_n$ and $AORC_n$ respectively. Both time values are examined separately by the Poisson regression models with robust standard error.

The results for the visits to Penang Botanic Gardens are shown in Table 25. Table 25 (1) shows the regression model that includes VC_n, OTC_n, and ORC_n. Table 25 (2) shows the regression model that includes VC_n, $AOTC_n$, and $AORC_n$. Comparing both models, it can be seen that $AOTC_n$ and $AORC_n$ have a better fit to the model compared to OTC_n and ORC_n, indicating that the value of time should be assigned to all individuals, without considering their employment status. Individuals with zero income might have a value of time of more than zero. A monetary value should be placed on time spent due to its scarcity.

Table 26 shows the results for the visits to Penang Youth Park. The regression model in Table 26 (1) includes VC_n, OTC_n, and ORC_n, while the model in Table 26 (2) includes VC_n, $AOTC_n$, and $AORC_n$. The results show that $AOTC_n$ and $AORC_n$ fit the model better as the log pseudolikelihood is higher. However, only $AORC_n$ is significant at the 1% level. This indicates that all individuals should be assigned a value of time, regardless of employment status and income level.

	(1) pbg visit	(2) pbg visit
VC	-0.0434	0.0714
	(-0.41)	(0.69)
OTC	-0.0413	
	(-0.81)	
ORC	0.0247	
	(0.75)	
AOTC		-0.135***
		(-2.67)
AORC		0.0538**
		(2.47)
Cons	1.916***	1.872***
	(13.18)	(9.89)
N	271	271
Log pseudolikelihood	-1541.538	-1486.974

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 25: Estimates of Travel Cost Components (Penang Botanic Gardens)

	(1) youth visit	(2) youth visit
VC	-0.207	-0.167
	(-1.50)	(-1.07)
OTC	0.0347	
	(0.60)	
ORC	0.0115	
	(0.51)	
AOTC		-0.0207
		(-0.37)
AORC		0.0517***
		(2.87)
Cons	1.924***	1.690***
	(12.83)	(8.76)
N	253	253
Log pseudolikelihood	-1441.807	-1415.605

t statistics in parentheses

* p <0.10, ** p <0.05, *** p <0.01

Table 26: Estimates of Travel Cost Components (Penang Youth Park)

Travel costs	Mean travel cost	CS	WTP
Poisson model			
TC1	1.01	12.82	13.83
TC2***	4.3	11.09	15.39
TC3	10.51	53.76	64.27
TC4*	2.64	17.70	20.34
TC5	5.62	64.10	69.72

t statistics in parentheses for TC coefficients
* p <0.10, ** p <0.05, *** p <0.01

Table 27: WTP estimations (Penang Botanic Gardens)

Travel costs	Mean travel cost	CS	WTP
Poisson model			
TC1	0.98	20.96	21.94
TC2	3.88	-230.95	-227.07
TC3**	10.21	-40.49	-30.28
TC4	2.4	131.06	133.46
TC5	5.48	-101.52	-96.04

t statistics in parentheses for TC coefficients
* p <0.10, ** p <0.05, *** p <0.01

Table 28: WTP estimations (Penang Youth Park)

5.4.6 Willingness to Pay and Consumer Surplus estimation in Travel Cost Model

The estimations of WTP and CS for a single visit are presented in Tables 27 and 28.

Table 27 presents the results of single-trip visits to Penang Botanic Gardens. The WTP estimates range from RM13.83 to RM69.72. Notably, *TC3* and *TC5* exhibit relatively higher WTP value estimates compared to other travel cost variables.

Moving to Table 28, there is a broad range in WTP values, ranging from RM-227.07 to RM133.46. Most CS values are negative due to the positive coefficient estimation of travel costs in the regression models.

5.5 Discussions

The results from the regression models highlight the critical role of travel cost calculations in estimating CS and WTP. Different travel cost components lead to significant variations in these estimates.

As depicted in Table 27, the CS estimates for visits to Penang Botanic Gardens are relatively lower for *TC1*s (which consider only vehicle operating costs) and *TC2*s and *TC4*s (which include vehicle operating costs and travel time value as indicators). This suggests that the demand curves are generally elastic when the value of travel time is incorporated. However, the CS significantly increases for *TC3* and *TC5*, which involve the value of recreational time, indicating that the demand curves are generally more inelastic. Despite the very low mean travel costs, the absolute coefficient estimation for travel costs is also minimal. This results in a highly inelastic demand function and, consequently, a high CS.

For visits to Penang Youth Park, the negative CS value estimates in *TC2*, *TC3*, and *TC5* suggest that the market price for the visit exceeds the maximum price that visitors are willing to pay.

In summary, these results emphasize that individuals not only incur direct monetary costs but also indirect non-monetary opportunity costs when visiting green spaces. Time is a finite resource, and people continually make choices among alternatives to allocate their limited time to their preferred activities. The alternative with the second-highest value represents the opportunity cost individuals incur when they choose to visit a green site. It should be assigned a value greater than zero due to the scarcity of time resources. By incorporating time costs into travel cost calculations in travel cost models, deeper insights into individuals' travel behaviour and preferences can be obtained.

In addition, as discussed in Section 5.4.5, the models demonstrate that $AOTC$ and $AORC$ are statistically significant at a 5% level, and their inclusion consistently results in a substantially better model fit compared to using OTC and ORC. These results align with the expectation that every individual assigns a nonzero value to their time, regardless of their employment status, which signals their market value of time. Assigning a zero value of time to unemployed individuals is unreasonable. Individuals, irrespective of employment status, allocate their time among various activities, such as work, leisure, and family engagements, using a subjective value based on individual opportunity cost. Estimating the value of time solely based on monetary income is inadequate, as individuals with lower or zero income also optimize their time allocation among different activities.

The positive coefficients of travel costs (*TC2*, *TC3*, and *TC5*) for visits to Penang

Youth Park may appear unexpected at first glance. Notably, *TC1*, which includes only vehicle operation costs, exhibits a negative sign. However, when travel costs encompass the values of travel time and recreational time, the coefficients turn positive. This suggests that respondents who incur higher time-related costs are more inclined to visit the park. This counter-intuitive observation can be explained by the idea that individuals tend to derive greater enjoyment from their visits when they spend more time travelling to the destination and subsequently spend an extended duration at the park.

For visits to Penang Youth Park, the negative sign on the coefficient of the education variable is surprising. It implies that respondents with lower levels of education visited both sites more frequently in the past year. This might be attributed to the fact that higher education levels are often linked to higher income, providing individuals in this group with more options for recreational activities, including those that may come at a higher cost. In contrast, individuals with lower education levels, possibly having limited financial resources, may prefer places with lower or no expenditure.

Another surprising finding is the positive sign for the number of visits to a substitute site. It suggests that individuals who visit Penang Botanic Gardens tend to visit Penang Youth Park more frequently. Consequently, the claim that Penang Botanic Gardens serves as the substitute site for Penang Youth Park may not be valid, as the results suggest that they are more complementary than substitutable. This phenomenon may be attributed to the short 1 km distance between both sites, making it convenient for individuals visiting one site to also visit the other during the same outing.

Additionally, the unexpected negative sign for the perceived importance of air quality in visits to Penang Youth Park contrasts with the best-worst scaling (BWS) analysis presented in Chapter 4, where air quality received the highest score. However, given that a majority of individuals selected this attribute as the most important, the resulting high BWS scores are distributed across a majority of respondents. This high homogeneity in responses could make it challenging to discern significant patterns in the Poisson model. Moreover, visitors might highly value the air quality ecosystem service regardless of their frequency of visits to the site, as air quality impacts the entire city, and people do not necessarily need to travel to the site to enjoy its benefits.

5.6 Conclusions

This paper demonstrates a case study application of the TCM in the context of urban Penang, Malaysia. It examines two major urban green spaces with distinct characteristics. The study also includes an analysis of various travel costs and argues that assigning a value of time to each individual is essential when estimating travel costs. Additionally, the study incorporates variables related to the perception of the importance of air quality and tree species to better understand their impact on the number of visits to these green spaces.

Urban green spaces offer significant recreational benefits to society. The high value attributed to Penang Botanic Gardens and Penang Youth Park underscores the importance of the natural environment and forestry ecosystem within the urban context of Penang Island, Malaysia. These findings provide valuable insights for urban planners, relevant authorities, and green space managers involved in decision-making regarding the development of urban green spaces. However, it is important to acknowledge several limitations in this paper. Firstly, this study is centred on a specific case study, which means that the results presented here are particular to the selected case. There is a question of external validity of these results when it applies to other study areas.

There are also considerations when assuming respondent income based on age groups without using real income data. People within the same age category may have significantly different incomes, and the lack of income data could distort the estimation results.

Another important consideration is that the values obtained through the TCM only represent a portion of the benefits provided by urban green spaces. Specifically, the TCM can only capture the direct use values associated with visits. To comprehensively understand the benefits, indirect use values should also be accounted for. Moreover, it is crucial to understand the relative importance of different attributes in urban green spaces. Chapter 6 addresses these aspects by employing the discrete choice experiment (DCE) to estimate the values that individuals place on various attributes of urban green spaces.

Chapter 6

Analysis of Time Value Effects in DCE

6.1 Introduction

In rapidly developing urban areas, urban green spaces have become important assets of cities as a high-quality urban green space contributes to urban residents' well-being and enhances the provision of urban ecosystem services. Urban green spaces are spaces that provide recreation and relaxation, and allow people to have contact with nature (Aronson et al., 2017; Haase et al., 2017; Fischer et al., 2018). The presence of urban green spaces has been shown to positively influence people's health and well-being (Tzoulas et al., 2007; Bowler et al., 2010). Moreover, urban green spaces are the main providers of urban ecosystem services such as carbon sequestration, noise level reduction, water quality improvement and biodiversity support (Breuste et al., 2013; Miller et al., 2015; Barbosa et al., 2007; Sanesi et al., 2009; Hendrey et al., 1999; Calfapietra et al., 2009; Gómez-Baggethun and Barton, 2013; Albert et al., 2014; Renterghem and Botteldooren, 2016; Jim and Chen, 2006a). These ecosystem services reduce the adverse environmental effects caused by the rapidly growing population in urban areas such as air pollution, noise, and increasing temperature (Bertram and Rehdanz, 2015a; Gidlöf-Gunnarsson and Öhrström, 2007; Janhäll, 2015; Dimoudi and Nikolopoulou, 2003). Urban green spaces also play an important role in maintaining habitats for plant and animal species (Knapp et al., 2020; Beninde et al., 2015), which have been destroyed due to the continuing development of urban areas. Although the benefits of urban green spaces are widely recognized by a substantial body of literature (Roberts et al., 2022), the allocation of land for green spaces in urban areas often faces challenges and competition from other land use purposes (Smith et al., 2012), and green spaces are always insufficiently provided especially at the most populated urban areas (Kabisch et al., 2015; Schetke et al., 2012). Therefore, for effective planning, efficient public and private financing, and administration of urban green spaces, it is essential to understand residents' preferences and their willingness to pay (WTP) for various attributes of these spaces, including factors like proximity, facilities, and the provision of ecosystem assets and services. By focusing on these aspects in the design of green spaces, it becomes possible to optimize costs and better meet residents' needs and expectations. However, estimating these WTP values can be challenging due to the heterogeneous nature of individual preferences, and other factors such as price sensitivity.

A substantial body of literature aimed at estimating the value of urban green spaces using various non-market valuation methods has been reviewed. Karunarathne and Gunawardena (2020) utilized the travel cost method (TCM) to estimate the value of an urban park in Colombo and found that household income and visitors' enjoyment positively affect their number of visits. Jim (2006) utilized the contingent valuation method to assess individuals' willingness to pay (WTP) for urban green spaces in Guangzhou, China. The study found a strong association between income and WTP, suggesting that urban greenspaces were treated as superior goods. The research revealed a high WTP, with 96.6% of respondents expressing a readiness to pay for the use of these spaces, and the estimated conservative average WTP was surpassing actual entrance-fee payments. The estimated monetary value of urban green spaces in Guangzhou was six times higher than the government's annual expenditures on these spaces. Chintantya and Maryono (2018) combined the TCM and CVM to estimate the value of urban green spaces by measuring WTP for ecosystem benefits. The research revealed that individuals with higher incomes were more likely to visit a location frequently, particularly those with tourist attractions. Moreover, urban green spaces with diverse supporting facilities were observed to attract a greater number of visitors. Some studies investigate the value of specific attributes of urban green spaces, particularly the ecosystem services they offer. Bronnmann et al. (2023) conducted a study in Germany, estimating the value of the naturalness of urban green spaces using the DCE. The findings revealed that local citizens had a mean WTP of 20.25 per month for an increase in the naturalness of the closest urban green space by one step. Aevermann and Schmude (2015) quantified and valued urban ecosystem services by focusing on the urban green space of Schlosspark Nymphenburg in Munich, Germany. The study specifically examined four ecosystem services—carbon sequestration and storage, air pollution removal, runoff reduction, and groundwater recharge. The analysis involved the classification of land cover types to calculate varying amounts of ecosystem services. Chang et al. (2017) employed the land rent method to assess the total ecosystem services in urban green spaces, including cultural and regulating services. The study revealed that the green spaces near the urban centre provided significantly higher cultural services compared to those located near the urban edge. Tyrväinen and Väänänen (1998) estimated the value of amenities provided by urban forests using a CVM. The study found that a majority of visitors were willing to pay for the use of wooded recreation areas. Roberts et al. (2022) estimated the value of urban green space attributes and functions through the DCE. The findings revealed a positive WTP for the maintenance of greenspace, and a higher WTP for larger greenspaces located closer to home, particularly those that were multifunctional, and incorporated both direct-use and biodiversity features. Another group of research investigates the impact of proximity to urban green spaces on various factors related to their valuation. Bertram and Rehdanz (2015b) explored the effect of proximity to urban green spaces on life satisfaction. The study identified an inverted U-shaped effect concerning both the amount of and distance to urban green space on life satisfaction. Notably, the results suggested that the optimal amount of green space in a 1km buffer for the most substantial positive effect on life satisfaction was 35 hectares, equivalent to 11% of the buffer area. Studies such as Engström and Gren (2017), Trojanek et al. (2018), Lategan et al. (2022), Panduro and Veie (2013), and Samad et al. (2020) examined the relationship between distance to urban green spaces and housing prices using the hedonic price method. del Saz Salazar and Menéndez (2007) investigated the connection between individuals' WTP for a new urban park in Valencia, Spain, and its proximity. The study found that the mean WTP is significantly higher for individuals who live closer to the park, suggesting a positive association between accessibility and the perceived value of the park. Despite these research efforts to estimate the values of urban green spaces and their attributes, varying results have been obtained. This variability can be attributed to differences in the methodologies employed during the valuation process. Therefore, it is important to consider these methodological differences when valuing the attributes of urban green spaces.

This study utilizes a distance-based DCE to estimate the value of urban green space attributes, particularly the ecosystem services provided by urban green spaces on Penang Island, Malaysia, by investigating the corresponding WTP values. Various choice experiment studies have similarly investigated preferences in the urban context. For instance, Van Dongen and Timmermans (2019) investigated individuals' preferences for various urban green space designs using a choice experiment in a virtual environment. The findings highlighted the significant influence of trees on preferences. In contrast, grass, often favoured by local governments, and vertical green had the smallest effects in residential streets according to the study's results.

Arnberger and Eder (2011) investigated people's preferences for recreational trails in urban open spaces using an image-based DCE in Vienna. The results indicated that visitor numbers and litter had the highest influence on trail preferences. Banfi et al. (2012) used a stated CE to estimate the WTP of residents in Zurich and Lugano for the reduction of urban externalities such as traffic noise, air pollution, and electromagnetic pollution. The results revealed a positive WTP for reducing air pollution and traffic noise. Respondents also expressed WTP for reducing electrosmog and removing mobile phone antennas from their view, although to a lesser extent. Heo and Kim (2013) investigated individuals' WTP for various attributes in an urban agriculture park using the DCE, including garden scales, learning and experience area, leisure and relaxation area, and fund. This research provides valuable insights into the preferences and valuation of different features within an urban agriculture setting.

In a DCE, an appropriate payment vehicle can mitigate individuals' price sensitivity when making choices. In this research, the distance to urban green space serves as the payment vehicle. The basic assumption is that travel distance is linked to travel costs, including out-of-pocket travel expenses and the opportunity cost, which, in this case, is represented by travel time. The total travel cost represents the price associated with accessing the urban green space. Therefore, the WTP for visiting an urban green space can be estimated (Chintantya and Maryono, 2018; Bishop, 2002).

A number of published studies have suggested using distance as a payment vehicle in the context of DCE (Christie et al., 2007; De Valck et al., 2017; Johnston et al., 2015; Faccioli et al., 2016; Luisetti et al., 2011). De Valck et al. (2017) argued that people's perception of distance to a recreational site is more likely to influence their choice of a recreational site compared to other direct costs. Johnston et al. (2015) asserted that individual WTP is a continuous function of the distance from an individual's house to the recreational site, with WTP decreasing as distance increases. In contrast, Cheshire and Stabler (1976) hold the opposite view, asserting that the journey itself for recreational activities may lead to a positive utility associated with travel. Despite this long-standing contradictory evidence, there remains a shortage of stated preference studies addressing this controversial issue (Ovaskainen et al., 2012).

Furthermore, while there exists an extensive body of literature investigating the validity of travel time in the context of the TCM, limited research has explored its validity within the framework of DCE (Cesario, 1976; Smith et al., 1983; McConnell, 1992; Sánchez et al., 2016; Lloyd-Smith et al., 2020b). Therefore, this study also seeks to determine whether travel time to urban green spaces is perceived as a travel cost within the framework of a distance-based DCE. The hypothetical nature of choices is plausibly considered consequential for informing public policy, especially when considering other plans for new green areas in Penang.

This research focuses on the decision-making process related to the selection of an urban green site. Two research questions have been identified:

1. Can a distance-based discrete choice model be employed as a valuable tool to comprehend individual sensitivities towards travel costs and preference heterogeneity for urban ecosystem services on the Island of Penang, Malaysia?

2. Should the time spent travelling to a recreational site be perceived as a cost of visitation?

The first research question is raised concerning the validity of distance-based discrete choice models in investigating the preference of individuals for site characteristics. The second research question is motivated by a number of published studies that explore the validity of travel time in travel cost models.

Section 3.1.1 provides a detailed discussion of the development of attributes and levels, the experimental design of the DCE, and an example of a choice task that was presented to the respondents (Figure 4). The details of the data collection process were elaborated upon in both Section 3.2 and Section 3.2.1. This study utilized sample sizes consisting of 404 observations, where the orthogonal design's DCE includes 215 samples, and the Bayesian design's DCE includes 189 samples. The methods of study were discussed in Section 2.5.

This study is structured as follows. Section 6.2 describes the study design. In this section, Section 6.2.1 introduces the model specification. Section 6.3 presents the findings of the study. Section 6.4 discusses the key findings of this study. Section 6.5 concludes the findings and discusses the limitations of the study.

6.2 Study design

A DCE was developed to investigate respondents' preferences for attribute levels in urban green spaces and to estimate their WTP for these attributes. During the survey, respondents were asked to select their preferred green site from two hypothetical green sites, each presented with different attribute levels. Each respondent encountered 12 repeated choice tasks, comprising two alternative sites labelled as *Green Space A* and *Green Space B*. The *distance* of the site from the place of residence served as the payment vehicle. In the survey, respondents traded off the distance they were willing to travel with other attribute levels presented in the DCE questions.

Several models were estimated, including the MNL models, uncorrelated MXL models, and correlated MXL models in preference space (shown in Appendix F), and uncorrelated MXL models, and correlated MXL models in WTP-space specification (shown in Section 6.3.1).

Moreover, as the *distance* attribute is linked to the respondent's travel costs, this attribute in the estimation of the DCE model was replaced with two types of travel costs. The first travel cost included only vehicle operating costs, while the second travel cost included both vehicle operating costs and the value of travel time. Longer distances incurred higher travel costs. Both travel costs were examined separately to determine whether travel time should be perceived as a component of travel costs in the context of the DCE. This was done by comparing the empirical results of both models.

6.2.1 Model specification

Converting ***distance*** attribute to travel costs

In the DCE, *distance* is not expressed in monetary cost units, whereas travel cost can be. Therefore, the *distance* from the home location to the hypothetical site is converted into travel costs for a round-trip visit for the purpose of model estimation. The WTPs from both models were compared to investigate how the results change when the value of travel time is a part of the travel costs.

The calculation of round-trip travel costs for each individual n follows the methods discussed in Section 5.3.1. In the hypothetical scenario of the DCE, all respondents were assumed to be driving a car alone to visit the green site, as stated in the question. Therefore, the mode of travel m for all respondents is assumed to be *car*. Consequently, the round-trip vehicle operating cost for each individual n for alternative j in choice situation t (VC_{njt}) is expressed as:

$$VC_{njt} = \frac{2 \times voc_{njt}}{1} \ , \tag{6.1}$$

where voc_{njt} represents the single-trip vehicle operating cost for each individual n for alternative j in choice situation t, derived from the formula $0.205 \times \text{Distance}_{njt}$. The variable $Distance_{njt}$ represents the distance from individual n's place of residence to the hypothetical green site for alternative j in choice situation t, measured in kilometres. The vehicle operating costs were estimated based on an average consumption rate of 0.1 litres per kilometre for driving a car, and the price of unleaded 95 petrol in Malaysia, which is RM2.05 per litre.

The travel cost, in this case study, also includes the value of travel time ($AOTC_{njt}$), which assumes that all individuals have a time value based on their expected hourly income, regardless of their employment status. Similar to the calculation method of $AOTC_n$ in Section 5.3.1, the expected hourly income is derived from the median income for different age groups provided by the Department of Statistics Malaysia, as shown in Table 13 in Chapter 5 (DOSM, 2021).

The expected travel time per kilometre is calculated based on the average travel time per kilometre from the sample I obtained in the survey. The travel cost information used for calculation includes only the travel time incurred while driving a car. During the survey, respondents were asked to provide their home address or the area of their home address. Travel information was then collected from Google Maps, which estimated the travel time and travel distance from the respondent's house location to Penang Botanic Gardens and Penang Youth Park. Based on the data collected from Google Maps, the average travel time is 1.82 minutes per kilometre to reach Penang Botanic Gardens and Penang Youth Park. Using real-world travel time data better reflects the respondent's expected travel time in a hypothetical situation. The variable $AOTC_{njt}$ is expressed as:

$$AOTC_{njt} = \frac{Expected\ Hourly\ Income_n}{60\ minutes} \times \frac{Distance_{njt} \times 2 \times 1.8213}{3}, \tag{6.2}$$

where $Distance_{njt}$ represents the distance from individual n's place of residence to the green site (alternative j in choice situation t).

Two travel costs were derived, including TC and TCA, which are shown in equation 6.3 and equation 6.4 respectively:

$$TC_{njt} = VC_{njt}, \tag{6.3}$$

$$TCA_{njt} = VC_{njt} + AOTC_{njt}, \tag{6.4}$$

where TC_{njt} and TCA_{njt} represent the round-trip travel costs for individual n for alternative j in choice situation t.

After the *distance* attribute is converted to travel costs, different respondents will have the same TC if they make similar choices. However, if respondents fall into different age groups, they will have different TCA values even if they choose the same alternative. Therefore, it is believed that converting travel distance into different travel costs will yield distinct empirical results, which can provide valuable insights for urban planners and authorities.

Model variables

The levels of five attributes—*air quality*, *facilities*, *noise levels*, *nursery habitat maintenance*, *tree species and ecosystems*—have been represented by numerical codes: 0, 1, and 2. Level 0 corresponds to the current status, and as such, no dummy variable is created for it, resulting in the creation of only two dummy variables for the remaining two levels. Consequently, a total of 10 dummy variables have been generated, and 10 coefficients have also been produced: (1) The coefficient for level 1 reflects the additional utility associated with increasing from the base level (level zero) to level one; (2) The coefficient for level 2 captures the additional utility associated with increasing from level one to level two. This is different from the standard coding capturing changes from a common baseline.

The distance attribute has values of 1, 5, 10, 20, 30, and 40 and is not treated as a dummy variable. An alternative-specific constant for alternative 1 (*alt 1*) is added to examine the changes in utility when the respondents choose site A over site B in the question sets. Table 29 shows an overview of these variables.

Utility specification

With reference to Table 29, the generic utility specification framework employed in this study is represented as follows:

$$\begin{aligned} U(A) &= \beta_1 x_{1A} + \beta_2 x_{2A} + \beta_3 x_{3A} + \beta_4 x_{4A} + \beta_5 x_{5A} + \beta_6 x_{6A} \\ &+ \beta_7 x_{7A} + \beta_8 x_{8A} + \beta_9 x_{9A} + \beta_{10} x_{10A} + \beta_{11} x_{11A} + \beta_{12} x_{12A} + \varepsilon, \\ U(B) &= \beta_1 x_{1B} + \beta_2 x_{2B} + \beta_3 x_{3B} + \beta_4 x_{4B} + \beta_5 x_{5B} + \beta_6 x_{6B} \\ &+ \beta_7 x_{7B} + \beta_8 x_{8B} + \beta_9 x_{9B} + \beta_{10} x_{10B} + \beta_{11} x_{11B} + \beta_{12} x_{12B} + \varepsilon \end{aligned} \tag{6.5}$$

In this study, I have developed several random utility choice models. To illustrate my analysis, I will provide an example. Specifically, I will describe a mixed logit (MXL) model using a preference space specification. An MXL model specifies the coefficients to be random in order to capture unobserved preference heterogeneity over respondents (Train, 1998; Revelt and Train, 1998; Hensher and Greene, 2003). In the following example, five attributes—*air quality*, *facilities*, *noise levels*, *nursery habitat maintenance*, *tree species and ecosystems*—were considered as random variables. Additionally, the *distance* attribute, representing cost, was treated as fixed. The *alt 1* was also treated as fixed. Consequently, Equation 6.5 can be expanded, which is expressed as follows:

Attribute	Level	Level (number)	Variable	Symbol in eq.
Alternative-specific constant			*alt1*	x_1
Distance	1KM	1	*dis*	x_2
	5KM	5		
	10KM	10		
	20KM	20		
	30KM	30		
	40KM	40		
Travel cost			*tc* *tca*	x_2
Air quality	Poor	0		
	Normal	1	*air1*	x_3
	Good	2	*air2*	x_4
Facilities	A few	0		
	Normal	1	*fac1*	x_5
	A lot	2	*fac2*	x_6
Noise levels	Loud	0		
	Normal	1	*noi1*	x_7
	Quiet	2	*noi2*	x_8
Nursery habitat maintenance	Poor	0		
	Normal	1	*nur1*	x_9
	Good	2	*nur2*	x_{10}
Tree species and ecosystems	A few	0		
	Normal	1	*tree1*	x_{11}
	A lot	2	*tree2*	x_{12}

Table 29: Variables used in choice experiment

$$
\begin{aligned}
U_{njt} = \beta^1 x^1 + \beta^2 x^2 + (\bar{\beta}^3 + \sigma^3 \eta_{njt}) x^3 + (\bar{\beta}^4 + \sigma^4 \eta_{njt}) x^4 \\
+ (\bar{\beta}^5 + \sigma^5 \eta_{njt}) x^5 + (\bar{\beta}^6 + \sigma^6 \eta_{njt}) x^6 \\
+ (\bar{\beta}^7 + \sigma^7 \eta_{njt}) x^7 + (\bar{\beta}^8 + \sigma^8 \eta_{njt}) x^8 \\
+ (\bar{\beta}^9 + \sigma^9 \eta_{njt}) x^9 + (\bar{\beta}^{10} + \sigma^{10} \eta_{njt}) x^{10} \\
+ (\bar{\beta}^{11} + \sigma^{11} \eta_{njt}) x^{11} + (\bar{\beta}^{12} + \sigma^{12} \eta_{njt}) x^{12} \\
+ \varepsilon_{njt}
\end{aligned}
\tag{6.6}
$$

where the coefficient β_n for each random variable is expanded into $\bar{\beta}^k + \sigma^k \eta_{njt}$ with $\eta_{njt} \sim N(0,1)$ to denote the estimation of the distributional moments of random parameter β^k for the k attribute. $\bar{\beta}^k$ denotes the mean value of β_n^k, σ^k is the standard deviation (SD), and η_{njt} is a random draw from a normal distribution with zero means (Scarpa et al., 2012). This is because respondents' preferences for these attributes could either be positive or negative, meaning they may either value or devalue these attributes (Carlsson and Martinsson, 2003). Conversely, if the coefficient for distance is set to be random, the random coefficient for the negative of the distance variable could follow a log-normal distribution, or any positively signed distribution. This choice is based on the understanding that this variable can only have negative effects on respondents' utilities, as increased travel distance is generally associated with decreased utility.

The MNL models were estimated using Stata 14.2, utilizing the `clogit` command. The MXL models in preference space were also estimated in Stata 14.2, employing the `mixlogit` command (Hole, 2007). The MXL models in WTP-space specification were estimated in R version 2023.06.1+524, utilizing the *apollo* package (Hess and Palma, 2019).

The estimation results of models in preference-space specification are shown in Appendix F.

6.3 Results

6.3.1 Models in WTP-space specification

Table 30 presents the estimation results for the MNL model (Model 1), uncorrelated MXL models in WTP-space specification (Models 2-4), and MXL models with correlation in WTP-space specification (Models 5-6). Model 6 constrains the insignificant Cholesky decomposition to be close to zero (0.0001). To improve estimation accuracy, an initial 100 Modified Latin Hypercube Sampling (MLHS) draws were employed for estimating random parameters. The estimated values from this stage were then used as starting values for a subsequent estimation employing 2000 MLHS draws. The table displays the final results.

Although the models in WTP-space provide direct estimations of WTP values, the WTP values are not directly presented in the table since the *distance* variable is expressed in kilometres (KM) rather than monetary units. Therefore, the WTP estimations, as shown in Table 30, were multiplied by $(RM0.205 \times 2)$ to account for the return trip travel cost, as indicated in Equation 6.3. The WTP values for attributes in Malaysian Ringgit are presented in Table 31.

When comparing Model 3 in this table with the uncorrelated MXL model and the correlated MXL model in preference space (Model 2 in Table 65 and Model 2 in Table 67), it is evident that Model 3 has the lowest log-likelihood (-2320.22), and the highest AIC and BIC values. These findings suggest that this model does not provide the best fit to the data. Moreover, Model 3 yields lower and more reasonable WTP estimates overall compared to the two models in preference space.

Model 1 in Table 30 reveals that to estimate respondents' WTP in an MNL model, the cost coefficient (represented by the negative distance, *n_dis*) is constrained to 1. The results indicate that all coefficient estimations are statistically significant at a 1% confidence level. In Table 31, the second column shows that among all attributes, *air quality* has the highest WTP. Respondents are willing to pay RM8.10 for the improvement of air quality from the base level to level 1 and RM5.55 for the improvement from level 1 to level 2. However, the improvement in facilities and nursery habitat maintenance from level 1 to level 2 receives negative values. Although the improvements in these two attributes from the base level to level 1 have positive WTP values, the WTP drops significantly, showing them insignificant for further improvements.

Models 2, 3, and 4 demonstrate that all mean WTPs are significant at a 5% level, and most SDs of WTP are significant at a 10% level, except for *nur2* and *tree2* in Model 2. In all three models, the mean coefficient for distance is negative, consistent with the results in preference space. In Model 2, the improvement of air quality at the first level has the highest WTP, with a mean WTP of RM14.36, followed by the second level of air quality improvement, with a mean WTP of RM9.73. More than 96% of respondents have positive coefficients for both levels of air quality improvement. The second-highest WTP is for the reduction of noise at the first level, with a mean WTP of RM5.28. However, the mean WTP for the second level of improvement drops significantly to only RM1.88, with only 76.95% of respondents having positive coefficients for this level. The third-highest WTP is for the improvement of facilities, with mean WTPs of RM2.98 and RM2.85 for the first and second levels, respectively. Approximately 92.87% and 77.95% of respondents place positive values on the improvement of facilities at levels 1 and 2, respectively. The improvement of tree species and ecosystems at levels 1 and 2 has WTPs of RM1.51 and RM1.82, respectively. Although the WTPs are lower compared to other attributes, virtually 100% of respondents place positive values on the level 2 improvement. However, only 64.19% of respondents place positive values on the level 1 improvement. The improvement of nursery habitat maintenance has the lowest average WTP. Respondents are willing to pay RM1.43 and RM1.08 for the improvement of this attribute at levels 1 and 2, respectively, with 68.61% and 99.38% of respondents having positive coefficients for both improvements.

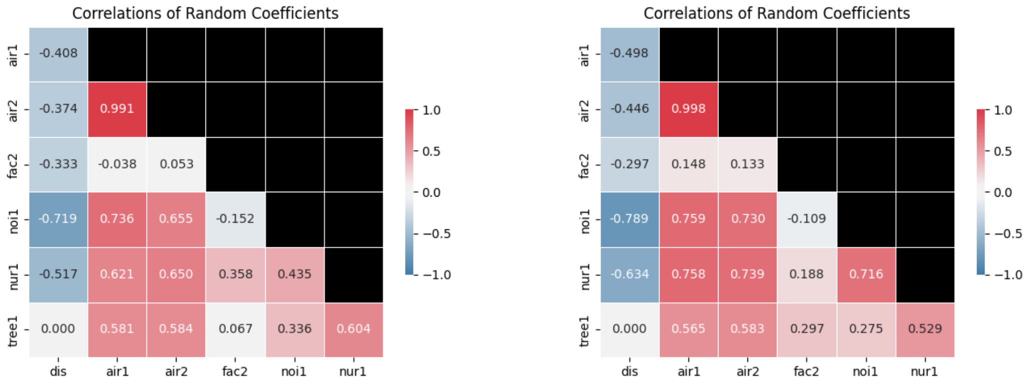
The mean WTP estimations in Models 3 and 4 do not show a significant difference compared to the mean WTPs in Model 2. However, the differences in the means and SDs of WTPs in Models 2 and 3 result in a difference in the percentage of respondents who have positive coefficients for the attributes. For instance, in model 3, 74.81% of respondents have positive coefficients for the improvement of nursery habitat maintenance at level 1, compared to only 68.61% in model 2. Moreover, 71.58% of respondents place positive values on the improvement of tree species and ecosystems at level 1 in model 3, showing a slight increase compared to 64.19% in model 2.

The upper part of Table 32 compares the model fit across Models 2-4. For the comparison between Models 2 (unconstrained) and 3 (constrained), the p-value associated with the chi-square test statistic is 0.09, indicating there is no strong evidence to prove that Model 2 is better than Model 3. However, when comparing Models 2 (unconstrained) and 4 (constrained), the p-value is given as 0.00, which provides strong evidence to indicate that the Model 2 is better than Model 4. Based on the given information, Model 3 provides a better fit to the data compared to Models 2 and 4.

As indicated in the lower part of Table 32, the p-value is reported as 0.00, suggesting that Model 5 (constrained) outperforms Model 3 (unconstrained). This is further supported by a lower log-likelihood of -2243.39 in Model 5 compared to that from Model 3. However, when comparing Model 5 (unconstrained) and Model 6 (constrained), the p-value of 0.98 indicates that Model 5 does not perform better than Model 6. Therefore, it can be concluded that Model 6 performs better than Model 5.

Based on this information, it can be assumed that the coefficients in the model should be correlated. The upper part of Table 30 displays the mean coefficient estimates. It reveals that all coefficients are positive except for distance. These results are consistent with the findings obtained in the previous MXL models.

Table 31 shows the WTP estimates expressed in Malaysian currency (RM).



(a) MXL with correlations (b) Constraints on Cholesky decomposition

Figure 10: Correlations of random coefficients in WTP-space models (Cost coefficient: Distance)

Compared to the models with no correlations, most of the WTP estimates are slightly higher in Models 5 and 6. For instance, the WTP estimates for *air1* increase from RM14.70 (Model 3) to approximately RM18.90 (Model 5). The WTP for *noi1* increases from RM5.67 in Model 3 to RM8.98 in Model 5, and the WTP for *nur1* increases from RM1.88 (Model 3) to RM3.45 (Model 5). The lower part of Table 30 presents the SDs of correlated random coefficients, all of which are significant at the 10% confidence level.

The heat plots in Figure 10 illustrate the correlations of attribute coefficients in Models 5 and 6. The coefficient for *air1* exhibits a strong and positive correlation with the coefficient for *air2*, with correlation values ranging from 0.991 to 0.998. Furthermore, *air1* and *air2* also display strong and positive correlations with *noi1*, indicated by correlation values ranging from 0.655 to 0.759. Additionally, the coefficient for *tree1* is strongly and positively correlated with the coefficients for *air1*, *air2*, and *nur1*. However, it exhibits weaker positive correlations with *fac2* and *noi1*.

The impact of travel time value on WTP

Table 33 presents the estimation results of MNL (Models 1-2) and uncorrelated MXL models in WTP-space, where the *distance* attribute was converted into travel costs (Models 3-8). The results reveal that replacing *n_tc* with *n_tca* as the negative of the individual travel cost variable does not improve the model fit, as evidenced by lower log-likelihood and higher AIC and BIC values. This suggests that the inclusion of

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Main/Mean						
alt1	0.499***	0.187***	0.17***	0.166***	0.195***	0.189***
	(0.102)	(0.047)	(0.045)	(0.039)	(0.046)	(0.046)
air1	19.766***	35.019***	35.853***	34.844***	46.103***	46.228***
	(0.179)	(1.329)	(1.419)	(1.455)	(2.569)	(2.482)
air2	13.524***	23.725***	23.689***	22.609***	29.939***	29.8***
	(0.158)	(0.883)	(1.112)	(1.240)	(1.903)	(1.904)
fac1	3.319***	7.051***	7.259***	6.97***	5.166***	5.462***
	(0.107)	(0.734)	(0.890)	(1.085)	(0.765)	(0.855)
fac2	-1.314***	6.674***	6.945***	5.406***	8.859***	8.806***
	(0.142)	(0.763)	(1.050)	(1.095)	(1.432)	(1.508)
noi1	0.817***	12.873***	13.829***	15.011***	21.896***	22.266***
	(0.195)	(0.757)	(1.047)	(1.226)	(1.587)	(1.826)
noi2	5.809***	4.597***	3.979***	3.775***	3.286***	2.757***
	(0.181)	(0.763)	(0.871)	(1.168)	(0.781)	(0.898)
nur1	6.558***	3.478***	4.574***	4.579***	8.418***	8.038***
	(0.235)	(0.867)	(1.003)	(1.161)	(1.478)	(1.378)
nur2	-2.462***	2.639***	1.842**	2.79**	2.269***	2.252**
	(0.190)	(0.910)	(0.884)	(1.107)	(0.810)	(0.889)
tree1	7.684***	3.679***	5.518***	5.666***	6.719***	7.072***
	(0.154)	(0.843)	(1.005)	(1.119)	(1.425)	(1.127)
tree2	0.799***	4.446***	4.122***	5.325***	4.561***	4.243***
	(0.230)	(0.775)	(0.897)	(1.161)	(0.820)	(0.869)
dis	1	-2.458***	-2.613***	-3***	-2.634***	-2.695***
	(.)	(0.117)	(0.097)	(0.067)	(0.106)	(0.103)
SD						
air1		15.918***	17.922***		19.61***	19.384***
		(1.238)	(1.958)		(1.613)	(1.760)
air2		13.165***	11.094***		2.104*	0
		(0.811)	(0.910)		(1.140)	
fac1		4.81***				
		(0.754)				
fac2		8.663***	8.623***		4.329***	4.623***
		(0.521)	(0.818)		(0.822)	(1.137)
noi1		4.65***	5.936***		2.197***	0
		(0.726)	(0.873)		(0.756)	
noi2		6.237***				
		(0.907)				
nur1		7.173***	6.841***		6.388***	4.506***
		(0.927)	(1.033)		(1.117)	(1.525)
nur2		1.056				
		(1.647)				
tree1		10.123***	9.674***		3.986***	4.282***
		(0.683)	(0.787)		(0.839)	(1.289)
tree2		1.29				
		(0.935)				
dis		1.552***	1.354***	0.989***	1.542***	1.456***
		(0.165)	(0.141)	(0.079)	(0.149)	(0.121)
N	9690	9690	9690	9690	9690	9690
Log-likelihood	-19232.84	-2316.18	-2320.22	-2366.34	-2243.39	-2242.56
AIC	38487.68	4678.35	4678.44	4758.69	4566.77	4551.11
BIC	38566.65	4827.54	4801.68	4843.01	4826.22	4765.16

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

Table 30: MNL and MXL models in WTP-space specification

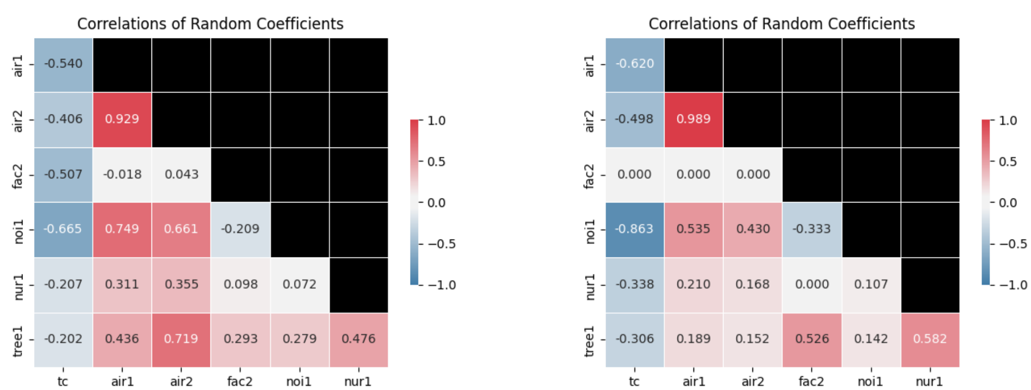
	Main/Mean WTP (Malaysian Ringgit)					
	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
air1	8.10	14.36	14.70	14.29	18.90	18.95
air2	5.55	9.73	9.71	9.27	12.28	12.22
fac1	1.36	2.89	2.98	2.86	2.12	2.24
fac2	-0.54	2.74	2.85	2.22	3.63	3.61
noi1	0.34	5.28	5.67	6.15	8.98	9.13
noi2	2.38	1.88	1.63	1.55	1.35	1.13
nur1	2.69	1.43	1.88	1.88	3.45	3.30
nur2	-1.01	1.08	0.76	1.14	0.93	0.92
tree1	3.15	1.51	2.26	2.32	2.75	2.90
tree2	0.33	1.82	1.69	2.18	1.87	1.74
	SD of WTP (Malaysian Ringgit)					
air1		6.53	7.35		8.04	7.95
air2		5.40	4.55		0.86	
fac1		1.97				
fac2		3.55	3.54		1.77	1.90
noi1		1.91	2.43		0.90	
noi2		2.56				
nur1		2.94	2.80		2.62	1.85
nur2		0.43				
tree1		4.15	3.97		1.63	1.76
tree2		0.53				

Table 31: WTP for attributes (models in WTP-space)

Model 2 to Model 3		Model 2 to Model 4	
Model 2	-2316.18	Model 2	-2316.18
Model 3	-2320.22	Model 4	-2366.34
Chi-sq test	8.08	Chi-sq test	100.32
Restrictions	4	Restrictions	10
P-value	0.09	P-value	0.00

Model 3 to Model 5		Model 5 to Model 6	
Model 5	-2243.39	Model 5	-2243.39
Model 3	-2320.22	Model 6	-2242.56
Chi-sq test	153.66	Chi-sq test	1.66
Restrictions	21	Restrictions	7
P-value	0.00	P-value	0.98

Table 32: Comparison of models in WTP-space



(a) MXL with correlations (b) Constraints on Cholesky decomposition

Figure 11: Correlations of random coefficients in WTP-space models (Cost coefficient: TC)

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7	Model 8
Main/Mean								
alt1	0.186**	0.313***	0.198***	0.203***	0.18***	0.167***	0.174***	0.168***
	(0.088)	(0.115)	(0.053)	(0.055)	(0.051)	(0.049)	(0.044)	(0.043)
air1	8.154***	12.980***	14.543***	23.249***	14.399***	23.493***	14.223***	22.784***
	(0.150)	(0.208)	(0.548)	(0.647)	(0.775)	(1.074)	(0.652)	(1.056)
air2	5.872***	8.831***	9.841***	15.576***	9.735***	15.706***	9.284***	15.124***
	(0.161)	(0.217)	(0.454)	(0.477)	(0.623)	(1.066)	(0.557)	(0.911)
fac1	1.269***	2.438***	2.927***	3.667***	2.721***	4.099***	2.896***	4.202***
	(0.110)	(0.153)	(0.313)	(0.396)	(0.516)	(0.724)	(0.491)	(0.789)
fac2	-0.309**	-0.692***	2.301***	4.354***	2.571***	4.161***	1.895***	3.162***
	(0.127)	(0.169)	(0.384)	(0.423)	(0.575)	(0.845)	(0.495)	(0.810)
noi1	0.530***	0.661***	5.766***	9.467***	5.586***	9.185***	5.954***	9.427***
	(0.146)	(0.204)	(0.403)	(0.468)	(0.448)	(0.789)	(0.557)	(0.895)
noi2	2.221***	4.139***	1.423***	2.58***	1.603***	2.686***	1.631***	2.94***
	(0.154)	(0.219)	(0.323)	(0.406)	(0.423)	(0.722)	(0.527)	(0.871)
nur1	1.931***	3.779***	1.799***	2.73***	1.305**	2.713***	1.776***	2.796***
	(0.167)	(0.236)	(0.338)	(0.406)	(0.526)	(0.802)	(0.521)	(0.819)
nur2	-0.467***	-1.182***	1.041***	1.514***	0.997**	1.598**	1.237**	1.683**
	(0.146)	(0.219)	(0.299)	(0.406)	(0.472)	(0.723)	(0.497)	(0.802)
tree1	2.610***	4.511***	2.169***	2.103***	1.696***	2.883***	2.081***	2.789***
	(0.136)	(0.176)	(0.285)	(0.410)	(0.503)	(0.806)	(0.499)	(0.762)
tree2	0.357**	0.622**	2.085***	3.195***	1.91***	3.048***	2.258***	3.819***
	(0.179)	(0.246)	(0.363)	(0.426)	(0.415)	(0.755)	(0.528)	(0.865)
tc	1		-1.518***		-1.713***		-2.095***	
	(.)		(0.157)		(0.110)		(0.072)	
tca		1		-1.852***		-2.271***		-2.602***
		(.)		(0.175)		(0.102)		(0.073)
SD								
air1			6.313***	10.476***	6.66***	9.712***		
			(1.069)	(0.563)	(0.752)	(1.473)		
air2			5.517***	10.113***	5.572***	8.159***		
			(0.477)	(0.503)	(0.737)	(1.310)		
fac1			1.857***	4.086***				
			(0.314)	(0.457)				
fac2			4.257***	6.576***	4.157***	6.65***		
			(0.386)	(0.403)	(0.449)	(0.893)		
noi1			1.922***	4.606***	2.079***	2.768***		
			(0.387)	(0.439)	(0.617)	(0.838)		
noi2			1.889***	5.107***				
			(0.347)	(0.344)				
nur1			2.942***	5.81***	2.263***	5.053***		
			(0.329)	(0.287)	(0.513)	(0.997)		
nur2			0.654***	0.188				
			(0.253)	(0.322)				
tree1			4.313***	6.027***	3.993***	5.44***		
			(0.330)	(0.359)	(0.480)	(1.030)		
tree2			1.207***	0.337				
			(0.337)	(0.423)				
tc			1.619***		1.297***		0.937***	
			(0.239)		(0.153)		(0.085)	
tca				1.781***		1.239***		0.952***
				(0.240)		(0.139)		(0.085)
N	7680	7680	7680	7680	7680	7680	7680	7680
ll	-6509.80	-10600.00	-1831.52	-1834.19	-1836.34	-1843.38	-1875.07	-1876.85
AIC	5898.09	4699.30	3709.05	3714.38	3710.68	3724.77	3776.14	3779.70
BIC	5977.06	4864.42	3852.87	3858.20	3829.49	3843.58	3857.43	3860.99

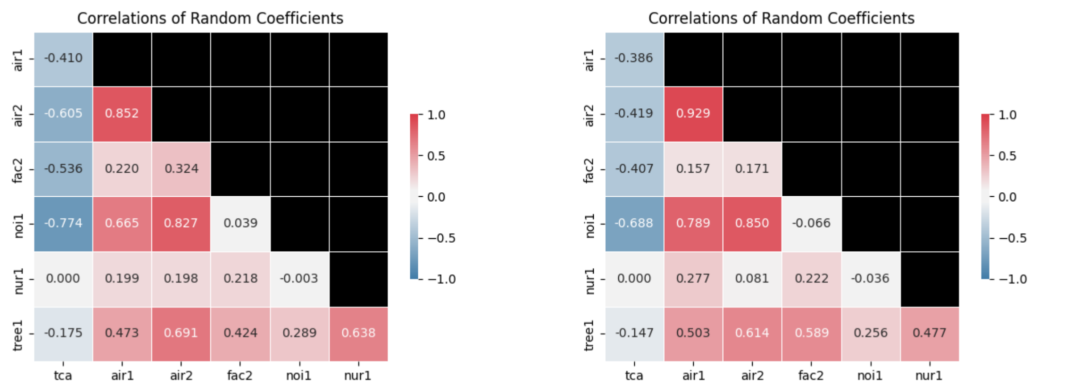
Standard errors in parentheses
* p<0.10, ** p<0.05, *** p<0.01

Table 33: MNL and uncorrelated MXL models in WTP-space specification (Travel costs)

	Model 1	Model 2	Model 3	Model 4
Main/Mean				
alt1	0.198***	0.188***	0.192***	0.186***
	(0.054)	(0.054)	(0.052)	(0.052)
air1	19.275***	28.134***	18.155***	27.56***
	(0.883)	(1.523)	(1.008)	(1.234)
air2	12.068***	20.81***	11.818***	19.666***
	(0.842)	(1.406)	(0.820)	(1.178)
fac1	2.295***	3.092***	2.12***	3.078***
	(0.295)	(0.367)	(0.300)	(0.301)
fac2	3.591***	5.903***	2.261***	4.483***
	(0.616)	(1.021)	(0.392)	(0.776)
noi1	8.881***	13.374***	8.035***	12.526***
	(0.741)	(1.121)	(0.737)	(1.014)
noi2	1.087***	2.178***	1.882***	1.929***
	(0.325)	(0.488)	(0.343)	(0.414)
nur1	2.488***	4.011***	2.645***	1.904***
	(0.676)	(1.094)	(0.561)	(0.424)
nur2	0.993***	1.17***	0.489	1.764***
	(0.287)	(0.418)	(0.321)	(0.341)
tree1	3.025***	4.549***	2.852***	4.968***
	(0.648)	(0.952)	(0.599)	(0.916)
tree2	1.52***	2.367***	1.092***	1.431***
	(0.308)	(0.412)	(0.375)	(0.429)
tc	-1.56***		-1.72***	
	(0.136)		(0.119)	
tca		-2.077***		-2.032***
		(0.126)		(0.118)
SD				
air1	8.353***	11.24***	5.802***	12.483***
	(0.612)	(0.697)	(0.574)	(0.700)
air2	2.718***	5.48***		5.391***
	(0.569)	(0.543)		(0.421)
fac1				
fac2	3.235***	5.774***	2.982***	4.998***
	(0.213)	(0.473)	(0.431)	(0.328)
noi1	1.114***	0.113	1.048***	
	(0.288)	(0.420)	(0.230)	
noi2				
nur1	0.954**	6.101***	1.524***	5.28***
	(0.402)	(0.438)	(0.386)	(0.416)
nur2				
tree1	0.073	0.152		
	(0.264)	(0.509)		
tree2				
tc	1.79***		1.549***	
	(0.184)		(0.156)	
tca		1.688***		1.84***
		(0.161)		(0.183)
N	7680	7680	7680	7680
ll	-1775.44	-1778.07	-1789.73	-1780.00
AIC	3630.87	3636.14	3635.45	3630.01
BIC	3881.00	3886.27	3810.54	3848.87

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

Table 34: MXL models with correlations in WTP-space specification (Travel costs)



(a) MXL with correlations

(b) Constraints on Cholesky decomposition

Figure 12: Correlations of random coefficients in WTP-space models (Cost coefficient: *TCA*)

travel time value in the travel cost does not result in an improved model fit.

In Models 1 and 2, where the travel costs are constrained to -1, the attribute coefficients are directly interpreted as WTP in currency units. The mWTPs are higher when *n_tca* is used as the cost variable, mainly due to its higher values compared to *n_tc*. Surprisingly, *fac2* and *nur2* have negative mWTP values in these models, indicating that respondents are not willing to pay for the improvement of these attributes at these levels, and may even require compensation for these improvements. However, the coefficients in the other models in the table do not exhibit negative signs.

In Models 3 and 4, all of the mean of the mWTPs are statistically significant at the 1% level. Model 3 shows that all of the SDs of the mWTP are statistically significant at the 1% level, while Model 4 indicates that most of the SDs of the mWTP are statistically significant at the 1% level, except for *nur2* and *tree2*, which are not significant. In Models 5 - 8, all of the means and SDs of WTP are statistically significant at the 5% level.

The estimation results of Models 3-8 reveal that the mean WTPs show a significant increase when *n_tc* is replaced by *n_tca*. For instance, in Model 4, the mean WTP for *air1* is RM23.25, which is significantly higher than the value of RM14.53 obtained in Model 3. Similarly, in Model 4, the estimated mean WTP for *air2* is RM15.58, representing a significant increase from the value of RM9.84 obtained in Model 3. Additionally, the mean WTPs for *fac2* and *noi1* in Model 4 also experience substantial increases compared to the values from Model 3. Overall, the mean WTPs for attribute levels exhibit an average increase of 52.56% when *n_tc* is replaced by *n_tca* (from Model 3 to Model 4). Comparing Models 5 and 6, the mean WTPs for attribute levels show an average increase of 66.67%, with the highest increase observed in the WTP for *nur1*, which experiences a 108% increment.

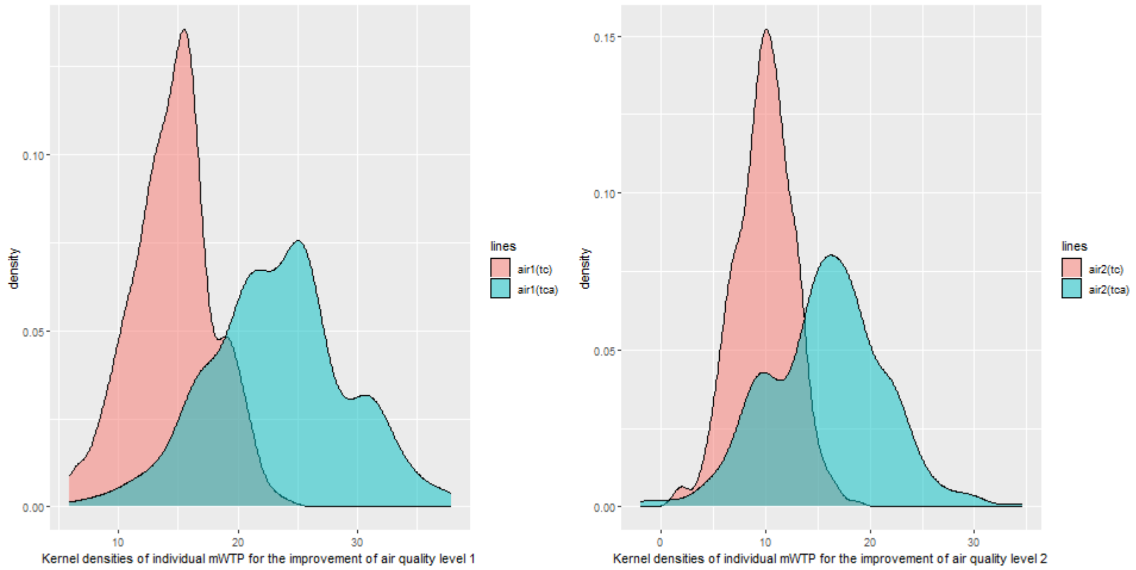
The results of Models 3 and 4 indicate that replacing *tc* with *tca* leads to changes in the proportion of positive coefficients. In Model 3, approximately 94.25% of respondents have positive values for *fac1*. However, this proportion decreases to 81.52% in Model 4. Similarly, the proportion of positive coefficients for *noi2* changes from 77.44% in Model 3 to 69.33% in Model 4. For attribute *nur1*, it changes from 72.95% to 68.08% between the two models. Conversely, for attribute *nur2*, the percentage of positive coefficients increases from 94.43% in Model 3 to 100% in Model 4.

The proportion of positive coefficients for *tree1* in Model 3 is 69.25%, but it decreases to 63.65% in Model 4. In contrast, for *tree2*, the percentage of positive coefficients increases from 95.79% in Model 3 to 100% in Model 4. These changes suggest that the inclusion of travel time value in the travel cost has an impact on respondent preferences.

Table 34 provides the estimation results of MXL models with correlations in WTP-space. Models 1 and 2 assume that seven attributes are correlated, while Models 3 and 4 constrain the insignificant Cholesky factors to be close to zero. In terms of the significance of the mean and SD of WTP estimates, most of the estimates are statistically significant at a 10% confidence level, with a few exceptions. Specifically, the mean WTP for *nur2* in Model 3 and the SD of WTP for *tree1* in Models 1 and 2 are not statistically significant.

When comparing the impact of replacing *tc* with *tca*, there are variations in model performance. Model 1, which uses *tc*, has a higher log-likelihood compared to Model 2, indicating that Model 1 fits the data better. However, Model 4, which uses *tca*, has a higher log-likelihood compared to Model 3, suggesting that the model using *tca* provides a better fit to the data.

Figures 11 (Models 1 and 3) and 12 (Models 2 and 4) display heat plots of the correlations between coefficients. Importantly, *air1* and *air2* exhibit strong positive correlations in both cases, indicating that preferences for air quality improvements are closely related. Additionally, the correlation between *noi1* and *air1*/*air2* is moderate, but it becomes stronger when *tca* is used as the cost attribute. Weak positive correlations are observed between *nur1* and *air1*/*air2* as well as between *tree1* and *noi1*, while moderate positive correlations exist between *tree1* and *fac2* and between *tree1* and *nur1*. Notably, the correlation between *tree1* and *air1*/*air2* strengthens significantly when *tca* is used as the cost attribute.



(a) Air quality level 1 (b) Air quality level 2

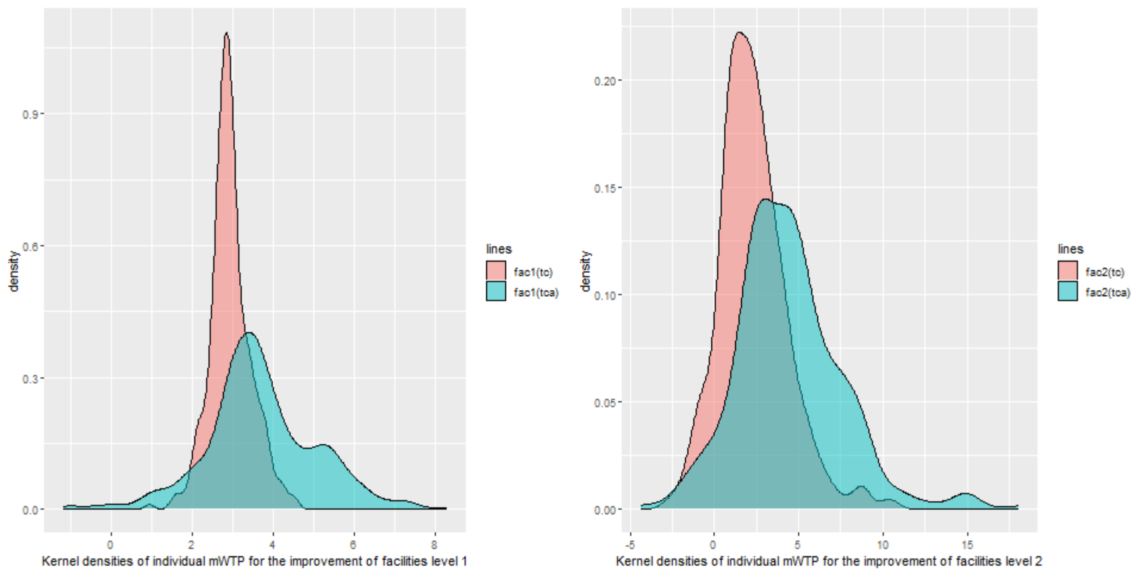
Figure 13: Kernel densities of $mWTP_i$ with and without value of travel time for the improvement of air quality

In addition, kernel density plots are employed to describe the probability density function of mWTPs, providing a visualization of the distribution of mWTP for each attribute level. Figures 13 - 17 illustrate the distribution of the mean of mWTPs for Models 3 and 4 in Table 33. Overall, the density curves for most models using *tc* are narrow, suggesting a small dispersion of mWTP values. Conversely, models employing *tca* exhibit wider density curves, indicating a larger spread of mWTP values, except for the mWTP for *nur2* and *tree2*. It is noted that the SD of mWTPs for these two attribute levels is statistically insignificant (Model 4 in Table 33). Therefore, the observed density curves in the kernel density plot may not accurately reflect the true dispersion of mWTP values for these specific attribute levels.

These results highlight the sensitivity of the estimation outcomes to the choice of cost attribute and model specification.

6.4 Discussions

Comparing the models that use distance as a cost attribute, it is evident that the MXL model estimations generally provide a better fit compared to the MNL models.



(a) Facilities level 1 (b) Facilities level 2

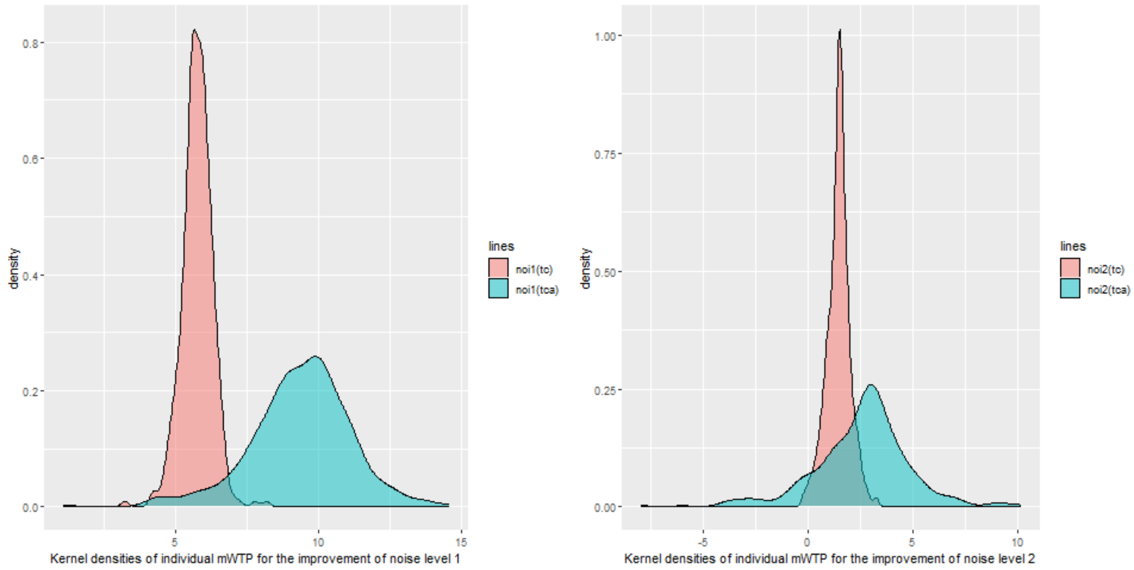
Figure 14: Kernel densities of $mWTP_i$ with and without value of travel time for the improvement of facilities

Among the MXL models, the one that assumes only some coefficients are random fits the data the best. Furthermore, the model demonstrates an even better fit when some random coefficients are assumed to be correlated.

The observation that the coefficient for distance is negatively correlated with coefficients for some attributes, such as *air quality*, *facilities*, and *noise levels*, implies that as the distance to green space increases, the perceived importance of these attributes decreases. While the results don't show explicit trade-offs between longer distances and better attribute qualities, they do indicate that distance to green space becomes a more influential factor in decision-making as it increases.

Regarding the models in WTP-space, they do not demonstrate a better fit compared to the MXL models with correlations in preference space (Table 67 in Appendix F). However, when comparing uncorrelated MXL models in preference space (Table 65 in Appendix F) and WTP-space, the model in WTP-space assuming all attribute coefficients are random has higher log-likelihood and lower AIC and BIC values compared to the model in preference space with the same assumption. However, no single model specification is definitively better, and the choice should be guided by the specific research objectives, and credibility assumptions.

In terms of individuals' mWTP for attributes, the mean and SD of mWTP values

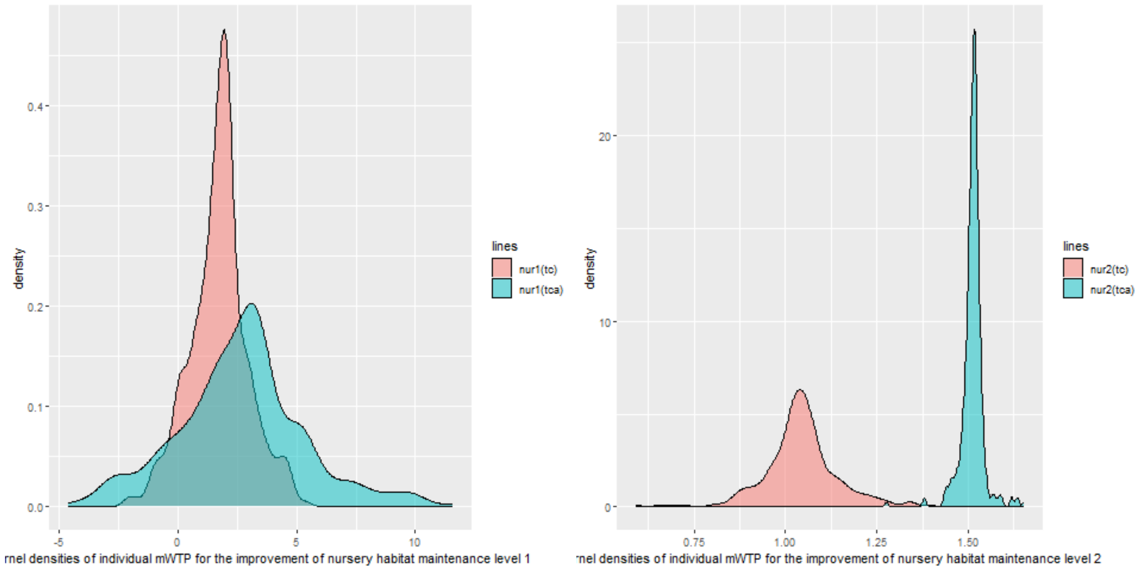


(a) Noise level 1 (b) Noise level 2

Figure 15: Kernel densities of $mWTP_i$ with and without value of travel time for the improvement of noise levels

from models in preference space are substantially higher compared to the ones from models in WTP-space. These findings raise doubts about the reasonability of the high mWTP values. In a real-world situation, the average travel cost for visits to Penang Botanic Gardens and Penang Youth Park is approximately RM1.00 when only the vehicle operating cost is included (refer to Table 18, and Figures 8a and 9a in Chapter 5). The mWTP values in WTP-space appear to be more reasonable, as they are closer to the travel costs in the real-world situation. Furthermore, the substantial difference between mWTP estimates also raises questions about individuals' actual mWTP for the attributes.

Several reasons for this discrepancy have been highlighted. First, the current method is an indirect valuation method, as individuals are not directly asked about their WTP. Second, the attributes used in the hypothetical scenarios may not accurately represent real-world situations that individuals consider when making choices, as real-world decisions are influenced by a wider range of factors. Third, individuals may exhibit hypothetical biases when making decisions in hypothetical scenarios. In such scenarios, individuals do not face the real financial consequences that they would in real-world situations, which can result in their WTP values being either understated or overstated.

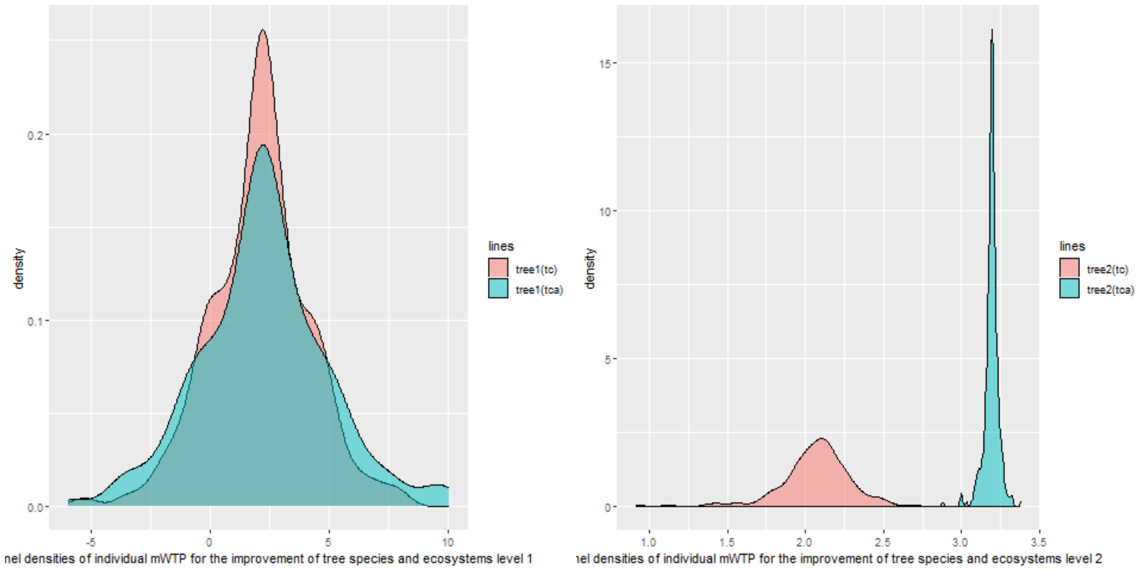


(a) Nursery habitat maintenance level 1 (b) Nursery habitat maintenance level 2

Figure 16: Kernel densities of $mWTP_i$ with and without value of travel time for the improvement of nursery habitat maintenance

Based on the results in WTP-space, it shows that individuals are more willing to pay for the improvement in air quality, as it has the highest mean WTP value. However, there is a slight decrease in WTP value for further improvement in this attribute. This decrease in WTP value for the second level of improvement in air quality could be attributed to diminishing marginal utility. As air quality improves, individuals may perceive that each additional improvement of the same attribute produces a lower benefit to them, leading to a decrease in their WTP. Similar observations can also be made for further improvements in noise levels and nursery habitat maintenance.

The investigation into the validity of the value of travel time suggests that including the value of travel time in the travel cost calculation does not result in a better model fit. Several factors may contribute to this outcome. First, the value of time calculated in the models may not accurately capture the true value of time as it occurs in the real world. The value of time is estimated based on individuals' ages, which simplifies the estimation process and may not reflect the actual income-related variations in the value of time. This estimation method limits the models' ability to effectively describe individuals' preferences for the attributes. This is also a challenge faced in other studies.



(a) Tree species and ecosystems level 1 (b) Tree species and ecosystems level 2

Figure 17: Kernel densities of $mWTP_i$ with and without value of travel time for the improvement of tree species and ecosystems

Second, the DCE questions explicitly assume that respondents are driving alone to the green site, which leads to the derived travel cost being based on this assumption. For instance, an individual who typically takes the bus to the recreational site may find it challenging to estimate the travel time of driving a car and might assume the travel time is similar to taking the bus. In reality, the travel time for taking a bus and driving a car can vary substantially. Therefore, failing to account for this heterogeneity can lead to an under- or over-representation of the value of travel time. To address these limitations, a more comprehensive estimation of the value of time should be considered. Additionally, the inclusion of the value of travel time in the travel cost calculation can introduce complexities in the model estimation process. Respondents have limited knowledge of how the value of time is estimated when making their choices, which introduces measurement errors.

However, it is important to note that the absence of model fit improvement does not necessarily invalidate the concept of the value of travel time. Instead, it highlights the need for more accurate methods to estimate the value of travel time. These results will contribute to informed decision-making in the planning, funding, design, and management of urban green spaces to maximize their benefits for the community, by investigating the value of attributes and ecosystem services provided by these spaces.

6.5 Conclusions

This study investigates the urban green site choices of local residents in the context of urban Penang Island, Malaysia. Using a data set of 404 respondents which includes a total of 9696 observations, the WTP estimations for several key attributes of urban green spaces are estimated. The significance of the coefficient estimates confirms the research question, as this model has provided us with insights into how individuals weigh various attributes against distance. These attributes are not just abstract concepts; they directly connect with the urban ecosystem services that Penang Island offers.

Additionally, the conversion of the *distance* attribute to travel costs intends to capture the expenses associated with travel, providing a richer understanding of individuals' decision-making processes. While the inclusion of the value of travel time in the travel cost calculation ensures that more aspects of participants' decision-making processes are captured. By incorporating the individual-specific value of travel time in the DCE context, this study extends the understanding of the monetary valuation of ecosystem services. This provides a better perspective on how people perceive the costs and benefits of urban green spaces.

Moreover, this study tests models which are parameterized in preference space and WTP-space, results indicate that the WTP estimations derived from models in WTP-space are more reasonable as the WTP values which use travel costs as indicators are closer to the real-world situations, a case study discussed in Chapter 5.

The findings have the potential to inform policy decisions regarding urban planning, green space preservation, and design. By understanding what attributes people value and how they weigh them against distance, which serves as an indicator for travel costs, policymakers can make informed choices that align with the preferences of the community. Furthermore, as cities continue to grow, the need for well-designed green spaces becomes increasingly important. The insights gained can help guide decisions about where to allocate resources, how to design green spaces, and how to ensure that these spaces are well-utilized by the communities. Importantly, this research bridges the gap between individual preferences, travel costs, and the valuation of ecosystem services, offering practical implications for urban planning in a sustainable way.

It is important to acknowledge several limitations in this paper. Firstly, this study focuses on a specific case study, which means that the results may not be applicable to other study areas, raising questions about their reliability. Additionally, when converting the *distance* attribute into travel costs, it is important to note that this transformation may not fully capture the complex trade-offs between distance and other attributes. Although the relationship between distance and travel cost has been stated at the beginning of the DCE question, different individuals may have varying preferences for distance, and they can be influenced by various factors, such as accessibility and environmental considerations. Lastly, individuals who have never driven a car may struggle to estimate travel costs when making choices, potentially leading to inaccurate decisions. Future research should consider individual modes of transportation and real income data to improve the accuracy of this conversion.

Chapter 7

Assessing Spatial Heterogeneity in Willingness to Pay for Urban Ecosystem Services

7.1 Introduction

Urban ecosystem services are primarily provided by urban green spaces, such as parks, gardens, and recreational sites, which are strategically located within urban areas. These spaces are important urban ecosystem assets in rapidly developing areas, which contribute to the environmental sustainability and quality of life of urban residents through the provision of urban ecosystem services.

The effective design and management of urban green spaces have the ability to control the delivery of various ecosystem services and hence maximize the benefits for the community. In order to assist authorities in making informed decisions, a straightforward approach involves estimating the non-market value of these services. The estimated value reflects individuals' preferences and willingness to pay (WTP) for these services, therefore providing valuable information for decision-making processes. However, due to the complex and heterogeneous nature of individuals' preferences and WTP, the estimation of urban green space ecosystem services can be challenging. Various non-market valuation methods have been used to measure the value of urban green spaces or their ecosystem services (Brander and Koetse, 2011; Aevermann and Schmude, 2015), such as the hedonic price method (Mansfield et al., 2005; Anderson and West, 2006; Palmquist et al., 2010), contingent valuation method (Jim and Chen, 2006b; Lindsey and Knaap, 1999; Lockwood and Tracy, 1995; Tyrväinen and Väänänen, 1998), and choice experiment (Mokas et al., 2019; Roberts et al., 2022).

The WTPs for various ecosystem services provided by urban green spaces in Penang, Malaysia were derived through the discrete choice experiment (DCE). This study seeks to advance the investigation by measuring the individual-specific marginal WTP (mWTP) for these services in a hypothetical urban green site, from the same discrete choice model. With these individual-specific mWTP data, this study proceeds to investigate the influence of individual spatial information on these data. This research aims to investigate the relationship between individual spatial data and preference heterogeneity. The individual spatial data refers to the spatial characteristics in an area where the individual is residing, it refers explicitly to the geographical characteristics within a 2-kilometre radius centred around individuals' house locations. The choice of a 2-kilometer radius is driven by the desire to capture the immediate neighborhood, recognizing that this distance includes the local surroundings that significantly influence individuals' daily experiences and perceptions. Preference heterogeneity refers to the variation in individual preferences for certain ecosystem services in urban green spaces within a sample size. The individual preferences are indicated by the individual mWTP for each attribute level, a stronger preference for an attribute level has higher WTP values. It aims to explore the influences of the diversities of spatial characteristics, including the land use areas of forests, open and recreational lands, commercial areas, agricultural lands and industrial areas, and individual socio-demographic characteristics on the variation in individual preferences.

A number of previous studies seek to establish a link between spatial characteristics and individual WTP for environmental improvements or changes. Yao et al. (2014) employed the distance to forest areas as one of the main independent variables to explain patterns of individual-specific WTP for biodiversity enhancement. Czajkowski et al. (2017) used data on forest land area and distance from places of residents to explain the variability in WTP for forest management. Vollmer et al. (2016) examined the influence of distance from the river on the WTP for urban river rehabilitation. Schaafsma and Brouwer (2020) employed a choice experiment to estimate the values of environmental quality improvements and further explored individual spatial substitution patterns. Roberts et al. (2022) investigated the influence of proximity to green spaces and their area or similar facilities around the place of residents on the patterns of variability in WTP for green space. Valck et al. (2017) used the spatial densities of nature substitute sites to explain individual preferences for nature restoration. Campbell et al. (2009) investigated the spatial dependence of individual-specific WTP for rural landscape improvements. Badura et al. (2020) utilized individualised choice maps to explain the effect of places of residents on the values concerning land use change. Liu et al. (2020) employed urban air pollution data which represented the spatial distribution of pollution levels to explain the heterogeneity in WTP for green space. While certain studies have investigated the impact of spatial characteristics on the heterogeneity of mWTP for urban green spaces and their attributes, there is currently no identified research that specifically examines the combined influence of the following spatial features, including total land areas of forests, commercial areas, agricultural lands, and industrial areas in a single model, as per the current state of knowledge. Therefore, exploring the link between these spatial features and mWTP for urban green space attributes could enhance urban green space planning and management.

In this context, two research questions have been identified:

1. How is the spatial distribution of individual-level mWTP for a specific attribute level across Penang Island characterized?

2. How do land use categories, such as forests, open and recreational lands, commercial areas, agricultural lands, and industrial areas in individual neighborhoods of residences, impact their mWTP for attributes associated with urban green spaces?

The first research question aims to visually represent the distribution of individual-level mWTP on a map, enabling an exploration of areas with relatively higher or lower mWTP values. The second research question seeks to investigate the relationship between land use patterns within a 2-kilometre radius (in the neighborhood) of the respondent's residence and their individual mWTP for attributes related to hypothetical urban green sites. To examine the validity of the effects of neighborhood land use patterns on individual mWTP, a Seemingly Unrelated Regression (SUR) model was developed.

This study is structured as follows. Section 7.2 describes the study design. The survey design for the DCE was outlined in Section 3.1.1. The details of the data collection process were discussed in both Section 3.2 and Section 3.2.1. Removing incomplete data, a total of 404 valid DCE samples were analysed. The study's methodologies were addressed in Chapter 2. More precisely, Section 2.5.5 covered the process of extracting individual-level mWTP from a MXL model, while Section 2.6 elaborated on the theoretical framework of the SUR model. Section 7.3 discusses the methods of deriving scaled mWTP values in order to better visualize the spatial distribution of mWTP for attribute levels. Section 7.4 describes the SUR models used for data analysis. Section 7.5 reports the study's results. Section 7.6 concludes the findings and discusses implications for policies and future research.

7.2 Study design

This study aims to explore the relationship between individual spatial information and individual mWTP for attribute levels in urban green spaces, extracted from the MXL model in the WTP-space specification (Model 2 in Table 30 in Chapter 6).

Building upon the research objective, the following hypotheses were formulated:

- H_0: There is no significant impact of individual spatial characteristics and socio-demographic factors on individual mWTP for urban green space attributes.

- H_1: The individual spatial characteristics and socio-demographic factors have a significant impact on individual mWTP for urban green space attributes.

The hypotheses assert that the combination of spatial characteristics (the land use areas of forests, open and recreational lands, commercial areas, agricultural lands, and industrial areas within a 2-kilometre radius area centred around an individual's house location) and socio-demographic factors influences individual mWTP for the 10 urban green space attributes (as stated in Model 2 in Table 30).

During the survey, participants were requested to provide their house location or the area where their residence is situated. A 2-kilometre radius area centred around an individual's house location was defined, and information regarding land use within this area was extracted.

The individual-level mWTP values for each attribute level were then mapped to visualize the distribution of these values across various areas within the study region. This analysis aims to discover whether mWTP values exhibit variations based on whether individuals reside in areas offering more ecosystem assets and services.

Lastly, this study used a SUR model to investigate the impacts of individual spatial characteristics and socio-demographic factors on individual mWTP for the various attribute levels (refer to Section 2.6 for an explanation of the rationale behind choosing SUR as the methodology).

7.3 Spatial distribution of individual-level mWTP

With regard to the cost attribute being represented by *distance* and not expressed in monetary terms, the individual-level mWTPs for individual n ($\hat{E}_s[WTP_n]$) have not been derived. Instead, they are expressed as the individual-level marginal willingness to travel to visit a green site (in kilometres). Therefore, to calculate the mWTPs expressed in monetary units, $\hat{E}_s[WTP_n]$ is multiplied by RM0.41, the estimated return-trip travel cost for each additional kilometre (km).

For an enhanced visualization of the mWTP distribution on the map, the individual mean and standard deviation (SD) of mWTP values for a specific attribute level are normalized to a range of 0 to 1. This normalization process enables meaningful comparisons between different levels within the same attribute. The scaled mean and SDs of mWTP are calculated using Equations 7.1 and 7.2.

$$scaled\,mean\,WTP_n = \frac{mean\,WTP_n - mean\,WTP_{min}}{mean\,WTP_{max} - mean\,WTP_{min}} \tag{7.1}$$

$$scaled\,sd\,WTP_n = \frac{sd\,WTP_n - sd\,WTP_{min}}{sd\,WTP_{max} - sd\,WTP_{min}} \tag{7.2}$$

where $mean\,WTP_n$ represents the means of mWTP for individual n, $mean\,WTP_{min}$ is the minimum means of mWTP across all individuals, $mean\,WTP_{max}$ is the maximum means of mWTP across all individuals. $sd\,WTP_n$ represents the SDs of mWTP for individual n, $sd\,WTP_{min}$ is the minimum SDs of mWTP across all individuals, $sd\,WTP_{max}$ is the maximum SDs of mWTP across all individuals.

The scaled means and SDs of mWTP for each individual will be depicted as coloured plot markers on the Penang Island land use map. These preference indicators range from yellow to red, with yellow denoting the lowest mWTP value and red signifying the highest. The colour intensity is meaningful only when comparing the two levels for the same attribute. The maps were generated using Python 3.10 and the `folium` package.

Figure 18 displays the land use map, where distinct land use categories are represented by various colours on the map.

7.4 SUR model specification

7.4.1 Derivation of spatial data

This study involves the analysis of individual spatial data concerning land use areas, encompassing forests, open and recreational lands, commercial areas, agricultural lands, and industrial areas around individuals' house locations. The process comprises several steps, such as plotting individuals' house locations on the land use map, establishing a radius of 2 kilometres around each house location, masking the area beyond the radius circle, computing the number of pixels with specific colours within the circle, and converting the aggregate pixel counts into land area measurements for each category. The estimation process was executed using Python 3.10.

In the initial step, individual house locations were plotted on the Penang land use

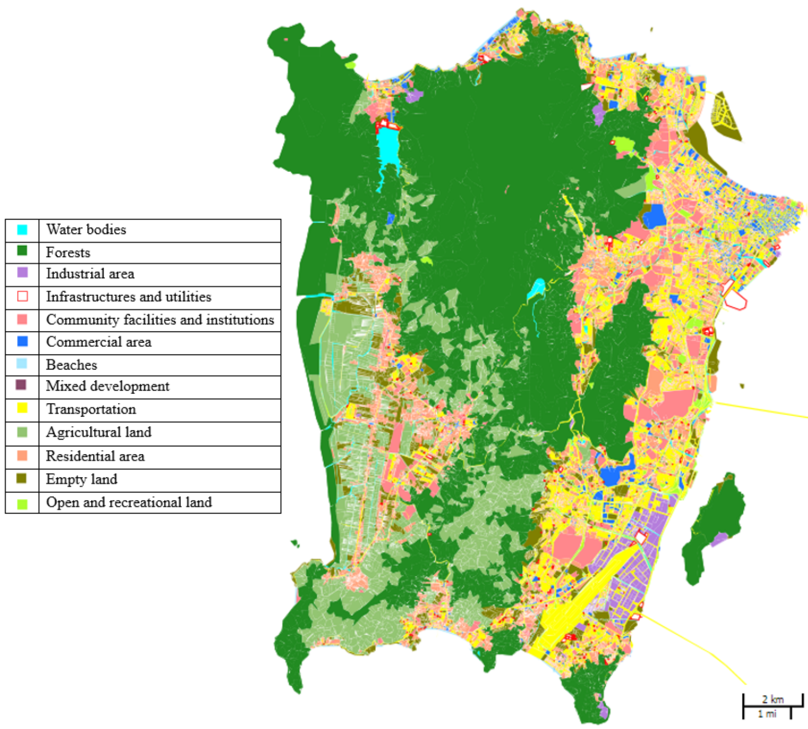


Figure 18: Penang land use map (Source: Peninsular Malaysia Town & Country Planning Department, 2023)

map, and the latitude and longitude coordinates for each location were determined. The subsequent step involved establishing a radius circle around each plotted location, with a radius of 2 kilometres defining the area of interest around each house location. This means that the study focused on the land use distribution within a 2-kilometre distance from each plotted location. The radius circle was visually represented by a circle outlined in red on the map. The first two steps were executed using Python's `folium` package.

The third step involved applying a masking technique to create a black mask outside the boundary of the radius circle, effectively removing unrelated land areas while preserving the land area within the circle. This process was achieved using Python's `cv2` and `numpy` packages. An example of the resulting image is depicted in Figure 19.

The subsequent step involved calculating the number of pixels of specific colours within the resulting image. An image is composed of a large number of pixels, and this step determines how many pixels correspond to specific land use categories within the image. Since different land use categories were represented by distinct colours on the land use map, these categories could be easily identified by detecting their respective RGB values. The colours of the land use categories were translated into RGB values, and a tolerance range of 30 RGB units was applied to account for minor variations in RGB values. A higher pixel count of a specific colour indicated a larger area of the corresponding land use category. The pixel count estimation was carried out using the `image` package in Python 3.10.

The final step involved converting the pixel counts for each land use category into land area measurements. A 2-kilometre radius circle has a total area of 12.56637 km^2, calculated using the equation: $Area\, of\, circle = \pi \times radius^2$). In this case, the total number of pixels within the circle is 8201 units. Consequently, one unit of the pixel corresponds to 0.0015 km^2 of land area. The total land use area for a specific category can be derived by:

$$Land\, use\, area_{wn} = Pixel\, count_{wn} \times 0.0015\, km^2 \tag{7.3}$$

where $Land\, use\, area_{wn}$ represents the total land use area of land use category w for individual n, $Pixel\, count_{wn}$ represents the total pixel count of land use category w for individual n.

7.4.2 Model variables

The dependent variables are the individual-level means of mWTP for improving attribute levels. The attribute levels include air quality level 1 (*air1*), air quality

Figure 19: Example of land use map area of an individual within a 2-KM radius from house location

level 2 (*air2*), facilities level 1 (*fac1*), facilities level 2 (*fac2*), noise level 1 (*noi1*), noise level 2 (*noi2*), nursery habitat maintenance level 1 (*nur1*), nursery habitat maintenance level 2 (*nur2*), tree species and ecosystems level 1 (*tree1*) and tree species and ecosystems level 2 (*tree2*).

The independent variables in this study are categorized into three groups:

1. Individual spatial data: the variables include the total land use areas of forests (*frst*), open and recreational lands (*recr*), commercial areas (*comm*), agricultural lands (*agri*) and industrial areas (*inds*).

2. Socio-demographic characteristics: the variables include age (*age*), education level (*educ*), employment status dummy variable (*employed*) and household size (*household*).

3. Visit patterns to urban green spaces: the variable includes visit frequency to urban green space (*visit freq*).

The *age* of the respondent is calculated as the difference between the year of completing the survey and the year of birth, as self-reported in the survey. *Educ* represents the education level of the respondent, categorized into seven distinct categories. *Employed* is a binary dummy variable with *1* denoting individuals who work either full-time or part-time, and *0* indicating others. *Household* records the household size of the respondent, reflecting the number of people residing in the household. Lastly, *visit freq* measures the frequency of visits to urban green sites in the past year and comprises five levels, ranging from *never* to *more than once a week*. Additional details about these variables used in the SUR model are provided in Table 35.

Variables	**Variables**	**Description**
$meanWTP_k$	*mean WTP*	Means of mWTP for improvement of attribute level k
X_1	*frst*	Total land use areas of forests
X_2	*recr*	Total land use areas of open and recreational lands
X_3	*comm*	Total land use areas of commercial areas
X_4	*agri*	Total land use areas of agricultural lands
X_5	*inds*	Total land use areas of industrial areas
X_6	*age*	Year of completing survey - Year of birth
X_7	*educ*	Education level: 1=primary school education; 2=secondary school education; 3=O-level or equivalent; 4=A-level/Diploma or equivalent; 5=Bachelor's degree or equivalent; 6=Master's degree or equivalent; 7=Doctoral degree or equivalent
X_8	*employed*	Employment status: 1=Employed either full-time or part-time; 0=Otherwise
X_9	*household*	Household size: number of people in the household
X_{10}	*visit freq*	Respondent's visit frequency to urban green sites in the past year: 0=never; 1=a few times a year; 2=a few times a month; 3=once a week; 4=more than once a week

Table 35: Description of variables used in SUR model

7.4.3 Correlation analysis

Correlation analysis techniques were employed to assess the presence of multicollinearity within the model. Multicollinearity occurs when two or more independent variables within a model exhibit high correlation, which can complicate the estimation and interpretation of the model's results. To evaluate multicollinearity, Pearson's correlation coefficient was utilized. The results of the correlation analysis, presented in Appendix G, revealed that none of the pairs of independent variables exhibited strong correlations, as evidenced by correlation coefficient values below 0.6. However, a moderate negative correlation of -0.496 was observed between *age* and *educ*. This moderate correlation suggests a heightened likelihood of detecting multicollinearity. Consequently, to determine which variable to exclude from the analysis, various factors such as theoretical relevance, research objectives, and correlation coefficients with other variables were considered. Given that *age* also demonstrates weak correlation with *visit freq* and lacks theoretical relevance to the dependent variable, the decision was made to eliminate the *age* variable to ensure a stable model estimation.

7.4.4 The seemingly unrelated regression models

The SUR model is specified as:

$$\begin{aligned} meanWTP_{attr,t} = \alpha_i + \beta_{1,i}X_{1,t} + \beta_{2,i}X_{2,t} + \beta_{3,i}X_{3,t} + \beta_{4,i}X_{4,t} + \beta_{5,i}X_{5,t} \\ + \beta_{7,i}X_{7,t} + \beta_{8,i}X_{8,t} + \beta_{9,i}X_{9,t} + \beta_{10,i}X_{10,t} + \mu_{it} \end{aligned} \tag{7.4}$$

where $meanWTP_{attr,t}$ represents the means of mWTP for attribute level *attr* by observation t, α_i is the constant term for each i^{th} equation. A constant term was added to each model except for the first model. $\beta_{1,i}$ represents the coefficient estimation of independent variable X_1 for each i^{th} equation and $X_{1,t}$ represents the independent variable X_1 for t^{th} observation, and so on. As there are 10 attribute levels *attr*, a SUR model that consists of 10 linear regression models was estimated. The μ_{it} is the error term for each equation i for each observation t.

The SUR model was estimated in R using the `systemfit` package.

7.5 Results

Statistics of individual spatial information

The individual land use areas for each category are expressed in km^2. An example of total land use areas for each category for the first ten respondents is shown in Table

Individual	Forests	Open and recreational lands	Commercial areas	Agricultural lands	Industrial areas
1	1.096	0.438	0.034	0.077	0.002
2	0.000	0.000	0.000	0.000	0.000
3	4.952	0.622	0.484	0.567	0.008
4	0.000	0.127	0.230	0.196	0.078
5	1.005	0.428	0.038	0.074	0.005
6	3.576	0.181	0.179	0.126	0.244
7	1.623	0.280	0.044	0.086	0.008
8	0.791	0.095	0.093	0.316	0.524
9	2.245	0.285	0.051	0.083	0.008
10	0.656	0.369	0.037	0.069	0.006

Table 36: Total land use areas within a 2-km radius of individual's house location in 2022 (km^2)

Var.	Obs	Mean	SD	Min	Max	0%	25%	50%	75%	100%
frst	317	1.987	2.169	0.000	9.83	0	0.011	1.405	3.216	9.8328
recr	317	0.201	0.121	0.000	0.637	0	0.127	0.184	0.261	0.637
comm	317	0.285	0.248	0.000	0.837	0	0.057	0.207	0.466	0.837
agri	317	0.488	0.863	0.052	6.578	0.052	0.101	0.202	0.345	6.578
inds	317	0.150	0.332	0.000	2.433	0	0.011	0.041	0.1	2.433

Table 37: Summary statistics of spatial data (km^2)

36. The table shows that Individual 2 had zero-value land use areas in all categories, indicating that the corresponding respondent's house was not located in the study area. As a result, the corresponding observation was removed from the data analysis. Table 37 shows the summary statistics of individual-level spatial data. A total of 317 valid samples were recorded. A total of 87 observations were removed, including 70 that did not provide their residential address and 17 who provided an address outside the research area. On average, the forest had the largest area within a 2 km radius of individuals' house location, which is consistent with the total land use distribution as shown in the land use map, where forests occupied almost half of the total area in Penang Island. The second largest land use category is agricultural lands, followed by commercial areas and open and recreational lands. Industrial areas had the smallest area on average, which was about 13 times less than the land use area of forests.

7.5.1 Individual-level mWTP from discrete choice model analysis

The summary statistics of means and SDs of conditional mWTPs for ten attribute levels are presented in Table 38, with values expressed in Malaysian currency (RM). The sample size consists of 404 respondents who completed the DCE survey, and the data was obtained from Model 2 in Table 30, as discussed in Section 6.3.1. The model estimation involved 2000 MLHS simulation draws for each individual. Consequently, the $mean\,WTP_{ki}$ for each attribute k and each individual i represents the average of 2000 simulated draws of mWTPs generated during the model estimation.

Among all respondents, air quality level 1 exhibited the highest mean individual-level mWTP, with a value of RM14.32. This attribute also demonstrated the widest range of means of mWTP, spanning from RM1.39 to RM23.12. The second-highest means of mWTP was observed for air quality level 2, with an average value of RM9.73, ranging from RM1.06 to RM16.85.

The improvement of noise level 1 had a means of mWTP value of RM5.27, ranking as the third-highest means of mWTP among all attribute levels. However, the means of mWTP for the improvement from noise level 1 to level 2 was RM1.88, indicating a significant decrease compared to the mWTP for the first level.

The enhancement of facilities from the base level to level 1 had means of mWTP of RM2.88, while the further improvement from level 1 to level 2 had means of mWTP value of RM2.71. Although these mWTP values were closely aligned, the SDs of means of mWTP were RM0.53 for level 1 and RM1.53 for level 2, illustrating a greater dispersion of the means of mWTP for the second level. This dispersion is corroborated by a wider range of mean mWTP values across respondents, with a minimum value of negative RM2.45 and a maximum value of RM9.23.

The means of mWTP for nursery habitat maintenance for the improvement of the first and second levels were RM1.40 and RM1.09, respectively. Similarly, the means of mWTP for the improvement of tree species and ecosystems for the first and second levels were RM1.62 and RM1.82. These two attributes exhibited relatively lower mWTP values compared to the other attributes. However, the means of mWTP for tree species and ecosystems level 1 had an SD of RM2.02 and ranged from negative RM5.71 to RM8.90, indicating substantial variability in mWTP values across individuals.

The individual-level SDs of mWTP for each attribute level measures the dispersion of mWTP values within an individual. A higher SD indicates greater variability in mWTP for a particular attribute within an individual.

Among the attribute levels, air quality level 1, which had the highest means of

Variable	Obs	Mean	Std. Dev.	Min	Max
$meanWTP_{air1}$	404	14.324	3.610	1.385	23.118
$meanWTP_{air2}$	404	9.733	2.670	1.060	16.853
$meanWTP_{fac1}$	404	2.879	0.531	1.272	4.933
$meanWTP_{fac2}$	404	2.713	1.528	-2.446	9.233
$meanWTP_{noi1}$	404	5.274	0.512	3.381	7.113
$meanWTP_{noi2}$	404	1.879	0.915	-1.384	4.541
$meanWTP_{nur1}$	404	1.398	1.123	-2.347	4.806
$meanWTP_{nur2}$	404	1.085	0.051	0.826	1.294
$meanWTP_{tree1}$	404	1.616	2.022	-5.706	8.899
$meanWTP_{tree2}$	404	1.819	0.070	1.418	2.142
$sdWTP_{air1}$	404	5.457	0.892	3.102	8.370
$sdWTP_{air2}$	404	4.640	0.603	3.026	6.849
$sdWTP_{fac1}$	404	1.875	0.133	1.251	2.501
$sdWTP_{fac2}$	404	3.196	0.306	2.210	3.955
$sdWTP_{noi1}$	404	1.831	0.135	1.271	2.476
$sdWTP_{noi2}$	404	2.374	0.201	1.521	3.170
$sdWTP_{nur1}$	404	2.706	0.276	1.902	3.754
$sdWTP_{nur2}$	404	0.426	0.027	0.276	0.525
$sdWTP_{tree1}$	404	3.598	0.426	2.419	5.302
$sdWTP_{tree2}$	404	0.520	0.033	0.414	0.780

Table 38: Summary statistics of means and SDs of conditional mWTP for attribute levels

mWTP, also exhibited the largest SDs of mWTP, with an average value of RM5.46. Following this, air quality level 2 had an average SD of RM4.64.

Interestingly, the average SDs of mWTP for *fac2* and *tree1* were RM3.20 and RM3.60, respectively, indicating a substantial dispersion of mWTP values within individuals for these attributes. This suggests that individuals are more uncertain about their mWTP for the corresponding attribute levels.

Conversely, the individual SDs of mWTP for *nur2* and *tree2* had average values of RM0.43 and RM0.52, respectively, indicating less variability in mWTP values within individuals on average. This can be interpreted as individuals being more certain about their perceived mWTP values for these two attributes.

Kernel density plot for the individual mean estimates of mWTP

In this section, kernel density plots are employed to describe the probability density function of mWTPs, providing a detailed visualization of the distribution of mWTP for each attribute level.

Figure 20 illustrates the distribution of means of mWTP for the improvement of air quality levels 1 and 2. The density curve for *air1* exhibits a slight left skew, indicating that the mean is less than the median, with a peak around RM16.00. In contrast, the density curve for *air2* is symmetric, peaking at approximately RM11.00, signifying a balanced distribution of mWTP values across individuals.

The distributions of means of mWTP for the improvement of facilities levels 1 and 2 are displayed in Figure 21. The density curve for *fac1* is narrow and centred at around RM2.70, suggesting a small dispersion of mWTP values. Conversely, *fac2* exhibits a wider density curve, peaking at around RM2.30, indicating a larger spread of mWTP values.

Figure 22 presents the distributions of means of mWTP for the improvement of noise levels 1 and 2. The density curve for *noi1* peaks at approximately RM5.50, characterized by a narrower curve, signifying a smaller dispersion and higher concentration of mWTP values around the centre point. In contrast, *noi2* features a slightly wider density curve with a lower concentration of mWTP values at the centre point (approximately RM2.00), indicating higher uncertainty in the distribution.

The distribution of means of mWTP for the improvement of nursery habitat maintenance at both levels is shown in Figure 23. The density curve for *nur1* is extremely narrow, centred at around RM1.10, implying a minimal dispersion of mWTP values. Conversely, the density curve for *nur2* peaks at approximately RM1.40 and is almost flat, indicating a very high degree of uncertainty in the distribution.

Finally, Figure 24 presents the distribution of means of mWTP for the improvement of tree species and ecosystems at both levels. *Tree1* is characterized by a very thin curve, peaking at around RM2.00, signifying a minimal spread of mWTP values. In contrast, *tree2* exhibits a wide curve, peaking at around RM2.00, indicating a highly dispersed distribution of mWTP values.

7.5.2 Spatial distribution of mWTPs

The spatial distribution of individual-level mWTP, as discussed in Section 7.3, is visualized using the land use maps. On the map, the scaled means and SDs of mWTPs for each individual, corresponding to each attribute, are represented by varying shades, transitioning from yellow to red.

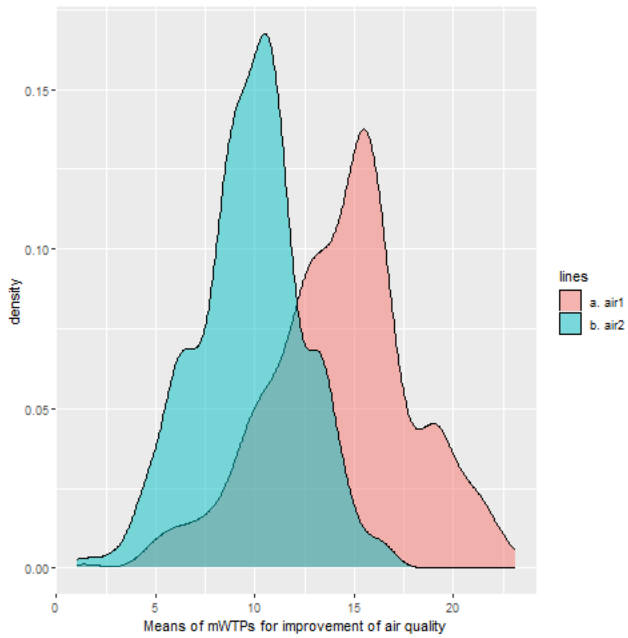


Figure 20: Means of mWTP for improvement of air quality

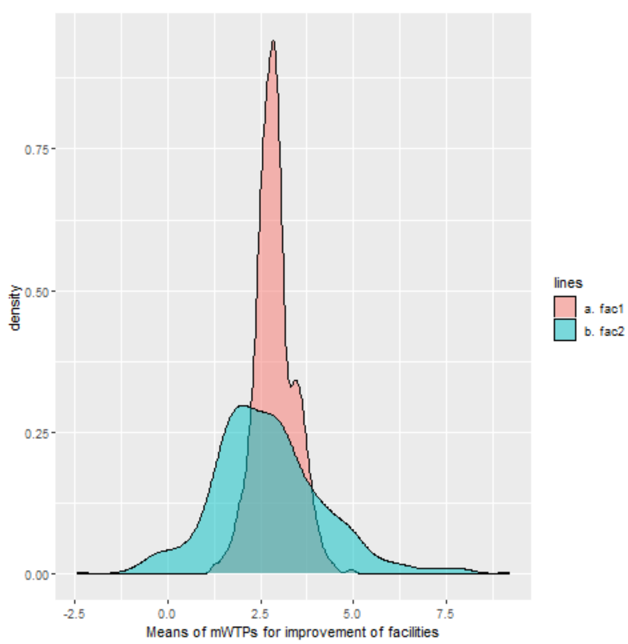


Figure 21: Means of mWTP for improvement of facilities

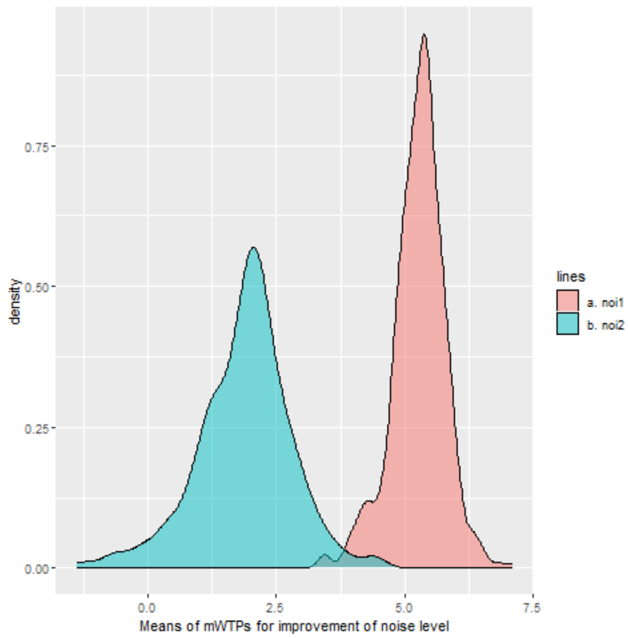


Figure 22: Means of mWTP for improvement of noise levels

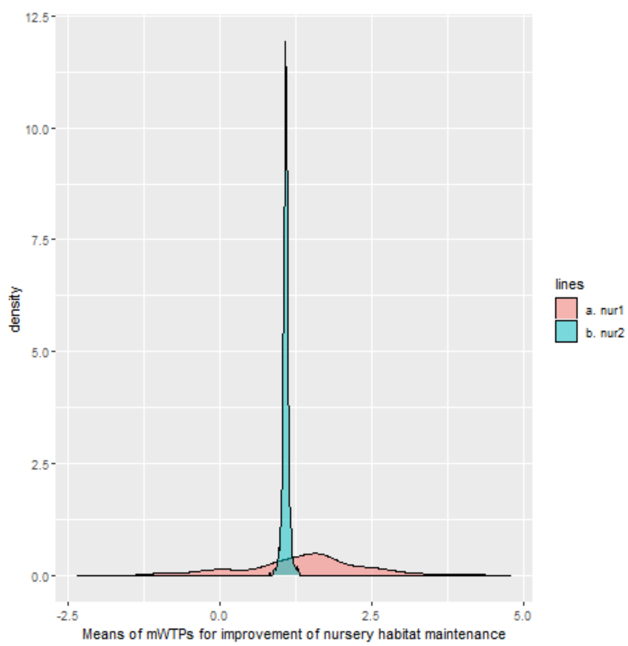


Figure 23: Means of mWTP for improvement of nursery habitat maintenance

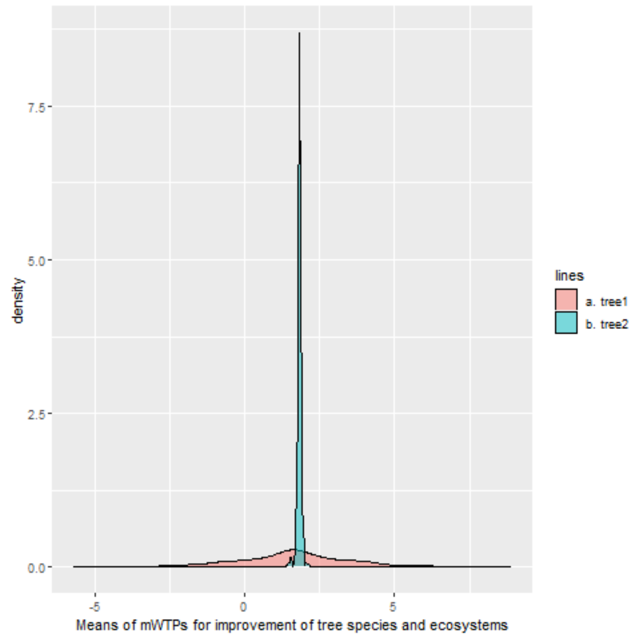


Figure 24: Means of mWTP for improvement of tree species and ecosystems

Individual-level means of mWTP

The spatial distribution of individual-level means of mWTP for air quality is depicted in Figure 25. In this figure, the left panel displays the distribution of individual means of mWTP for the first level, while the right panel illustrates those for moving from the first to the second level of air quality. The mWTP for air quality level 1 exhibits relatively higher values in comparison to the second level, as indicated by the stronger intensity of colours. In these plots, the red areas signify higher mWTP values. It is evident from the figure that a concentration of red plots is located near the forests and along the northeastern part of Penang Island, which also corresponds to the city centre, George Town. Despite the mWTPs for air quality level 2 being relatively lower, the figure demonstrates that individuals residing near the forests and the city centre are more willing to pay compared to others.

Figure 26 provides insight into the spatial distribution of individual-level means of mWTP for facilities. The plots for the means of mWTP concerning the first-level facilities exhibit a more balanced colour intensity compared to the second-level facilities. Conversely, the plots for the second-level facilities display a wider range of colour gradients, indicating greater diversity in individual means of mWTP. The spatial information suggests that individuals residing near commercial areas (depicted in blue) and city centres are more willing to pay for both levels of facility improvements.

Figure 27 illustrates the spatial distribution of mWTP for noise reduction. The mWTP for noise reduction level 1 exhibits relatively higher values when compared to level 2, as indicated by the stronger colour gradient for the first level. The colour gradient of the plots related to mWTP for noise reduction level 1 is more balanced, with a slightly weaker colour intensity in the southeastern part of the island, which is surrounded by commercial and industrial areas. On average, the mWTPs for noise reduction level 2 are relatively higher for individuals residing near forests and agricultural areas.

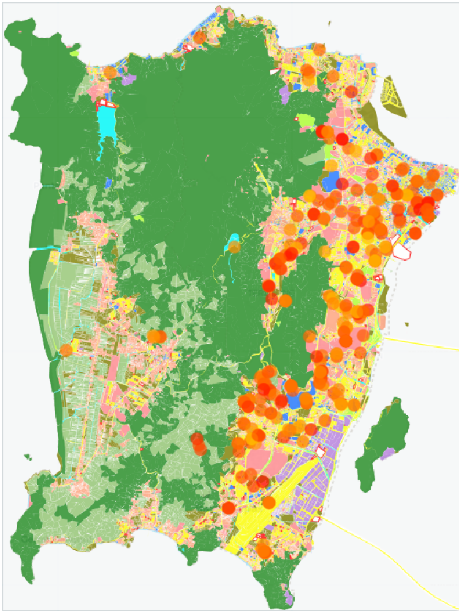
Figure 28 provides insights into the spatial distribution of means of mWTP for nursery habitat maintenance. In comparison to the second-level improvement, the colour plots for the first-level improvement exhibit stronger colour intensities on average and a wider range of colour gradients. Darker colour plots are concentrated in the city centres and southeastern parts of the island, particularly in commercial areas and areas adjacent to forests. The means of mWTP for second-level nursery habitat maintenance is more balanced in comparison, although a notable number of darker colour plots are observed around commercial areas.

Figure 29 illustrates the spatial distribution of means of mWTP for tree species and ecosystems. The plots for means of mWTP concerning the first-level improvement display a larger gradient, indicating greater diversity in mWTP values across individuals. A significant observation from the figure is that individuals residing in the city centre are willing to pay a higher price for the first level of improvement. Conversely, the second level of improvement exhibits a more balanced colour intensity in the plots, suggesting that individual means of mWTP do not vary significantly across individuals for this attribute level.

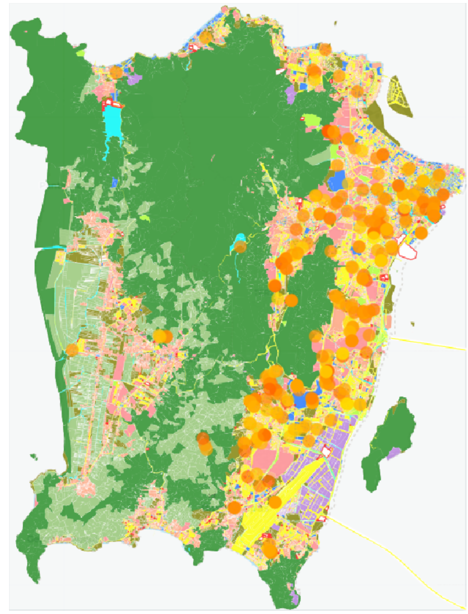
Individual-level SDs of mWTP

The spatial distribution of individual-level SDs of mWTP for attributes is also examined. The figure on the left depicts the individual-level SDs of mWTP for the first level of attribute improvement, while the figure on the right represents the SDs of mWTP for the second level of attribute improvement.

Figure 30 illustrates the SDs of mWTP for air quality. The SDs of mWTP for the first level are relatively larger compared to the second level, as evidenced by the stronger colour intensities. Moreover, the individual SDs of mWTP for the first level exhibit greater diversity, as reflected in the varying intensities of colours. However, this diversity in SDs is more likely to be attributed to stochastic individual preferences, as there is insufficient evidence to suggest that individuals residing near

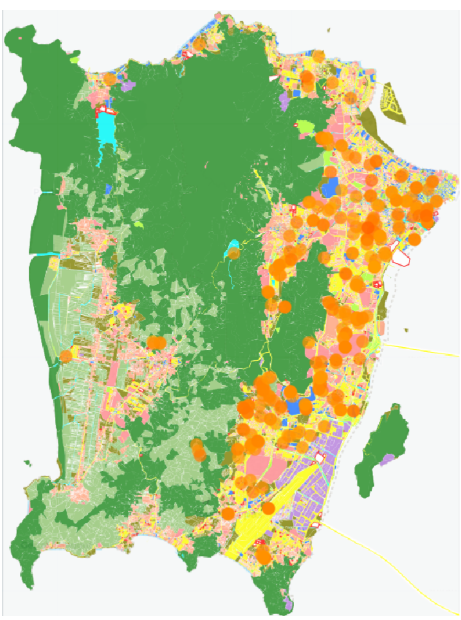


(a) Air quality level 1

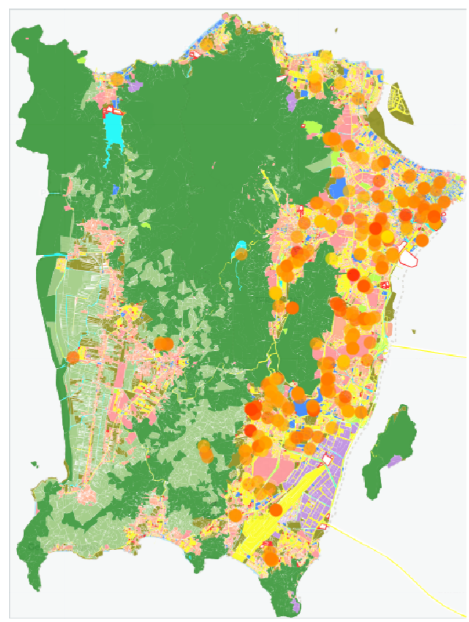


(b) Air quality level 2

Figure 25: Means of mWTP for improvement of air quality

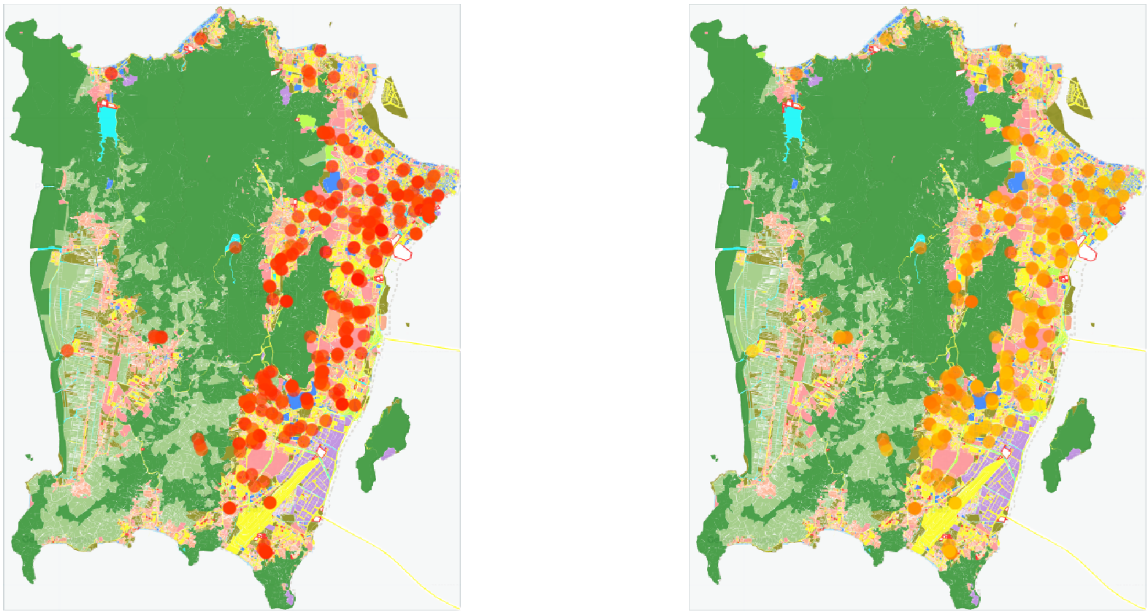


(a) Facilities level 1



(b) Facilities level 2

Figure 26: Means of mWTP for improvement of facilities



(a) Noise reduction level 1

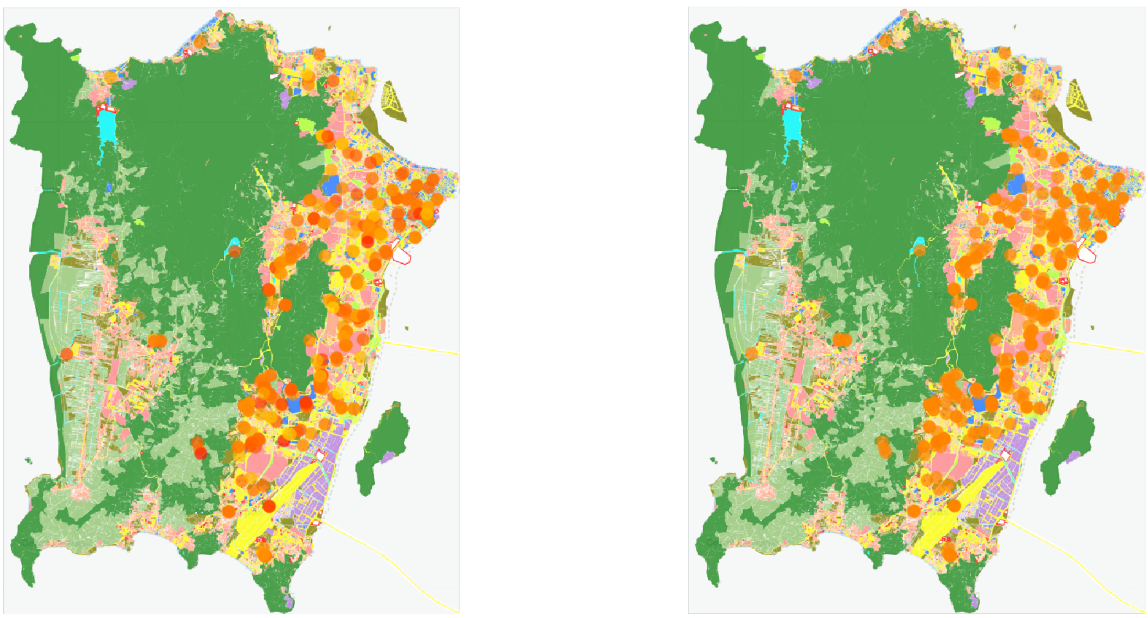
(b) Noise reduction level 2

Figure 27: Means of mWTP for reduction of noise levels

green areas have more consistent mWTP values. In contrast, the SDs of mWTP for the second level are smaller in comparison to the first level, but they demonstrate more consistency, as indicated by the relatively balanced colour intensity.

Figure 31 illustrates the spatial distribution of individual-level SDs of mWTP for facilities. The colour gradient of the plots for the first level of improvement is relatively balanced, indicating that there is no specific group of respondents residing in a particular area with consistently high or low mWTP values. The SDs of mWTP for the second level, on the other hand, exhibit more red plots compared to the first level, suggesting that mWTP values are more scattered around the mean. Additionally, the plots for the second level display a wider range of colour gradients, indicating greater diversity in individual SDs of mWTP. It is evident that individuals residing near forests and commercial areas (indicated by blue areas) tend to have more dispersed mWTP values for the corresponding attribute level.

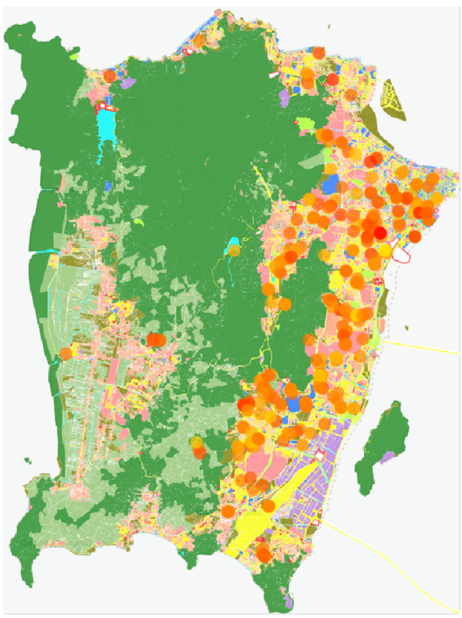
Figure 32 presents the spatial distribution of SDs of mWTP for noise reduction. The plots for the SDs of mWTP for the first-level improvement have lighter colours, while those for the second level have substantially darker colours. This suggests that mWTP values within individuals for the first-level improvement are more consistent compared to the second-level improvement. Notably, individuals living near green



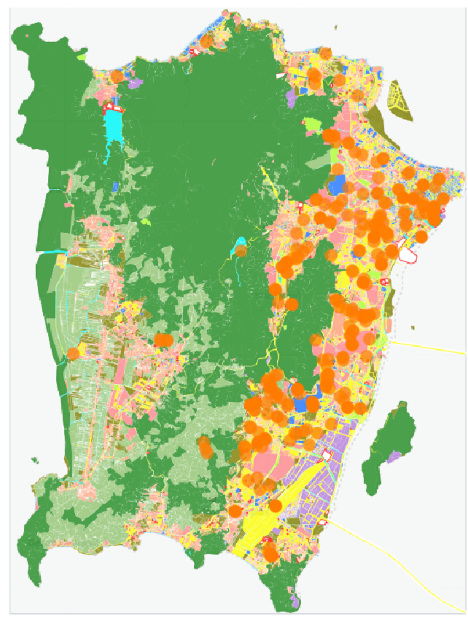
(a) Nursery habitat maintenance level 1

(b) Nursery habitat maintenance level 2

Figure 28: Means of mWTP for improvement of nursery habitat maintenance



(a) Tree species and ecosystems level 1



(b) Tree species and ecosystems level 2

Figure 29: Means of mWTP for improvement of tree species and ecosystems

areas or in commercial areas tend to have larger SDs of mWTP for both levels of noise reduction. A larger SD indicates that individuals' mWTP for the corresponding attribute level is more diverse.

Figure 33 presents the spatial distribution of the SDs of mWTP for nursery habitat maintenance. The plots for the first-level improvement exhibit substantially stronger colour intensities compared to the second-level improvement, and they display a wider range of colour tones. This suggests two observations: first, the mWTP values within individuals for the first-level improvement are more diverse, and second, the individual SDs of mWTP values are more scattered. In the case of the first-level improvement, the darker colour plots are concentrated in commercial areas and areas near forests. Conversely, for the second-level improvement, the colour gradient is more balanced, and the SDs are smaller, indicating a higher level of certainty in individuals' mWTP for the attribute level.

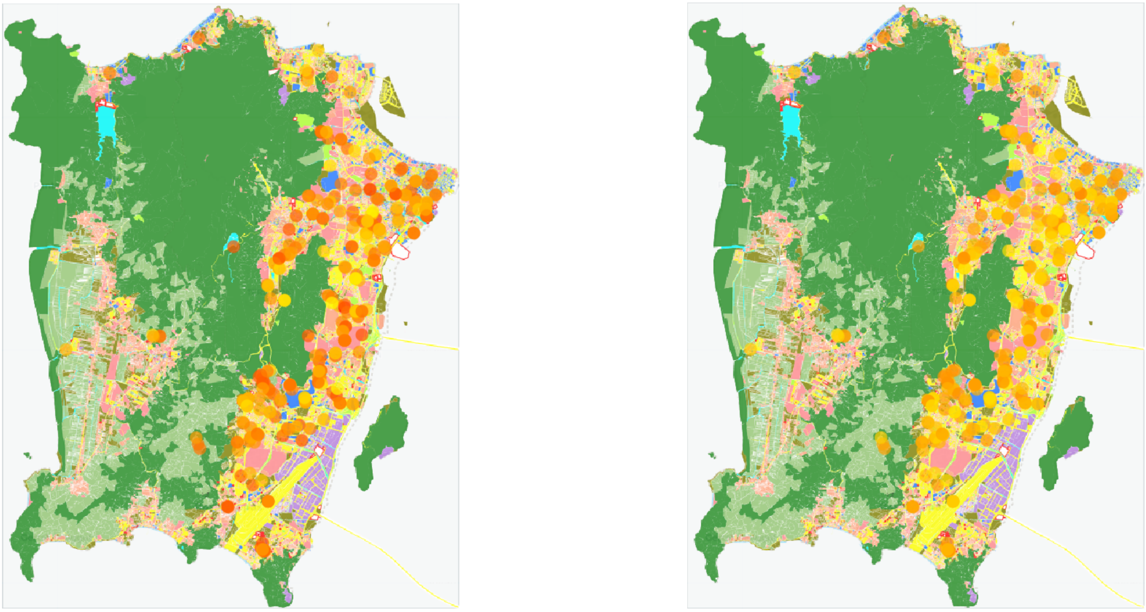
The spatial distribution of the SDs of mWTP for tree species and ecosystems is shown in Figure 34. The SDs of mWTP for tree species and ecosystems level 1 have relatively higher values compared to the second-level improvement, as indicated by the darker colours for the first level. The colour gradient of the plots at this level is also more diverse, with darker colour plots concentrated in commercial areas and forests. In contrast, the individual SDs of mWTP for the second-level improvement are relatively lower and exhibit less variability, as evidenced by the plots' nearly uniform colour tones.

7.5.3 SUR model results

The SUR results are presented in Table 39. The codes used for estimation and results in R are shown in Section H.1 in Appendix H.

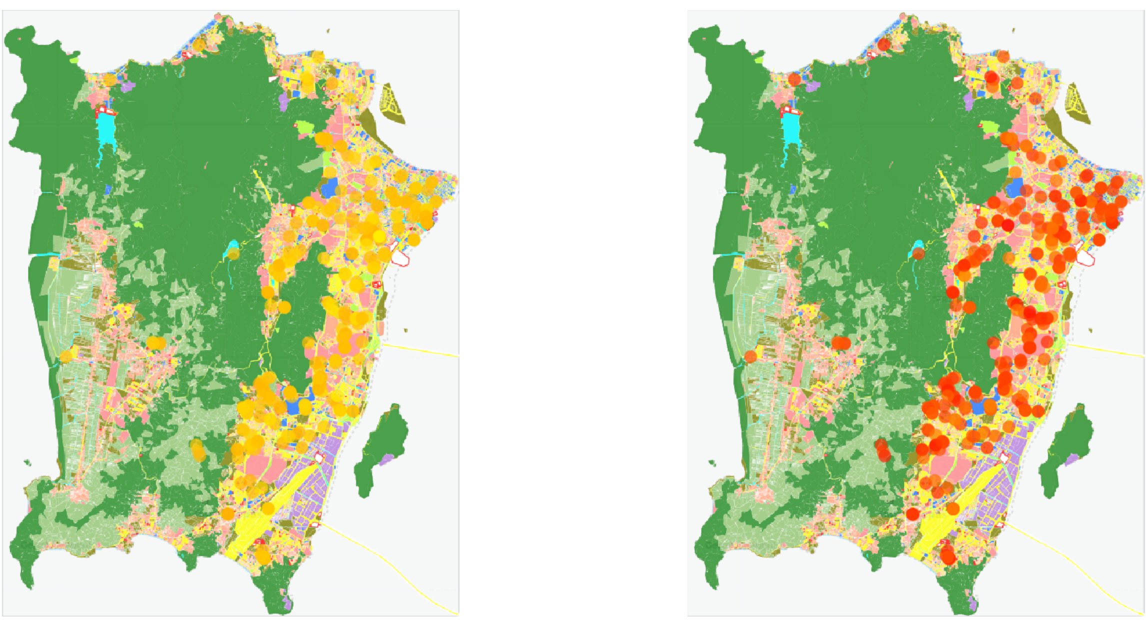
In presenting the results of the SUR model, it is important to note that the overall explanatory power of the model is limited, and most of the variables emerges as statistically insignificant, commonly referred to as 'negative results'. I have chosen to report these results to offer a complete picture of the model's performance. It's crucial to recognize that these insignificant findings provide valuable insights into the complexity of the research context. Simultaneously, the significant variables contribute positively to the understanding of the factors influencing the mWTP for attribute levels.

All models exhibit very low R-squared values, suggesting that the independent variables in these models do not explain a significant amount of variance in the mWTPs. In Model 1, only the coefficient estimation of *employed* is statistically significant at the 5% level. While the coefficient carries a negative sign, suggesting



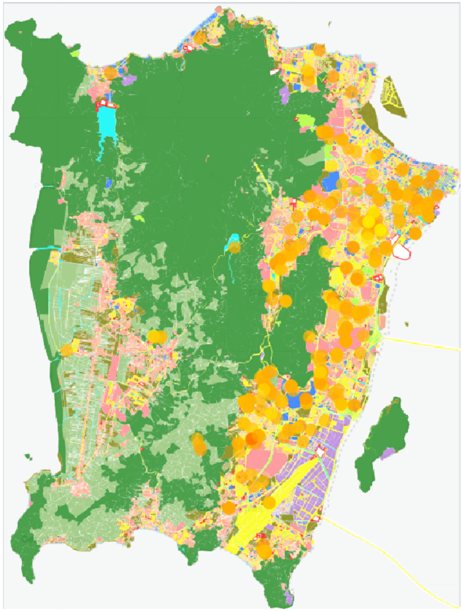
(a) Air quality level 1 (b) Air quality level 2

Figure 30: SDs of mWTPs for improvement of air quality

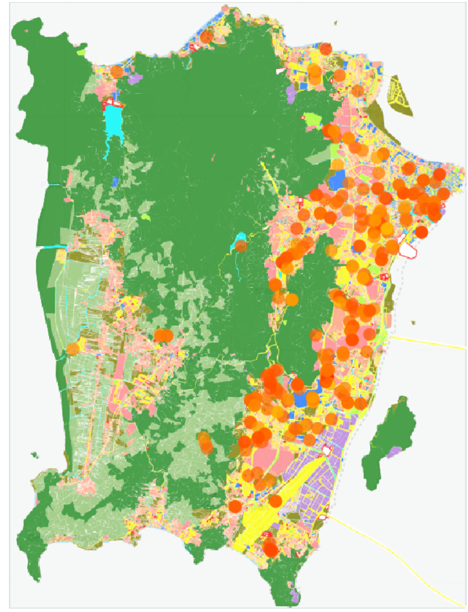


(a) Facilities level 1 (b) Facilities level 2

Figure 31: SDs of mWTPs for improvement of facilities



(a) Noise reduction level 1



(b) Noise reduction level 2

Figure 32: SDs of mWTPs for reduction of noise levels

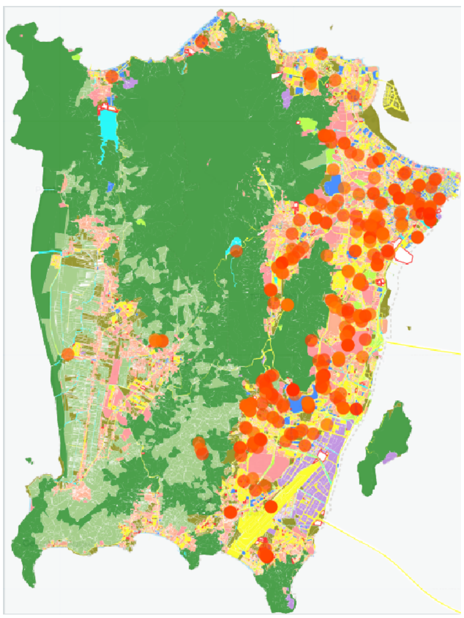
that employed respondents were willing to pay less for air quality improvement from level 0 to 1, compared to unemployed respondents. In Model 2, the coefficient estimate of *visit freq* carries a negative sign, suggesting that individuals who visited urban green spaces more frequently were less willing to pay for air quality improvement from level 1 to 2.

In Model 3, the coefficient estimate of *recr* carries a negative sign and is statistically significant at the 10% level. This implies that the land use area of open and recreational lands surrounding respondents' houses of residence had a negative effect on individuals' mWTP for the improvement of facilities from level 0 to 1 in a green space.

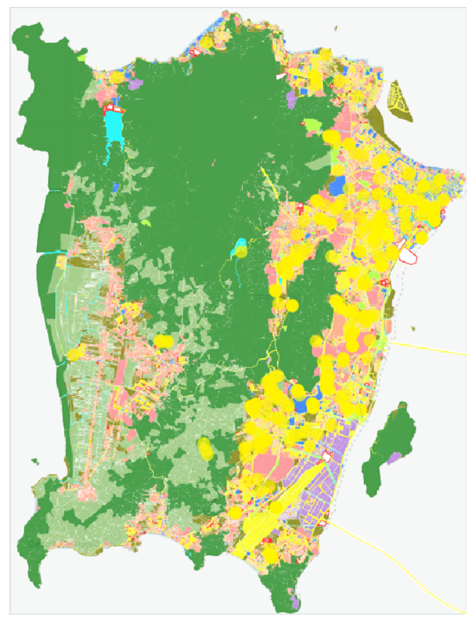
In Model 7, the coefficient estimate of *educ* is positive and statistically significant at the 5% level. This implies that individuals with higher education levels were willing to pay higher prices for the improvement of nursery habitat maintenance from level 0 to 1.

The coefficient estimate of *comm* in Model 9 is negative and statistically significant at the 5% level, suggesting that the land use area of commercial areas surrounding respondents' houses of residence had a negative effect on individuals' mWTP for the improvement of tree species and ecosystem from level 0 to 1.

While an attempt was made to estimate the spatial models, the obtained results

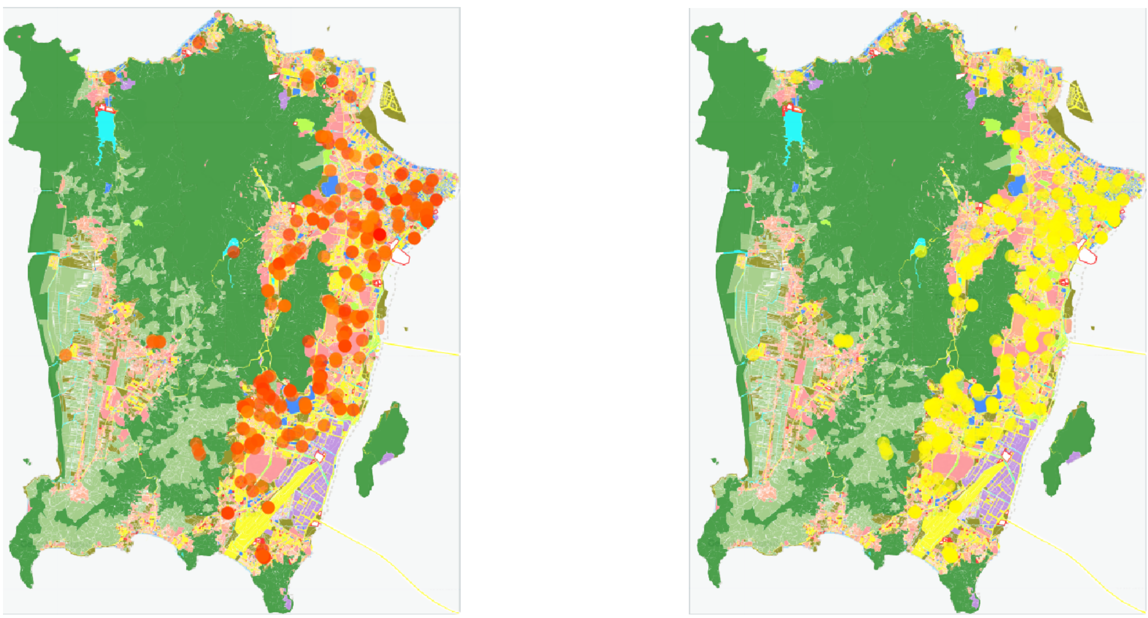


(a) Nursery habitat maintenance level 1



(b) Nursery habitat maintenance level 2

Figure 33: SDs of mWTPs for improvement of nursery habitat maintenance



(a) Tree species and ecosystems level 1

(b) Tree species and ecosystems level 2

Figure 34: SDs of mWTPs for improvement of tree species and ecosystems

revealed a lack of explanatory power. The spatial models failed to capture significant spatial dependencies or provide valid insights into the respondents' preferences. This suggests that spatial factors might not play an important role in this context or the data collected was insufficient to accurately identify these effects. Therefore, I have decided not to include detailed results in the main text. However, the complete results can be found in Section H.2 in Appendix H for those interested in further exploration.

Model	1	2	3	4	5	6	7	8	9	10
Ind. Var.	$mWTP_{air1}$	$mWTP_{air2}$	$mWTP_{fac1}$	$mWTP_{fac2}$	$mWTP_{noi1}$	$mWTP_{noi2}$	$mWTP_{nur1}$	$mWTP_{nur2}$	$mWTP_{tree1}$	$mWTP_{tree2}$
Coefficient										
frst	0.038	0.08	-0.009	-0.02	<0.000	0.034	0.005	<0.000	-0.087	<0.000
	(0.101)	(0.075)	(0.015)	(0.044)	(0.015)	(0.027)	(0.032)	(0.001)	(0.059)	(0.002)
recr	1.436	1.185	-0.535.	-0.021	-0.122	0.117	-0.295	-0.019	0.403	0.053
	(1.862)	(1.374)	(0.279)	(0.804)	(0.274)	(0.492)	(0.587)	(0.027)	(1.091)	(0.037)
comm	0.553	-0.551	0.031	-0.448	0.102	-0.176	0.154	0.009	-1.043*	0.021
	(0.870)	(0.642)	(0.130)	(0.376)	(0.128)	(0.230)	(0.274)	(0.013)	(0.510)	(0.017)
agri	-0.029	-0.018	-0.022	0.149	-0.041	0.011	0.09	0.001	0.094	0.001
	(0.254)	(0.187)	(0.038)	(0.110)	(0.037)	(0.067)	(0.080)	(0.004)	(0.149)	(0.005)
inds	-1.042	-0.583	-0.051	0.251	-0.095	-0.217	0.247	-0.001	0.109	-0.011
	(0.684)	(0.505)	(0.103)	(0.296)	(0.101)	(0.181)	(0.216)	(0.010)	(0.401)	(0.014)
educ	-0.053	0.06	0.004	-0.077	0.009	-0.003	0.099*	<0.000	-0.025	0.001
	(0.128)	(0.095)	(0.019)	(0.055)	(0.019)	(0.034)	(0.040)	(0.002)	(0.075)	(0.003)
employed	-0.894*	-0.064	-0.08	0.122	0.06	0.129	-0.008	0.003	0.017	0.001
	(0.424)	(0.313)	(0.064)	(0.183)	(0.062)	(0.112)	(0.134)	(0.006)	(0.249)	(0.008)
household	0.082	0.103	0.003	0.048	0.012	-0.012	-0.014	-0.001	0.044	0.004
	(0.135)	(0.100)	(0.020)	(0.058)	(0.020)	(0.036)	(0.043)	(0.002)	(0.079)	(0.003)
visit freq	-0.275	-0.291*	0.031	0.026	-0.04	-0.03	-0.055	0.001	-0.043	-0.008*
	(0.168)	(0.124)	(0.025)	(0.073)	(0.025)	(0.044)	(0.053)	(0.002)	(0.098)	(0.003)
cons	14.772***	9.525***	2.952***	2.703***	5.274***	1.921***	1.108***	1.082***	1.906**	1.802***
	(1.046)	(0.772)	(0.157)	(0.452)	(0.154)	(0.276)	(0.330)	(0.015)	(0.613)	(0.021)
N	316	316	316	316	316	316	316	316	316	316
DF	306	306	306	306	306	306	306	306	306	306
SSR	3985.246	2170.139	89.561	743.621	86.364	278.345	396.430	0.853	1369.489	1.556
MSE	13.024	7.092	0.293	2.430	0.282	0.910	1.296	0.003	4.475	0.005
RMSE	3.609	2.663	0.541	1.559	0.531	0.954	1.138	0.053	2.116	0.071
R-sq	0.036	0.040	0.020	0.022	0.021	0.028	0.037	0.006	0.024	0.032
Adj R-sq	0.007	0.012	-0.009	-0.007	-0.008	-0.001	0.009	-0.023	-0.005	0.003

Standard errors in parentheses
. p<0.1, * p<0.05, ** p<0.01, *** p<0.001

Table 39: SUR model results ($mWTP$=$meanWTP$)

7.6 Discussions and conclusions

The analysis of the relationship between spatial characteristics and mWTP heterogeneity offers valuable insights for authorities tasked with efficient land use management and resource allocation. By visualizing the spatial distribution of individual mWTP values for improvements in urban green space characteristics, areas with higher mWTP values can be identified as priority targets, thereby enhancing environmental management efforts. Given the constraints of funding for urban green space improvements, analyzing how individual-level spatial information influences preferences for these spaces allows authorities to make informed decisions regarding fund allocation. This paper begins by detailing the process of deriving individual-level spatial data from the land use map and obtaining conditional mWTP values for improvements in urban green space characteristics through DCE. Subsequently, a SUR model is employed to explore the relationship between these mWTP values and the derived spatial information.

Although the model exhibits little explanatory power, there are two important findings worth discussing. The results indicate that individuals residing in areas with a greater abundance of open and recreational lands within a 2-kilometer radius were willing to pay less for improvements in facilities within urban green spaces. It could be related to the perception of existing facilities. If the surrounding areas already have abundant open and recreational spaces, individuals may perceive the need for additional improvements as less important, leading to a lower WTP for further improvements.

The results also indicate that individuals residing in areas with larger land areas of commercial zones were willing to pay less for improvements in tree species and ecosystems within urban green spaces. This could be attributed to residents in areas with substantial commercial land prioritizing economic activities over environmental enhancements. They may perceive the commercial spaces as more important to them compared to enhancements in tree species and ecosystems, leading to a lower mWTP, especially if residents view the current state of tree species and ecosystems as less critical than economic considerations.

These findings highlight the influence of specific land use characteristics on individual mWTP for various urban green space improvements, providing valuable insights for environmental management and resource allocation decisions.

Chapter 8

Analyzing the Impact of Improved Urban Green Space Attributes: Simulated Changes in Consumer Surplus

8.1 Introduction

Many cities worldwide are currently experiencing rapid urbanization and economic growth, resulting in a surge in population and economic activities. While this urban expansion brings significant benefits, it also presents several challenges, including environmental deterioration, climate change, and a decrease in green spaces. Urban green spaces (UGS), which encompass various forms like parks, gardens, forests, vegetable allotments, and fields, serve as areas for recreational activities, ultimately enhancing the well-being of urban residents.

However, as the demand for land in urban areas rises to accommodate economic activities and infrastructure development, the availability of green spaces significantly reduces. This leads to a disruption of ecological balance, a reduction in biodiversity, and a decline in the provision of ecosystem services within urban areas. Consequently, it has a detrimental impact on the well-being of inhabitants and the sustainability of urban development. Unfortunately, these effects are not considered in the System of National Accounts (SNA), which primarily relies on GDP as an indicator, recording economic transactions occurring within the market. Neglecting these aspects can introduce bias into decision-making processes.

Weitzman (1976) has emphasized the importance of assessing economic well-being in terms of sustained and future consumption levels. Consequently, efforts have been made to propose the inclusion of environmental assets and services within the SNA (Caparrós et al., 2017; Howarth and Farber, 2002; Heal and Kriström, 2005). Nordhaus and Kokkelenberg (1999) proposed the inclusion of ecosystem assets and services in national accounts as part of a broader effort to develop more comprehensive economic indicators. This approach recognises that economic and social welfare extends beyond traditional market transactions and includes various 'near-market' and non-market activities.

In response to this recognition, the United Nations has introduced the System of Environmental-Economic Accounting Central Framework (SEEA-CF) (United Nations et al., 2014b), and the SEEA Experimental Ecosystem Accounting framework (SEEA-EEA) (United Nations et al., 2014a). SEEA is the first well-established international standard framework for environmental accounting, initially adopted by the UN Statistical Commission in 2012 (Brouwer et al., 2013; Tapsuwan et al., 2021). This framework provides a comprehensive methodology for measuring both the physical units and the monetary costs and benefits associated with ecosystem assets and services, in alignment with the SNA (Obst and Vardon, 2014). The United Nations has recently introduced an updated version of the SEEA Ecosystem Accounting (SEEA-EA) framework. This framework is spatially based and is designed as an integrated statistical framework for organizing biophysical information related to ecosystems, measuring ecosystem services, monitoring changes in ecosystem extent and condition, valuing ecosystem services and assets, and establishing connections between this data and indicators of economic and human activities (United Nations et al., 2021).

In the context of the monetary valuation of ecosystem assets and services within the SEEA framework, the primary focus is on recording exchange values (Comte et al., 2022; Fenichel and Obst, 2019). When considering goods or services acquired without a direct monetary transaction, the SEEA suggests using the prices of equivalent goods or services in alternative markets as a reference point. However, it is worth noting that the economists' traditional approach to valuing ecosystem assets and services, which assesses the willingness to pay (WTP) and consumer surplus (CS), falls under the category of welfare values (United Nations et al., 2014a).

Welfare values, as noted in ecosystem accounting (Caparrós et al., 2003; Obst et al., 2016), are excluded in the SEEA framework. While Discrete Choice Experiments (DCE), which ultimately derive WTP and CS, are deemed unsuitable for monetary valuation in the SEEA framework. However, to capture the true value of urban green spaces and their ecosystem services, it is crucial to incorporate welfare values. Therefore, this study aims to estimate the welfare changes resulting from changes in the provision of ecosystem assets and services within the UGS on Penang Island. These changes include improvements in attributes like air quality, noise levels, nursery habitat maintenance, and tree species and ecosystems.

This chapter introduces a simulation framework designed to estimate individual CS, individual number of visits, individual travel costs, and derive exchange values of the UGS. The study focuses on Penang Island as its research area. While recognizing the abundance of UGS, the analysis assumes the existence of only five urban green spaces on Penang Island, all offering free access. Additionally, it considers that no new entries are expected in the short term. Therefore, the annual number of visits reported by respondents will be distributed among these five sites exclusively.

Within this context, a key research question arises: How do changes in the provision of ecosystem assets and services across different UGS impact welfare values under different scenarios? This research question provides a basic understanding of the effects of changes in the flow of ecosystem services and the stock of ecosystem assets. This lays the foundation for understanding the relationship between the quality and extent of ecosystem services and assets and the welfare values attributed to these spaces. The results provided by this analysis will enable authorities, urban planners, and stakeholders to make informed decisions and conduct comprehensive cost-benefit analyses. Ultimately, this contributes to the promotion of sustainable development within Penang's urban green spaces.

This study is structured as follows. Section 8.2 describes the design of the study. Section 8.3 presents and discusses the findings of the study. Section 8.4 concludes the findings and discusses implications for policies and future research.

8.2 Study design

A framework performing the simulation analysis to compute changes in welfare values in response to variations in attribute levels within selected urban green spaces is established. It includes modifying the attribute levels, initially represented by dummy codes, to match the scenario descriptions of interest. When attributes are enhanced to higher levels at a particular green space, it results in alterations in individual-level total utility for that specific site. Consequently, this impacts the individual CS estimates. The changes in CS estimates capture the marginal benefits as individuals make choices within different scenarios.

To conduct the analysis, the individual-level coefficient estimates from the DCE were used. Section 3.1.1 discusses the DCE's survey design, while Section 3.2.1 provides details on the data collection process.

The DCE assumes hypothetical sites in urban Penang Island in the country of Malaysia, without referring to any specific urban green space. Nonetheless, for this study, five actual UGS located in various urban areas on Penang Island are selected. These green spaces include The Esplanade, Sia Boey Urban Archaeological Park, Areca Park, Bukit Dumbar Park, and Relau Metropolitan Park, with each offering unique characteristics. To simplify the analysis, this study focuses solely on the six attributes used in the DCE. Except for *distance* attribute, it assumes consistent attribute levels across individuals for the five selected urban green spaces. The *distance* variable is allowed to be uniquely estimated for each individual, computing the distance from the individual's house of residence to the five UGS of interest, determined based on driving distances obtained from Google Maps data. The characteristics of these green spaces are discussed in Section 8.2.1.

It is expected that the parameter estimates for the *distance* attribute will generally be negative for most individuals, signifying a preference for shorter distances. Conversely, the parameter estimates for the other attributes are expected to be positive, indicating their positive influence on utility. After deriving individual-level conditional parameter estimates using a mixed logit (MXL) model with correlation in the WTP-space specification, employing constrained Cholesky decomposition, these estimates can be utilized to calculate individual-level utilities for visits to the respective green spaces. These individual utilities will vary as the attribute levels change according to scenario settings, with the extent of these changes depending on individual parameter estimates and explanatory variables.

Upon calculating individual utilities for all UGS in the status quo and in response to improvements in attribute levels, it becomes feasible to subsequently estimate the changes in individual-level CS, which approximates the improvement in welfare. In other words, the changes in individual-level CS are calculated by determining the difference between an individual's utility for a suggested scenario and their utility for the present state (status quo). To explore potential welfare improvements, 24 scenarios of interest were suggested.

The subsequent sections introduce the characteristics of the urban green spaces of interest (Section 8.2.1), the scenarios of interest (Section 8.2.2), and the simulation framework (Section 8.2.3).

8.2.1 Urban green spaces of interest

Five green spaces that are selected for the case study include Esplanade, Sia Boey Urban Archaeological Park, Areca Park, Bukit Dumbar Park, and Relau Metropolitan Park. Figure 35 shows the location of urban green spaces of interest.

For simulation purposes within this study, the present-state attribute levels of these urban green spaces are defined. These attributes include air quality, facilities, noise level, nursery habitat maintenance, as well as tree species and ecosystems. These attributes are categorized into three levels: the base level, level 1, and level 2. The status quo levels for each attribute in the scenario analysis were determined through a comparative assessment of the urban green spaces included in the study. Specifically, the attributes were adjusted based on observable differences among the urban green spaces. For example, the urban green spaces chosen as having the most abundance of tree species in the study area was designated as representing the highest level of the *tree species and ecosystems* attribute. Conversely, UGS with the least abundance of tree species were assigned the lowest level, while those with intermediate abundance were classified accordingly. This approach ensured that the status quo levels reflected the range of conditions observed among the UGS included in the study, providing a realistic basis for scenario analysis and interpretation of results.

To represent these attributes in the analysis, they are transformed into dummy-coded variables. When an attribute is at the base level, the dummy variables for levels 1 and 2 are assigned a value of zero. Upon achieving the first-level quality, the dummy variable for level 1 is set to one, while the dummy variable for level 2 remains at zero. If the attribute reaches the highest standard, which is the second level in this context, both the dummy variables for levels 1 and 2 are set to one. The attribute levels reflecting the status quo of these urban green spaces are detailed in Table 40.

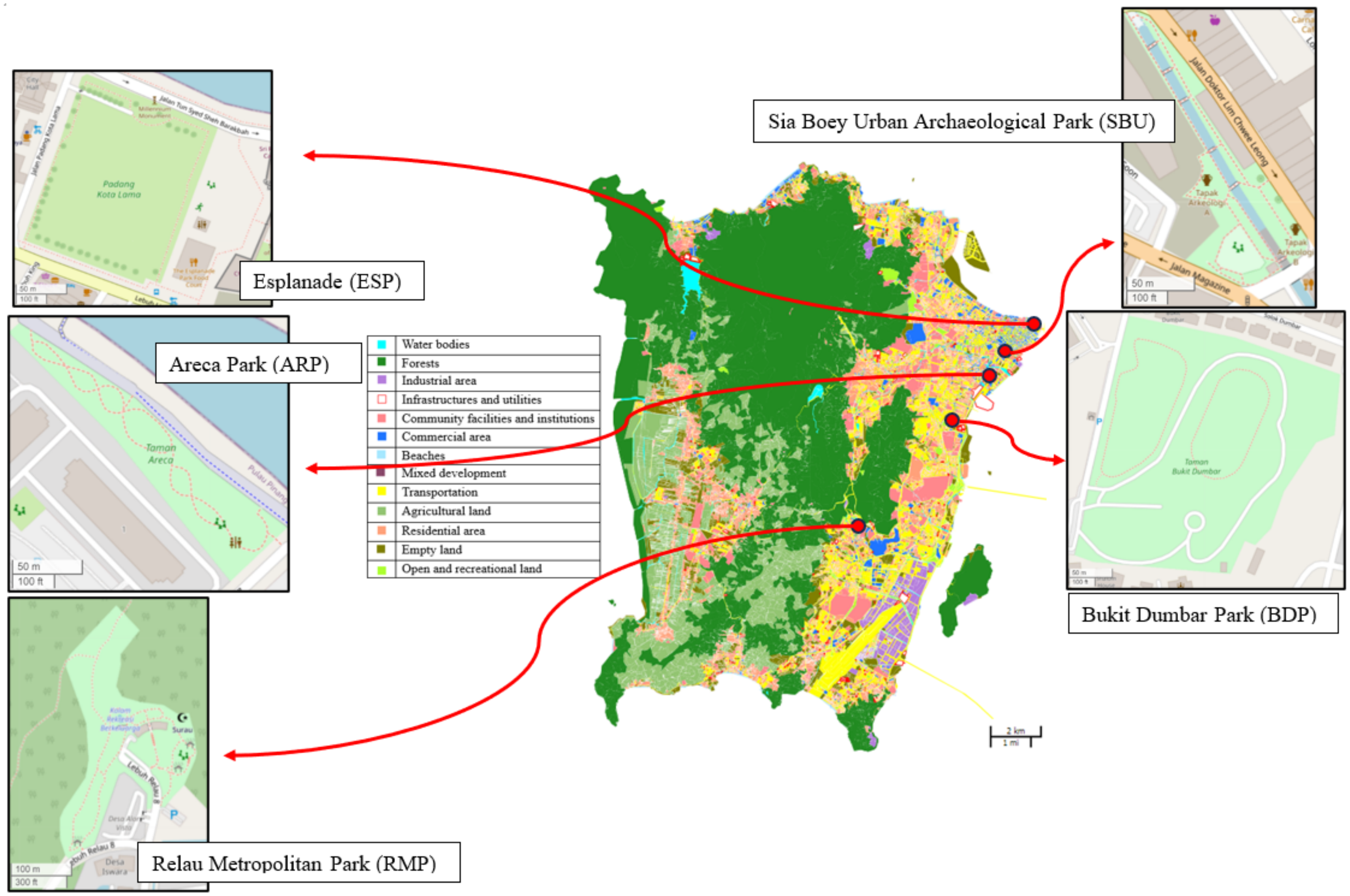


Figure 35: Urban green spaces of interest on Penang Island
Source: OpenStreetMaps (2023a)

Attribute level												
UGS	Avg. dist (km)	Avg. travel time (min)	Dummy variable									
			Air1	Air2	Fac1	Fac2	Noi1	Noi2	Nur1	Nur2	Tree1	Tree2
ESP	10.98	19.85	1	0	1	0	0	0	0	0	0	0
SBU	9.05	15.87	1	0	1	0	0	0	0	0	1	0
ARP	8.60	16.72	1	0	1	0	0	0	1	0	1	0
BDP	6.94	12.91	1	1	1	1	1	0	1	0	1	1
RMP	9.69	18.43	1	1	1	0	1	1	1	1	1	1

Table 40: Status quo of urban green spaces (UGS) of interest

The Esplanade

The Esplanade (ESP) is a large green field waterfront area located in the centre of George Town, Penang, Malaysia. The area served as a military parade ground before being repurposed for recreational and sports activities in the mid-19th century. Surrounded by Fort Cornwallis, the seaside promenade, the Penang City Hall, and the Cenotaph of Penang, the approximately 6.1-acre site is currently a popular venue for celebrations and cultural events. The field features a playground and a walking track along its edge (Penang Global Tourism, 2023a).

Sia Boey Park

Sia Boey Urban Archaeological Park (SBU) is the country's first urban archaeology park located in the heart of the city of George Town on Penang Island. The park was opened to the public in November 2019 and covers a land area of 2.53 acres. Before the park was built, the place where it is located, namely Sia Boey, served as an early trading hub in George Town. The 19th-century market hall and the shop houses are gathered in this area. It also served as a main place for community engagement for Penang urban dwellers. In 2018, the government of Penang decided to transform Sia Boey into an area that facilitates the coexistence of sustainable development, cultural heritage education, urban greens, and heritage conservation, following the discovery of the Old Prangin Canal Basin in 2015. Because of its historical significance, the park has unique characteristics such as the canal basin, aged-old trees, and a playground (Buletin Mutiara, 2019; Penang Global Tourism, 2023b).

Areca Park

Areca Park (ARP) is akin to a small Botanic Garden situated in the centre of George Town. This 3.2-acre park was established in 2021, showcasing 404 trees and 25,059 shrubs. Within the park, there are 30 species of trees, including Alma, Jelutong, Gelugor, Ara, and mangrove trees, as well as 45 species of shrubs. The park features a silver garden, a butterfly garden, a family square, outdoor fitness equipment, and a playground. Moreover, it offers a space for urban residents to engage in urban farming (Buletin Mutiara, 2021).

Bukit Dumbar Park

Bukit Dumbar Park (BDP) is a public recreational area developed upon a man-made hill located in another urban area on Penang Island, namely Jelutong. The park was established in 1958 and covers a land area of approximately 16.3 acres. It has become a favoured destination among Penang Island residents for engaging in recreational activities. The park's landscape is covered with casuarina trees, and it offers an abundance of recreational facilities, including a 0.5-mile hill slope walking trail, a playground, as well as public badminton and squash courts situated within the park.

Relau Metropolitan Park

Relau Metropolitan Park (RMP) is a recreational zone situated in another urban area in the southern region of Penang Island, known as Relau. It was opened for public use in 2003 and spans a land area of approximately 10.0 acres. This park serves as a popular spot for residents from the southern part of Penang Island to involve in various recreational activities. It encompasses a 1-mile trail designed for cycling, running, and walking, along with playgrounds and a swimming pool.

8.2.2 Scenarios of interest

Considering the current state of the urban green spaces of interest, as outlined in Table 40, the attribute levels associated with each urban green space, coded as *0*, indicate that these sites do not meet the conditions specified by the attribute levels in the DCE, will undergo the first and/or second level of enhancements. For instance, if the *fac1* variable is initially coded as zero, indicating facilities at level 0, it will be replaced with the value *1* if the facilities are enhanced from level 0 to 1. This research introduces an analysis containing 24 distinct scenarios of interest. Comprehensive information detailing these scenarios can be found in Table 41.

Currently, at ESP, SBU, and ARP, the air quality has already achieved level 1, eliminating the need for an upgrade from level 0 to 1. However, there is a need for an upgrade from level 1 to 2. Therefore, the first three scenarios (S1-S3) involve enhancing air quality from level 1 to 2 at these three sites, respectively. On the other hand, at BDP and RMP, the current air quality has already reached level 2, and as a result, no improvements are planned for these two sites.

Scenarios S4-S7 involve upgrading facilities from level 1 to 2 at ESP, SBU, ARP, and RMP, respectively. Scenarios S8-S10 focus on improving noise levels from level 0 to 1 at ESP, SBU, and ARP, while S11-S13 address noise level enhancements from level 0 to 2 at ESP, SBU, and ARP. Additionally, S14 centres on improving noise levels from level 1 to 2 at BDP.

Scenarios S15-S20 concentrate on enhancing nursery habitat maintenance from either level 0 or 1 to 2, while S21-S24 aim to improve tree species and ecosystems from level 0 or 1 to 2. Given the existing attribute level conditions, there is substantial room for improvement within these green spaces.

There are several ways to enhance these attribute levels within these green spaces, each requiring distinct strategies and considerations. For instance, improving air quality can involve the strategic planting of more trees or selecting tree species known for their superior air filtration capabilities. Another approach is to address the primary source of air pollution—motor vehicles—by implementing measures to reduce their presence around these green sites. Moreover, it is crucial for government authorities to take notice and prevent the construction of factories, which are significant contributors to air pollution, in close proximity to these green sites, or impose pollution abatement measures.

Enhancing facilities within these green spaces is a relatively straightforward task. Authorities can expand the array of essential amenities available, but this may be constrained by space limitations. Consequently, further research could focus on identifying and prioritizing specific facilities that hold greater importance for the residents and visitors.

To mitigate noise levels in these areas, a priority should be given to reducing both the number of motor vehicles and the establishment of new factories, which are the major sources of noise pollution. Additionally, implementing soundproofing measures such as noise abatement barriers and vegetation on noise-absorbent panels can contribute to effective noise reduction.

For improving nursery habitat maintenance in these green spaces, better planning is required to preserve the existing plant life. This may involve targeted planting, regular maintenance, and the introduction of plant species that contribute positively to the ecosystem.

Enhancing tree species and ecosystems within these green spaces requires a comprehensive approach. It involves the selection of suitable tree species, habitat restoration, and ecosystem management practices aimed at fostering biodiversity and ecological balance. This may include actions such as protecting native species, controlling invasive species, and creating favourable conditions for various tree species to thrive.

Scenario	Attribute	UGS	Changes in level	Scenario	Attribute	UGS	Changes in level
S1	Air quality	ESP	L1 to L2	S4	Facilities	ESP	L1 to L2
S2	Air quality	SBU	L1 to L2	S5	Facilities	SBU	L1 to L2
S3	Air quality	ARP	L1 to L2	S6	Facilities	ARP	L1 to L2
				S7	Facilities	RMP	L1 to L2
S8	Noise level	ESP	L0 to L1	S11	Noise level	ESP	L0 to L2
S9	Noise level	SBU	L0 to L1	S12	Noise level	SBU	L0 to L2
S10	Noise level	ARP	L0 to L1	S13	Noise level	ARP	L0 to L2
				S14	Noise level	BDP	L1 to L2
S15	Nursery habitat m.	ESP	L0 to L1	S17	Nursery habitat m.	ESP	L0 to L2
S16	Nursery habitat m.	SBU	L0 to L1	S18	Nursery habitat m.	SBU	L0 to L2
				S19	Nursery habitat m.	ARP	L1 to L2
				S20	Nursery habitat m.	BDP	L1 to L2
S21	Tree species & ecos.	ESP	L0 to L1	S22	Tree species & ecos.	ESP	L0 to L2
				S23	Tree species & ecos.	SBU	L1 to L2
				S24	Tree species & ecos.	ARP	L1 to L2

Table 41: Scenarios of interest

8.2.3 Simulation framework

A simulation framework for estimating the welfare values of Scenario 1 for the first five respondents is depicted in Figures 36 and 37. In Figure 36, the individual parameter estimates for attribute k for individual n, denoted as $E[\beta_{nk}|\text{ind}_n]$, are shown in columns 2-11. To simplify the framework, these conditional parameter estimates are represented as β_{nk}.

Next, it illustrates the multiplication of these individual parameter estimates by variables corresponding to each attribute level of every UGS, denoted as $\beta_{nk}x_{nk}$ (columns 12-23). The attribute level variables (x_{nk}) for air quality, facilities, noise level, nursery habitat maintenance, and tree species and ecosystems remain constant across individuals. However, the *distance* variable, measured in kilometres, varies among individuals as it is calculated based on the actual distance from an individual's place of residence to each green space. For clarity, this figure exclusively presents the estimation process for *The Esplanade (ESP)*. The same procedures are repeated for all the green spaces of interest. Depending on the description of each scenario, the dummy code for a specific variable will be replaced with the appropriate code that corresponds to the scenario's description. The change is highlighted in the red cell at column 15, row 3. While the attribute was '0' in the status quo, it has been updated to '1' to examine variations in CS estimates, as per the description provided for Scenario 1.

In Figure 37, columns 2-6 present the total utilities for each green space, obtained by summing all $\beta_{nk}x_{nk}$ values across individuals (n) and attribute levels (k). Column 7 presents the reported visit frequency to all UGS in the study area in the past year, as provided by individuals in the survey. In Column 9 (yellow cells), changes in individual CS estimates resulting from the improvement of *air2* at ESP are depicted. By multiplying the visit frequency information in Column 7, Column 10 illustrates the changes in individual CS estimates per year.

	Attribute												The Esplanade (ESP)										
	Individual parameter estimates												Dis	Air1	Air2	Fac1	Fac2	Noi1	Noi2	Nur1	Nur2	Tree1	Tree2
id	Dis	Air1	Air2	Fac1	Fac2	Noi1	Noi2	Nur1	Nur2	Tree1	Tree2	*Dis (km)*		1	1	1	0	0	0	0	0	0	0
1	-0.42	11.92	7.11	2.30	3.15	5.92	1.16	0.67	0.95	3.23	1.79	9.5	-4.00	11.92	7.11	2.30	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	-0.03	1.80	1.20	0.16	0.27	0.94	0.08	0.38	0.07	0.28	0.12	6.6	-0.19	1.80	1.20	0.16	0.00	0.00	0.00	0.00	0.00	0.00	0.00
3	-0.03	1.70	1.10	0.17	0.34	0.94	0.09	0.39	0.07	0.30	0.13	1.5	-0.05	1.70	1.10	0.17	0.00	0.00	0.00	0.00	0.00	0.00	0.00
4	-0.43	17.42	11.26	2.37	3.50	6.40	1.20	1.07	0.98	2.70	1.84	9.6	-4.17	17.42	11.26	2.37	0.00	0.00	0.00	0.00	0.00	0.00	0.00
5	-0.20	11.09	7.45	1.10	1.22	4.64	0.56	1.73	0.45	2.27	0.86	9	-1.82	11.09	7.45	1.10	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Figure 36: Simulated values of changes in individual utilities when the attribute level *air2* is achieved

	Total Utilities for each UGS						St.Dev.CS(S1)	0.620	29.746
							Average.CS(S1)	0.275	7.650
id	ESP	SBU	ARP	BDP	RMP	Visit freq/yr	log-sum(S1)	delta_CS	delta_CS/year
1	17.33	14.34	15.56	35.00	30.96	96	35.013	0.000	0.000
2	2.97	2.07	2.41	4.93	4.54	12	5.603	0.726	8.714
3	2.93	2.14	2.44	4.90	4.37	12	5.527	0.662	7.940
4	26.88	19.24	20.83	45.40	41.30	24	45.415	0.000	0.000
5	17.82	12.66	13.50	27.95	25.82	24	28.066	0.000	0.002

Figure 37: Simulated values of changes in individual utilities when the attribute level *air2* is achieved (continued)

	Attribute (k)										
	Dis	Air1	Air2	Fac1	Fac2	Noi1	Noi2	Nur1	Nur2	Tree1	Tree2
id (n)	**Conditional parameter estimates ($E[\beta_{nk}\|\text{ind}_n]$)**										
1	-0.42	11.92	7.11	2.30	3.15	5.92	1.16	0.67	0.95	3.23	1.79
2	-0.03	1.80	1.20	0.16	0.27	0.94	0.08	0.38	0.07	0.28	0.12
3	-0.03	1.70	1.10	0.17	0.34	0.94	0.09	0.39	0.07	0.30	0.13
4	-0.43	17.42	11.26	2.37	3.50	6.40	1.20	1.07	0.98	2.70	1.84
5	-0.20	11.09	7.45	1.10	1.22	4.64	0.56	1.73	0.45	2.27	0.86
...											
Avg.	-0.2	6.628	4.1	1.13	1.35	2.577	0.57	0.586	0.47	0.981	0.88

Table 42: Individual parameter estimates from MXL model of the first 5 respondents

Individual parameter estimates

In this study, individual parameter estimates are derived using an MXL model with correlation in the WTP-space, employing constrained Cholesky decomposition. The detailed estimation results can be found in Section 6.3.1, specifically under Model 6 presented in Table 30 within that section.

Section 2.5.5 within Chapter 2 covered the process of extracting individual-level coefficients from a MXL model. The conditional parameter estimates for individual n and attribute k ($E[\beta_{nk}|\text{ind}_n]$) for the first five respondents are shown in Table 42.

Annual visit frequency

The individual-level annual visit frequency are determined using data collected from a question involving five visit frequency categories. A conversion of the categorical measure of visit frequency from the survey into a continuous variable (the number of visits per year) are applied. Each category of visit frequency in the survey is translated into a fixed number of annual visits, providing a standardized unit for estimating individual-level visit frequencies, as presented in Table 43. This method facilitates statistical modelling and allows for a better understanding of visitor behaviour within urban green spaces.

Welfare measures

The study's analysis of welfare measures, as determined by the Compensating Variation (CV) resulting from changes in attribute levels, is detailed in Section 2.5.7 within Chapter 2.

Categories	No. of yearly visits assumed
More than once a week	96
Once a week	48
A few times a month	24
A few times a year	12
Never	0

Table 43: Assumed conversion of visit frequency categories into no. of yearly visits

Initially, the price coefficient is defined in terms of travel distance (in kilometres). However, it is later adjusted by a factor of $RM0.205 \times 2$ to account for the round-trip travel cost. After obtaining the estimated monetary unit CV for individual n ($\widehat{CV_n}$), this value is then multiplied by the estimated total annual visit frequency for individual n ($\widehat{Visit\ freq_n}$). The average individual CV across the sample ($\bar{CV}$) was then obtained:

$$\bar{CV} = \frac{\sum_{n=1}^{N} \widehat{CV_n} \times \widehat{Visit\ freq_n}}{N} \tag{8.1}$$

The $\bar{CV}$ is then multiplied by the total population of Penang Island (790,200 units) to calculate the CV estimates for the entire population.

8.3 Results and discussions

8.3.1 Consumer surplus estimates

This section presents the results of welfare measures, focusing on estimates that simulate consumer surplus (CS). These values represent an approximation to the economic benefits of the investigated scenarios. The CS estimates for the scenarios are presented in Table 44.

When comparing scenarios S1, S2, and S3, where air quality improvements from level 1 to level 2 are implemented at ESP, SBU, and ARP respectively, it becomes evident that the improvement at ARP is expected to yield the highest total CS estimate, with a median value of RM12.55 million. Following closely is SBU, with a median CS of RM9.21 million. In contrast, the improvement at ESP generates the lowest CS, with a median amount of RM6.05 million. In cases where budget constraints necessitate choosing only one site for air quality enhancement, ARP emerges as the most favourable option, as it is predicted to generate the highest CS.

In the scenarios involving the enhancement of facilities from level 1 to level 2 at

RMP (S7), the CS estimate for the entire population is expected to reach RM37.27 million, which not only surpasses the other sites but also stands as the highest across all investigated scenarios. Conversely, the median CS estimates for this improvement at ESP, SBU, and ARP are notably lower, amounting to RM4.76 million, RM3.20 million, and RM3.30 million, respectively.

Regarding the reduction of noise levels in these green spaces, a two-level improvement at ESP, SBU, and ARP is projected to yield median CS estimates of RM7.40 million, RM9.30 million, and RM11.69 million respectively. In contrast, a one-level improvement at BDP results in a median value of RM18.96 million. However, it is important to consider that the cost of noise reduction for two levels is likely higher than that for a one-level reduction. If the cost remains consistent across levels and these sites, BDP would be the preferred option.

The one-level enhancement of nursery habitat maintenance at BDP results in a significantly higher CS estimate, amounting to RM15.37 million. This stands in contrast to the CS estimate generated by enhancing the same attribute at ESP, SBU, and ARP.

Conversely, the improvement of tree species and ecosystems generates comparatively lower CS estimates when compared to the enhancements of other attributes. To elevate this attribute to level 2, the median CS estimates range from RM0.33 million to RM1.06 million. It is worth noting that focusing exclusively on improving tree species and ecosystems can potentially trigger a positive cascade effect. This effect may lead to improvements in air quality, noise levels, and nursery habitat maintenance, as indicated by previous studies (Pathak et al., 2011; Cohen et al., 2014; Pandey et al., 2015; Liquete et al., 2016; da Silva et al., 2021). Consequently, comprehensive research is essential to clarify the mechanisms and outcomes of this ecological process. Capital investments directed towards enhancing tree species and ecosystems in urban green spaces should be prioritized, as this approach may potentially mitigate the overall costs associated with improving these attributes.

Another crucial factor to consider is that an increase in the number of improved attributes within urban green spaces may lead to a higher overall visitation rate, resulting in an increase in total CS. However, it is important to note that if the improvements in urban green spaces exceed certain levels, it can result in a heightened visitation rate, leading to congestion during site visits and potentially creating negative experiences for visitors.

Moreover, the location of these green spaces should be taken into account. Since these sites are situated within urban areas and may be surrounded by essential public amenities such as hospitals, schools, and temples, a surge in visitation rates leading to traffic congestion can have adverse effects on the area. Therefore, a mere comparison of CS estimates may not provide a comprehensive understanding of the situation. It necessitates conducting more comprehensive research to estimate the overall effects, especially the changes in visitation rates, in urban green spaces.

While enhancing attributes across all green spaces can involve significant costs, the results obtained from this simulation analysis offer valuable guidance to decision-makers when allocating limited resources to optimize welfare within an urban environment. Nevertheless, it is essential to consider that concurrently enhancing attributes across multiple urban green spaces may lead to economies of scale, potentially reducing the costs associated with attribute improvements at each individual location.

It is important to acknowledge that the CS estimates provided in this study are valuable for policymaking but not useful for the SEEA. Additionally, it is crucial to acknowledge the limitations of this study. The assumption of only five UGS in the market serves an illustrative purpose, and therefore, drawing conclusions about the overall impacts of implementing these scenarios is challenging.

Sce.	**Attribute**	**UGS**	ΔCS **Percentiles**			**Sce.**	**Attribute**	**UGS**	ΔCS **Percentiles**		
			2.5	**Med.**	**97.5**				**2.5**	**Med.**	**97.5**
S1	Air quality	ESP	3.46	6.05	8.63	S4	Facilities	ESP	1.80	4.76	7.71
S2	L1 to L2	SBU	7.01	9.21	11.40	S5	L1 to L2	SBU	1.82	3.20	4.57
S3		ARP	9.73	12.55	15.38	S6		ARP	2.27	3.30	4.32
						S7		RMP	31.56	37.27	42.98
S8	Noise level	ESP	2.94	5.56	8.18	S11	Noise level	ESP	4.72	7.40	10.08
S9	L0 to L1	SBU	6.00	8.04	10.08	S12	L0/1 to L2	SBU	6.91	9.30	11.70
S10		ARP	7.82	10.29	12.77	S13		ARP	8.94	11.69	14.44
						S14		BDP	16.65	18.96	21.27
S15	Nursery habitat m.	ESP	-1.47	0.32	2.11	S17	Nursery habitat m.	ESP	-0.18	1.22	2.62
S16	L0 to L1	SBU	1.74	2.41	3.08	S18	L0/1 to L2	SBU	2.14	2.93	3.72
						S19		ARP	0.39	0.52	0.65
						S20		BDP	13.49	15.37	17.25
S21	Tree species & ecos.	ESP	-3.08	-1.11	0.85	S22	Tree species & ecos.	ESP	-1.17	0.33	1.83
	L0 to L1					S23	L0/1 to L2	SBU	0.62	1.06	1.50
						S24		ARP	0.76	1.01	1.26

Table 44: Changes in CS estimates of scenarios of interest in Ringgit Malaysia (million)

8.3.2 Kernel density plots for changes in CS estimates

In this section, kernel density plots are used to describe the probability density function of changes in individual CS estimates, providing a detailed visualization of the distribution of changes in CS for each attribute level improvement.

Figure 38 illustrates the distribution of changes in CS for the improvement of air quality level 2. The density curves for S1, S2, and S3 exhibit a right skew, indicating that there are some high CS values pulling the mean to the right, causing the mean to be greater than the median. All density curves peak at zero CS, indicating that a significant proportion of individual CS changes is centred around zero. Furthermore, S1 exhibits the highest peak among the three scenarios, suggesting that a larger proportion of individual CS changes in S1 is concentrated around zero compared to S2 and S3. Conversely, S3 displays the lowest peak. This observation aligns with the CS estimates presented in Table 44, where S1 has the lowest CS values, while S3 has the highest. Moreover, it is also observed that S3 exhibits a wider curve compared to the other two, indicating a larger spread of CS changes.

The distributions of changes in CS for the improvement of facilities level 2 are displayed in Figure 39. The density curves for S4, S5, and S6 are narrow and centred at zero, suggesting a small dispersion of CS changes. Conversely, S7 exhibits a wider density curve, indicating a larger spread of changes in CS values.

Figure 40 presents the distributions of changes in CS for the improvement of noise level 1. The density curves for S8, S9, and S10 exhibit a right skew, and all density curves peak at zero, with S8 having the highest peak, followed by S9 and S10. S10 exhibits a slightly broader curve compared to the other two, suggesting a more extensive range of changes in CS.

Figure 41 presents the distributions of changes in CS for the improvement of noise level 2. Similarly, the density curves for S11, S12, and S13 show a right skewness, and are peak at zero. S11 has the highest peak, followed by S12 and S13. In comparison to the previous three, the curve of S14 is characterized by a broader distribution, peaking around 0.5. It is important to note that while S11, S12, and S13 represent two-level improvements in noise, S14 only exhibits a one-level improvement.

Figures 42 and 43 present the distributions of changes in CS for the improvement of nursery habitat maintenance levels 1 and 2. The density curves for S15 to S19 are narrow, indicating a less extensive range of changes in CS, while the curve for S20 is wider, suggesting a larger spread of changes in CS values.

Finally, Figures 44 and 45 present the distributions of changes in CS for the improvement of tree species and ecosystems levels 1 and 2. All density curves (S21 to S24) are characterized by a very thin curve, peaking at zero, signifying a minimal spread of changes in CS values.

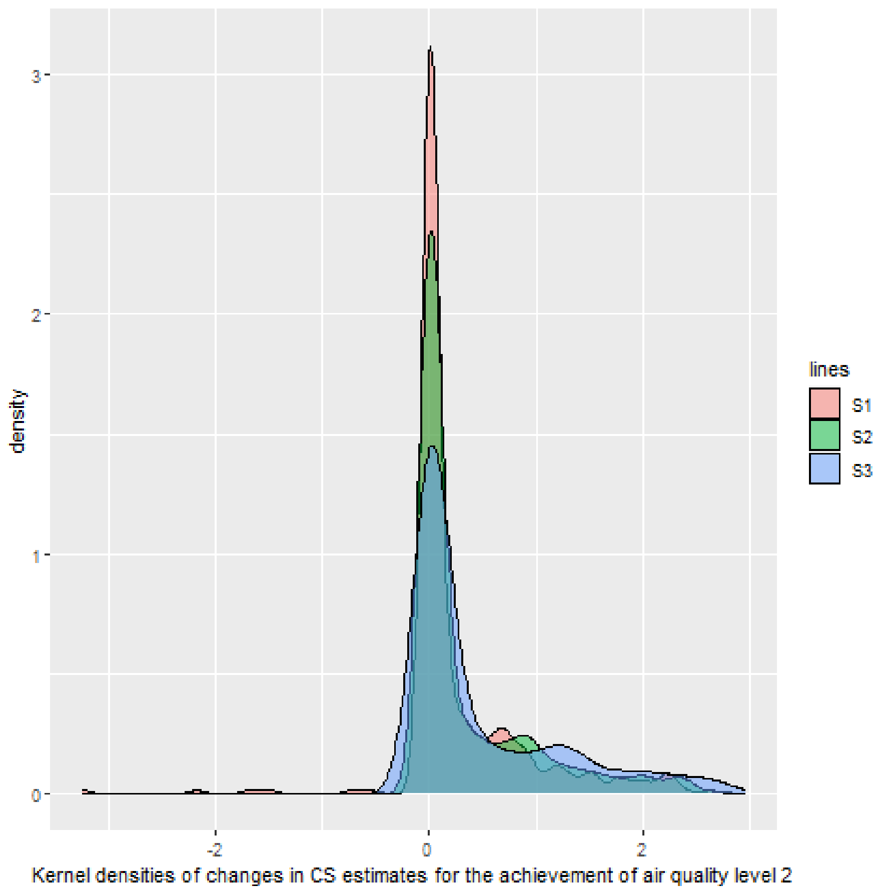


Figure 38: Kernel densities of changes in CS estimates for the achievement of air quality level 2

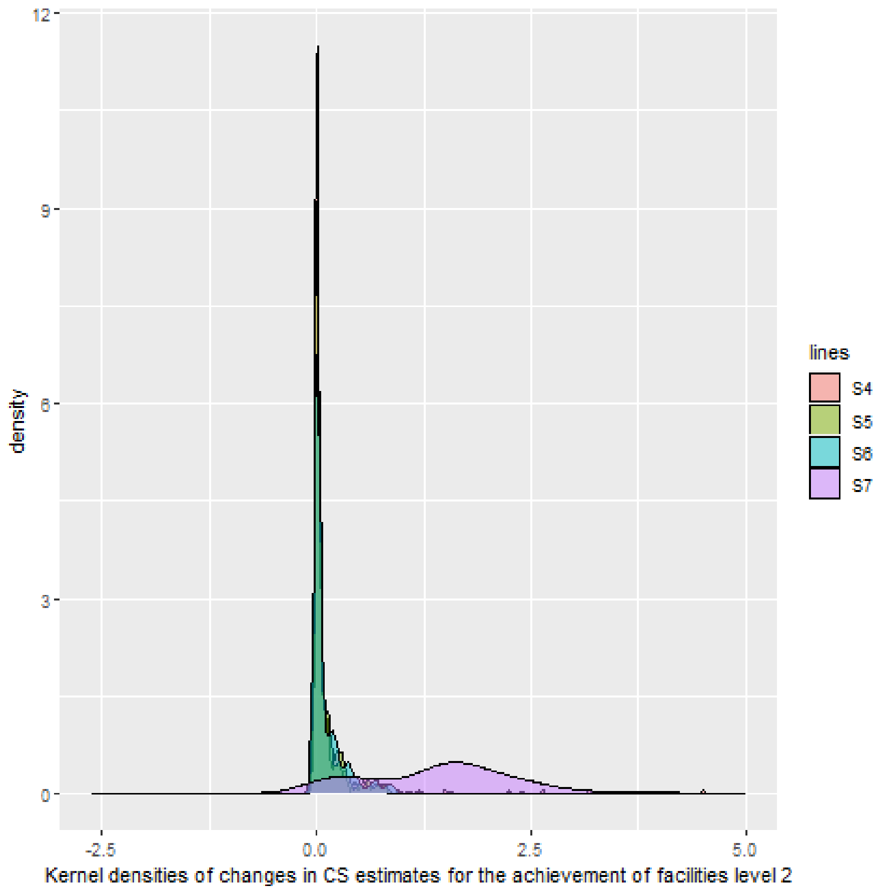


Figure 39: Kernel densities of changes in CS estimates for the achievement of facilities level 2

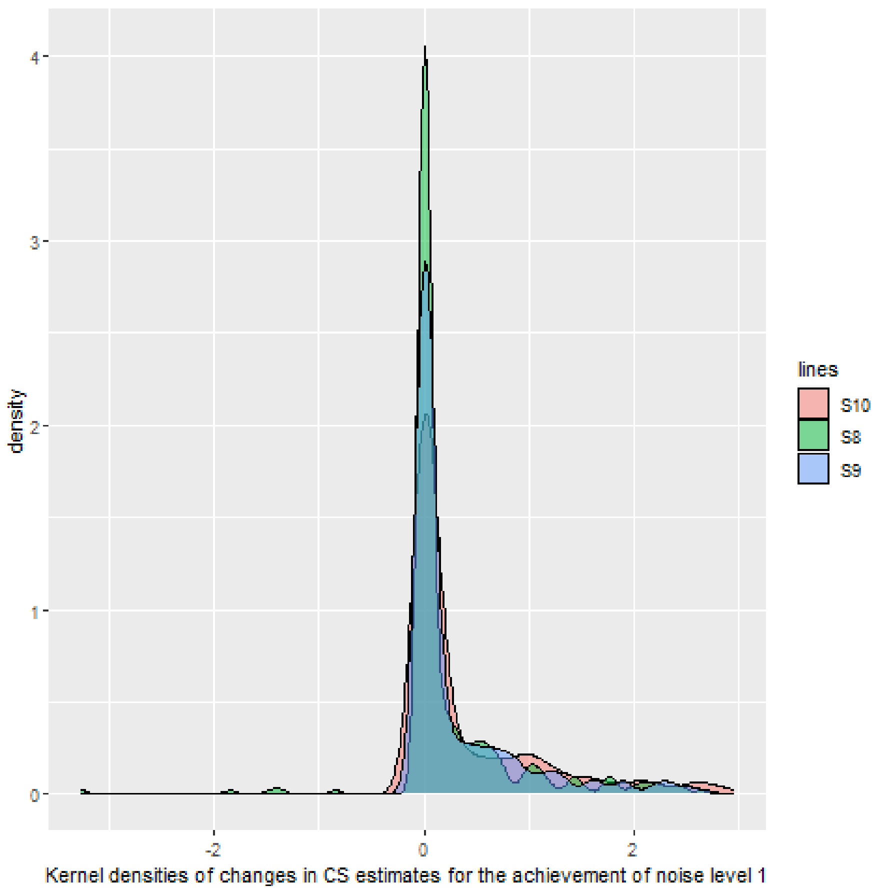


Figure 40: Kernel densities of changes in CS estimates for the achievement of noise level 1

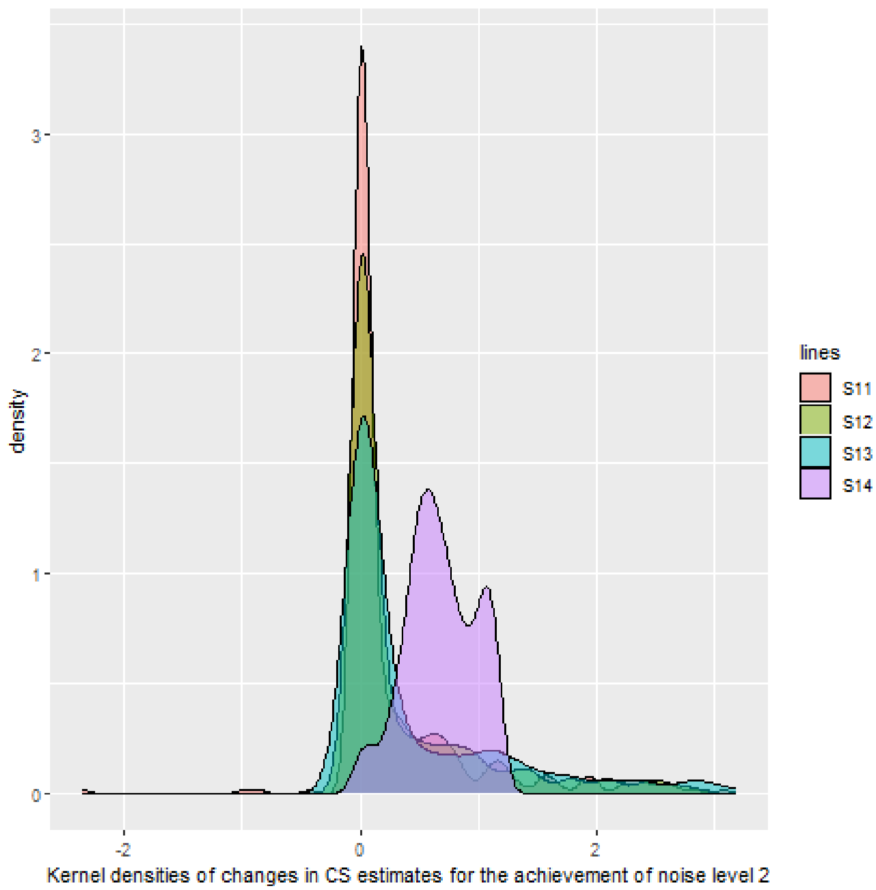


Figure 41: Kernel densities of changes in CS estimates for the achievement of noise level 2

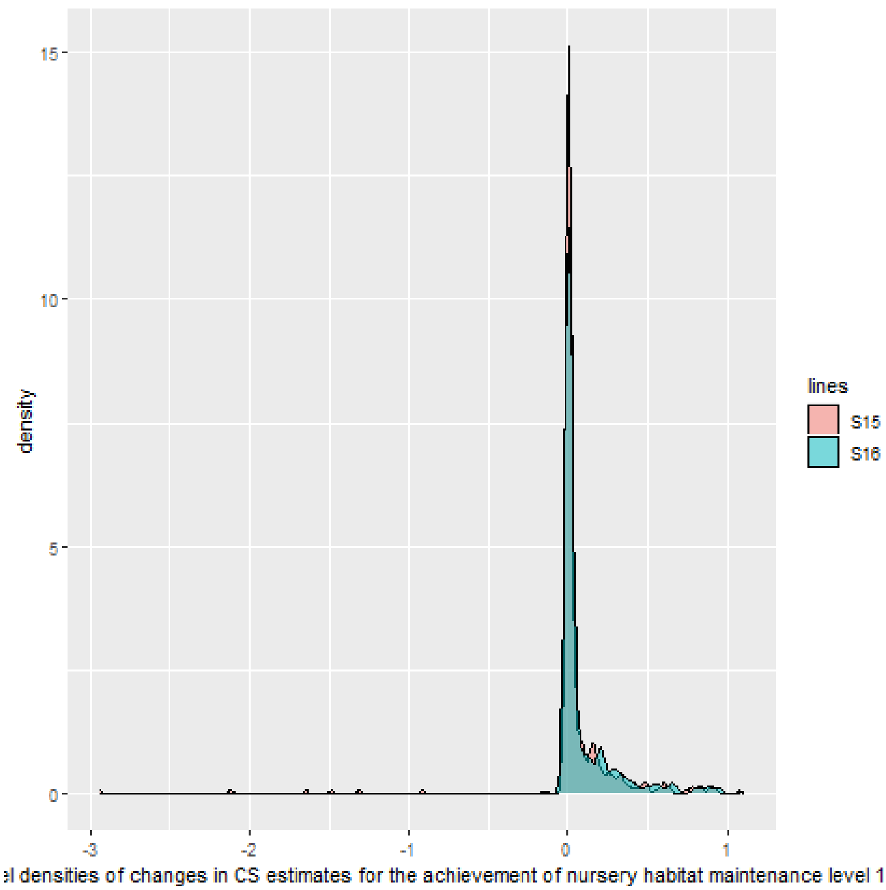


Figure 42: Kernel densities of changes in CS estimates for the achievement of nursery habitat maintenance level 1

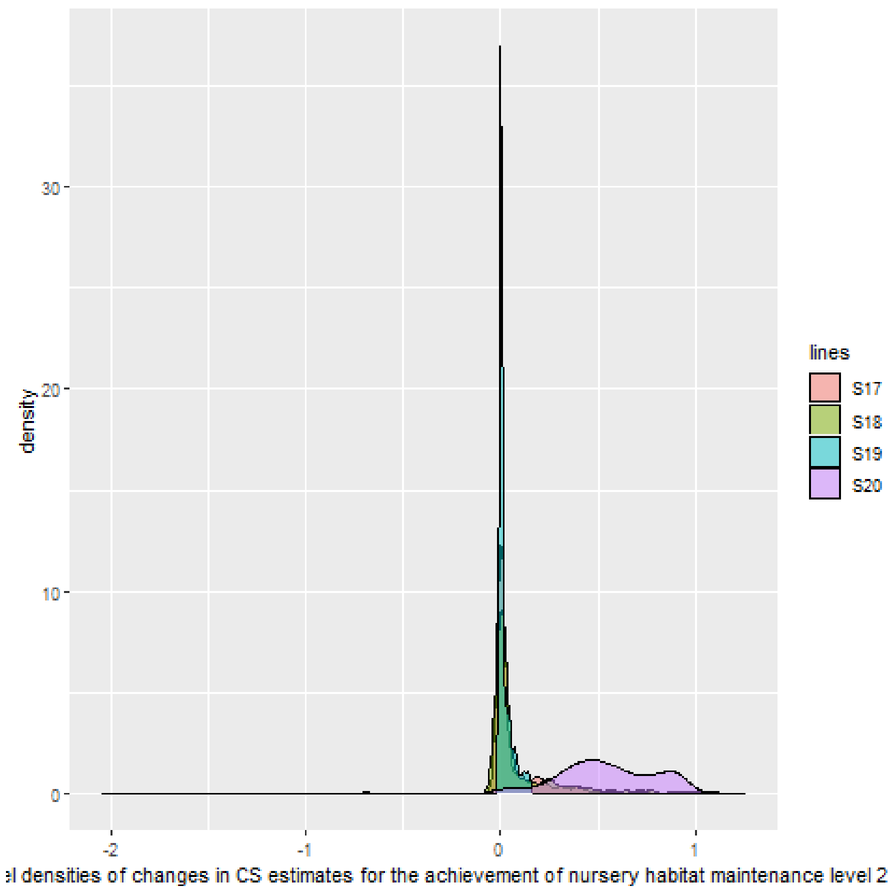


Figure 43: Kernel densities of changes in CS estimates for the achievement of nursery habitat maintenance level 2

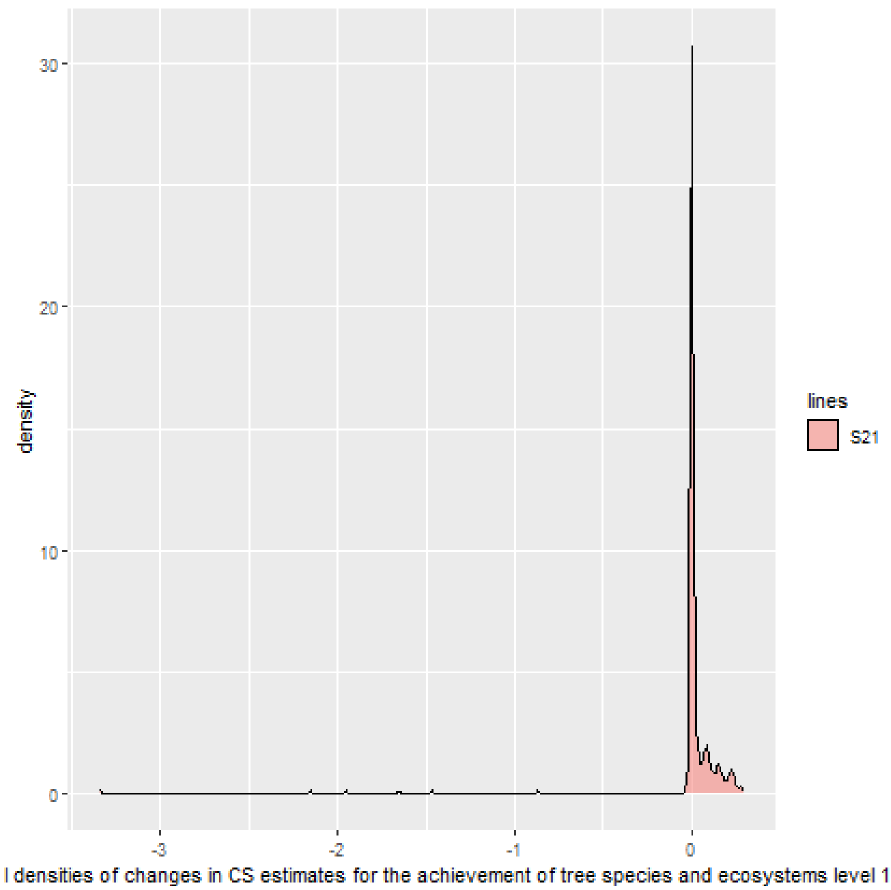


Figure 44: Kernel densities of changes in CS estimates for the achievement of tree species and ecosystems level 1

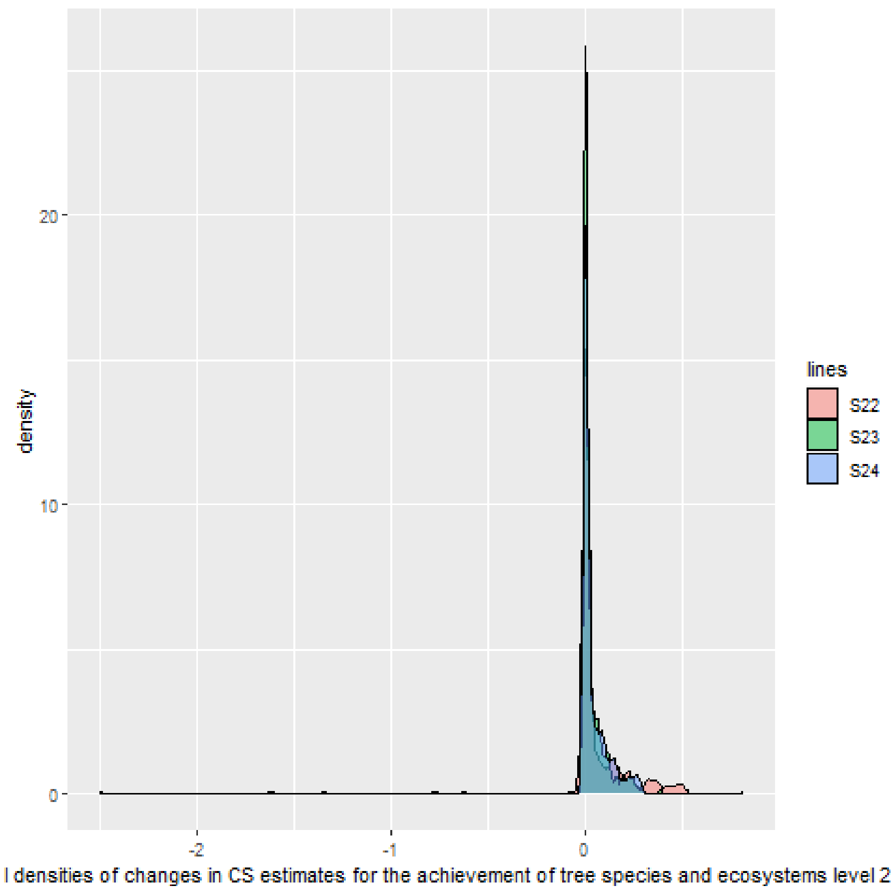


Figure 45: Kernel densities of changes in CS estimates for the achievement of tree species and ecosystems level 2

8.4 Conclusions

This study reveals an important finding. The improved attributes contribute to the increase of perceived value experienced by visitors, and it is reflected in the increased CS estimates. This simulation framework, which assumes the simulation of the entire market, emerges as a valuable tool for predicting changes in welfare value under various policy implications. In contrast to the conventional method of directly estimating total population CS by multiplying the average population number of visits, this framework suggests a different approach. It involves calculating individual-level total CS estimates by multiplying each individual's reported number of visits. Subsequently, the total population CS can be derived by aggregating these individual-level estimations. This approach introduces an additional step to explicitly estimate welfare values at the individual level, offering a better understanding of the overall welfare impact of changes in the attributes of interest.

Although the CS estimates were increased, the transformation of the CS estimates into observable changes in individual behaviours, such as an immediate increase in demand (number of visits), does not happen quickly. The speed of this change is affected by different factors, especially limitations related to travel costs. In economic theory, the prediction of demand is often done by analysing the demand functions. By examining the relationships between the quantity demanded of a good or service and factors like price and income, the demand functions can be constructed. Including these demand functions in my future studies will provide a more detailed understanding of how improvements in these attributes affect changes in demand.

Moreover, the CS estimates are based on a fixed conversion factor, assuming that each individual drives alone to UGS. However, this assumption may not accurately reflect real-world behaviour, where some individuals might walk or cycle. If the conversion factor accounted for different modes of transport, the simulated CS would likely be affected. Specifically, if travel costs were adjusted according to actual transport modes, the CS estimates might be lower. This is because individuals often optimize their travel choices to minimize costs, which would result in lower WTP for certain attributes when better transport options are available. Therefore, incorporating a variable conversion factor into the simulation framework could provide a more realistic estimation of welfare changes. This adjustment would provide a better understanding of the true economic value of UGS improvements. This could be an important direction for future research, providing deeper insights into the sensitivity of welfare estimates to varying transportation behaviours.

The results presented in this study also provide valuable insights for authorities and urban green space planners, serving as a guide for informed decision-making and the formulation of effective policies. Furthermore, these results offer relevant information that aims to strike a balance between urban development and the preservation of urban green spaces. Despite the lasting connection between urban development and land scarcity, the study provides data concerning the expected welfare values when urban lands are allocated to urban greens, facilitating the provision of ecosystem assets and services that are often overlooked by authorities. Given that land use decisions in urban areas frequently centre around profitability, the study offers stakeholders a method to capture the real value of urban greens, fostering a better understanding of the benefits associated with urban green spaces.

Moreover, this study could be further expanded to explore the complex interconnections between ecological and socioeconomic factors within urban green space planning. Expanding research in this direction could provide more comprehensive information for urban green space development.

My recommendation is to use this simulation framework as a complement to the SEEA-EA framework. This approach enables the pre-conditioning of various factors, including the competitive environment, the maximum available budget, and the preferred improvements achievable within these budget constraints. Through the simulation of these scenarios, it becomes possible to compute the net expected welfare values. This framework offers immediate responses when new policies are proposed, assisting in the identification of potential policies that could yield greater net benefits than the ones initially proposed.

Chapter 9

Conclusions

In the journey of employing non-market valuation techniques to assess the value of ecosystem services in urban green spaces (UGS) in Penang, Malaysia, this study has revealed the complex relationship between the economy and the environment. The primary objective of this research was to explore the economic significance of urban green spaces and their attributes, particularly in the context of developing countries. Moreover, this thesis also aimed to integrate non-market values as a complementary addition to the SEEA framework. By integrating non-market values, such as those derived from willingness to pay and consumer surplus, this research provides a more comprehensive view of the UGS benefits for better informed policy-making and urban planning, thereby promoting the UGS conservation and sustainable development in urban Penang.

This journey went across five main research topics, each with distinct objectives, offering valuable insights not only for academia but also for urban planners and park managers. The first topic employed the best-worst scaling (BWS) method to identify the most and least preferred attributes of UGS. The findings highlighted that air quality is the most valued attribute among urban residents, reflecting a high demand for cleaner air. However, inconsistencies in respondent preferences were identified when varying the set of attributes, suggesting challenges in maintaining stable attribute hierarchies with additional elements in survey designs.

The second topic used the travel cost method (TCM) to analyse visit patterns and preferences for two UGS in Penang. The study found that travel costs significantly influence visit behaviours, and incorporating time-related costs provides a comprehensive understanding of travel preferences. Additionally, environmental preferences played an important role in visit decisions, emphasizing the importance of ecosystems and its services in UGS.

The third topic used a discrete choice experiment (DCE) to explore public preferences for UGS attributes, with travel distance converted into travel costs as the payment vehicle. Improved air quality was again found to be highly valued, indicating residents' willingness to pay for better environmental conditions. Although attempts to include travel time in cost estimations were made, there was no significant improvement in model fit, highlighting challenges in accurately valuing travel time.

The fourth topic assessed spatial heterogeneity in welfare estimates using a Seemingly Unrelated Regression model. Although the spatial model had limited explanatory power, it provided insights into how spatial characteristics affect preferences.

The fifth topic involved a simulation framework to estimate welfare values in response to changes in attribute levels based on a blend of real-world conditions and certain assumptions. The results showed that improvements in UGS attributes lead to higher consumer surplus, demonstrating the economic benefits of investing in UGS. These improvements in attributes positively impact welfare values beyond direct travel costs. While this framework cannot entirely replicate the complexities of the real world due to study limitations, it holds the potential for further development based on real-world conditions, contingent on the availability of support and funding from local authorities. This framework also has the capacity to inform relevant stakeholders in making informed decisions about resource allocation and the preservation of urban green spaces. Specifically, the research aimed to integrate non-market values as a complement to the SEEA framework.

Throughout these research topics, I believe that these studies have made substantial contributions to the field of the environmental valuation of urban green spaces and their ecosystem services. These findings have several important implications for urban planning and policy-making in Penang. Firstly, the highly valued air quality suggests that policies focusing on pollution reduction and air quality improvement should be prioritized. Secondly, the significant role of environmental preferences in visit behaviours emphasizes the need for considering the ecological aspects during UGS planning. Thirdly, understanding travel costs, including time-related costs, is important for effective UGS planning. Fourthly, the spatial analysis highlights the needs of designing UGS policies to specific areas, taking into account existing land use. Lastly, the positive relationship between UGS attribute improvements and consumer surplus emphasizes the broader economic benefits of investing in UGS.

While this thesis provides valuable insights into resident preferences and the economic implications of UGS, there are several limitations. Time and budget constraints during data collection limited the sample size and potentially the diversity of respondents. A larger sample size might have provided more robust results and a better understanding of the preferences across different demographic groups. Moreover, the TCM relies on the assumption that travel costs can fully capture the value of UGS visits. However, this method might not be able to account for all factors affecting visit behaviour, such as personal preferences or peer influences. In addition, the DCE assumes respondents make choices based on the attributes presented in the choice tasks, but this might not always reflect the real-world decision-making processes. The spatial models tested in this study showed limited explanatory power, indicating that the spatial factors considered might not be comprehensive enough. This limitation suggests that more detailed spatial data in the future research is needed.

In particular, in the context of integrating non-market values to the SEEA framework, welfare values pose limitations in accurately defining equilibrium prices and generating exchange values within the SEEA framework. The interplay between demand and supply of UGS is central to SEEA's objectives. However, in the short term, the quantitative adjustments of UGS are restricted. Therefore, the supply is fixed. However, the qualitative improvements, such as improving services and accessibility can be made. In short term, the existing UGS features can be modified, such as tree compositions and facilities, rather than increasing UGS volume or area. In the medium and long term, UGS expansion is feasible. However, both medium and long term may not align with SEEA's immediate exchange value generation requirements. Therefore, estimating changes in monetary values of ecosystem services due to quantitative supply changes is challenging.

Future research directions may include refining the framework to better align with the SEEA framework and exploring additional valuation methods to enhance its accuracy and applicability. This could involve integrating both the supply and demand sides of UGS valuation, establishing connection with planners and park managers to obtain accurate data on UGS attributes, which can be used to improve the accuracy of non-market valuations. Additionally, future research might consider modelling the time dimension of UGS attributes, such as the growth and maturation of trees, income growth and demographic changes in the city, versus the immediate costs incurred during visits, providing a more realistic estimate of consumer surplus over time.

With this chapter, I bring this exciting journey to a close, concluding this thesis with the aspiration of reinforcing urban sustainability and enhancing the well-being of urban dwellers.

Appendix A

Individualized travel cost-based DCE

A.1 From *distance* to travel costs

This section converts the *distance* attribute to individualized travel costs by imitating the individuals' actual visit behaviour to Penang Botanic Gardens and Penang Youth Park, to investigate the effects of travel costs to preference heterogeneity as when the values of travel and recreational time are parts of the travel costs. It converts the distance attribute in the DCE questions to *TC1, TC2* and *TC3* to examine the effects of different travel costs to visit utilities in the DCE context. Three travel costs (TC1-TC3) were derived for every individual n in every choice situation t, based on the actual visit information provided by individual n. Table 45 shows the modified travel cost equations.

Where $TC1_{nt}$, $TC2_{nt}$ and $TC3_{nt}$ are the return-trip travel costs for individual n in choice situation t. VC_{nt} represents the vehicle operation cost for individual n in choice situation t based on the transport mode used by individual n to visit Penang Botanic Gardens and Penang Youth Park. If the transport mode used by individual n is different for both green sites, the transport mode used to visit Penang Botanic

Travel costs	Equations
TC1	$TC1_{nt} = VC_{nt}$
TC2	$TC2_{nt} = VC_{nt} + AOTC_{nt}$
TC3	$TC3_{nt} = VC_{nt} + AOTC_{nt} + AORC_{nt}$

Table 45: Equations of travel costs used in DCE

Gardens will be employed. The equation of VC_{nt} is expressed as

$$VC_{nt} = \frac{2 \times voc_m}{1} \tag{A.1}$$

where voc_m is the single-trip vehicle operating cost, which varies depending on the transport mode m (equation 5.2), and the equation is divided by 1 as it assumes individual n are travelling alone. $AOTC_{nt}$ is the value of travel time which assumes that all individuals have a time value. It is expressed as

$$AOTC_{nt} = \frac{awc_n}{60} \times (distance_{nt} \times travel\,time_n) \times 2 \tag{A.2}$$

where awc_n is one-third of the per-hour value of time for individual n, it is divided by 60 to derive the per-minute value of time. While $(distance_{nt} \times travel\,time_n)$ is the single-trip travel time in minutes for individual n in choice situation t. It is multiplied by 2 to derive the return-trip travel time in minutes. The $distance_{nt}$ is the distance shown in the choice situation t for individual n, The expected travel time per km $(traveltime_n)$ is derived from the real travel time for each individual n divided by the travel distance from respondent n's house location to Penang Botanic Gardens. If the data for Penang Botanic Gardens is not available, the data for Penang Youth Park is used. Therefore, an individual who walks to the green site will have more expected travel time compared to one who drives a car.

The recreational time opportunity costs $AORC_{nt}$ represents the recreation time opportunity costs which assumes that all individuals have a time value. It is expressed as

$$AORC_{nt} = \frac{awc_n}{60} \times recreation\ time_n \tag{A.3}$$

where the expected $recreation\ time_n$ is derived from the average real recreation time spent for each individual n at Penang Botanic Gardens and Penang Youth Park. If the data for one site is missing, only the data for another site is used.

After the *distance* is converted to travel costs, $TC1_{nt}$, $TC2_{nt}$ and $TC3_{nt}$ are different across respondent and choice situation, due to the different values in VC_{nt}, $AOTC_{nt}$ and $AORC_{nt}$. Therefore, the conversion of travel distance to different travel costs is expected to generate different empirical results which will draw useful implications for urban planners and authorities.

A.2 Results

A.2.1 Mixed logit models in WTP space specification

Table 46 shows the estimation results for the mixed logit models in willingness to pay (WTP) space specification. All models use 2000 Modified Latin Hypercube Sampling (MLHS) draws in the estimation process. Models 1-3 assume all attribute coefficients are random, and models 4-6 assume only some coefficients of attribute levels are random. The models in WTP space estimate the WTP values directly, therefore, separate WTP estimation is not required. The WTP is expressed in Malaysian currency (Malaysian Ringgit). Comparing models 1-3, model 1 has the highest log-likelihood (-1831.79) and lowest AIC and BIC values (3709.57 and 3853.40), indicating that the model has a better fit to the data set compared to models 2 and 3. This suggests that *TC1* should be used for estimation, and there is no strong evidence to prove that both value of travel time and recreational time should be included in the travel cost for the estimation. However, the means and standard deviations of *TC2* and *TC3* coefficients are statistically significant at 1% level, indicating that these coefficients are not deemed invalid in the estimation. Therefore, the value of travel time and recreational time should not be excluded in the model although the models do not perform better when *TC2* and *TC3* are used. Apart from the cost coefficients, all of the mean WTPs in models 1-3 are statistically significant at 5% level, and most of the standard deviations of WTP are significant at 10% level. A significant standard deviation denotes that preference heterogeneity across individuals is present. Therefore, most of the coefficients in models 1-3 vary across individuals. Model 3 has the highest WTP for attributes in most cases, as compared to the other two models.

Among models 4-6, model 4 has the highest log-likelihood (-1836.20) and lowest AIC and BIC values (3710.41 and 3829.22). A large improvement of model fit is shown if the cost variable (travel costs) does not account for value of travel time and recreational time. However, *TC2* and *TC3* should not be ignored as their mean and standard deviation of coefficient are statistically significant at 1% level. It suggests that both value of travel and recreational time should be considered as a part of travel cost, although models 5 and 6 do not perform better than model 4. The estimation results also show that all of the means and standard deviations of WTP are significant 10% level, except the mean WTP for *nur2* in models 5 and 6. In most cases, model 6 shows the highest WTP for most attributes, and model 4 shows the lowest WTP for attributes as the cost attribute - *TC1* includes only vehicle operation costs. All WTP values for attribute levels increase significantly from model 4 to model 5. However, it is noted that the WTP for attributes estimates in model 5 and 6 are very close to each other. The WTP values for *air1* are the highest among the three models (RM16.46 in model 6), followed by *air2* and *noi1* (RM10.98 and RM6.45 in model 6).

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Main/Mean						
alt1	0.196***	0.201***	0.184***	0.173***	0.166***	0.166***
	(0.052)	(0.055)	(0.052)	(0.052)	(0.049)	(0.049)
air1	7.107***	16.641***	17.109***	7.185***	16.46***	16.458***
	(0.282)	(0.469)	(0.654)	(0.306)	(0.757)	(0.758)
air2	5.018***	11.191***	11.548***	4.886***	10.978***	10.978***
	(0.223)	(0.277)	(0.475)	(0.358)	(0.742)	(0.742)
fac1	1.398***	2.628***	2.793***	1.212***	2.78***	2.781***
	(0.168)	(0.289)	(0.393)	(0.198)	(0.538)	(0.538)
fac2	1.148***	3.179***	3.218***	1.348***	2.866***	2.866***
	(0.246)	(0.302)	(0.451)	(0.262)	(0.596)	(0.596)
noi1	2.458***	6.508***	6.39***	2.864***	6.453***	6.452***
	(0.180)	(0.337)	(0.452)	(0.281)	(0.611)	(0.611)
noi2	0.927***	1.77***	1.933***	0.728***	1.86***	1.861***
	(0.225)	(0.322)	(0.379)	(0.251)	(0.515)	(0.515)
nur1	0.503**	2.21***	1.476***	0.902***	2.039***	2.038***
	(0.214)	(0.295)	(0.477)	(0.263)	(0.555)	(0.555)
nur2	0.714***	1.105***	1.398***	0.417*	0.902	0.902
	(0.180)	(0.305)	(0.432)	(0.234)	(0.563)	(0.563)
tree1	0.916***	1.657***	2.267***	0.849***	1.994***	1.995***
	(0.216)	(0.286)	(0.403)	(0.233)	(0.594)	(0.594)
tree2	0.832***	2.391***	1.974***	0.96***	2.025***	2.025***
	(0.165)	(0.263)	(0.433)	(0.243)	(0.537)	(0.537)
TC1	-0.897***			-0.917***		
	(0.129)			(0.125)		
TC2		-1.524***			-1.935***	
		(0.163)			(0.101)	
TC3			-1.762***			-1.935***
			(0.140)			(0.101)
SD						
air1	3.001***	7.54***	7.905***	3.4***	7.052***	7.051***
	(0.197)	(0.332)	(0.586)	(0.262)	(0.923)	(0.923)
air2	2.806***	7.252***	6.435***	2.728***	5.83***	5.828***
	(0.159)	(0.378)	(0.426)	(0.335)	(0.942)	(0.941)
fac1	0.243	2.796***	2.507***			
	(0.180)	(0.279)	(0.451)			
fac2	1.864***	5.009***	4.481***	1.953***	4.581***	4.581***
	(0.148)	(0.267)	(0.323)	(0.248)	(0.598)	(0.598)
noi1	0.418**	3.116***	1.904***	1.724***	2.006***	2.005***
	(0.180)	(0.236)	(0.514)	(0.176)	(0.668)	(0.669)
noi2	1.566***	3.783***	3.621***			
	(0.196)	(0.244)	(0.372)			
nur1	1.225***	4.086***	4.065***	1.694***	3.672***	3.671***
	(0.206)	(0.311)	(0.333)	(0.197)	(0.610)	(0.610)
nur2	0.133	0.222	0.355			
	(0.177)	(0.193)	(0.412)			
tree1	1.808***	4.666***	3.943***	1.623***	3.75***	3.75***
	(0.269)	(0.334)	(0.449)	(0.232)	(0.779)	(0.779)
tree2	0.526**	0.221	0.452			
	(0.216)	(0.327)	(0.506)			
TC1	1.53***			1.363***		
	(0.189)			(0.169)		
TC2		1.781***			1.234***	
		(0.221)			(0.137)	
TC3			1.545***			1.234***
			(0.209)			(0.137)
N	7680	7680	7680	7680	7680	7680
ll	-1831.79	-1840.43	-1847.76	-1836.20	-1852.23	-1852.21
AIC	3709.57	3726.85	3741.53	3710.41	3742.46	3742.43
BIC	3853.40	3870.68	3885.35	3829.22	3861.27	3861.24

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

Table 46: Uncorrelated mixed logit models in WTP space specification

Appendix B

Focus Group Results

B.1 Focus group results

B.1.1 Socio-demographic characteristics of focus group participants

The socio-demographic characteristics of focus group participants are detailed in Table 47. Of the 30 focus group participants, 29 completed the survey, with one participant partially completing only the socio-demographic questions. However, this incomplete survey was considered valid for focus group analysis, as it provided valuable socio-demographic information. Table 47 summarizes the socio-demographic characteristics of the survey respondents who also participated in the focus groups.

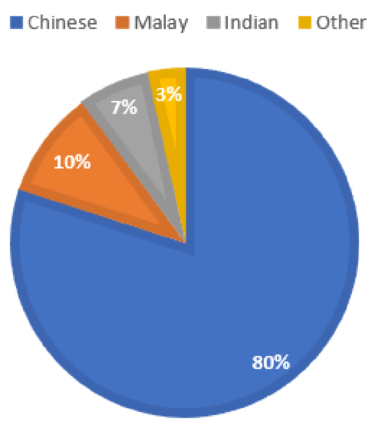
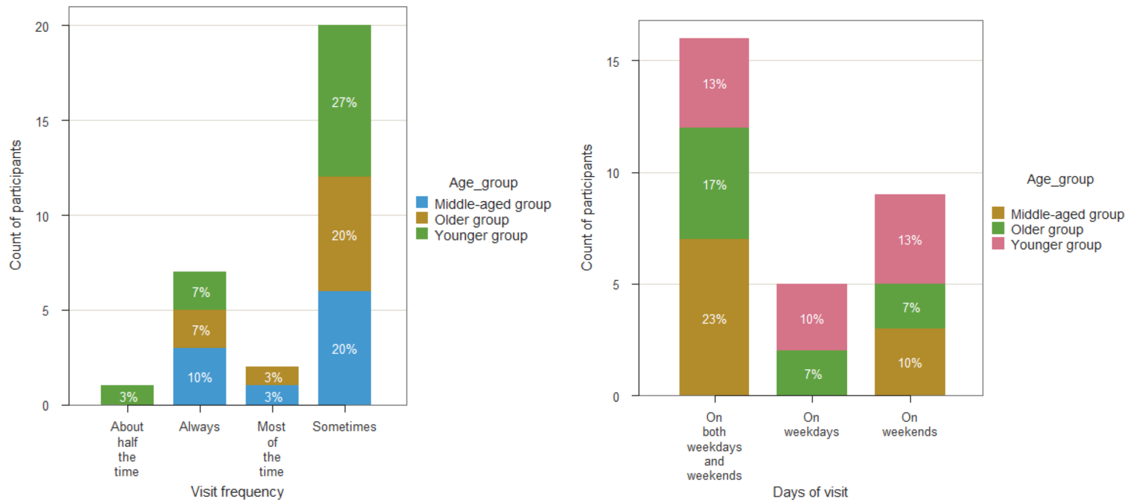


Figure 46: Ethnic composition of focus groups

Focus group	Gender		Age			Education level	
	Female	**Male**	**Mean**	**Range**			
Group 1	6	4	62.9	40-49	1	Secondary	2
				50-59	2	O-level or equivalent	1
				60-69	4	A-level/Diploma or equivalent	3
				70-79	3	Bachelor's Degree or Equivalent	2
						Master Degree or Equivalent	2
Group 2	5	5	28.5	20-29	6	O-level or equivalent	1
				30-39	4	A-level/Diploma or equivalent	2
						Bachelor's Degree or Equivalent	4
						Master Degree or Equivalent	3
Group 3	3	7	38.5	20-29	5	Secondary	2
				30-39	1	O-level or equivalent	1
				40-49	1	A-level/Diploma or equivalent	1
				50-59	1	Bachelor's Degree or Equivalent	5
				60-69	1	Master Degree or Equivalent	1
				80-89	1		
Total	14	16	43.3				

Table 47: Socio-demographic characteristics of focus group participants



(a) Visit frequency (b) Days of visit

Figure 47: Visit patterns of focus group participants

B.1.2 Visit behaviour of focus group participants

The visit patterns of the focus group participants were collected through a follow-up survey and are presented in Figure 47. The bar chart on the left displays the frequency of visits based on age group, while the bar chart on the right illustrates the days of visits based on age group. The majority of participants reported visiting urban green spaces *sometimes*, with a higher proportion among younger participants. Approximately one-fourth of the participants reported always visiting urban green spaces. It was noted that all participants had visited urban green spaces at least once. In addition, slightly more than half of the participants reported visiting urban green spaces on both weekdays and weekends, while 30% of participants only visited on weekends and 17% of participants only on weekdays.

B.1.3 Text analysis of focus group data

The results of text mining are presented in three different ways: word clouds, word counts, and text networks. Figure 48 displays the word clouds of data that were derived from each of the three focus groups separately. The word *tree* was the most frequently mentioned in all groups, followed by *time, car, government, facility,* and so on. The word *time* specifically referred to driving time, and *car* was mentioned frequently because there was a session that discussed the costs associated with

(a) Focus group 1 (b) Focus group 2 (c) Focus group 3

Figure 48: Word clouds of focus groups

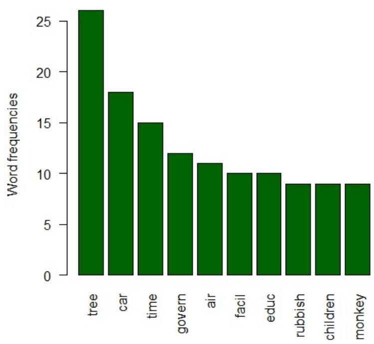
travelling to urban green spaces in each focus group. *Government* was also frequently mentioned, as urban green spaces are often associated with the government. Other high-frequency words that appeared in Group 1 included *air, facility, education, rubbish, children, and monkey.* Meanwhile, Group 2 mentioned *facility, distance, car, near, jog, walk,* and *buses* most frequently, and Group 3 mentioned *exercise, nature, improve, cafeteria, lighting, group,* and *run* the most (Figure 49).

Figure 50 shows the text network of all three focus groups combined. The word *park* was the central word, as people often used *park* interchangeably with *urban green space.* In line with the word clouds and word count analysis, *trees, time, facilities, car, distance, walk,* and *exercise* were mentioned with *park. Youth* was also linked very strongly to *park*, as Penang Youth Park was frequently mentioned during the discussions. Additionally, the text network showed that *trees* were often mentioned together with words like *lot, many,* and *good*, while *time* was linked with phrases like *important* and *lot.*

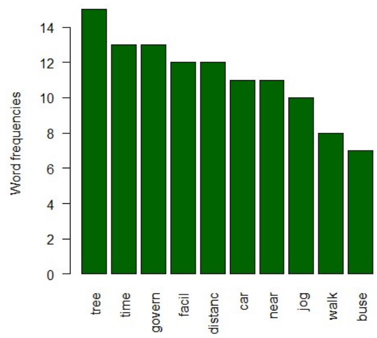
B.1.4 Thematic analysis of focus group data

During the discussions, participants were asked to identify the urban green space characteristics that were important to them. Overall, six main themes were identified, and a summary of themes and sub-themes is shown in Table 48.

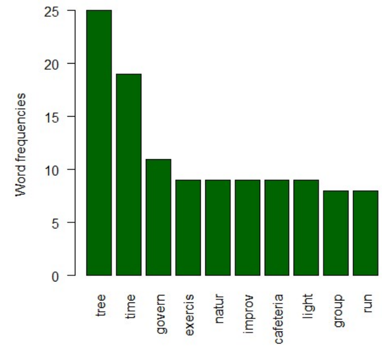
1. **Facilities**
   The discussion among participants frequently focused on the facilities provided



(a) Focus group 1



(b) Focus group 2



(c) Focus group 3

Figure 49: Word frequencies of focus groups

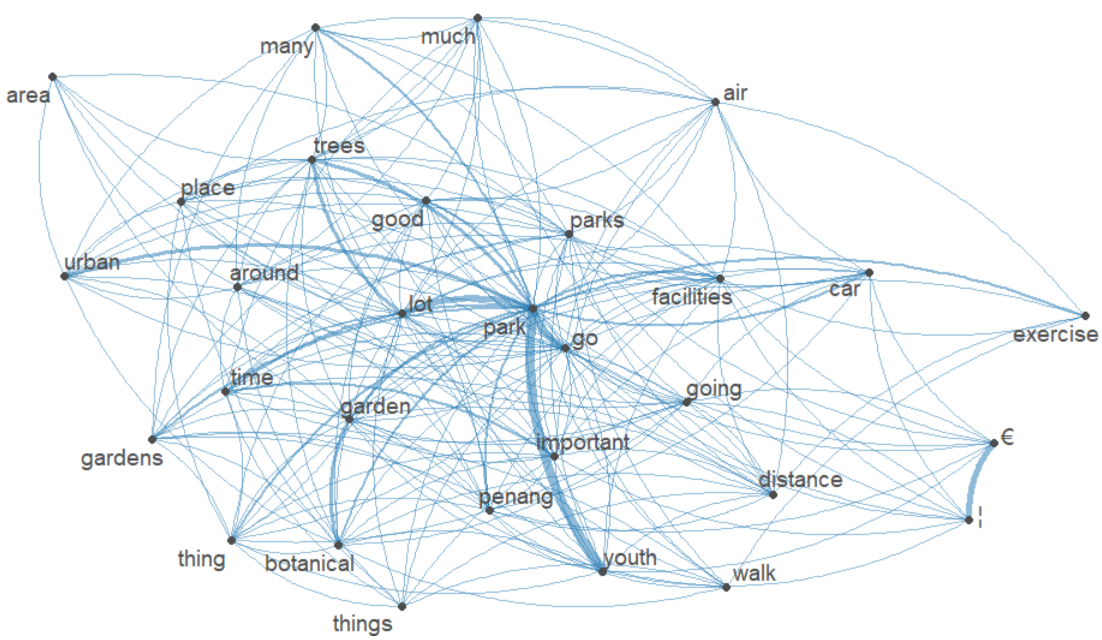


Figure 50: Text network of focus groups

in urban green spaces. Basic facilities such as toilets, car park spaces, medical stations, shelters, benches, and lighting were identified as crucial by the respondents. For example:
"Toilet is very important."
"I realize that a lot of parks, don't have toilets."
"These are the things where proper lighting to ensure the safety of these users."
"Certain strategic places, need benches, especially near slopes because people running up the slope may feel very tired."
"I was thinking about AED, automated external defibrillator...if you see now the Penang state government has set up quite a few. If there's anything happens, then someone can use the AED to maybe save some precious time before the ambulance arrives."

On the other hand, car parking facilities were perceived negatively by participants as they often caused pollution and raised safety concerns, as reflected in their discussions:
"Cars are going inside there, they stop their cars there and the car park is in the middle of the compound, so the pollution is there, it can be avoided to allow cars to go in there."
"Many cars and many motorbikes just go in the way and no control...I have nothing to stop them from causing trouble."

During the discussions, participants also frequently mentioned sports facilities as an important aspect of urban green spaces. Exercise equipment, jogging tracks, game fields, and bicycle paths were among the specific facilities mentioned. For instance:
"There is a jogging track there. So, the purpose I go to a park normally is to jog."
"We have a skating rink and we can jog there, we can play badminton."
"We also have a basketball court so it is a very lively place."

During the discussions, playground equipment such as slides, swings, and carousels were frequently mentioned, particularly by participants who have children:
"I prefer Youth Park because of the light of my kids...my daughter, she can go and play in the playground and everything is safe."

In addition, the condition and maintenance of facilities were also discussed:
"Maintenance of facilities, I think that is another important thing, We have all kinds of machines, but this one is broken and that one... you know, I can fall, we are older people."
"The facilities the gym equipment in there is a bit broken maybe it is not that safe for maybe like the old person."

2. **Ecosystem and species**

   The ecosystem and species theme included discussions on trees and plants, animals, landscapes, and the size of the green spaces. Participants generally preferred the presence of more trees and plants in urban green spaces, and they also mentioned specific tree species that they would like to see planted. For example:
   "I think it is quite a good park because ... got a lot of trees, but the tree is still...all of the trees is not big enough, so I think the shade is...I think it is very hot after around 8 am or 9 am."
   "Due to the hustle and bustle of people living in urban areas, there are a lot of you know...pollution and all that, so having a park with a lot of trees and all that, it would create less pollution... and so in a way health wise is very productive."
   "Yeah, and then the rain tree you know rain tree is not good because the Mosquito likes to go there and hide. So the trees what types of trees are also very important. Angsana (tree species) will be any better than rain tree...they are not considered because expensive, very difficult to maintain."

   Participants recognized the importance of animals in urban green spaces but expressed concern about the presence of an excessive number of monkeys and stray dogs:
   "There are a lot of monkeys in Youth Park, but the children who were holding onto a little bag, the monkeys are curious, they want to take away the bags, so try to go empty-handed."
   "The worst experience I've had was winter or morning jog only to be chased by a stray dog."

   Landscape emerged as a significant characteristic of urban green spaces, with participants highlighting the importance of landscape variety and naturalness. These attributes were particularly emphasized by younger participants, for

instance:
"Difficult terrain is pretty good. Because near my house there are a few parks and gardens. But I just don't want to go there because it is too easy."
"You cross over you see a jungle, you cross over you see a mountain you cross the street. So that is the combination of everything."
"They should have a well-planned landscape. Okay, and then of course, not forgetting the scenic views."

Participants also discussed the size of urban green spaces, for instance:
"Another one is it should be spacious, also I think already said yeah, so people usually avoid crowded places."

3. **Ecosystem services**
   Participants discussed two themes related to ecosystem services: air quality and noise levels. Air quality was a more commonly discussed topic compared to noise levels. For example:
   "The air is fresher, so I feel that it is a good thing to visit the park."
   "I think it is quite a good park because the air quality is not bad."
   "It has to be free and safe and nature has to be quiet, has to be far from the city."

4. **Accessibility**
   Participants pointed out that accessibility plays a crucial role in their decision to visit an urban green space. Distance from their residence and the availability of public transportation were identified as two sub-themes that affect accessibility.
   "The distance just affects us because we have a traffic jam problem in Penang. I would prefer to go to the nearer park affected by the traffic jam, and then time is of the essence."
   "Yes, in Penang, for example, the botanical garden, which I liked the most, but for me I prefer to go to Raffel Park because is nearer. But to the Botanical Park, maybe I need to spend 30 minutes because of the traffic. So I prefer the nearer park."
   "I would prefer a park that is within walking distance."
   "People like to say without transportation, let's say maybe like buses or maybe like a public transport, they will need more accessible to people to these places."

5. **Security**

Participants expressed concerns about the security of urban green spaces and discussed safety issues related to crime and wildfires. Some participants suggested measures to improve safety, such as hiring park rangers and installing CCTV cameras. Examples of participants' comments on this topic include:
"The toilet has to be open and should not be set at the corner, it is very dangerous for the ladies and people can do all kinds of nonsense there."
"We can upgrade the parks let's say better lighting and security wise for safety purposes for these people to go in a night." "Robbery and raping cases."
"The only thing that I could put in the paper is we need qualified park rangers."
"You can put so many CCTV but it still..."

6. **Cleanliness**
   Cleanliness emerged as a significant issue as participants pointed out that littering is a common problem in urban green spaces. Respondents expressed their preference for visiting sites that offer scheduled clean-up services and have sufficient recycling bins:
   "People sometimes quite often throw their rubbish there, ..., and I feel so sad, they are beautiful, but they are dirty, very dirty, with rubbish."
   "The government of Penang did the clean-up, they got the students, and then cleaned up all the litter that was left behind by people, and seriously it cannot be done in a day, there was a lot."
   "I find that a lot of people go through...tend to have garbage...lack of dustbins is also a need to be addressed, to ease people to not to litter."

B.1.5 Perceptions of the importance of urban green space characteristics

The perceived importance of urban green space characteristics was obtained from a follow-up survey where participants ranked nine different characteristics based on their perceptions. Figure 51 shows that *air quality* was consistently ranked as the most important characteristic by respondents of all age groups, followed by *distance* and *trees and plants*. The *abundance of animals* received the lowest rank. Interestingly, older people ranked *trees and plants* as the second most important characteristic but younger people ranked this characteristic as the fifth most important. *Playgrounds and exercise equipment* were lowly ranked by older people (6.50) but were highly ranked by the younger group (4.27). *Landscape naturalness* was ranked lower by

Theme	Subthemes
Facilities	Basic facilities Sports facilities Playground equipment
Ecosystem and species	Trees and plants Animals Landscape Size
Ecosystem services	Air quality Noise levels
Accessibility	Distance Public transportation
Security	Personal safety
Cleanliness	Scheduled clean up Recycling bin

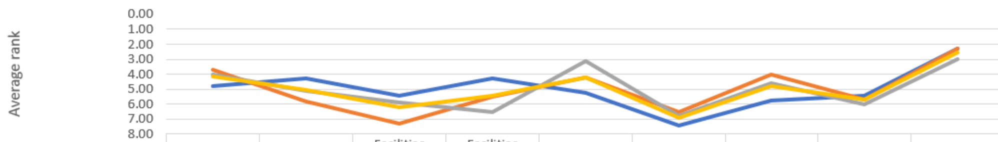
Table 48: Characteristics of urban green spaces identified by focus group participants

the younger group (5.73) but was ranked higher by the middle-aged group and older group.

When respondents were asked to state the levels of importance of each urban green space characteristic, *air quality* was perceived as the most important overall, with a mean value of 4.52 out of 5 (Figure 52), followed by *trees and plants* (4.11), *landscape naturalness* (3.69) and *noise levels* (3.69). In line with the rank data mentioned above, the *abundance of animals* was perceived as the least important characteristic (2.54). However, *distance* which was ranked high in the rank data, was perceived as the second least important characteristic (2.99). While *noise levels* which was ranked seventh in the rank data, was perceived as the third most important characteristic. In general, *facilities* were shown as less important compared to most characteristics based on both results. The importance score of the characteristics did not show a huge difference between age groups.

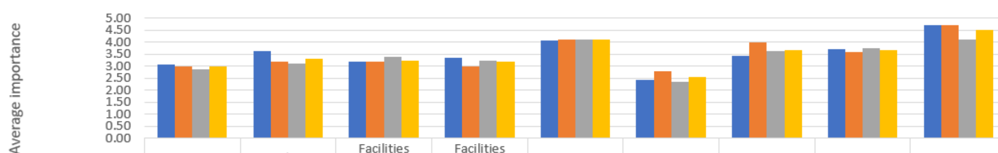
B.1.6 Perceptions of the importance of urban ecosystem services

In the follow-up survey, respondents were also asked to select the five most important ecosystem services provided by urban green spaces. Respondents were asked to choose



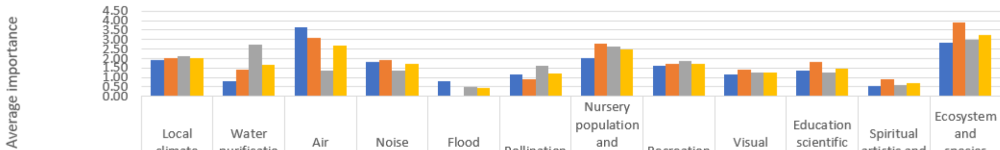
	Distance	Facilities walking paths car parks	Facilities restroom shelters and benches	Facilities playgrounds and exercise equipment	Trees and plants	Abundance of animals	Landscape naturalness	Noise level	Air quality
Younger group	4.82	4.27	5.45	4.27	5.27	7.45	5.73	5.45	2.27
Middle-aged group	3.70	5.80	7.30	5.50	4.20	6.50	4.00	5.70	2.30
Older group	4.00	5.13	5.88	6.50	3.13	6.75	4.63	6.00	3.00
Overall	4.17	5.07	6.21	5.42	4.20	6.90	4.78	5.72	2.52

Figure 51: Average rank of urban green space characteristics



	Distance	Facilities walking paths car parks	Facilities restroom shelters and benches	Facilities playgrounds and exercise equipment	Trees and plants	Abundance of animals	Landscape naturalness	Noise level	Air quality
Younger group	3.09	3.64	3.18	3.36	4.09	2.45	3.45	3.73	4.73
Middle-aged group	3.00	3.20	3.20	3.00	4.10	2.80	4.00	3.60	4.70
Older group	2.88	3.13	3.38	3.25	4.13	2.38	3.63	3.75	4.13
Overall	2.99	3.32	3.25	3.20	4.11	2.54	3.69	3.69	4.52

Figure 52: Average importance of urban green space characteristics



	Local climate regulation	Water purificatio n	Air filtration	Noise regulation	Flood control	Pollination	Nursery population and habitat maintenan ce	Recreation	Visual amenity	Education scientific and research	Spiritual artistic and symbolic	Ecosystem and species appreciati on
Younger group	1.91	0.82	3.64	1.82	0.82	1.18	2.00	1.64	1.18	1.36	0.55	2.82
Middle-aged group	2.00	1.40	3.10	1.90	0.00	0.90	2.80	1.70	1.40	1.80	0.90	3.90
Older group	2.13	2.75	1.38	1.38	0.50	1.63	2.63	1.88	1.25	1.25	0.63	3.00
Overall	2.01	1.66	2.70	1.70	0.44	1.24	2.48	1.74	1.28	1.47	0.69	3.24

Figure 53: Average importance of urban ecosystem services

five important ecosystem services from the set of twelve ecosystem services listed, the chosen five were assigned values based on respondents' perceived importance levels, while the ecosystem services that were not chosen were assigned the value of zero. As a result, the average importance values were lower than those for the urban green space characteristics.

Overall, *ecosystem and species appreciation* was perceived as the most important ecosystem service (3.24), followed by *air filtration* (2.70), *nursery population and habitat maintenance* (2.48). *Flood control* was ranked the least important compared to other ecosystem services (0.44). Interestingly, *water purification*, which ranked seventh overall, was perceived as the second most important by older people (2.75). Conversely, *air filtration*, which ranked second overall, was not as important as most ecosystem services for them (1.38). *Air filtration* was the most important ecosystem service for younger people (3.64), instead of *ecosystem and species appreciation* which was ranked the highest overall. *Pollination* has an importance score of 0.90 by the middle-aged group, but a significant increase in the score value was shown among the older group (1.63). The middle-aged group had a score of 2.80 for *nursery population and habitat maintenance*. Still, the younger group had a score of only 2.00 for that ecosystem service (see Figure 53).

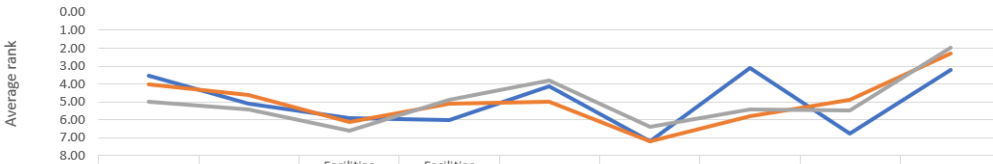
Reviewing results from Figure 51 to 53, although *ecosystem and species appreciation*, which was perceived as the most important, could refer to trees, plants and animals, the *abundance of animals* was ranked the lowest according to Figure 51 and 52. This circumstance was caused by the negative impression of monkeys and stray dogs in the green spaces, which was explained in detail during the discussions.

B.1.7 Comparison between focus groups

Figure 54, 55 and 56 present the results of the estimates for perceptions of urban green space characteristics and ecosystem services. Figure 54 reveals that there were variations in the ranking of characteristics across the different groups, with Group 1's participants, who had an average age of 62.9, giving *distance* an average rank of 3.56, while Group 3's participants, who had an average age of 38.5, gave it an average rank of 5.00. The participants in Group 2, with an average age of 28.5, gave *trees and plants* an average rank of 5.00, whereas Group 3's participants gave it an average rank of 3.80. *Landscape naturalness* was highly ranked by Group 1's participants, with an average rank of 3.11, but lower rankings were given by Groups 2 and 3 participants, who had average ranks of 5.80 and 5.40 respectively. *Noise levels* was ranked 4.90 on average by participants from Group 2, but a substantially lower ranking was given by Group 1's participants (6.78). Finally, *Air quality* was highly ranked by all focus groups.

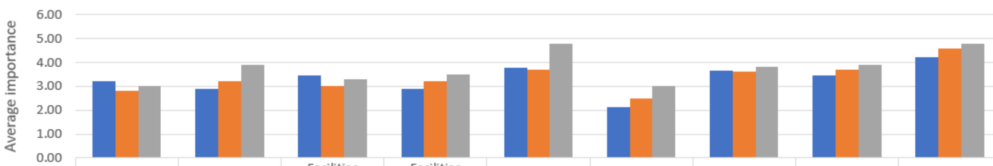
Figure 55 shows the average importance scores for urban green space characteristics, as rated by groups. The score for *walking paths and car parks* was 3.90 from Group 3 but was significantly lower at 2.89 from Group 1. Similarly, the score for *trees and plants* was 4.80 from Group 3, but only 3.78 from Group 1. In contrast, the *abundance of animals* received an average score of 2.11 from Group 1, but a score of 3.00 from Group 3. This characteristic had the lowest score among all the characteristics across all groups.

Figure 56 presents the average importance scores for urban ecosystem services by groups. *Ecosystem and species appreciation* was perceived as the most important ecosystem service in Groups 1 and 2. However, a substantial difference in scores can be seen between these two groups (4.33 and 2.40). *Flood control* was ranked the least important in Groups 1 and 2. Surprisingly, *water purification*, which had an average score of 2.20 in Group 2, had a score of only 0.90 in Group 3. In turn, *air filtration*, which had a score of 4.00 in Group 3, received a score of 1.10 from participants in Group 2. Similar to *air filtration*, *noise regulation* received a score of 1.10 from Group 2. However, this ecosystem service had scores of 2.00 and 2.11 from Groups 3 and 1 respectively. The participants in Group 1 had a score of 3.11 for *nursery population and habitat maintenance*, whereas participants in Group 2 had a score of only 2.10.



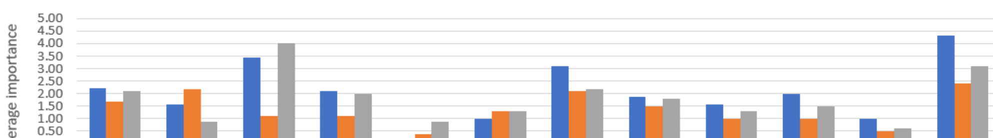
	Distance	Facilities walking paths car parks	Facilities restroom shelters and benches	Facilities playgrounds and exercise equipment	Trees and plants	Abundance of animals	Landscape naturalness	Noise level	Air quality
Group 1	3.56	5.11	5.89	6.00	4.11	7.22	3.11	6.78	3.22
Group 2	4.00	4.60	6.10	5.10	5.00	7.20	5.80	4.90	2.30
Group 3	5.00	5.40	6.60	4.90	3.80	6.40	5.40	5.50	2.00

Figure 54: Average rank of urban green space characteristics (focus group)



	Distance	Facilities walking paths car parks	Facilities restroom shelters and benches	Facilities playgrounds and exercise equipment	Trees and plants	Abundance of animals	Landscape naturalness	Noise level	Air quality
Group 1	3.22	2.89	3.44	2.89	3.78	2.11	3.67	3.44	4.22
Group 2	2.80	3.20	3.00	3.20	3.70	2.50	3.60	3.70	4.60
Group 3	3.00	3.90	3.30	3.50	4.80	3.00	3.80	3.90	4.80

Figure 55: Average importance of urban green space characteristics (focus group)



	Local climate regulation	Water purification	Air filtration	Noise regulation	Flood control	Pollination	Nursery population and habitat maintenance	Recreation	Visual amenity	Education scientific and research	Spiritual artistic and symbolic	Ecosystem and species appreciation
Group 1	2.22	1.56	3.44	2.11	0.00	1.00	3.11	1.89	1.56	2.00	1.00	4.33
Group 2	1.70	2.20	1.10	1.10	0.40	1.30	2.10	1.50	1.00	1.00	0.50	2.40
Group 3	2.10	0.90	4.00	2.00	0.90	1.30	2.20	1.80	1.30	1.50	0.60	3.10

Figure 56: Average importance of urban ecosystem services (focus group)

Appendix C

R Code to Derive Individual Coefficient Estimates from the DCE

```
rm(list = ls())
    library(apollo)
    setwd("")
    apollo _initialise()
    apollo _control = list(
    modelName = "1 _WTP _sp _allnrm _MLHS2000",
    modelDescr = "Mixed logit model WTP-space All NRM (2000 MLHS)",
    indivID = "ID",
    mixing = TRUE,
    nCores = 6,
    outputDirectory = "output"
    )
    database = read.csv("data _all _b1 _short.csv",header=TRUE)
    apollo _beta = c(asc _1 = 0.155 ,
    asc _2 = 0.000 ,
    mu _b _air1 = 34.537 ,
    mu _b _air2 = 22.030 ,
    mu _b _fac1 = 5.832 ,
    mu _b _fac2 = 6.541 ,
    mu _b _noi1 = 12.239 ,
    mu _b _noi2 = 4.175 ,
```

mu _b _nur1 = 3.446 ,
mu _b _nur2 = 2.350 ,
mu _b _tree1 = 4.222 ,
mu _b _tree2 = 4.074 ,
mu _log _lambda = -2.538 ,
intersigma _b _air1 = 14.618 ,
intersigma _b _air2 = 13.906 ,
intersigma _b _fac1 = 5.289 ,
intersigma _b _fac2 = 8.126 ,
intersigma _b _noi1 = 2.433 ,
intersigma _b _noi2 = 3.280 ,
intersigma _b _nur1 = 5.886 ,
intersigma _b _nur2 = 0.133 ,
intersigma _b _tree1 = 8.318 ,
intersigma _b _tree2 = -0.654 ,
intersigma _log _lambda = 1.324
)
apollo _fixed = c(”asc _2”)
apollo _draws = list(
interDrawsType = ”MLHS”,
interNDraws = 2000,
interUnifDraws = c(),
interNormDraws = c(”interdraws _lambda”,”interdraws _air1”,”interdraws _air2”,”interdraws _fac1”,”interdraws _fac2”,”interdraws _noi1”, ”interdraws _noi2”,”interdraws _nur1”,”interdraws _nur2”,”interdraws _tree1”,”interdraws _tree2”)
)
apollo _randCoeff = function(apollo _beta, apollo _inputs)
randcoeff = list()
randcoeff[[”b _lambda”]] = exp(mu _log _lambda + intersigma _log _lambda * interdraws _lambda) randcoeff[[”b _air1”]] = mu _b _air1 + intersigma _b _air1 * interdraws _air1
randcoeff[[”b _air2”]] = mu _b _air2 + intersigma _b _air2 * interdraws _air2
randcoeff[[”b _fac1”]] = mu _b _fac1 + intersigma _b _fac1 * interdraws _fac1
randcoeff[[”b _fac2”]] = mu _b _fac2 + intersigma _b _fac2 * interdraws _fac2
randcoeff[[”b _noi1”]] = mu _b _noi1 + intersigma _b _noi1 * interdraws _noi1
randcoeff[[”b _noi2”]] = mu _b _noi2 + intersigma _b _noi2 * interdraws _noi2
randcoeff[[”b _nur1”]] = mu _b _nur1 + intersigma _b _nur1 * interdraws _nur1
randcoeff[[”b _nur2”]] = mu _b _nur2 + intersigma _b _nur2 * interdraws _nur2

```
randcoeff[["b _tree1"]] = mu _b _tree1 + intersigma _b _tree1 * interdraws _tree1
randcoeff[["b _tree2"]] = mu _b _tree2 + intersigma _b _tree2 * interdraws _tree2
return(randcoeff)
apollo _inputs = apollo _validateInputs()
apollo _probabilities=function(apollo _beta, apollo _inputs, functionality="estimate")
apollo _attach(apollo _beta, apollo _inputs)
on.exit(apollo _detach(apollo _beta, apollo _inputs))
P = list()
V = list()
V[["alt1"]] = asc _1 + b _lambda * ( b _air1 * air1 _1 + b _air2 * air2 _1 + b _fac1 * fac1 _1 + b _fac2 * fac2 _1 + b _noi1 * noi1 _1 + b _noi2 * noi2 _1 + b _nur1 * nur1 _1 + b _nur2 * nur2 _1 + b _tree1 * tree1 _1 + b _tree2 * tree2 _1 - dis _1)
V[["alt2"]] = asc _2 + b _lambda * ( b _air1 * air1 _2 + b _air2 * air2 _2 + b _fac1 * fac1 _2 + b _fac2 * fac2 _2 + b _noi1 * noi1 _2 + b _noi2 * noi2 _2 + b _nur1 * nur1 _2 + b _nur2 * nur2 _2 + b _tree1 * tree1 _2 + b _tree2 * tree2 _2 - dis _2)
mnl _settings = list(
alternatives = c(alt1=1, alt2=2),
avail = list(alt1=1, alt2=1),
choiceVar = choice,
utilities = V
)
P[["model"]] = apollo _mnl(mnl _settings, functionality)
P = apollo _panelProd(P, apollo _inputs, functionality)
P = apollo _avgInterDraws(P, apollo _inputs, functionality)
P = apollo _prepareProb(P, apollo _inputs, functionality)
return(P)
model = apollo _estimate(apollo _beta, apollo _fixed,apollo _probabilities, apollo _inputs)
apollo _modelOutput(model)
apollo _saveOutput(model)
apollo _sink()
unconditionals ¡- apollo _unconditionals(model,apollo _probabilities, apollo _inputs)
conditionals ¡- apollo _conditionals(model,apollo _probabilities, apollo _inputs)
mean(unconditionals[["b _air1"]])
sd(unconditionals[["b _air1"]])
summary(conditionals[["b _air1"]])
income _n = apollo _firstRow(database $hh _inc _abs, apollo _inputs)
write.csv(conditionals,paste0(model $apollo _control $outputDirectory,
```

model $apollo _control $modelName, " _conditionals.csv"))
apollo _sink()

Appendix D

Correlation Analysis for Variables in Travel Costs Analysis

D.1 Correlation Analysis - Penang Botanic Gardens

	pbg_tc1	pbg_tc2	pbg_tc3	pbg_tc4	pbg_tc5	age	educ	employed	household	youth_visit	tree	air	years_pen	visit_freq
pbg_tc1	1													
pbg_tc2	0.8164	1												
	0													
pbg_tc3	0.5546	0.7703	1											
	0	0												
pbg_tc4	0.7336	0.7437	0.5838	1										
	0	0	0											
pbg_tc5	0.4429	0.5001	0.5678	0.8656	1									
	0	0	0	0										
age	-0.0422	-0.0749	-0.0125	-0.012	0.0011	1								
	0.4994	0.2305	0.8416	0.848	0.9862									
educ	0.1579	0.17	0.162	0.22	0.2599	-0.5307	1							
	0.0111	0.0062	0.0091	0.0004	0	0								
employed	0.0895	0.1389	0.1639	0.6056	0.8002	0.0776	0.1541	1						
	0.1516	0.0257	0.0084	0	0	0.166	0.0057							
household	-0.0773	0.0041	-0.0422	-0.0732	-0.1144	-0.1717	-0.0877	-0.1159	1					
	0.2157	0.9476	0.5001	0.2415	0.0666	0.0021	0.1176	0.0382						
youth_visit	-0.0665	-0.0583	0.0036	-0.0187	0.0021	0.1457	-0.1189	0.0907	0.0093	1				
	0.2869	0.3509	0.9541	0.7646	0.9734	0.0093	0.034	0.1066	0.8692					
tree	-0.0841	-0.0204	-0.0666	-0.0373	-0.0215	0.0367	0.0008	0.0232	-0.0158	0.0463	1			
	0.1783	0.7445	0.2868	0.5511	0.7306	0.5133	0.9888	0.6799	0.7777	0.4105				
air	0.0115	0.0503	0.0816	0.0298	0.0207	-0.0721	0.0395	-0.0735	0.0027	-0.0176	0.0182	1		
	0.8543	0.4212	0.1913	0.6335	0.7412	0.198	0.4818	0.1897	0.961	0.7548	0.7404			
years_pen	-0.228	-0.223	-0.0863	-0.152	-0.0724	0.7952	-0.4934	0.0538	-0.1059	0.238	0.0313	-0.0077	1	
	0.0002	0.0003	0.1668	0.0145	0.2464	0	0	0.3376	0.0585	0	0.5767	0.8902		
visit_freq	-0.0512	-0.1045	-0.0384	-0.0583	-0.0279	0.2446	-0.0394	0.046	-0.0907	0.3602	-0.0446	-0.1005	0.2944	1
	0.4126	0.0941	0.5392	0.3513	0.6553	0	0.4824	0.4119	0.1055	0	0.4268	0.0726	0	

Table 49: Pearson's Correlation Analysis (Penang Botanic Gardens)

	pbg_tc1	pbg_tc2	pbg_tc3	pbg_tc4	pbg_tc5	age	educ	employed	household	youth_visit	tree	air	years_pen	visit_freq
pbg_tc1	1													
pbg_tc2	0.7206	1												
	0													
pbg_tc3	0.4148	0.6833	1											
	0	0												
pbg_tc4	0.639	0.6398	0.4382	1										
	0	0	0											
pbg_tc5	0.5425	0.5438	0.4769	0.9609	1									
	0	0	0	0										
age	-0.0566	-0.0569	0.1032	0.0487	0.0646	1								
	0.3649	0.3623	0.098	0.4358	0.3013									
educ	0.2381	0.1988	0.1169	0.2623	0.2678	-0.4058	1							
	0.0001	0.0013	0.0608	0	0	0								
employed	0.1246	0.1829	0.1684	0.7686	0.8147	0.1616	0.1824	1						
	0.0456	0.0032	0.0067	0	0	0.0093	0.0033							
household	-0.1423	0.0228	-0.042	-0.1105	-0.1218	-0.1995	-0.1654	-0.0902	1					
	0.0223	0.7159	0.5023	0.0765	0.0507	0.0013	0.0078	0.1486						
youth_visit	-0.2776	-0.1962	-0.0102	-0.1338	-0.1137	0.3419	-0.3518	0.0293	0.0281	1				
	0	0.0015	0.871	0.0317	0.0682	0	0	0.6392	0.6528					
tree	-0.0547	-0.0897	-0.0914	-0.0564	-0.0304	0.0427	-0.0296	-0.0049	-0.0006	0.0408	1			
	0.3815	0.1508	0.1432	0.367	0.6266	0.4947	0.6365	0.9372	0.9921	0.514				
air	-0.0627	-0.0242	0.0668	-0.0336	-0.0395	-0.0651	0.0684	-0.0826	-0.006	-0.0836	-0.0681	1		
	0.3157	0.6986	0.2849	0.5909	0.528	0.2975	0.2738	0.1861	0.9241	0.1806	0.276			
years_pen	-0.1602	-0.1134	0.021	-0.0753	-0.0553	0.7091	-0.4376	0.0471	-0.0561	0.4143	0.0462	-0.0265	1	
	0.01	0.0689	0.737	0.2282	0.3766	0	0	0.4514	0.3695	0	0.4601	0.6714		
visit_freq	-0.1344	-0.1153	0.0454	-0.0814	-0.0344	0.2822	-0.0777	0.016	-0.0177	0.4017	0.0169	-0.0209	0.2906	1
	0.031	0.0644	0.4679	0.1924	0.5824	0	0.2135	0.7982	0.7775	0	0.7875	0.7378	0	

Table 50: Spearman's Correlation Analysis (Penang Botanic Gardens)

D.2 Correlation Analysis - Penang Youth Park

	youth_tc1	youth_tc2	youth_tc3	youth_tc4	youth_tc5	age	educ	employed	household	pbg_visit	tree	air	years_pen	visit_freq
youth_tc1	1													
youth_tc2	0.8508	1												
	0													
youth_tc3	0.5343	0.7264	1											
	0	0												
youth_tc4	0.8327	0.7978	0.6033	1										
	0	0	0											
youth_tc5	0.4329	0.4869	0.6067	0.7988	1									
	0	0	0	0										
age	-0.0053	-0.0575	-0.0355	0.0119	0.0144	1								
	0.9345	0.3671	0.5788	0.8519	0.8214									
educ	0.0834	0.0984	0.124	0.1668	0.2386	-0.5307	1							
	0.1912	0.1222	0.0515	0.0085	0.0002	0								
employed	0.0695	0.0974	0.1586	0.4939	0.7754	0.0776	0.1541	1						
	0.2763	0.1262	0.0125	0	0	0.166	0.0057							
household	-0.0424	0.0279	0.0043	-0.0523	-0.0829	-0.1717	-0.0877	-0.1159	1					
	0.5076	0.6624	0.9459	0.4124	0.1939	0.0021	0.1176	0.0382						
pbg_visit	-0.0941	-0.1409	-0.0739	-0.0753	-0.0369	0.1815	-0.0743	0.057	-0.047	1				
	0.1409	0.0268	0.2484	0.2385	0.5646	0.0011	0.1861	0.3111	0.4034					
tree	-0.0798	-0.0523	-0.0906	-0.0626	-0.0392	0.0367	0.0008	0.0232	-0.0158	0.0483	1			
	0.2115	0.4126	0.1559	0.3265	0.5398	0.5133	0.9888	0.6799	0.7777	0.391				
air	0.0093	0.0491	0.0553	0.0191	0.0092	-0.0721	0.0395	-0.0735	0.0027	0.0541	0.0182	1		
	0.885	0.4419	0.3869	0.7652	0.8851	0.198	0.4818	0.1897	0.961	0.3361	0.7404			
years_pen	-0.2008	-0.2081	-0.0569	-0.1519	-0.0567	0.7952	-0.4934	0.0538	-0.1059	0.258	0.0313	-0.0077	1	
	0.0015	0.001	0.3732	0.0166	0.3752	0	0	0.3376	0.0585	0	0.5767	0.8902		
visit_freq	-0.038	-0.0907	-0.0033	-0.0446	-0.0065	0.2446	-0.0394	0.046	-0.0907	0.3957	-0.0446	-0.1005	0.2944	1
	0.5517	0.1546	0.9583	0.4849	0.9185	0	0.4824	0.4119	0.1055	0	0.4268	0.0726	0	

Table 51: Pearson's Correlation Analysis (Penang Youth Park)

	youth_tc1	youth_tc2	youth_tc3	youth_tc4	youth_tc5	age	educ	employed	household	pbg_visit	tree	air	years_pen	visit_freq
youth_tc1	1													
youth_tc2	0.727	1												
	0													
youth_tc3	0.3681	0.6592	1											
	0	0												
youth_tc4	0.6307	0.6445	0.4226	1										
	0	0	0											
youth_tc5	0.5063	0.5337	0.4679	0.9527	1									
	0	0	0	0										
age	-0.0533	-0.0946	0.0447	0.0545	0.0677	1								
	0.4049	0.139	0.485	0.3947	0.2899									
educ	0.2191	0.1994	0.0672	0.2798	0.2735	-0.4125	1							
	0.0005	0.0017	0.2935	0	0	0								
employed	0.0993	0.1686	0.1517	0.76	0.8116	0.1677	0.1756	1						
	0.1202	0.0081	0.0173	0	0	0.0084	0.0057							
household	-0.1487	0.0277	0.0374	-0.1209	-0.108	-0.1796	-0.1464	-0.0689	1					
	0.0197	0.6654	0.559	0.0583	0.0911	0.0047	0.0216	0.2815						
pbg_visit	-0.2327	-0.2638	-0.0399	-0.1895	-0.1589	0.3778	-0.3049	-0.0559	0.0353	1				
	0.0002	0	0.5329	0.0029	0.0126	0	0	0.3828	0.5814					
tree	-0.0643	-0.0845	-0.0923	-0.073	-0.0482	0.0631	-0.0245	-0.0205	-0.0054	0.1014	1			
	0.3153	0.1867	0.1491	0.2543	0.4517	0.3246	0.7026	0.7496	0.9331	0.1128				
air	-0.1027	-0.0267	0.0367	-0.0402	-0.0616	-0.0818	0.04	-0.0898	0.0079	-0.0347	-0.0671	1		
	0.1083	0.6768	0.5667	0.5299	0.3358	0.201	0.5323	0.1602	0.902	0.5876	0.2943			
years_pen	-0.1478	-0.1228	0.0412	-0.0687	-0.0489	0.7164	-0.4204	0.0464	-0.0524	0.3639	0.0483	-0.0289	1	
	0.0204	0.0544	0.5201	0.2829	0.4453	0	0	0.4691	0.4134	0	0.4506	0.6516		
visit_freq	-0.1662	-0.1261	0.0565	-0.1016	-0.0524	0.2897	-0.0782	0.0089	-0.0306	0.4443	0.0141	-0.0446	0.2871	1
	0.009	0.0483	0.3775	0.1121	0.4131	0	0.2217	0.8894	0.6325	0	0.8259	0.4858	0	

Table 52: Spearman's Correlation Analysis (Penang Youth Park)

Appendix E

Travel Cost Model Regression Results

E.1 Penang Botanic Gardens

E.1.1 Poisson regression analysis

E.1.2 Zero-inflated Poisson regression analysis

Five zero-inflated Poisson regression models were estimated, each corresponding to one of the five methods used to compute travel costs. The results of these models are presented in Table 54. Model 2 has the highest log-likelihood, while Model 5 has the lowest log-likelihood. All travel costs exhibit a negative sign and are statistically significant, except for *TC1*.

For respondents who reported visiting urban green spaces only a few times in the past year, all the travel costs have negative signs and are significant. This suggests that lower travel costs make Penang Botanic Gardens more attractive to this group of respondents.

The coefficient estimates for the *age* variable have positive signs and are significant across all models. Conversely, the coefficient estimates for the *educ* variable have negative signs and are statistically insignificant. The *employed* dummy variable shows positive signs but lacks significance in some models, indicating a weak relationship with the dependent variable.

The coefficient estimates for the *household* variable are consistently negative and significant in all models, implying that households with more members tend to visit Penang Botanic Gardens less frequently.

	(1) pbg_visit	(2) pbg_visit	(3) pbg_visit	(4) pbg_visit	(5) pbg_visit
pbg_visit					
pbg_tc1	-0.0780***				
	(3.32)				
pbg_tc2		-0.0902***			
		(8.88)			
pbg_tc3			-0.0186***		
			(4.30)		
pbg_tc4				-0.0565***	
				(4.64)	
pbg_tc5					-0.0156***
					(2.58)
age	0.0121***	0.0120***	0.0119***	0.0124***	0.0120***
	(6.36)	(6.34)	(6.35)	(6.54)	(6.34)
educ	0.0325	0.0548***	0.0319	0.0375*	0.0284
	(1.59)	(2.67)	(1.57)	(1.84)	(1.40)
employed	0.0503	0.0844*	0.0485	0.236***	0.190**
	(0.99)	(1.66)	(0.96)	(3.61)	(2.42)
household	-0.0225	-0.00544	-0.0189	-0.0188	-0.0229
	(1.32)	(0.33)	(1.12)	(1.11)	(1.35)
youth_visit	0.0374***	0.0398***	0.0385***	0.0374***	0.0374***
	(32.83)	(33.18)	(32.69)	(32.85)	(32.87)
air	0.00911***	0.00909***	0.00942***	0.0101***	0.00962***
	(4.40)	(4.43)	(4.57)	(4.85)	(4.61)
tree	0.0157***	0.0154***	0.0156***	0.0156***	0.0163***
	(3.06)	(3.02)	(3.04)	(3.04)	(3.18)
_cons	0.440*	0.537**	0.526**	0.321	0.375
	(1.92)	(2.31)	(2.28)	(1.38)	(1.62)
N	257	257	257	257	257
ll	-1087.6	-1045.7	-1084.4	-1081.4	-1090.5

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 53: Ordinary Poisson (Penang Botanic Gardens)

Interestingly, the coefficient estimates for the *youth visit* variable are positive and statistically significant, indicating that individuals who visit Penang Botanic Gardens more frequently are also more likely to visit Penang Youth Park.

On the other hand, *air1* exhibits a negative relationship with the dependent variable, while *air2* shows a positive relationship. However, the coefficient estimates are statistically insignificant, except for *air1* in Model 4. This is an interesting finding, suggesting that respondents assigning higher importance to air quality tend to make fewer visits.

Finally, the coefficient estimates for *tree1* and *tree2* are positive and statistically significant, indicating that respondents who place greater importance on tree species and ecosystems are more likely to visit Penang Botanic Gardens.

Zero-inflated Poisson regression analysis - Equidispersion tests

The zero-inflated Poisson regression models were also estimated using Stata version 14.2. The estimation results are reported in Table 54 in Appendix E.

An equidispersion test was also conducted for the zero-inflated Poisson model. As shown in Table 55, the coefficient estimates of $\hat{\mu}$ in five models are not zero and have a negative sign, indicating that the count data are overdispersed. Therefore, as suggested by Cameron and Trivedi (1990), this study uses the quasi-Poisson model, which maximizes the Poisson maximum likelihood estimation but employs the robust estimate of the variance-covariance estimator (VCE) to account for overdispersion.

Robust Estimate of VCE for Zero-inflated Poisson Model

The results of the robust estimate of VCE for the zero-inflated Poisson model are presented in Table 56. As expected, the robust standard errors are generally larger. Compared to the ordinary zero-inflated Poisson model, more coefficient estimations of independent variables are statistically insignificant. Model 2 has the highest log-likelihood, while Model 5 has the lowest log-likelihood.

Among the five travel costs, only the coefficient estimates of *TC2* is statistically significant at 5 % level. Most of the coefficient estimates are statistically insignificant at the 10% level. However, the coefficients of *age* and the number of visits to the substitute park - Penang Youth Park (*youth visit*) are positive and significant at the 10% level.

The estimations of the marginal effects of variables in the model are presented in Table 57. A one-unit decrease in travel costs, expressed in Malaysian currency (Ringgit Malaysia), increases the expected visit frequency by a range between 0.05 and 0.42. An increase of one unit in the number of visits to the substitute park -

	(1) pbg_visit	(2) pbg_visit	(3) pbg_visit	(4) pbg_visit	(5) pbg_visit
pbg_visit					
pbg_tc1	-0.0614***				
	(2.72)				
pbg_tc2		-0.0752***			
		(7.51)			
pbg_tc3			-0.0144***		
			(3.33)		
pbg_tc4				-0.0416***	
				(3.41)	
pbg_tc5					-0.00866
					(1.43)
age	0.00916***	0.00946***	0.00903***	0.00951***	0.00895***
	(4.76)	(4.92)	(4.74)	(4.93)	(4.65)
educ	0.0224	0.0421**	0.0204	0.0251	0.0159
	(1.09)	(2.03)	(1.00)	(1.22)	(0.77)
employed	0.0989*	0.122**	0.0942*	0.231***	0.172**
	(1.94)	(2.38)	(1.85)	(3.55)	(2.19)
household	-0.0542***	-0.0346**	-0.0501***	-0.0499***	-0.0546***
	(3.13)	(2.01)	(2.90)	(2.88)	(3.15)
youth_visit	0.0354***	0.0375***	0.0362***	0.0354***	0.0354***
	(30.19)	(30.41)	(29.96)	(30.24)	(30.20)
air	0.00900***	0.00922***	0.00916***	0.00981***	0.00921***
	(4.30)	(4.43)	(4.39)	(4.65)	(4.37)
tree	0.0138***	0.0126**	0.0136***	0.0131**	0.0145***
	(2.65)	(2.44)	(2.61)	(2.52)	(2.78)
_cons	0.809***	0.839***	0.882***	0.709***	0.782***
	(3.51)	(3.59)	(3.82)	(3.04)	(3.37)
inflate					
_cons	-2.263***	-2.337***	-2.262***	-2.280***	-2.259***
	(10.20)	(9.89)	(10.19)	(10.11)	(10.21)
N	257	257	257	257	257
ll	-1042.2	-1012.6	-1040.6	-1039.7	-1045.3

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 54: Zero-inflated Poisson model (Penang Botanic Gardens)

	(1)	(2)	(3)	(4)	(5)
	y*	y*	y*	y*	y*
$\hat{\mu}$	-0.104***	-0.105***	-0.105***	-0.104***	-0.105***
	(15.28)	(14.50)	(15.45)	(15.14)	(15.41)
N	257	257	257	257	257

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 55: Equidispersion test

Penang Youth Park, results in an expected increase of approximately 0.21 visits to Penang Botanic Gardens. Lastly, a one year increase in respondents' age raises the expected number of visits by a maximum of 0.06.

E.2 Penang Youth Park

E.2.1 Poisson regression analysis

E.2.2 Zero-inflated Poisson regression analysis

Zero-inflated Poisson regression analysis - Equidispersion tests

The results of the zero-inflated Poisson (ZTP) regression models for visits to Penang Youth Park are presented in Table 59 in Appendix E.

The results of the equidispersion test in Table 60 show that the coefficients of $\hat{\mu}$ in all models are negative, indicating that the count data are underdispersed. Therefore, the robust estimate of the variance-covariance estimator (VCE) for the model is employed.

Robust estimate of VCE for zero-inflated Poisson model

The results of the robust estimate of VCE for the zero-inflated Poisson model are presented in Table 61. Among the models, Model 3 has the highest log-likelihood, while Model 2 has the lowest log-likelihood. Among the five travel costs, only the coefficient estimate for *TC3* is statistically significant at 10 % level. Surprisingly, this variable has a positive sign when it was expected to be negative.

The coefficient estimates for *educ* (education level), *pbg visit* (number of visits to Penang Botanic Gardens), and *air* are statistically significant. However, the

	(1) pbg_visit	(2) pbg_visit	(3) pbg_visit	(4) pbg_visit	(5) pbg_visit
pbg_visit					
pbg_tc1	-0.0614				
	(1.14)				
pbg_tc2		-0.0752**			
		(2.40)			
pbg_tc3			-0.0144		
			(1.26)		
pbg_tc4				-0.0416	
				(1.43)	
pbg_tc5					-0.00866
					(0.55)
age	0.00916*	0.00946**	0.00903*	0.00951**	0.00895*
	(1.91)	(1.98)	(1.91)	(2.01)	(1.87)
educ	0.0224	0.0421	0.0204	0.0251	0.0159
	(0.35)	(0.65)	(0.32)	(0.40)	(0.25)
employed	0.0989	0.122	0.0942	0.231	0.172
	(0.65)	(0.82)	(0.61)	(1.46)	(0.89)
household	-0.0542	-0.0346	-0.0501	-0.0499	-0.0546
	(0.93)	(0.62)	(0.86)	(0.86)	(0.93)
youth_visit	0.0354***	0.0375***	0.0362***	0.0354***	0.0354***
	(4.45)	(5.26)	(4.83)	(4.48)	(4.50)
air	0.00900	0.00922	0.00916	0.00981	0.00921
	(1.50)	(1.63)	(1.52)	(1.63)	(1.51)
tree	0.0138	0.0126	0.0136	0.0131	0.0145
	(1.20)	(1.20)	(1.19)	(1.16)	(1.24)
_cons	0.809	0.839	0.882	0.709	0.782
	(1.26)	(1.29)	(1.38)	(1.09)	(1.19)
inflate					
_cons	-2.263***	-2.337***	-2.262***	-2.280***	-2.259***
	(10.04)	(9.56)	(10.03)	(9.87)	(10.06)
N	257	257	257	257	257
Log pseudolikelihood	-1042.2	-1012.6	-1040.6	-1039.7	-1045.3

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 56: Robust estimate of VCE for Zero-inflated Poisson model (Penang Botanic Gardens)

	(1)	(2)	(3)	(4)	(5)
	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx
pbg_tc1	-0.36				
pbg_tc2		-0.42**			
pbg_tc3			-0.08		
pbg_tc4				-0.24	
pbg_tc5					-0.05
age	0.05*	0.05**	0.05*	0.06**	0.05*
educ	0.13	0.24	0.12	0.15	0.09
employed	0.58	0.69	0.55	1.34	1.01
household	-0.32	-0.19	-0.29	-0.29	-0.32
youth_visit	0.21***	0.21***	0.21***	0.21***	0.21***
air	0.05	0.05	0.05	0.06	0.05
tree	0.08	0.07	0.08	0.08	0.08

t statistics in parentheses

* p <0.10, ** p <0.05, *** p <0.01

Table 57: Marginal effects of variables from Zero-inflated Poisson model (Penang Botanic Gardens)

coefficient estimates for *age*, *employed*, *household*, *air1*, and *tree* are found to be statistically insignificant.

The estimation results of marginal effects of independent variables are presented in Table 62. Specifically, for every one unit increase in *TC3*, there is an associated increase in the expected visit frequency by 0.1. On average, for every one-level increase in education level, the expected number of visits decreases within the range of 0.55 to 0.65. Additionally, a unit increase in the number of visits to Penang Botanic Gardens leads to an expected increase in the number of visits to Penang Youth Park by approximately 0.26.

	(1) youth_visit	(2) youth_visit	(3) youth_visit	(4) youth_visit	(5) youth_visit
youth_visit					
youth_tc1	-0.0477**				
	(2.12)				
youth_tc2		0.00433			
		(0.59)			
youth_tc3			0.0247***		
			(5.80)		
youth_tc4				-0.00763	
				(0.74)	
youth_tc5					0.00985*
					(1.70)
age	0.000119	0.00000795	-0.000746	0.000185	-0.000452
	(0.06)	(0.00)	(0.38)	(0.09)	(0.23)
educ	-0.145***	-0.152***	-0.161***	-0.148***	-0.157***
	(6.90)	(7.22)	(7.69)	(7.00)	(7.33)
employed	0.0860	0.0642	0.0245	0.0933	-0.0255
	(1.56)	(1.16)	(0.44)	(1.46)	(0.32)
household	0.00978	0.00801	-0.00312	0.0106	0.00524
	(0.57)	(0.47)	(0.18)	(0.62)	(0.31)
pbg_visit	0.0534***	0.0542***	0.0548***	0.0538***	0.0540***
	(34.39)	(34.46)	(35.26)	(34.76)	(34.98)
air	-0.00806***	-0.00843***	-0.00950***	-0.00815***	-0.00868***
	(3.73)	(3.87)	(4.29)	(3.74)	(3.95)
tree	0.000396	0.000967	0.00329	0.000667	0.000872
	(0.07)	(0.18)	(0.61)	(0.12)	(0.16)
_cons	2.210***	2.207***	2.130***	2.185***	2.275***
	(9.47)	(9.46)	(9.15)	(9.30)	(9.60)
N	245	246	245	246	245
ll	-845.6	-851.1	-832.4	-851.0	-847.1

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 58: Ordinary Poisson (Penang Youth Park)

	(1) youth_visit	(2) youth_visit	(3) youth_visit	(4) youth_visit	(5) youth_visit
youth_visit					
youth_tc1	-0.0535**				
	(2.28)				
youth_tc2		0.00425			
		(0.56)			
youth_tc3			0.0208***		
			(4.72)		
youth_tc4				-0.00827	
				(0.77)	
youth_tc5					0.00903
					(1.54)
age	-0.000736	-0.000938	-0.00168	-0.000680	-0.00135
	(0.37)	(0.47)	(0.83)	(0.34)	(0.66)
educ	-0.112***	-0.120***	-0.132***	-0.115***	-0.125***
	(5.10)	(5.45)	(6.06)	(5.23)	(5.62)
employed	0.0305	0.0167	-0.00204	0.0449	-0.0658
	(0.53)	(0.29)	(0.04)	(0.68)	(0.81)
household	0.0179	0.0133	0.00188	0.0166	0.0114
	(1.02)	(0.74)	(0.10)	(0.93)	(0.64)
pbg_visit	0.0532***	0.0537***	0.0540***	0.0535***	0.0537***
	(33.23)	(33.32)	(33.77)	(33.44)	(33.68)
air	-0.00641***	-0.00675***	-0.00776***	-0.00643***	-0.00707***
	(2.89)	(3.03)	(3.43)	(2.87)	(3.14)
tree	0.00162	0.00202	0.00408	0.00167	0.00220
	(0.30)	(0.37)	(0.73)	(0.30)	(0.40)
_cons	2.084***	2.098***	2.064***	2.064***	2.161***
	(8.64)	(8.72)	(8.63)	(8.48)	(8.83)
inflate					
_cons	-2.566***	-2.498***	-2.534***	-2.512***	-2.536***
	(8.63)	(8.78)	(8.82)	(8.71)	(8.79)
N	245	246	245	246	245
ll	-806.8	-812.2	-799.5	-812.0	-809.0

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 59: Zero-inflated Poisson model (Penang Youth Park)

	(1)	(2)	(3)	(4)	(5)
	y*	y*	y*	y*	y*
$\hat{\mu}$	-0.0924***	-0.0932***	-0.0925***	-0.0929***	-0.0926***
	(11.41)	(11.54)	(11.41)	(11.51)	(11.54)
N	245	246	245	246	245

t statistics in parentheses

* p <0.10, ** p <0.05, *** p <0.01

Table 60: Equidispersion test

	(1) youth_visit	(2) youth_visit	(3) youth_visit	(4) youth_visit	(5) youth_visit
youth_visit					
youth_tc1	-0.0535				
	(1.42)				
youth_tc2		0.00425			
		(0.26)			
youth_tc3			0.0208*		
			(1.95)		
youth_tc4				-0.00827	
				(0.72)	
youth_tc5					0.00903
					(1.08)
age	-0.000736	-0.000938	-0.00168	-0.000680	-0.00135
	(0.19)	(0.24)	(0.43)	(0.17)	(0.33)
educ	-0.112**	-0.120**	-0.132***	-0.115**	-0.125***
	(2.49)	(2.51)	(2.76)	(2.49)	(2.64)
employed	0.0305	0.0167	-0.00204	0.0449	-0.0658
	(0.23)	(0.12)	(0.02)	(0.34)	(0.47)
household	0.0179	0.0133	0.00188	0.0166	0.0114
	(0.64)	(0.45)	(0.06)	(0.56)	(0.37)
pbg_visit	0.0532***	0.0537***	0.0540***	0.0535***	0.0537***
	(16.39)	(16.75)	(16.41)	(16.11)	(16.29)
air	-0.00641*	-0.00675*	-0.00776*	-0.00643*	-0.00707*
	(1.68)	(1.71)	(1.93)	(1.66)	(1.84)
tree	0.00162	0.00202	0.00408	0.00167	0.00220
	(0.16)	(0.20)	(0.42)	(0.17)	(0.22)
_cons	2.084***	2.098***	2.064***	2.064***	2.161***
	(4.69)	(4.62)	(4.88)	(4.55)	(4.74)
inflate					
_cons	-2.566***	-2.498***	-2.534***	-2.512***	-2.536***
	(8.21)	(8.33)	(8.46)	(8.28)	(8.39)
N	245	246	245	246	245
Log pseudolikelihood	-806.8	-812.2	-799.5	-812.0	-809.0

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 61: Robust estimate of VCE for Zero-inflated Poisson model (Penang Youth Park)

	(1)	(2)	(3)	(4)	(5)
	dy/dx	dy/dx	dy/dx	dy/dx	dy/dx
youth_tc1	-0.26				
youth_tc2		0.02			
youth_tc3			0.1*		
youth_tc4				-0.04	
youth_tc5					0.04
age	0	0	-0.01	0	-0.01
educ	-0.55**	-0.59**	-0.65***	-0.56**	-0.61**
employed	0.15	0.08	-0.01	0.22	-0.32
household	0.09	0.06	0.01	0.08	0.06
pbg_visit	0.26***	0.26***	0.26***	0.26***	0.26***
air	-0.03	-0.03*	-0.04*	-0.03	-0.03*
tree	0.01	0.01	0.02	0.01	0.01

t statistics in parentheses
* p <0.10, ** p <0.05, *** p <0.01

Table 62: Marginal effects of variables from Zero-inflated Poisson model (Penang Youth Park)

Appendix F

Models in Preference-space Specification

F.1 Multinomial logit model

Table 63 presents the results of the MNL models estimated on the full sample. In Model 1, *distance* serves as the cost variable, while Models 2 and 3 use *travel cost*s as the cost variable. In Model 1, the coefficient for *alt 1* is positive and statistically significant at the 1% confidence level. This suggests a preference for choosing site A over site B, indicating slight left-to-right bias, which is common in population that write left-to-right.

Besides the alternative-specific constant, all coefficient estimates exhibit positive signs and high levels of significance, except for *distance*. This indicates that people place value on the included characteristics of the green sites, in addition to considering other personal factors. The *distance* variable, treated as the cost variable, has a negative coefficient that is statistically significant at the 1% confidence level. This aligns with economic theory, where each additional kilometre of travel from the respondent's place of residence to the hypothetical green site reduces the marginal utility of the visit by a factor of 0.041. These results indicate that even when the travel distance is less than 40 kilometres, respondents are sensitive to additional travel distances. Furthermore, all dummy-coded variables have positive coefficients, but the increment from level 1 to level 2 is smaller. This suggests that higher attribute levels lead to greater utility, but there is diminishing marginal utility as the attributes improve, implying a degree of satiation.

Air quality is the attribute that received the highest rating from respondents. An improvement in air quality from poor to intermediate (represented by *air1*), which corresponds to an average increase of 50 units in the air quality index, leads to a significant increase in marginal utility, with a factor of 1.437. Furthermore, a further enhancement of air quality from intermediate to good (represented by *air2-air1*), equivalent to an additional 50 units in the air quality index beyond level 1, results in a marginal utility increase with a factor of 0.994. Consequently, the combined effect of improving air quality from poor to good leads to the highest incremental factor of 2.431 among all attribute levels. This demonstrates that respondents highly value improvements in air quality.

Noise reduction at the site is also highly valued by respondents. The improvement in noise levels from loud to intermediate (represented by *noi1*) leads to a substantial increase in utility, with a factor of 0.643. Additionally, further enhancing noise levels from intermediate to quiet (represented by *noi2-noi1*) results in a marginal utility increase by a factor of 0.227. This level of improvement corresponds to a 30 dB reduction in noise on average. This indicates that respondents place a greater value on the initial noise reduction but assign less value to further noise reduction. This observation suggests that for most respondents, an intermediate noise level ranging from 40 dB to 60 dB (similar to noise from conversation or refrigerator) is satisfactory.

The presence of facilities at the green site also holds substantial importance for respondents. Transitioning from a few facilities to intermediate-level facilities (represented by *fac1*) results in an increase in marginal utility, with a factor of 0.270. Furthermore, further elevating the level of facilities to a lot of facilities (represented by *fac2-fac1*) leads to an additional increase in marginal utility, with a factor of 0.291. This observation suggests that the presence of specific amenities, such as cycling tracks and playgrounds, significantly enhances the site's appeal. Additionally, the inclusion of other amenities like a football field, baseball court, hiking track, and rock-climbing wall contributes to the site's attractiveness at almost a similar level.

Furthermore, respondents exhibit a preference for greater diversity in tree species and ecosystems compared to lower diversity. The transition from low to intermediate diversity (represented by *tree1*) results in a utility increase with a factor of 0.135. A more substantial increase in utility is observed when transitioning from intermediate to higher diversity (represented by *tree2-tree1*), yielding an increase in marginal utility by a factor of 0.227. These findings highlight that the presence of a high diversity of tree species serves as a significant factor influencing respondents' decisions on which site to visit. In contrast, the presence of intermediate diversity does not show as much influence on respondents' decision-making regarding their choice of site to visit.

Regarding nursery habitat maintenance, respondents do display positive preferences for improvements in this attribute. However, the coefficient estimates for both levels of improvement are relatively lower compared to those for other attributes. An improvement from poor to intermediate maintenance (represented by *nur1*) leads to a utility increase with a factor of 0.183, while further improvement from intermediate to good maintenance (represented by *nur2-nur1*) results in a utility increase with a factor of 0.117. These findings suggest that respondents not highly value on-site nurseries contributing to the reproduction of specific tree and plant species compared to other attributes. This could be attributed to the possibility that new and young trees may come from off-site nurseries.

Model 2 introduces the *tc* variable, and Model 3 introduces the *tca* variable. In these models, the travel cost coefficients are consistently negative, while the coefficients for the other attributes are positive.

To assess model fit, log-likelihood, AIC, and BIC were employed. These metrics are commonly used to evaluate how well a model fits the data accounting for the number of parameters and vice versa. The empirical results indicate that the inclusion of the value of time in the travel cost variable does not improve the fit of the data.

F.1.1 Marginal effects

Table 64 displays the average marginal effects (ME) of attribute levels on the choice for Models 1 - 3. The results indicate that an additional kilometre of *distance* from the respondent's residence to the hypothetical green site reduces the probability of the green site being visited by 0.5% (Model 1), 1.4% (Model 2) and 0.8% (Model 3). This suggests that small changes in travel distance have a relatively minor impact on the choice probability.

Air quality exhibits the most substantial influence on choice probability. An improvement from the base level to level 1 increases the probability of the site being chosen by 24.8% (Model 1), 25.3% (Model 2) and 24.9% (Model 3), while a two-level improvement raises the probability by 33.7% (Model 1), 34.2% (Model 2) and 33.7% (Model 3). Among all the attributes, *air quality* has the most significant effect on choice probability.

The presence of facilities at the site is also important. An upgrade from the base level to level 1 increases the probability of the site being chosen by 3.8% (Model 1), 3.7% (Model 2) and 3.6% (Model 3), while a two-level improvement elevates the probability by 7.4% (Model 1) and 7.1% (Models 2 and 3).

Noise level improvements have a greater impact than facilities. A reduction in noise levels by one level increases the probability of the site being chosen by 9.3% - 9.4%, while a two-level noise reduction boosts the choice probability by 12% - 12.1%.

Regarding nursery habitat maintenance, upgrading from level 0 to level 1 increases the choice probability by 2.5% - 2.6%, and the two-level improvement raises it by 4%

- 4.2%.

Lastly, enhancing tree species diversity and ecosystems from the base level to level 1 increases the probability of the site being chosen by 1.9% (Model 1) and 1.5% (Models 2 and 3), while a two-level improvement increases the choice probability by 4.8% (Models 1 and 2) and 4.7% (Model 3).

Comparing Models 2 and 3, the introduction of either *tc* or *tca* does not show a significant difference in marginal effects. However, in most cases, the marginal effects of attribute levels in Model 2 are higher than in Model 3, even though *tca* takes into account the value of travel time.

F.2 Uncorrelated coefficients in mixed logit model

In this section, the MXL models assumed uncorrelated coefficients among random coefficients. Three models were developed:

1. The MXL model that assumed the coefficients for the dummy variables of five attributes, including *air quality*, *facilities*, *noise levels*, *nursery habitat maintenance*, and *tree species and ecosystems*, to be random with a normal distribution, and the coefficients for the negative of the distance variable (*n_dis*) and *alt1* were assumed to be fixed;

2. The MXL model that assumed the coefficient for the *distance* variable to be random with a log-normal distribution, and the coefficients for the attributes that had significant SDs at 1% level in the first model to be random with a normal distribution. These attributes include *air1, air2, fac2, noi1, nur1,* and *tree1*, while the other variables were assumed to be fixed;

3. The MXL model that assumed the coefficient for the *distance* variable to be random with a log-normal distribution, and other variables were assumed to be fixed.

The MXL models in preference space were estimated using maximum simulated likelihood. To obtain a better estimation, 50 Halton draws were initially used in the estimation of random parameters. The estimated values were then employed as starting values in estimation with larger draws. For the final estimation with starting values, 1000 Halton draws were used. The results obtained using 1000 Halton draws are presented in Table 65.

For Models 1, 2, and 3, the coefficient estimate for *alt1* is positive and significant at the 1% confidence level, indicating that respondents prefer choosing site A over site

	Model 1	Model 2	Model 3
dis	-0.041		
	(0.002)		
tc		-0.103	
		(0.005)	
tca			-0.061
			(0.003)
air 1	1.437	1.462	1.448
	(0.062)	(0.070)	(0.070)
air 2	2.431	2.466	2.440
	(0.081)	(0.093)	(0.092)
fac 1	0.270	0.266	0.259
	(0.052)	(0.059)	(0.059)
fac 2	0.561	0.542	0.536
	(0.055)	(0.062)	(0.061)
noi 1	0.643	0.653	0.648
	(0.057)	(0.065)	(0.064)
noi 2	0.871	0.879	0.873
	(0.055)	(0.062)	(0.062)
nur 1	0.183	0.186	0.186
	(0.055)	(0.062)	(0.062)
nur 2	0.300	0.310	0.314
	(0.053)	(0.060)	(0.060)
tree 1	0.135	0.106*	0.107*
	(0.052)	(0.058)	(0.058)
tree 2	0.362	0.370	0.363
	(0.055)	(0.061)	(0.061)
alt 1	0.123	0.129	0.127
	(0.035)	(0.040)	(0.039)
N	9696	7680	7680
Log-likelihood	-2505.42	-1976.15	-1981.95
AIC	5034.83	13041.61	21208.91
BIC	5120.984	13118.02	21285.32

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

Table 63: MNL model estimations

	Model 1	Model 2	Model 3
dis	-0.005		
tc		-0.014	
tca			-0.008
air 1	0.248	0.253	0.249
air 2	0.337	0.342	0.337
fac 1	0.038	0.037	0.036
fac 2	0.074	0.071	0.071
noi 1	0.093	0.094	0.093
noi 2	0.120	0.121	0.120
nur 1	0.025	0.026	0.025
nur 2	0.040	0.041	0.042
tree 1	0.019	0.015*	0.015*
tree 2	0.048	0.048	0.047
alt 1	0.016	0.017	0.017
N	9696	7680	7680

Standard errors in parentheses
* p<0.10, ** p<0.05, ‘ ’ p<0.01

Table 64: Marginal effects from MNL model estimations

B, confirming left-to-right bias. In Model 1, the coefficient for the negative of distance variable (*n_dis*) is positive and significant, suggesting that the longer the distance, the lower the utility of the site visit. This result highlights respondents' sensitivity to distance, even when it is less than 40 kilometres. In Models 2 and 3, the coefficients for the negative of distance variable (*n_dis*) have positive signs because they were assumed to be random with a log-normal distribution in these models. Thus, it can be interpreted that a longer distance results in a lower utility for the visit.

The coefficient estimates for all dummy-coded random variables for site attributes consistently exhibit a positive sign at the mean in all models and are significant at the 10% confidence level. The SD of coefficients for random coefficients is shown in the lower part of the table. In Model 1, the SD of coefficients for *air1, air2, fac2, noi1, nur1*, and *tree1* are significant at the 1% confidence level, while that of *fac2* is significant at the 10% confidence level. Additionally, the SD of coefficients for *tree1* is significant in Model 1 but becomes insignificant in Model 2.

The results clearly demonstrate that respondents have strong preferences for improving air quality, as indicated by the significantly positive coefficient estimates for *air1* and *air2*. According to Models 1 and 2, the mean and SD estimates of the *air1* random coefficient suggest that there is 91.74% and 93.33% probability of drawing a respondent with the current coefficient with value of more than zero. Furthermore, as many as 93.3% (Model 1) and 96.12% (Model 2) of respondents would prefer the air quality to be improved from level 1 to 2. On average, the utility increases by a factor of 1.969 (Model 1), 2.292 (Model 2), and 1.618 (Model 3) when the air quality improves from level 0 to 1, while the improvement from level 1 to 2 increases the average utility by a factor of 1.338 (Model 1), 1.475 (Model 2), and 1.099 (Model 3). This suggests that respondents place a higher value on the level-one improvement of air quality than on the level-two improvement.

The reduction of noise is also strongly preferred by respondents, as both levels of noise reduction are estimated to have significantly positive coefficients. The improvement from level 0 to 1 is highly valued by respondents, with as much as 93.55% (Model 1) and 92.45% (Model 2) of respondents placing a positive value on this improvement. It increases the marginal utility by a factor of 0.855 (Model 1), 0.897 (Model 2), and 0.665 (Model 3). However, the improvement from level 1 to 2 is not as highly valued, with an increase in marginal utility of only 0.249 (Model 1), 0.329 (Model 2), and 0.294 (Model 3). This indicates that for most respondents, a normal noise level is important at the green site, but a quiet level is not necessary. The SD of coefficients for the first-level improvement is significant, suggesting that preferences for the first-level noise reduction vary across individuals, while there is no significant variation in preferences for the second-level improvement.

The third highly valued attribute that influences the probability of a choice being made is the presence of facilities at the green site. Both levels of facility improvement have nearly the same effect on increasing marginal utility. The changes in the presence of facilities from level 0 to 1 result in a marginal utility increase ranging from 0.308 to 0.428, and 92.97% of respondents prefer this level of facility improvement. The improvement from level 1 to 2 increases utility by a factor ranging from 0.309 to 0.405, with 87.25% (Model 1) and 76.07% (Model 2) of respondents having positive coefficients for this dummy variable. This suggests that on average the combined effects of having cycling tracks and a playground are roughly equivalent to the combined effects of having a football field, baseball court, hiking track, and rock-climbing wall, but possibly would benefit different visitors.

The coefficient estimates for *tree species and ecosystems* at both levels are positive and significant, although they have lower mean values compared to other attributes. The first-level improvement in tree species and ecosystems increases utility by a factor of 0.212 (Model 1), 0.211 (Model 2), and 0.142 (Model 3). Additionally, 68.28% of respondents prefer this level of improvement in Model 1, but this percentage significantly increases to 99.53% in Model 2. The second-level improvement in this attribute increases utility by a factor ranging between 0.235 and 0.316. Although the coefficient estimates at the mean are significant for both levels, the SD for the second-level improvement is insignificant. This indicates that there is preference heterogeneity across individuals for the first level of improvement, but not for the second.

Compared to the previously discussed attributes, nursery habitat maintenance appears to be of lesser concern to the respondents in general. The improvement from poor maintenance (base level) to intermediate maintenance has coefficient estimates of 0.260 (Model 1), 0.279 (Model 2), and 0.191 (Model 3). While a significant portion of respondents, as much as 72.72% (Model 1) and 72.28% (Model 2), place a positive value on this improvement. The increase in utility from the improvement in this attribute from the first level to the second level of maintenance (considered good maintenance) is less significant, with a coefficient estimate of only 0.124. This suggests that an intermediate level of maintenance is sufficient for most respondents, and they do not highly value the presence of higher levels of maintenance in the nursery habitat. The SD of the coefficient for the first-level improvement is significant, indicating the presence of preference heterogeneity at this level, but it is statistically insignificant for the second-level improvement, suggesting that preference heterogeneity is absent for the second level.

Model fit indicators reveal that among the three models, Model 2 provides the best fit to the dataset. The log-likelihood of Model 1 is -2409.93, which increases sharply in Model 2 to a log-likelihood of -2303.96. Model 3 is also a slightly better fit to the dataset compared to Model 1, with a log-likelihood of -2383.11. Both the AIC and BIC tests support these findings, indicating that Model 2 is the best fit for the dataset due to its lower AIC and BIC values.

F.2.1 Model comparison if travel costs were used

The estimation results of the MXL model in which the *distance* variable was transformed into an individual travel cost are presented in Table 66. In Models 1, 3, and 5, the cost variables were represented as negative *tc* (*n_tc*), while in Models 2, 4, and 6, they were represented as negative *tca* (*n_tca*), which includes the value of travel time. Most of the coefficient estimates in Models 2, 4, and 6 are lower in magnitude compared to the estimates in Models 1, 3, and 5, respectively.

Regarding the model fit tests, replacing *n_tc* with *n_tca* as a cost variable does not lead to an improvement in model fit. In fact, the models using *n_tca* generally perform a little worse. The log-likelihood of Model 2 is approximately 5 units lower than that of Model 1, and Models 4 and 6 have log-likelihoods that are 3 units lower than Models 3 and 5, respectively. Similarly, the AIC and BIC tests support these findings, with higher AIC and BIC values observed in Models 2, 4, and 6 compared to Models 1, 3, and 5, respectively.

The empirical results suggest that considering the value of individual travel time as a cost in this context may be inappropriate. To further validate these results and gain a deeper understanding of travel cost estimations, additional research can be undertaken. The current study assumed that all respondents were driving a car alone, and it considered the time spent on the road for each kilometre to be consistent across respondents. The main variation accounted for in the travel time value was the expected hourly income, which contributed to differences in travel time value among respondents.

To provide a more robust rationale for the validity of travel time value and its inclusion as a cost in future studies, further research could explore the following aspects: mode of transportation, travel time sensitivity, and travel time variability.

F.3 Correlated coefficients in MXL model

To explore the presence of correlated unobserved effects among choices in a choice situation, MXL models that allow for the correlation of random parameters were developed. This correlation model produces a covariance matrix that contains off-diagonal elements, which imply the existence of correlations among random utility

	Model 1	Model 2	Model 3
Mean			
n_dis	0.053	-3.124	-3.403
	(0.003)	(0.090)	(0.089)
alt1	0.145	0.133	0.116
	(0.041)	(0.044)	(0.038)
air1	1.970	2.292	1.618
	(0.122)	(0.143)	(0.069)
air2	1.338	1.475	1.099
	(0.095)	(0.104)	(0.064)
fac1	0.381	0.428	0.308
	(0.066)	(0.070)	(0.058)
fac2	0.404	0.405	0.309
	(0.068)	(0.079)	(0.059)
noi1	0.855	0.897	0.665
	(0.081)	(0.089)	(0.063)
noi2	0.249	0.329	0.294
	(0.062)	(0.068)	(0.058)
nur1	0.260	0.279	0.191
	(0.070)	(0.077)	(0.061)
nur2	0.124*	0.118*	0.100*
	(0.066)	(0.071)	(0.059)
tree1	0.212	0.211	0.142**
	(0.067)	(0.071)	(0.059)
tree2	0.284	0.316	0.235
	(0.067)	(0.073)	(0.061)
SD			
air1	1.420	1.527	
	(0.126)	(0.142)	
air2	0.893	0.836	
	(0.114)	(0.135)	
fac1	0.259*		
	(0.154)		
fac2	0.355	0.571	
	(0.136)	(0.120)	
noi1	0.563	0.625	
	(0.110)	(0.124)	
noi2	0.007		
	(0.224)		
nur1	0.430	0.472	
	(0.118)	(0.134)	
nur2	0.099		
	(0.294)		
tree1	0.446	0.081	
	(0.107)	(0.393)	
tree2	0.012		
	(0.142)		
n_dis		1.012	1.010
		(0.088)	(0.084)
N	9696	9696	9696
Log-likelihood	-2409.93	-2303.96	-2383.11
AIC	4863.857	4645.925	4792.215
BIC	5021.805	4782.335	4885.548

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

Table 65: MXL model estimations

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Mean						
n_tc	0.134		-2.170		-2.456	
	(0.007)		(0.093)		(0.091)	
n_tca		0.080		-2.665		-2.956
		(0.004)		(0.095)		(0.093)
alt1	0.153	0.151	0.143	0.143	0.123	0.123
	(0.046)	(0.046)	(0.050)	(0.050)	(0.043)	(0.043)
air1	2.019	2.004	2.319	2.301	1.618	1.615
	(0.138)	(0.137)	(0.160)	(0.157)	(0.077)	(0.077)
air2	1.376	1.368	1.509	1.505	1.086	1.083
	(0.110)	(0.111)	(0.119)	(0.119)	(0.071)	(0.071)
fac1	0.384	0.375	0.428	0.424	0.297	0.295
	(0.074)	(0.074)	(0.079)	(0.079)	(0.064)	(0.064)
fac2	0.393	0.398	0.397	0.398	0.291	0.291
	(0.080)	(0.080)	(0.090)	(0.090)	(0.066)	(0.066)
noi1	0.881	0.877	0.918	0.915	0.661	0.660
	(0.093)	(0.093)	(0.101)	(0.101)	(0.070)	(0.070)
noi2	0.250	0.247	0.324	0.324	0.287	0.287
	(0.071)	(0.071)	(0.076)	(0.076)	(0.065)	(0.065)
nur1	0.269	0.268	0.293	0.291	0.191	0.191
	(0.080)	(0.080)	(0.087)	(0.087)	(0.068)	(0.068)
nur2	0.136*	0.142*	0.13	0.131	0.110*	0.111*
	(0.075)	(0.075)	(0.080)	(0.080)	(0.065)	(0.065)
tree1	0.189**	0.188**	0.214	0.210	0.130**	0.129*
	(0.077)	(0.077)	(0.081)	(0.081)	(0.066)	(0.066)
tree2	0.332	0.327	0.353	0.351	0.259	0.256
	(0.077)	(0.077)	(0.083)	(0.083)	(0.068)	(0.068)
SD						
air1	1.396	1.376	1.488	1.473		
	(0.139)	(0.138)	(0.157)	(0.156)		
air2	0.957	0.979	0.888	0.898		
	(0.129)	(0.128)	(0.143)	(0.143)		
fac1	0.252	0.253				
	(0.189)	(0.188)				
fac2	0.458	0.461	0.611	0.616		
	(0.131)	(0.131)	(0.132)	(0.132)		
noi1	0.579	0.585	0.640	0.635		
	(0.125)	(0.124)	(0.135)	(0.137)		
noi2	0.011	0.02				
	(0.245)	(0.243)				
nur1	0.470	0.474	0.510	0.510		
	(0.128)	(0.128)	(0.149)	(0.149)		
nur2	0.033	0.044				
	(0.283)	(0.274)				
tree1	0.459	0.470	0.174	0.141		
	(0.126)	(0.123)	(0.266)	(0.294)		
tree2	0.012	0.006				
	(0.167)	(0.164)				
n_tc			0.890		0.905	
			(0.083)		(0.086)	
n_tca				0.933		0.940
				(0.088)		(0.088)
N	7680	7680	7680	7680	7680	7680
Log-likelihood	-1899.27	-1904.64	-1830.31	-1833.08	-1897.98	-1900.6
AIC	3842.545	3853.282	3698.622	3704.164	3821.952	3827.196
BIC	3995.365	4006.102	3830.603	3836.145	3912.254	3917.498

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

Table 66: MXL model estimations (Travel costs)

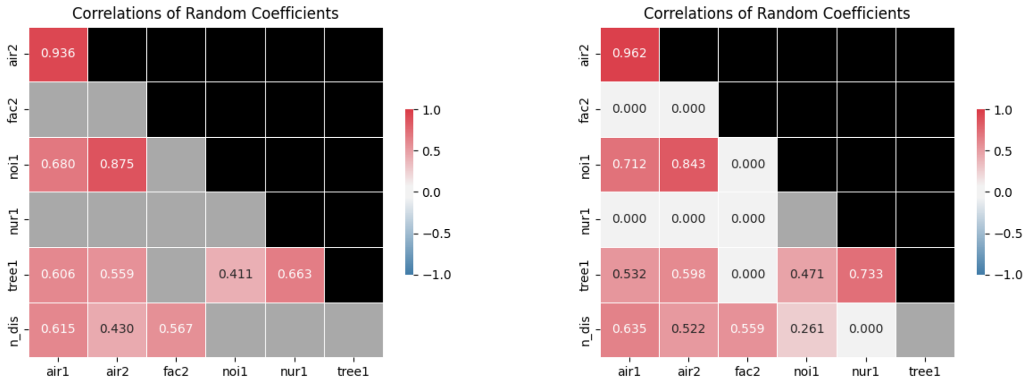
coefficients, helping to examine the effects of one attribute on another attribute (Hensher and Greene, 2003). Instead of assuming full correlation among all random variables, this model assumes only the attribute coefficients that were assumed to be random in Model 2 in Table 65 to be random in this section.

The correlation model includes seven random and correlated coefficients for the attributes, which are *air1*, *air2*, *fac2*, *noi1*, *nur1*, *tree1*, and *n_dis*. The coefficients for *air1*, *air2*, *fac2*, *noi1*, *nur1*, and *tree1* are assumed to follow a normal distribution, and the coefficient for *n_dis* is assumed to follow a log-normal distribution. As a result, the WTP is distributed as the ratio of a normal distribution to a log-normal distribution. Four coefficients are assumed to be fixed and uncorrelated, including *fac1*, *noi2*, *nur2*, and *tree2*.

The estimation results are presented in Table 67. Model 1 represents the estimation results for an MXL model with correlations and no constraints on the Cholesky factors, while Model 2 constrains the insignificant Cholesky decomposition values to be near zero (0.0001). Model 1 has a log-likelihood of -2219.328, while Model 2 has a log-likelihood of -2228.322. In comparison, the log-likelihood for the model with uncorrelated coefficients (Model 2 in Table 65) was -2303.96. The lower log-likelihood values in Models 1 and 2, when compared to the model with uncorrelated coefficients, suggest that there is evidence of correlation among the coefficients, assuming uncorrelated coefficients may lead to misleading results. Although Model 2 has a lower log-likelihood than Model 1, the AIC and BIC values for Model 2 are lower, because of the penalty on parameters. Therefore, Model 2 is considered better than Model 1 based on the log-likelihood criteria. Moreover, the estimation results of Model 2 are considered more accurate as the insignificant Cholesky factors are constrained to be close to zero.

The means of the coefficient estimates are positive, except for *n_dis*, which has a negative coefficient estimate. These results are consistent with the findings obtained in the previous MXL model with no correlation. The lower part of the table displays the SDs of correlated random coefficients, and all of them are significant at the 1% confidence level.

The model estimation involves not only the coefficient estimates but also the elements of the Cholesky decomposition of the variance-covariance matrix for the seven correlated parameters. These elements constitute the lower triangle matrix denoted as L. The matrix L is instrumental in deriving the variance-covariance matrix of the random coefficients. Specifically, the variance-covariance matrix can be expressed as $\sum = LL^T$, where L^T signifies the transpose of matrix L. It includes both the variances of random parameters and the covariances between different random parameters.



(a) MXL with correlations (b) Constraints on Cholesky decomposition

Figure 57: Correlations of random coefficients (Cost coefficient: Distance)

The results reveal that 16 out of 28 coefficient estimates are significant, and the majority of these coefficient estimates exhibit positive correlation values. A positive covariance value implies that two random parameters tend to move in the same direction, either increasing or decreasing together. Conversely, a negative covariance suggests that two random parameters tend to behave in opposite directions. The covariances of *air2* and *nur1*, *fac2* and *noi1*, *fac2* and *nur1*, as well as *nur1* and *n_dis* have negative values. However, the p-values indicate that these four pairs are statistically insignificant at the 10% confidence level.

Heat plots were generated using Python 3.1 to visualize the strength of correlations among random coefficients. Figure 57 (a) displays the heat map illustrating the correlations of random coefficients for Model 1. Out of the 21 correlation estimations, only 10 are statistically significant at the 10% or below significance level.

Regarding the improvement in air quality, a robust and positive correlation exists between the coefficients associated with *air1* and *air2* (0.936). This suggests that individuals who favour the first-level improvement in air quality also tend to prefer the second-level improvement. Another strong positive correlation is observed between the coefficients for *air2* and *noi1* (0.875), followed by the correlation between the coefficients for *air1* and *noi1* (0.680). This indicates that individuals who show a preference for enhancing air quality at both levels also tend to favour improving noise levels at the first level. A stronger correlation effect is observed among the coefficients for *air2* compared to those for the first-level air quality improvement.

Furthermore, the coefficients related to improvements in air quality at the first and second levels exhibit moderate and positive correlations with the coefficients

	Model 1	Model 2
Mean		
n_dis	-3.028	-3.063
	(0.095)	(0.101)
alt1	0.157	0.155
	(0.046)	(0.045)
fac1	0.435	0.412
	(0.075)	(0.072)
noi2	0.314	0.312
	(0.072)	(0.071)
nur2	0.180**	0.182**
	(0.073)	(0.072)
tree2	0.360	0.348
	(0.076)	(0.075)
air1	2.807	2.765
	(0.172)	(0.169)
air2	1.848	1.849
	(0.130)	(0.128)
fac2	0.483	0.478
	(0.092)	(0.083)
noi1	1.148	1.138
	(0.106)	(0.105)
nur1	0.319	0.315
	(0.092)	(0.082)
tree1	0.349	0.353
	(0.090)	(0.086)
SD		
n_dis	1.135	1.175
	(0.070)	(0.096)
air1	1.800	1.728
	(0.160)	(0.154)
air2	1.086	1.062
	(0.137)	(0.131)
fac2	0.662	0.633
	(0.110)	(0.118)
noi1	0.854	0.839
	(0.119)	(0.116)
nur1	0.630	0.576
	(0.116)	(0.120)
tree1	0.598	0.551
	(0.122)	(0.120)
N	9696	9696
Log-likelihood	-2219.328	-2228.322
AIC	4518.656	4514.644
BIC	4805.835	4722.849

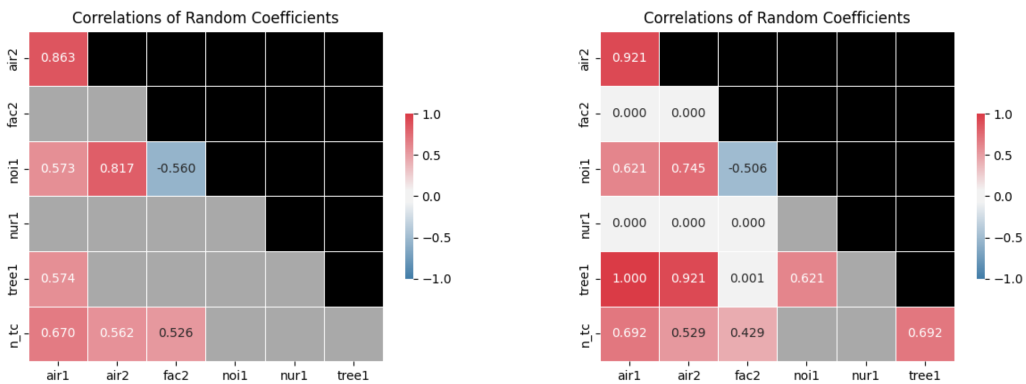
Standard errors in parentheses
* $p<0.10$, ** $p<0.05$, ‘ ’ $p<0.01$

Table 67: MXL models with correlations

	Model 1	Model 2	Model 3	Model 4
Mean				
n_tc	-2.077		-2.110	
	(0.101)		(0.098)	
n_tca		-2.599		-2.592
		(0.105)		(0.101)
alt1	0.172	0.173	0.163	0.162
	(0.052)	(0.052)	(0.050)	(0.051)
fac1	0.452	0.446	0.437	0.441
	(0.083)	(0.084)	(0.079)	(0.081)
noi2	0.293	0.298	0.295	0.290
	(0.081)	(0.081)	(0.078)	(0.079)
nur2	0.181**	0.181**	0.177**	0.175**
	(0.083)	(0.082)	(0.080)	(0.081)
tree2	0.381	0.391	0.381	0.382
	(0.085)	(0.086)	(0.084)	(0.084)
air1	2.831	2.808	2.754	2.770
	(0.202)	(0.196)	(0.187)	(0.193)
air2	1.879	1.853	1.854	1.852
	(0.156)	(0.147)	(0.145)	(0.148)
fac2	0.427	0.403	0.398	0.396
	(0.103)	(0.103)	(0.095)	(0.097)
noi1	1.161	1.152	1.111	1.122
	(0.119)	(0.119)	(0.115)	(0.117)
nur1	0.309	0.303	0.311	0.313
	(0.104)	(0.101)	(0.089)	(0.090)
tree1	0.329	0.320	0.330	0.317
	(0.101)	(0.099)	(0.091)	(0.093)
SD				
n_tc	1.080		1.063	
	(0.097)		(0.090)	
n_tca		1.082		1.086
		(0.082)		(0.099)
air1	1.840	1.814	1.772	1.791
	(0.185)	(0.180)	(0.184)	(0.183)
air2	1.171	1.166	1.160	1.154
	(0.157)	(0.154)	(0.155)	(0.156)
fac2	0.705	0.714	0.702	0.703
	(0.120)	(0.120)	(0.119)	(0.116)
noi1	0.854	0.874	0.850	0.839
	(0.129)	(0.131)	(0.137)	(0.127)
nur1	0.592	0.595	0.532	0.531
	(0.134)	(0.131)	(0.140)	(0.145)
tree1	0.479	0.523	0.234	0.353
	(0.153)	(0.150)	(0.100)	(0.221)
N	7680	7680	7680	7680
Log-likelihood	-1763.3	-1765.93	-1770.97	-1773.52
AIC	3606.608	3611.854	3593.935	3601.032
BIC	3884.463	3889.709	3774.541	3788.584

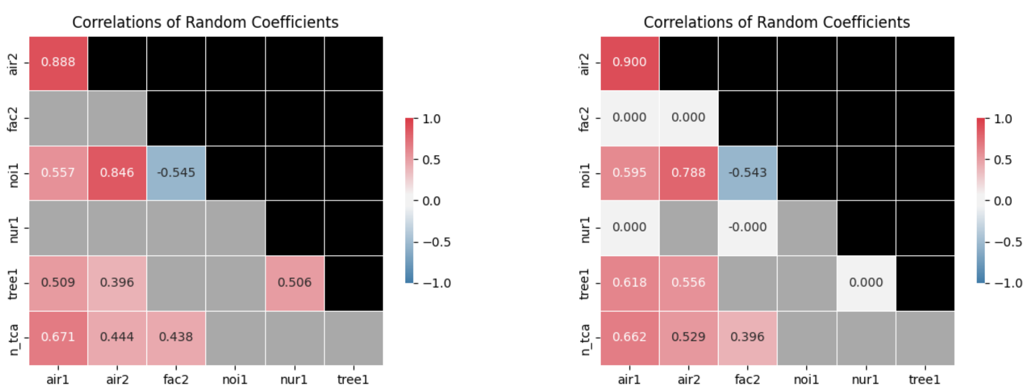
Standard errors in parentheses
* p<0.10, ** p<0.05, ‘ ’ p<0.01

Table 68: MXL models with correlations (Travel costs)



(a) MXL with correlations (b) Constraints on Cholesky decomposition

Figure 58: Correlations of random coefficients (Cost coefficient: TC)



(a) MXL with correlations (b) Constraints on Cholesky decomposition

Figure 59: Correlations of random coefficients (Cost coefficient: TCA)

associated with an enhancement in tree species and ecosystems at the first level, with correlation values of 0.606 and 0.559, respectively. This suggests that people who prioritize better air quality also tend to prefer a greater diversity of tree species and ecosystems at the green site.

Additionally, a positive correlation is observed between the random coefficient for *air quality* and the random coefficient for *n_dis*. This suggests that individuals who value improved air quality also tend to be less willing to travel longer distances to reach the green site. However, this negative correlation effect is weaker among those who prefer the second level of air quality improvement.

The improvement of green site facilities at the second level is moderately and positively correlated to *n_dis*. This indicates that respondents prefer to have more and better facilities at the green site but at the same time, they are not willing to travel long distances. The coefficients for *fac2* are uncorrelated with other attributes. The coefficients for *tree1* are positively correlated with the coefficients for *noi1* and *nur1*, with correlation values of 0.411 and 0.663 respectively. This indicates that people who prefer to see more variety of tree species tend to enjoy more at a quieter green site. Moreover, people who like a higher variety of tree species also like an intermediate level of nursery habitat maintenance.

Figure 57 (b) shows the correlations of random coefficients for Model 2. Although more pairs of random coefficients are found to be statistically significant and correlated (19 out of 21), eight pairs of correlated coefficients have values that are close to zero. The coefficients for both *air1* and *air2* are strongly correlated with the coefficients for *noi1*, *tree1* and *n_dis*. Similar to Figure 57 (a), the coefficients for *fac2* are negatively correlated with coefficients for *n_dis*, and the coefficient for *tree1* is positively correlated with *noi1* and *nur1*. The coefficient for *nur1* is uncorrelated with that for *n_dis* in Figure 57 (a) but is weakly correlated with *n_dis* in Figure 57 (b), with a correlation value of 0.261.

F.3.1 Mixed logit comparison if individual travel costs were used

Table 68 presents the estimation results for models with cost coefficients represented as travel costs (TC and TCA) instead of *distance*. Models 1 and 2 correspond to MXL models with correlations, while Models 3 and 4 are MXL models with correlations that constrain the insignificant Cholesky factors to be close to zero.

The models with correlated coefficients demonstrate a better fit to the data, as evidenced by higher log-likelihood values (found towards the bottom of Table 68) when compared to the log-likelihood values of -1830.31 and -1833.08 obtained from uncorrelated MXL models that use *TC* and *TCA* as cost variables (Models 3 and 4 in Table 66). However, the inclusion of the value of travel time does not lead to an improvement in model fit, as indicated by lower log-likelihood values and higher AIC and BIC values. Therefore, there is no strong evidence to support the notion that the value of travel time should be considered as part of travel costs in this context. Nevertheless, it is important to note that the mean and SD of the coefficient for *n_tca* are significant at the 1% confidence level. This suggests that, in addition to vehicle operating costs, the value of travel time should not be neglected when calculating travel costs.

Figure 58 displays the heat plot for correlations of random coefficients as estimated by Models 1 and 3 in Table 68, while Figure 59 shows the correlations of random coefficients as estimated by Models 2 and 4 in Table 68. In all of these figures, several patterns and correlations among random coefficients are evident. Firstly, there is a negative correlation between the coefficients for *noi1* and *fac2*, indicating that individuals who prefer the improvement of noise levels at the first level are less likely to prefer the improvement of facilities at the second level. Secondly, it shows that individuals who prefer the improvement of air quality at either level (*air1* or *air2*) are more likely to prefer the reduction of noise levels at the first level (*noi1*) and the improvement in tree species and ecosystems at the first level (*tree1*). Thirdly, there is a consistent negative correlation with travel costs. As expected, the coefficients representing *air1*, *air2*, and *fac2* are negatively correlated with the travel cost coefficients. This suggests that individuals are less willing to pay higher travel costs for better air quality or improved facilities. the positive correlation between the coefficients for *tree1* and *noi1* in Figure 58(b), which is not evident in other models, highlights the variability in correlation patterns among coefficients across different models. This variability suggests that the relationships between certain attributes may not be consistent across all contexts.

F.4 WTP estimation for models in preference space

Table 69 provides a comparison of the mean and SD of WTP estimates across models.

The significant variation in mean WTP estimates across different models shows the importance of carefully considering various model specifications to ensure the robustness and accuracy of the WTP estimates. Interpreting WTP estimates can be challenging, especially when multiple model specifications are involved.

F.4.1 The effect of travel time value on WTP

Table 70 provides an overview of the marginal WTP (mWTP) estimates derived from different model specifications when travel cost variables are used. The results illustrate how incorporating the value of travel time into travel cost calculations can significantly influence mWTP estimates in both MNL and MXL models. Before including the value of travel time, the mWTP estimates tend to be lower. However, after accounting for the value of travel time, the mWTP estimates experience a substantial increase. This emphasizes the importance of considering the value of travel time as it has a significant impact on mWTP estimates. The magnitude of change in mWTP estimates across different attribute levels varies across models, providing an understanding of the trade-offs that individuals make between the travel time value and other attribute levels.

The mWTP estimates exhibit an increase ranging from 64.38% to 71.3% across different attribute levels when the value of travel time is incorporated into travel cost calculations, as demonstrated by the MNL models. Notably, *nur2* experiences the most substantial percentage increase at 71.03%, while *fac1* records the smallest increase at 64.38%. On average, mean of the mWTP estimates experience increments ranging from 59.33% to 84.47% across various attribute levels in the MXL models. Simultaneously, the SDs of mWTP estimates reflect an average increase ranging from 20.43% to 104.45%. This heightened SD implies that individual mWTP estimates display greater diversity and are less consistent around the mean.

When accounting for the inclusion of travel time cost, the mean of the marginal mWTP for *air1* experiences the most substantial percentage increase, standing at 69.77%, as evident in MXL (Model 3). Furthermore, the SD of the mWTP for *air1* undergoes the most significant percentage increase, reaching 104.45%, as observed in the MXL correlations (Constrained) model. The MXL correlations (Constrained) model records the highest percentage increase in both the mean and SD of the mWTP for *air2* when travel time cost is considered, showcasing increases of 84.47% and 82.78%, respectively.

The mWTP estimation results reveal that when incorporating the travel time cost into the MXL model, the most significant percentage increase of 71.01% in mean of the mWTP for *fac1* is observed, as indicated by the MXL correlations (Constrained) model. Simultaneously, the highest percentage increase in the SD of WTP, reaching 78.57%, is observed in MXL (Model 2). Furthermore, the mean of the mWTP for *fac2* undergoes an average increase of approximately 70% across all MXL models, while the SD of the mWTP for this attribute level experiences the most substantial increase of 78.20% in MXL (Model 3).

In the case of *noi1*, there is a significant percentage increase of 74.80% from the

MXL correlations (Constrained) model, accompanied by a 31.1% increase in the SD of the mWTP within the same model. Similarly, the MXL correlations (Constrained) model registers the highest percentage increase in mean of the mWTP for *noi2* at 75.93% across all models, while MXL (Model 2) exhibits the highest percentage increase in the SD of the mWTP at 80.42% across all models.

The mean of the mWTP for *nur1* demonstrates an average increase ranging from 59.33% to 69.76%, with the highest increase observed in MXL (Model 3). In contrast, the SD of the mWTP for *nur1* shows an increase ranging from 27.22% to 77.68%. As for *nur2*, its mean of the mWTP experiences the most significant increase of 75.25% in MXL (Model 1), surpassing all other models. Concurrently, the SD of the mWTP for *nur2* experiences the highest increase of 81.76% across all models, as evident in MXL (Model 2).

When examining *tree1*, the most significant percentage increase in both mean and SD of the mWTP, at 75.94% and 99.14% respectively, is observed in the MXL correlations (Constrained) model. A substantial 77.52% increase in mean of the mWTP for *tree2* is also reported by the MXL correlations (Constrained) model, while MXL (Model 2) records the highest percentage increase in the SD of the mWTP for *tree2* at 79.44%.

In conclusion, incorporating the value of travel time into travel costs proves to be a valuable tool for understanding variations in preferences and mWTP estimates.

	Means (standard deviations) of WTP (Malaysian Ringgit)				
	Estimation models				
Attributes	MNL	MXL (Model 1)	MXL (Model 2)	MXL (Model 3)	MXL correlations (constrained)
air1	14.32	15.40	36.91	32.82	26.70
		(11.25)	(55.57)	(38.82)	(38.04)
air2	24.23	10.77	22.78	22.30	21.25
		(6.79)	(33.36)	(26.38)	(27.49)
fac1	2.69	2.96	6.59	6.25	7.13
		(0.00)	(7.82)	(7.39)	(10.54)
fac2	5.59	3.12	6.06	6.28	0.77
		(2.85)	(15.48)	(7.42)	(21.62)
noi1	6.41	6.32	14.57	13.49	15.36
		(4.41)	(27.46)	(15.96)	(26.95)
noi2	8.68	1.93	5.07	5.97	5.41
		(0.00)	(6.01)	(7.06)	(7.99)
nur1	1.83	2.16	4.67	3.87	6.40
		(3.54)	(11.36)	(4.58)	(20.16)
nur2	2.99	0.96	1.82	2.03	3.15
		(0.00)	(2.15)	(2.39)	(4.65)
tree1	1.34	1.83	3.25	2.89	4.30
		(3.43)	(3.85)	(3.41)	(15.77)
tree2	3.61	2.20	4.87	4.76	6.03
		(0.00)	(5.78)	(5.63)	(8.91)

Table 69: WTP estimations from different models

	Means (standard deviations) of WTP (Ringgit Malaysia)									
	Estimation models									
	MNL		MXL (Model 1)		MXL (Model 2)		MXL (Model 3)		MXL Correlations (Constraints)	
	Cost variables									
Attributes	*TC*	*TCA*	*TC*	*TCA*	*TC*	*TCA*	*TC*	*TCA*	*TC*	*TCA*
air1	14.19	23.74	15.16	25.18	31.26	52.88	28.22	47.91	21.52	36.17
			(10.63)	(17.54)	(40.82)	(72.22)	(29.13)	(51.84)	(30.53)	(62.42)
air2	23.94	40.00	10.67	17.77	19.56	33.27	18.95	32.11	16.87	31.12
			(7.00)	(11.98)	(25.69)	(45.86)	(19.56)	(34.75)	(23.40)	(42.77)
fac1	2.58	4.25	2.86	4.67	5.54	9.33	5.18	8.74	6.21	10.62
			(0.00)	(0.00)	(5.60)	(10.00)	(5.34)	(9.46)	(9.16)	(14.16)
fac2	5.26	8.79	2.91	4.92	4.96	8.47	5.07	8.62	-0.13	1.10
			(3.53)	(5.94)	(12.49)	(22.10)	(5.23)	(9.32)	(21.54)	(25.94)
noi1	6.34	10.62	6.26	10.42	12.42	21.14	11.53	19.58	14.88	26.01
			(4.36)	(7.37)	(20.18)	(36.13)	(11.90)	(21.19)	(31.00)	(40.66)
noi2	8.53	14.31	1.86	3.09	4.19	7.14	5.00	8.51	4.03	7.09
			(0.00)	(0.00)	(4.24)	(7.65)	(5.17)	(9.21)	(5.93)	(9.45)
nur1	1.81	3.05	2.16	3.61	4.11	6.97	3.34	5.67	5.95	9.48
			(3.73)	(6.28)	(9.38)	(16.45)	(3.45)	(6.13)	(20.46)	(26.03)
nur2	3.01	5.15	1.01	1.77	1.68	2.89	1.92	3.29	2.48	4.31
			(0.00)	(0.00)	(1.70)	(3.09)	(1.98)	(3.56)	(3.66)	(5.75)
tree1	1.03	1.75	1.59	2.66	2.77	4.62	2.26	3.82	1.87	3.29
			(3.40)	(5.81)	(2.81)	(4.95)	(2.34)	(4.13)	(9.27)	(18.46)
tree2	3.59	5.95	2.48	4.08	4.56	7.74	4.51	7.58	5.25	9.32
			(0.00)	(0.00)	(4.62)	(8.29)	(4.66)	(8.20)	(7.73)	(12.43)

Table 70: WTP estimations from different models (*TC* and *TCA*)

Appendix G

Correlation Analysis for Variables in Seemingly Unrelated Regression Model Analysis

	frst	recr	comm	agri	inds	age	educ	employed	household	visit freq
frst	1.000									
recr	-0.180	1.000								
	0.001									
comm	-0.085	0.073	1.000							
	0.129	0.195								
agri	0.213	-0.274	-0.074	1.000						
	0.000	0.000	0.189							
inds	-0.258	-0.102	0.313	-0.142	1.000					
	0.000	0.069	0.000	0.012						
age	-0.072	0.155	0.020	-0.048	-0.018	1.000				
	0.204	0.006	0.723	0.398	0.756					
educ	0.056	0.056	-0.057	0.032	-0.043	-0.496	1.000			
	0.318	0.320	0.316	0.576	0.448	0.000				
employed	0.068	-0.074	-0.012	-0.087	-0.067	0.117	0.125	1.000		
	0.230	0.187	0.838	0.120	0.232	0.021	0.014			
household	-0.029	-0.137	-0.017	0.031	0.041	-0.188	-0.100	-0.143	1.000	
	0.607	0.015	0.765	0.586	0.463	0.000	0.049	0.005		
visit freq	-0.058	0.157	0.051	-0.027	-0.051	0.243	-0.047	0.046	-0.055	1.000
	0.303	0.005	0.368	0.636	0.364	0.000	0.361	0.362	0.281	

Table 71: Pearson's Correlation Analysis (First stage)

Appendix H

Results of Seemingly Unrelated Regression Model Analysis in R

H.1 SUR model estimation in R

```
> # Seemingly Unrelated Regressions in R
> # Copyright 2013 by Ani Katchova
>
> # Clear memory
> rm(list = ls())
>
> # install.packages("systemfit")
> library(systemfit)
>
> mydata <- read.csv(" ")
> attach(mydata)
>
> # Defining variables
> Y1 <- air1_mean
> Y2 <- air2_mean
> Y3 <- fac1_mean
> Y4 <- fac2_mean
> Y5 <- noi1_mean
> Y6 <- noi2_mean
> Y7 <- nur1_mean
> Y8 <- nur2_mean
```

```
> Y9 <- tree1_mean
> Y10 <- tree2_mean
> X1 <- cbind(frst, recr, comm, agri, inds, educ, employed, household, visit_freq)
> eq1 <- Y1 ~ X1
> eq2 <- Y2 ~ X1
> eq3 <- Y3 ~ X1
> eq4 <- Y4 ~ X1
> eq5 <- Y5 ~ X1
> eq6 <- Y6 ~ X1
> eq7 <- Y7 ~ X1
> eq8 <- Y8 ~ X1
> eq9 <- Y9 ~ X1
> eq10 <- Y10 ~ X1
> system <- list (eq1 = eq1, eq2 = eq2, eq3 = eq3, eq4 = eq4, eq5 = eq5, eq6 = eq6, e
>
> # SUR
> sur <- systemfit(system, method = "SUR", data = mydata)
> summary(sur)

systemfit results
method: SUR

N DF SSR detRCov OLS-R2 McElroy-R2
system 3160 3060 9121.6 0.00044 0.033331 0.027208

N DF SSR MSE RMSE R2 Adj R2
eq1 316 306 3985.246305 13.023681 3.608834 0.035531 0.007165
eq2 316 306 2170.139426 7.091959 2.663073 0.039990 0.011754
eq3 316 306 89.560515 0.292681 0.541000 0.019552 -0.009285
eq4 316 306 743.621360 2.430135 1.558889 0.021875 -0.006893
eq5 316 306 86.363831 0.282235 0.531258 0.021202 -0.007586
eq6 316 306 278.345196 0.909625 0.953743 0.027843 -0.000750
eq7 316 306 396.429622 1.295522 1.138210 0.037252 0.008936
eq8 316 306 0.852708 0.002787 0.052789 0.006486 -0.022735
eq9 316 306 1369.488638 4.475453 2.115527 0.023979 -0.004728
eq10 316 306 1.556091 0.005085 0.071311 0.031550 0.003066

The covariance matrix of the residuals used for estimation
```

```
eq1 eq2 eq3 eq4 eq5 eq6 eq7 eq8
eq1 13.0236807 3.44231753 0.25122657 -0.25560607 -0.07147908 0.14394476 0.12903698 -0
eq2 3.4423175 7.09195891 0.06763553 -0.45973740 0.08771501 0.35035100 0.02188495 -0.0
eq3 0.2512266 0.06763553 0.29268142 0.35961723 -0.03496761 0.00368653 0.05891849 -0.0
eq4 -0.2556061 -0.45973740 0.35961723 2.43013516 -0.16632449 -0.09059697 -0.11165617
eq5 -0.0714791 0.08771501 -0.03496761 -0.16632449 0.28223474 0.25002131 -0.08720150 0
eq6 0.1439448 0.35035100 0.00368653 -0.09059697 0.25002131 0.90962482 -0.06528149 0.0
eq7 0.1290370 0.02188495 0.05891849 -0.11165617 -0.08720150 -0.06528149 1.29552164 0.
eq8 -0.0114108 -0.00797577 -0.00156414 -0.00479496 0.00337985 0.00435289 0.02177282 0
eq9 -0.1461294 0.16650734 0.06722031 0.17326869 0.01860896 0.33638280 0.12304742 0.01
eq10 0.0540374 0.02354193 -0.00187970 0.00405485 -0.00733586 -0.00638942 0.00147831 -
eq9 eq10
eq1 -0.1461294 0.054037404
eq2 0.1665073 0.023541931
eq3 0.0672203 -0.001879703
eq4 0.1732687 0.004054851
eq5 0.0186090 -0.007335863
eq6 0.3363828 -0.006389423
eq7 0.1230474 0.001478308
eq8 0.0120494 -0.000621633
eq9 4.4754531 -0.011030401
eq10 -0.0110304 0.005085266

The covariance matrix of the residuals
eq1 eq2 eq3 eq4 eq5 eq6 eq7 eq8
eq1 13.0236807 3.44231753 0.25122657 -0.25560607 -0.07147908 0.14394476 0.12903698 -0
eq2 3.4423175 7.09195891 0.06763553 -0.45973740 0.08771501 0.35035100 0.02188495 -0.0
eq3 0.2512266 0.06763553 0.29268142 0.35961723 -0.03496761 0.00368653 0.05891849 -0.0
eq4 -0.2556061 -0.45973740 0.35961723 2.43013516 -0.16632449 -0.09059697 -0.11165617
eq5 -0.0714791 0.08771501 -0.03496761 -0.16632449 0.28223474 0.25002131 -0.08720150 0
eq6 0.1439448 0.35035100 0.00368653 -0.09059697 0.25002131 0.90962482 -0.06528149 0.0
eq7 0.1290370 0.02188495 0.05891849 -0.11165617 -0.08720150 -0.06528149 1.29552164 0.
eq8 -0.0114108 -0.00797577 -0.00156414 -0.00479496 0.00337985 0.00435289 0.02177282 0
eq9 -0.1461294 0.16650734 0.06722031 0.17326869 0.01860896 0.33638280 0.12304742 0.01
eq10 0.0540374 0.02354193 -0.00187970 0.00405485 -0.00733586 -0.00638942 0.00147831 -
eq9 eq10
eq1 -0.1461294 0.054037404
eq2 0.1665073 0.023541931
```

```
eq3 0.0672203 -0.001879703
eq4 0.1732687 0.004054851
eq5 0.0186090 -0.007335863
eq6 0.3363828 -0.006389423
eq7 0.1230474 0.001478308
eq8 0.0120494 -0.000621633
eq9 4.4754531 -0.011030401
eq10 -0.0110304 0.005085266

The correlations of the residuals
eq1 eq2 eq3 eq4 eq5 eq6 eq7 eq8 eq9
eq1 1.0000000 0.35817969 0.12867706 -0.0454348 -0.0372827 0.04182133 0.03141412 -0.05
eq2 0.3581797 1.00000000 0.04694553 -0.1107418 0.0619991 0.13793965 0.00722005 -0.056
eq3 0.1286771 0.04694553 1.00000000 0.4264103 -0.1216643 0.00714478 0.09568232 -0.054
eq4 -0.0454348 -0.11074178 0.42641031 1.0000000 -0.2008333 -0.06093507 -0.06292818 -0
eq5 -0.0372827 0.06199913 -0.12166429 -0.2008333 1.0000000 0.49344714 -0.14421033 0.1
eq6 0.0418213 0.13793965 0.00714478 -0.0609351 0.4934471 1.00000000 -0.06013628 0.086
eq7 0.0314141 0.00722005 0.09568232 -0.0629282 -0.1442103 -0.06013628 1.00000000 0.36
eq8 -0.0598979 -0.05673490 -0.05476963 -0.0582681 0.1205184 0.08645836 0.36237055 1.0
eq9 -0.0191405 0.02955505 0.05873330 0.0525396 0.0165576 0.16671863 0.05110127 0.1078
eq10 0.2099765 0.12396588 -0.04872309 0.0364756 -0.1936374 -0.09394499 0.01821318 -0.
eq10
eq1 0.2099765
eq2 0.1239659
eq3 -0.0487231
eq4 0.0364756
eq5 -0.1936374
eq6 -0.0939450
eq7 0.0182132
eq8 -0.1651345
eq9 -0.0731166
eq10 1.0000000


SUR estimates for 'eq1' (equation 1)
Model Formula: Y1 ~ X1

Estimate Std. Error t value Pr(>|t|)
```

```
(Intercept) 14.7719210 1.0461719 14.11997 < 2e-16 ***
X1frst 0.0381844 0.1009638 0.37820 0.705545
X1recr 1.4359663 1.8615421 0.77139 0.441074
X1comm 0.5529079 0.8696942 0.63575 0.525415
X1agri -0.0285596 0.2537227 -0.11256 0.910451
X1inds -1.0419100 0.6842971 -1.52260 0.128892
X1educ -0.0526906 0.1283923 -0.41039 0.681809
X1employed -0.8940434 0.4242692 -2.10726 0.035909 *
X1household 0.0820706 0.1350056 0.60791 0.543701
X1visit_freq -0.2747246 0.1679551 -1.63570 0.102930
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 3.608834 on 306 degrees of freedom
Number of observations: 316 Degrees of Freedom: 306
SSR: 3985.246305 MSE: 13.023681 Root MSE: 3.608834
Multiple R-Squared: 0.035531 Adjusted R-Squared: 0.007165


SUR estimates for 'eq2' (equation 2)
Model Formula: Y2 ~ X1

Estimate Std. Error t value Pr(>|t|)
(Intercept) 9.5245038 0.7720036 12.33738 < 2e-16 ***
X1frst 0.0797326 0.0745044 1.07017 0.285385
X1recr 1.1846868 1.3736911 0.86241 0.389137
X1comm -0.5514987 0.6417750 -0.85933 0.390829
X1agri -0.0180236 0.1872300 -0.09626 0.923373
X1inds -0.5827545 0.5049646 -1.15405 0.249380
X1educ 0.0598623 0.0947448 0.63183 0.527972
X1employed -0.0640590 0.3130817 -0.20461 0.838015
X1household 0.1025396 0.0996249 1.02926 0.304172
X1visit_freq -0.2908408 0.1239394 -2.34664 0.019581 *
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.663073 on 306 degrees of freedom
Number of observations: 316 Degrees of Freedom: 306
```

```
SSR: 2170.139426 MSE: 7.091959 Root MSE: 2.663073
Multiple R-Squared: 0.03999 Adjusted R-Squared: 0.011754


SUR estimates for 'eq3' (equation 3)
Model Formula: Y3 ~ X1

Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.95177306 0.15683167 18.82128 < 2e-16 ***
X1frst -0.00921633 0.01513549 -0.60892 0.54303
X1recr -0.53522593 0.27906384 -1.91793 0.05605 .
X1comm 0.03066443 0.13037589 0.23520 0.81421
X1agri -0.02184977 0.03803557 -0.57446 0.56608
X1inds -0.05073796 0.10258300 -0.49460 0.62123
X1educ 0.00429885 0.01924730 0.22335 0.82341
X1employed -0.07975554 0.06360221 -1.25397 0.21081
X1household 0.00255447 0.02023869 0.12622 0.89964
X1visit_freq 0.03127978 0.02517816 1.24234 0.21506
---
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.541 on 306 degrees of freedom
Number of observations: 316 Degrees of Freedom: 306
SSR: 89.560515 MSE: 0.292681 Root MSE: 0.541
Multiple R-Squared: 0.019552 Adjusted R-Squared: -0.009285


SUR estimates for 'eq4' (equation 4)
Model Formula: Y4 ~ X1

Estimate Std. Error t value Pr(>|t|)
(Intercept) 2.7028429 0.4519094 5.98094 6.1839e-09 ***
X1frst -0.0196551 0.0436128 -0.45067 0.65254
X1recr -0.0212958 0.8041206 -0.02648 0.97889
X1comm -0.4478414 0.3756773 -1.19209 0.23415
X1agri 0.1494395 0.1095993 1.36351 0.17372
X1inds 0.2509217 0.2955922 0.84888 0.39661
X1educ -0.0772448 0.0554610 -1.39278 0.16470
```

```
X1employed 0.1222398 0.1832694 0.66700 0.50528
X1household 0.0477569 0.0583177 0.81891 0.41348
X1visit_freq 0.0258493 0.0725507 0.35629 0.72187
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.558889 on 306 degrees of freedom
Number of observations: 316 Degrees of Freedom: 306
SSR: 743.62136 MSE: 2.430135 Root MSE: 1.558889
Multiple R-Squared: 0.021875 Adjusted R-Squared: -0.006893


SUR estimates for 'eq5' (equation 5)
Model Formula: Y5 ~ X1

Estimate Std. Error t value Pr(>|t|)
(Intercept) 5.274029607 0.154007346 34.24531 < 2e-16 ***
X1frst 0.000404339 0.014862915 0.02720 0.97831
X1recr -0.121883372 0.274038269 -0.44477 0.65680
X1comm 0.101782017 0.128027993 0.79500 0.42723
X1agri -0.040953104 0.037350602 -1.09645 0.27374
X1inds -0.095335603 0.100735613 -0.94639 0.34469
X1educ 0.008826914 0.018900678 0.46702 0.64082
X1employed 0.059723344 0.062456820 0.95623 0.33971
X1household 0.011749221 0.019874218 0.59118 0.55484
X1visit_freq -0.039810744 0.024724734 -1.61016 0.10839
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.531258 on 306 degrees of freedom
Number of observations: 316 Degrees of Freedom: 306
SSR: 86.363831 MSE: 0.282235 Root MSE: 0.531258
Multiple R-Squared: 0.021202 Adjusted R-Squared: -0.007586


SUR estimates for 'eq6' (equation 6)
Model Formula: Y6 ~ X1
```

```
Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.92126583 0.27648231 6.94896 2.2222e-11 ***
X1frst 0.03438924 0.02668271 1.28882 0.19843
X1recr 0.11673945 0.49196831 0.23729 0.81259
X1comm -0.17616013 0.22984277 -0.76644 0.44401
X1agri 0.01074438 0.06705382 0.16024 0.87280
X1inds -0.21740666 0.18084602 -1.20216 0.23023
X1educ -0.00330658 0.03393152 -0.09745 0.92243
X1employed 0.12934545 0.11212586 1.15357 0.24958
X1household -0.01159987 0.03567927 -0.32512 0.74532
X1visit_freq -0.02969254 0.04438718 -0.66894 0.50404
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.953743 on 306 degrees of freedom
Number of observations: 316 Degrees of Freedom: 306
SSR: 278.345196 MSE: 0.909625 Root MSE: 0.953743
Multiple R-Squared: 0.027843 Adjusted R-Squared: -0.00075


SUR estimates for 'eq7' (equation 7)
Model Formula: Y7 ~ X1

Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.10784907 0.32995790 3.35755 0.00088572 ***
X1frst 0.00490034 0.03184352 0.15389 0.87779942
X1recr -0.29481963 0.58712195 -0.50214 0.61592753
X1comm 0.15440851 0.27429762 0.56292 0.57389935
X1agri 0.08973422 0.08002298 1.12136 0.26301569
X1inds 0.24693905 0.21582420 1.14417 0.25344837
X1educ 0.09873390 0.04049435 2.43821 0.01532880 *
X1employed -0.00813824 0.13381259 -0.06082 0.95154370
X1household -0.01408795 0.04258015 -0.33086 0.74097883
X1visit_freq -0.05480322 0.05297229 -1.03456 0.30168948
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.13821 on 306 degrees of freedom
```

Number of observations: 316 Degrees of Freedom: 306
SSR: 396.429622 MSE: 1.295522 Root MSE: 1.13821
Multiple R-Squared: 0.037252 Adjusted R-Squared: 0.008936

SUR estimates for 'eq8' (equation 8)
Model Formula: Y8 ~ X1

```
Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.081687487 0.015302966 70.68483 < 2e-16 ***
X1frst 0.000459970 0.001476856 0.31145 0.75567
X1recr -0.018966169 0.027229859 -0.69652 0.48663
X1comm 0.008922750 0.012721523 0.70139 0.48359
X1agri 0.000964525 0.003711349 0.25989 0.79513
X1inds -0.001166593 0.010009611 -0.11655 0.90730
X1educ 0.000492735 0.001878069 0.26236 0.79322
X1employed 0.002643795 0.006206033 0.42600 0.67040
X1household -0.000946927 0.001974805 -0.47950 0.63192
X1visit_freq 0.000853540 0.002456777 0.34742 0.72851
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Residual standard error: 0.052789 on 306 degrees of freedom
Number of observations: 316 Degrees of Freedom: 306
SSR: 0.852708 MSE: 0.002787 Root MSE: 0.052789
Multiple R-Squared: 0.006486 Adjusted R-Squared: -0.022735

SUR estimates for 'eq9' (equation 9)
Model Formula: Y9 ~ X1

```
Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.9059500 0.6132742 3.10783 0.0020617 **
X1frst -0.0874616 0.0591858 -1.47775 0.1405039
X1recr 0.4026907 1.0912506 0.36902 0.7123698
X1comm -1.0432085 0.5098216 -2.04622 0.0415889 *
X1agri 0.0942215 0.1487342 0.63349 0.5268873
X1inds 0.1090232 0.4011403 0.27178 0.7859719
```

```
X1educ -0.0248803 0.0752646 -0.33057 0.7411945
X1employed 0.0168099 0.2487099 0.06759 0.9461573
X1household 0.0438925 0.0791413 0.55461 0.5795675
X1visit_freq -0.0428209 0.0984566 -0.43492 0.6639257
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 2.115527 on 306 degrees of freedom
Number of observations: 316 Degrees of Freedom: 306
SSR: 1369.488638 MSE: 4.475453 Root MSE: 2.115527
Multiple R-Squared: 0.023979 Adjusted R-Squared: -0.004728


SUR estimates for ’eq10’ (equation 10)
Model Formula: Y10 ~ X1

Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.802256595 0.020672501 87.18135 < 2e-16 ***
X1frst -0.000339178 0.001995058 -0.17001 0.865115
X1recr 0.052785317 0.036784327 1.43499 0.152310
X1comm 0.021438958 0.017185277 1.24752 0.213161
X1agri 0.000505303 0.005013594 0.10079 0.919786
X1inds -0.010871169 0.013521804 -0.80397 0.422037
X1educ 0.001136578 0.002537050 0.44799 0.654476
X1employed 0.000735446 0.008383618 0.08772 0.930153
X1household 0.003527842 0.002667729 1.32241 0.187018
X1visit_freq -0.007546399 0.003318816 -2.27382 0.023669 *
---
Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.071311 on 306 degrees of freedom
Number of observations: 316 Degrees of Freedom: 306
SSR: 1.556091 MSE: 0.005085 Root MSE: 0.071311
Multiple R-Squared: 0.03155 Adjusted R-Squared: 0.003066
```

H.2 Spatial model analysis in Stata

```
. set more off

. import excel " ",sheet("Sheet1") first

.
. global socio age educ employed household visit_freq
. global landuse frst recr comm agri inds
. global ylist1 air1_mean air2_mean fac1_mean fac2_mean noi1_mean noi2_mean nu
> r1_mean nur2_mean tree1_mean tree2_mean
. global xlist1 frst recr comm agri inds educ employed household visit_freq

. gen lat_new = lat if !missing(lat)

. gen lon_new = lon if !missing(lon)

.
. *Spatial weight matrix (matrix list W)
. spatwmat, name(W) xcoord(lat_new) ycoord(lon_new) band(0 16) standardize eig
> enval(E)


The following matrices have been created:

1. Inverse distance weights matrix W (row-standardized)
   Dimension: 312x312
   Distance band: 0 < d <= 16
   Friction parameter: 1
   Minimum distance: 0.0
   1st quartile distance: 0.0
   Median distance: 0.1
   3rd quartile distance: 0.1
   Maximum distance: 0.2
   Largest minimum distance: 0.04
   Smallest maximum distance: 0.10

2. Eigenvalues matrix E
```

```
   Dimension: 312x1
.
. *Spatial error model
. spatreg air1_mean \$xlist1, weights(W) eigenval(E) model(error)

initial:       log likelihood = -838.34359
rescale:       log likelihood = -838.34359
rescale eq:    log likelihood = -838.34359
Iteration 0:   log likelihood = -838.34359
Iteration 1:   log likelihood = -837.66331
Iteration 2:   log likelihood = -837.66191
Iteration 3:   log likelihood = -837.66191


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial error model                                   Number of obs   =        312
                                                      Variance ratio  =      0.033
                                                      Squared corr.   =      0.033
Log likelihood = -837.66191                           Sigma           =       3.54

------------------------------------------------------------------------------
   air1_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
air1_mean    |
        frst |   .0000957   .0001412     0.68   0.498     -.000181    .0003724
        recr |   .0022661   .0026014     0.87   0.384    -.0028327    .0073648
        comm |   .0005001   .0012511     0.40   0.689    -.0019521    .0029522
        agri |  -.0000628   .0003612    -0.17   0.862    -.0007707     .000645
        inds |  -.0014064   .0009971    -1.41   0.158    -.0033606    .0005478
        educ |  -.0642738   .1269726    -0.51   0.613    -.3131355    .1845879
    employed |  -.8567027    .414908    -2.06   0.039    -1.669907   -.0434981
   household |   .0799662   .1343029     0.60   0.552    -.1832627    .3431951
```

```
  visit_freq |  -.2592166    .164742    -1.57   0.116    -.5821049    .0636717
       _cons |   14.74857   1.002695    14.71   0.000     12.78332    16.71381
-------------+----------------------------------------------------------------
      lambda |  -.2176984   .1986208    -1.10   0.273     -.606988    .1715913
------------------------------------------------------------------------------
Wald test of lambda=0:                   chi2(1) =   1.201 (0.273)
Likelihood ratio test of lambda=0:       chi2(1) =   1.200 (0.273)
Lagrange multiplier test of lambda=0:   chi2(1) =   1.233 (0.267)

Acceptable range for lambda: -1.280 < lambda < 1.000



. spatreg air2_mean \$xlist1, weights(W) eigenval(E) model(error)

initial:       log likelihood = -742.16672
rescale:       log likelihood = -742.16672
rescale eq:    log likelihood = -742.16672
Iteration 0:   log likelihood = -742.16672
Iteration 1:   log likelihood = -741.24237
Iteration 2:   log likelihood =  -741.2395
Iteration 3:   log likelihood =  -741.2395


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial error model                               Number of obs   =        312
                                                  Variance ratio  =      0.041
                                                  Squared corr.   =      0.038
Log likelihood = -741.2395                        Sigma           =       2.60

------------------------------------------------------------------------------
   air2_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
air2_mean    |
        frst |   .0001484   .0001018     1.46   0.145    -.0000512     .000348
        recr |   .0018749   .0018794     1.00   0.318    -.0018086    .0055584
        comm |  -.0010284   .0009097    -1.13   0.258    -.0028113    .0007546
        agri |  -.0000644   .0002617    -0.25   0.806    -.0005773    .0004486
        inds |  -.0009045   .0007263    -1.25   0.213     -.002328     .000519
        educ |   .0296222   .0945952     0.31   0.754    -.1557809    .2150253
    employed |  -.0587364   .3034164    -0.19   0.847    -.6534216    .5359488
   household |   .0930029   .0975852     0.95   0.341    -.0982605    .2842663
  visit_freq |  -.2713972   .1207743    -2.25   0.025    -.5081104   -.0346839
       _cons |   9.644576   .7301109    13.21   0.000     8.213585    11.07557
-------------+----------------------------------------------------------------
      lambda |  -.2651699   .2034054    -1.30   0.192    -.6638372    .1334973
------------------------------------------------------------------------------
Wald test of lambda=0:                chi2(1) =   1.700 (0.192)
Likelihood ratio test of lambda=0:    chi2(1) =   1.691 (0.194)
Lagrange multiplier test of lambda=0: chi2(1) =   1.649 (0.199)

Acceptable range for lambda: -1.280 < lambda < 1.000



. spatreg fac1_mean \$xlist1, weights(W) eigenval(E) model(error)

initial:       log likelihood = -246.15997
rescale:       log likelihood = -246.15997
rescale eq:    log likelihood = -246.15997
Iteration 0:   log likelihood = -246.15997
Iteration 1:   log likelihood = -245.70251
Iteration 2:   log likelihood = -245.70086
Iteration 3:   log likelihood = -245.70086


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
```

```
 Row-standardized: Yes


Spatial error model                                 Number of obs   =       312
                                                    Variance ratio  =     0.020
                                                    Squared corr.   =     0.019
Log likelihood = -245.70086                         Sigma           =      0.53

------------------------------------------------------------------------------
   fac1_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
fac1_mean    |
        frst |  -.0000179   .0000249    -0.72   0.472    -.0000667    .0000309
        recr |  -.0007902   .0004503    -1.75   0.079    -.0016728    .0000924
        comm |   .0000317   .0002142     0.15   0.882    -.0003881    .0004516
        agri |  -.0000281   .0000606    -0.46   0.644    -.0001469    .0000908
        inds |  -.0000828   .0001606    -0.52   0.606    -.0003975    .0002318
        educ |   .0071164    .019143     0.37   0.710    -.0304033     .044636
    employed |  -.0809092   .0632438    -1.28   0.201    -.2048648    .0430463
   household |    .006126   .0204303     0.30   0.764    -.0339167    .0461687
  visit_freq |   .0298219   .0249607     1.19   0.232    -.0191002     .078744
       _cons |    2.93311   .1618359    18.12   0.000     2.615917    3.250302
-------------+----------------------------------------------------------------
      lambda |   .1830265   .2077816     0.88   0.378     -.224218    .5902711
------------------------------------------------------------------------------
Wald test of lambda=0:                 chi2(1) =   0.776 (0.378)
Likelihood ratio test of lambda=0:     chi2(1) =   0.754 (0.385)
Lagrange multiplier test of lambda=0:  chi2(1) =   0.672 (0.412)

Acceptable range for lambda: -1.280 < lambda < 1.000



. spatreg fac2_mean \$xlist1, weights(W) eigenval(E) model(error)

initial:       log likelihood = -577.88741
rescale:       log likelihood = -577.88741
rescale eq:    log likelihood = -577.88741
```

```
Iteration 0:   log likelihood = -577.88741
Iteration 1:   log likelihood = -577.80508
Iteration 2:   log likelihood = -577.80493
Iteration 3:   log likelihood = -577.80493


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial error model                                Number of obs   =       312
                                                   Variance ratio  =     0.021
                                                   Squared corr.   =     0.021
Log likelihood = -577.80493                        Sigma           =      1.54

------------------------------------------------------------------------------
   fac2_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
fac2_mean    |
        frst |  -.0000323    .000067    -0.48   0.629    -.0001637     .000099
        recr |  -.0000276   .0012273    -0.02   0.982     -.002433    .0023777
        comm |  -.0006584   .0005836    -1.13   0.259    -.0018023    .0004854
        agri |   .0002327    .000167     1.39   0.163    -.0000946      .00056
        inds |   .0003658   .0004518     0.81   0.418    -.0005197    .0012513
        educ |  -.0737323   .0552061    -1.34   0.182    -.1819342    .0344697
    employed |   .1223426   .1822548     0.67   0.502    -.2348702    .4795553
   household |   .0494499   .0584591     0.85   0.398    -.0651278    .1640277
 visit_freq |   .0211484   .0720906     0.29   0.769    -.1201466    .1624434
       _cons |    2.69497   .4503931     5.98   0.000     1.812216    3.577725
-------------+----------------------------------------------------------------
      lambda |   .0072136   .2083708     0.03   0.972    -.4011856    .4156128
------------------------------------------------------------------------------
Wald test of lambda=0:                 chi2(1) =   0.001 (0.972)
Likelihood ratio test of lambda=0:     chi2(1) =   0.001 (0.972)
Lagrange multiplier test of lambda=0:  chi2(1) =   0.001 (0.973)
```

```
Acceptable range for lambda: -1.280 < lambda < 1.000


. spatreg noi1_mean \$xlist1, weights(W) eigenval(E) model(error)

initial:       log likelihood = -239.59006
rescale:       log likelihood = -239.59006
rescale eq:    log likelihood = -239.59006
Iteration 0:   log likelihood = -239.59006
Iteration 1:   log likelihood = -238.91899
Iteration 2:   log likelihood = -238.91724
Iteration 3:   log likelihood = -238.91724


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial error model                                   Number of obs   =        312
                                                      Variance ratio  =      0.021
                                                      Squared corr.   =      0.021
Log likelihood = -238.91724                           Sigma           =       0.52

------------------------------------------------------------------------------
   noi1_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
noi1_mean    |
        frst |  -1.20e-06   .0000206    -0.06   0.953    -.0000417    .0000393
        recr |  -.0002256   .0003821    -0.59   0.555    -.0009744    .0005233
        comm |   .0001826   .0001843     0.99   0.322    -.0001786    .0005439
        agri |  -.0000628   .0000528    -1.19   0.234    -.0001662    .0000406
        inds |   -.000181   .0001471    -1.23   0.219    -.0004692    .0001073
        educ |   .0070321   .0184832     0.38   0.704    -.0291944    .0432585
```

```
    employed |   .0468047   .0609625     0.77   0.443    -.0726795     .166289
   household |   .0171878    .019819     0.87   0.386    -.0216567    .0560323
  visit_freq |   -.037198   .0242369    -1.53   0.125    -.0847014    .0103054
       _cons |   5.265237   .1459631    36.07   0.000     4.979154    5.551319
-------------+----------------------------------------------------------------
      lambda |  -.2201667   .2023865    -1.09   0.277    -.6168369    .1765035
------------------------------------------------------------------------------
Wald test of lambda=0:                   chi2(1) =   1.183 (0.277)
Likelihood ratio test of lambda=0:       chi2(1) =   1.182 (0.277)
Lagrange multiplier test of lambda=0:    chi2(1) =   1.171 (0.279)

Acceptable range for lambda: -1.280 < lambda < 1.000



. spatreg noi2_mean \$xlist1, weights(W) eigenval(E) model(error)

initial:       log likelihood = -424.65571
rescale:       log likelihood = -424.65571
rescale eq:    log likelihood = -424.65571
Iteration 0:   log likelihood = -424.65571
Iteration 1:   log likelihood = -421.95067
Iteration 2:   log likelihood = -421.92349
Iteration 3:   log likelihood = -421.92347


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial error model                                 Number of obs   =       312
                                                    Variance ratio  =     0.027
                                                    Squared corr.   =     0.025
Log likelihood = -421.92347                         Sigma           =      0.93
```

```
------------------------------------------------------------------------------
   noi2_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
noi2_mean    |
        frst |   .0000393   .0000337     1.17   0.243    -.0000267    .0001054
        recr |    .000025   .0006318     0.04   0.968    -.0012133    .0012633
        comm |  -.0003003   .0003055    -0.98   0.326     -.000899    .0002984
        agri |    .000026   .0000883     0.29   0.769    -.0001471    .0001991
        inds |  -.0004458    .000254    -1.76   0.079    -.0009437     .000052
        educ |  -.0102786   .0328222    -0.31   0.754     -.074609    .0540518
    employed |   .0863355   .1089377     0.79   0.428    -.1271785    .2998496
   household |  -.0024686   .0347556    -0.07   0.943    -.0705883    .0656511
  visit_freq |  -.0173043   .0432242    -0.40   0.689    -.1020221    .0674136
       _cons |   1.953956    .251281     7.78   0.000     1.461454    2.446458
-------------+----------------------------------------------------------------
      lambda |  -.4810741   .2038444    -2.36   0.018    -.8806018   -.0815464
------------------------------------------------------------------------------
Wald test of lambda=0:                 chi2(1) =   5.570 (0.018)
Likelihood ratio test of lambda=0:     chi2(1) =   5.301 (0.021)
Lagrange multiplier test of lambda=0:  chi2(1) =   4.789 (0.029)

Acceptable range for lambda: -1.280 < lambda < 1.000


. spatreg nur1_mean \$xlist1, weights(W) eigenval(E) model(error)

initial:       log likelihood = -476.55203
rescale:       log likelihood = -476.55203
rescale eq:    log likelihood = -476.55203
Iteration 0:   log likelihood = -476.55203
Iteration 1:   log likelihood = -471.75227
Iteration 2:   log likelihood = -471.64259
Iteration 3:   log likelihood = -471.64235
Iteration 4:   log likelihood = -471.64235


Weights matrix
```

```
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial error model                                   Number of obs   =       312
                                                      Variance ratio  =     0.037
                                                      Squared corr.   =     0.034
Log likelihood = -471.64235                           Sigma           =      1.09

------------------------------------------------------------------------------
   nur1_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
nur1_mean    |
        frst |   .0000249     .000038     0.66   0.511    -.0000494    .0000993
        recr |  -.0006757    .0007169    -0.94   0.346    -.0020807    .0007293
        comm |   .0001114    .0003481     0.32   0.749    -.0005708    .0007936
        agri |   .0000976    .0001011     0.97   0.334    -.0001005    .0002957
        inds |   .0004274     .000289     1.48   0.139     -.000139    .0009938
        educ |   .0851924     .038328     2.22   0.026      .010071    .1603139
    employed |   .0318659     .125353     0.25   0.799    -.2138214    .2775532
   household |  -.0241131    .0403175    -0.60   0.550    -.1031339    .0549077
 visit_freq |  -.0642717    .0502417    -1.28   0.201    -.1627435    .0342002
       _cons |   1.231093    .2908301     4.23   0.000     .6610762    1.801109
-------------+----------------------------------------------------------------
      lambda |  -.5798098    .1793232    -3.23   0.001    -.9312768   -.2283428
------------------------------------------------------------------------------
Wald test of lambda=0:                   chi2(1) =  10.454 (0.001)
Likelihood ratio test of lambda=0:       chi2(1) =   9.656 (0.002)
Lagrange multiplier test of lambda=0:    chi2(1) =  10.645 (0.001)

Acceptable range for lambda: -1.280 < lambda < 1.000


. spatreg nur2_mean \$xlist1, weights(W) eigenval(E) model(error)
```

```
initial:       log likelihood =   477.4689
rescale:       log likelihood =   477.4689
rescale eq:    log likelihood =   477.4689
Iteration 0:   log likelihood =   477.4689
Iteration 1:   log likelihood =  478.29699
Iteration 2:   log likelihood =   478.2991
Iteration 3:   log likelihood =   478.2991


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial error model                                 Number of obs   =        312
                                                    Variance ratio  =      0.010
                                                    Squared corr.   =      0.008
Log likelihood = 478.2991                           Sigma           =       0.05

------------------------------------------------------------------------------
   nur2_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
nur2_mean    |
        frst |   1.20e-06   2.03e-06     0.59   0.555    -2.79e-06    5.19e-06
        recr |  -.0000377   .0000377    -1.00   0.318    -.0001115    .0000362
        comm |   .0000125   .0000181     0.69   0.489    -.0000229     .000048
        agri |   1.07e-06   5.22e-06     0.21   0.837    -9.15e-06    .0000113
        inds |  -5.03e-07   .0000145    -0.03   0.972     -.000029     .000028
        educ |   .0004633   .0018673     0.25   0.804    -.0031965     .004123
    employed |   .0036841   .0061096     0.60   0.546    -.0082904    .0156586
   household |  -.0013674   .0019578    -0.70   0.485    -.0052047    .0024698
 visit_freq |     .000958   .0024224     0.40   0.692    -.0037898    .0057058
       _cons |    1.08351   .0145455    74.49   0.000     1.055002    1.112019
-------------+----------------------------------------------------------------
      lambda |  -.2794809   .2274689    -1.23   0.219    -.7253117    .1663499
------------------------------------------------------------------------------
```

```
Wald test of lambda=0:                  chi2(1) =   1.510 (0.219)
Likelihood ratio test of lambda=0:      chi2(1) =   1.497 (0.221)
Lagrange multiplier test of lambda=0:  chi2(1) =   1.162 (0.281)

Acceptable range for lambda: -1.280 < lambda < 1.000



. spatreg tree1_mean \$xlist1, weights(W) eigenval(E) model(error)

initial:       log likelihood = -671.97874
rescale:       log likelihood = -671.97874
rescale eq:    log likelihood = -671.97874
Iteration 0:   log likelihood = -671.97874
Iteration 1:   log likelihood = -670.44638
Iteration 2:   log likelihood = -670.42196
Iteration 3:   log likelihood = -670.42195


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial error model                                   Number of obs   =       312
                                                      Variance ratio  =     0.024
                                                      Squared corr.   =     0.022
Log likelihood = -670.42195                           Sigma           =      2.07

------------------------------------------------------------------------------
  tree1_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
tree1_mean   |
        frst |  -.0001696    .000105    -1.62   0.106    -.0003754    .0000362
        recr |   .0009497    .001871     0.51   0.612    -.0027175    .0046169
        comm |  -.0012775   .0008853    -1.44   0.149    -.0030127    .0004576
```

```
        agri |   .0001656   .0002472     0.67   0.503    -.0003189    .0006501
        inds |   .0001158   .0006442     0.18   0.857    -.0011469    .0013785
        educ |   .0023671   .0752764     0.03   0.975     -.145172    .1499062
    employed |   .0237618   .2473536     0.10   0.923    -.4610425     .508566
   household |   .0509624   .0791331     0.64   0.520    -.1041357    .2060604
  visit_freq |  -.0468964   .0973465    -0.48   0.630     -.237692    .1438992
       _cons |   1.746772   .6567184     2.66   0.008     .4596275    3.033916
-------------+----------------------------------------------------------------
      lambda |   .3415143   .1912422     1.79   0.074    -.0333136    .7163422
------------------------------------------------------------------------------
Wald test of lambda=0:                 chi2(1) =   3.189 (0.074)
Likelihood ratio test of lambda=0:     chi2(1) =   2.950 (0.086)
Lagrange multiplier test of lambda=0:  chi2(1) =   2.839 (0.092)

Acceptable range for lambda: -1.280 < lambda < 1.000



. spatreg tree2_mean \$xlist1, weights(W) eigenval(E) model(error)

initial:       log likelihood =  384.40205
rescale:       log likelihood =  384.40205
rescale eq:    log likelihood =  384.40205
Iteration 0:   log likelihood =  384.40205
Iteration 1:   log likelihood =  384.69512
Iteration 2:   log likelihood =   384.6958
Iteration 3:   log likelihood =   384.6958


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial error model                                   Number of obs   =        312
                                                      Variance ratio  =      0.030
```

```
                                                Squared corr.   =     0.030
Log likelihood = 384.6958                       Sigma           =      0.07

------------------------------------------------------------------------------
  tree2_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
tree2_mean   |
        frst |  -5.89e-08   3.25e-06    -0.02   0.986    -6.42e-06    6.30e-06
        recr |   .0000792   .0000588     1.35   0.178    -.0000361    .0001944
        comm |    .000033   .0000279     1.18   0.237    -.0000217    .0000877
        agri |   2.71e-07   7.93e-06     0.03   0.973    -.0000153    .0000158
        inds |  -.0000165   .0000212    -0.78   0.437     -.000058    .0000251
        educ |   .0011795    .002544     0.46   0.643    -.0038065    .0061656
    employed |   .0002854    .008369     0.03   0.973    -.0161176    .0166884
   household |   .0036187   .0026967     1.34   0.180    -.0016667     .008904
  visit_freq |  -.0073344   .0033301    -2.20   0.028    -.0138612   -.0008076
       _cons |   1.801203   .0214439    84.00   0.000     1.759174    1.843232
-------------+----------------------------------------------------------------
      lambda |   .1436914   .2185595     0.66   0.511    -.2846773    .5720602
------------------------------------------------------------------------------
Wald test of lambda=0:                chi2(1) =   0.432 (0.511)
Likelihood ratio test of lambda=0:    chi2(1) =   0.424 (0.515)
Lagrange multiplier test of lambda=0: chi2(1) =   0.346 (0.556)

Acceptable range for lambda: -1.280 < lambda < 1.000



.
. *Spatial lag model
. spatreg air1_mean \$xlist1, weights(W) eigenval(E) model(lag)

initial:       log likelihood = -838.34359
rescale:       log likelihood = -838.34359
rescale eq:    log likelihood = -838.34359
Iteration 0:   log likelihood = -838.34359
Iteration 1:   log likelihood = -837.62639
Iteration 2:   log likelihood = -837.62499
```

```
Iteration 3:   log likelihood = -837.62499


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial lag model                                   Number of obs   =       312
                                                    Variance ratio  =     0.036
                                                    Squared corr.   =     0.039
Log likelihood = -837.62499                         Sigma           =      3.54

------------------------------------------------------------------------------
   air1_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
air1_mean    |
        frst |   .0000937   .0001535     0.61   0.542    -.0002071    .0003945
        recr |   .0024125   .0028016     0.86   0.389    -.0030787    .0079036
        comm |   .0004503   .0013377     0.34   0.736    -.0021716    .0030723
        agri |  -.0000516   .0003829    -0.13   0.893    -.0008021    .0006989
        inds |  -.0014421   .0010353    -1.39   0.164    -.0034712     .000587
        educ |  -.0586865   .1272305    -0.46   0.645    -.3080537    .1906806
    employed |  -.8725622   .4183202    -2.09   0.037    -1.692455   -.0526696
   household |   .0904912   .1343877     0.67   0.501    -.1729039    .3538863
  visit_freq |  -.2569175   .1655238    -1.55   0.121    -.5813381    .0675031
       _cons |   17.81163   3.019297     5.90   0.000     11.89392    23.72935
-------------+----------------------------------------------------------------
         rho |  -.2186499   .1934648    -1.13   0.258    -.5978339    .1605341
------------------------------------------------------------------------------
Wald test of rho=0:                 chi2(1) =   1.277 (0.258)
Likelihood ratio test of rho=0:     chi2(1) =   1.273 (0.259)
Lagrange multiplier test of rho=0:  chi2(1) =   1.360 (0.244)

Acceptable range for rho: -1.280 < rho < 1.000
```

```
. spatreg air2_mean \$xlist1, weights(W) eigenval(E) model(lag)

initial:       log likelihood = -742.16672
rescale:       log likelihood = -742.16672
rescale eq:    log likelihood = -742.16672
Iteration 0:   log likelihood = -742.16672
Iteration 1:   log likelihood =  -741.4426
Iteration 2:   log likelihood = -741.44115
Iteration 3:   log likelihood = -741.44115


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial lag model                                  Number of obs   =        312
                                                   Variance ratio  =      0.041
                                                   Squared corr.   =      0.045
Log likelihood = -741.44115                        Sigma           =       2.60

------------------------------------------------------------------------------
   air2_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
air2_mean    |
        frst |   .0001501   .0001137     1.32   0.187    -.0000727     .000373
        recr |   .0019026   .0020604     0.92   0.356    -.0021358     .005941
        comm |  -.0010699   .0009933    -1.08   0.281    -.0030167    .0008768
        agri |  -.0000565   .0002818    -0.20   0.841    -.0006088    .0004958
        inds |  -.0008984   .0007613    -1.18   0.238    -.0023905    .0005938
        educ |   .0425896   .0939872     0.45   0.650     -.141622    .2268012
    employed |  -.0637964   .3073123    -0.21   0.836    -.6661174    .5385245
   household |     .09525   .0986013     0.97   0.334     -.098005    .2885049
  visit_freq |  -.2730642   .1216359    -2.24   0.025    -.5114661   -.0346623
```

```
       _cons |   11.70836   2.069798     5.66   0.000     7.651627    15.76509
-------------+----------------------------------------------------------------
         rho |  -.2183822    .192178    -1.14   0.256    -.5950442    .1582798
------------------------------------------------------------------------------
Wald test of rho=0:                      chi2(1) =   1.291 (0.256)
Likelihood ratio test of rho=0:          chi2(1) =   1.287 (0.257)
Lagrange multiplier test of rho=0:       chi2(1) =   1.390 (0.238)

Acceptable range for rho: -1.280 < rho < 1.000



. spatreg fac1_mean \$xlist1, weights(W) eigenval(E) model(lag)

initial:       log likelihood = -246.15997
rescale:       log likelihood = -246.15997
rescale eq:    log likelihood = -246.15997
Iteration 0:   log likelihood = -246.15997
Iteration 1:   log likelihood = -245.76193
Iteration 2:   log likelihood = -245.76094
Iteration 3:   log likelihood = -245.76094


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial lag model                                  Number of obs   =        312
                                                   Variance ratio  =      0.020
                                                   Squared corr.   =      0.022
Log likelihood = -245.76094                        Sigma           =       0.53

------------------------------------------------------------------------------
   fac1_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
```

```
fac1_mean    |
        frst |  -.0000155   .0000228    -0.68   0.497    -.0000603    .0000293
        recr |   -.000757   .0004262    -1.78   0.076    -.0015924    .0000784
        comm |   .0000416    .000201     0.21   0.836    -.0003524    .0004355
        agri |  -.0000289   .0000575    -0.50   0.615    -.0001415    .0000838
        inds |  -.0000792   .0001554    -0.51   0.610    -.0003838    .0002253
        educ |   .0068972    .019061     0.36   0.717    -.0304617     .044256
    employed |  -.0787768   .0627687    -1.26   0.209    -.2018012    .0442476
   household |    .004783   .0201816     0.24   0.813    -.0347723    .0443383
 visit_freq |   .0292998   .0248619     1.18   0.239    -.0194286    .0780283
       _cons |   2.451378   .6354791     3.86   0.000     1.205862    3.696894
-------------+----------------------------------------------------------------
         rho |   .1668105   .2073367     0.80   0.421    -.2395619    .5731829
------------------------------------------------------------------------------
Wald test of rho=0:                    chi2(1) =   0.647 (0.421)
Likelihood ratio test of rho=0:        chi2(1) =   0.634 (0.426)
Lagrange multiplier test of rho=0:     chi2(1) =   0.570 (0.450)

Acceptable range for rho: -1.280 < rho < 1.000



. spatreg fac2_mean \$xlist1, weights(W) eigenval(E) model(lag)

initial:       log likelihood = -577.88741
rescale:       log likelihood = -577.88741
rescale eq:    log likelihood = -577.88741
Iteration 0:   log likelihood = -577.88741
Iteration 1:   log likelihood = -577.80353
Iteration 2:   log likelihood = -577.80338
Iteration 3:   log likelihood = -577.80338


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes
```

```
Spatial lag model                                   Number of obs   =       312
                                                    Variance ratio  =     0.021
                                                    Squared corr.   =     0.021
Log likelihood = -577.80338                         Sigma           =      1.54

------------------------------------------------------------------------------
   fac2_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
fac2_mean    |
        frst |  -.0000322   .0000663    -0.49   0.627    -.0001621    .0000977
        recr |  -.0000254   .0012227    -0.02   0.983    -.0024218    .0023709
        comm |  -.0006566    .000582    -1.13   0.259    -.0017972     .000484
        agri |    .000232   .0001664     1.39   0.163    -.0000942    .0005582
        inds |   .0003653    .000451     0.81   0.418    -.0005186    .0012493
        educ |  -.0736854      .0552    -1.33   0.182    -.1818754    .0345046
    employed |   .1224593   .1820459     0.67   0.501    -.2343442    .4792627
   household |   .0494192   .0584123     0.85   0.398    -.0650668    .1639051
  visit_freq |   .0211547   .0720548     0.29   0.769      -.12007    .1623794
       _cons |    2.65888   .7107589     3.74   0.000     1.265818    4.051942
-------------+----------------------------------------------------------------
         rho |   .0133101   .2029634     0.07   0.948    -.3844909     .411111
------------------------------------------------------------------------------
Wald test of rho=0:                 chi2(1) =   0.004 (0.948)
Likelihood ratio test of rho=0:     chi2(1) =   0.004 (0.948)
Lagrange multiplier test of rho=0:  chi2(1) =   0.004 (0.948)

Acceptable range for rho: -1.280 < rho < 1.000


. spatreg noi1_mean \$xlist1, weights(W) eigenval(E) model(lag)

initial:       log likelihood = -239.59006
rescale:       log likelihood = -239.59006
rescale eq:    log likelihood = -239.59006
Iteration 0:   log likelihood = -239.59006
```

```
Iteration 1:   log likelihood = -238.91107
Iteration 2:   log likelihood = -238.90981
Iteration 3:   log likelihood = -238.90981


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial lag model                                     Number of obs   =       312
                                                      Variance ratio  =     0.024
                                                      Squared corr.   =     0.027
Log likelihood = -238.90981                           Sigma           =      0.52

------------------------------------------------------------------------------
   noi1_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
noi1_mean    |
        frst |  -3.86e-07   .0000224    -0.02   0.986    -.0000444    .0000436
        recr |  -.0002412    .000413    -0.58   0.559    -.0010506    .0005682
        comm |   .0002068   .0001962     1.05   0.292    -.0001777    .0005913
        agri |  -.0000655   .0000562    -1.17   0.244    -.0001756    .0000446
        inds |  -.0001782   .0001526    -1.17   0.243    -.0004773     .000121
        educ |   .0071448    .018609     0.38   0.701    -.0293281    .0436178
    employed |   .0492495   .0614165     0.80   0.423    -.0711245    .1696236
   household |   .0151572    .019746     0.77   0.443    -.0235443    .0538587
  visit_freq |  -.0381033   .0243132    -1.57   0.117    -.0857564    .0095497
       _cons |   6.401702   1.043916     6.13   0.000     4.355664     8.44774
-------------+----------------------------------------------------------------
         rho |  -.2144772   .1957987    -1.10   0.273    -.5982356    .1692812
------------------------------------------------------------------------------
Wald test of rho=0:                     chi2(1) =   1.200 (0.273)
Likelihood ratio test of rho=0:         chi2(1) =   1.197 (0.274)
Lagrange multiplier test of rho=0:      chi2(1) =   1.252 (0.263)
```

```
Acceptable range for rho: -1.280 < rho < 1.000


. spatreg noi2_mean \$xlist1, weights(W) eigenval(E) model(lag)

initial:       log likelihood = -424.65571
rescale:       log likelihood = -424.65571
rescale eq:    log likelihood = -424.65571
Iteration 0:   log likelihood = -424.65571
Iteration 1:   log likelihood = -422.21041
Iteration 2:   log likelihood = -422.19826
Iteration 3:   log likelihood = -422.19826


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial lag model                                  Number of obs   =       312
                                                   Variance ratio  =     0.035
                                                   Squared corr.   =     0.051
Log likelihood = -422.19826                        Sigma           =      0.93

------------------------------------------------------------------------------
   noi2_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
noi2_mean    |
        frst |   .0000531   .0000401     1.33   0.185    -.0000254    .0001317
        recr |    .000122   .0007381     0.17   0.869    -.0013247    .0015687
        comm |  -.0003249   .0003532    -0.92   0.358    -.0010173    .0003674
        agri |   .0000195   .0001007     0.19   0.847    -.0001778    .0002168
        inds |  -.0004176   .0002751    -1.52   0.129    -.0009568    .0001216
        educ |  -.0083347   .0334142    -0.25   0.803    -.0738254     .057156
    employed |   .1095407   .1104027     0.99   0.321    -.1068446     .325926
```

```
   household |  -.0064086   .0354059    -0.18   0.856     -.075803    .0629858
  visit_freq |  -.0220085   .0436843    -0.50   0.614    -.1076281     .063611
       _cons |    2.75669    .462591     5.96   0.000     1.850029    3.663352
-------------+----------------------------------------------------------------
         rho |  -.4385117   .1975153    -2.22   0.026    -.8256346   -.0513887
------------------------------------------------------------------------------
Wald test of rho=0:                    chi2(1) =   4.929 (0.026)
Likelihood ratio test of rho=0:        chi2(1) =   4.751 (0.029)
Lagrange multiplier test of rho=0:     chi2(1) =   4.593 (0.032)

Acceptable range for rho: -1.280 < rho < 1.000



. spatreg nur1_mean \$xlist1, weights(W) eigenval(E) model(lag)

initial:       log likelihood = -476.55203
rescale:       log likelihood = -476.55203
rescale eq:    log likelihood = -476.55203
Iteration 0:   log likelihood = -476.55203
Iteration 1:   log likelihood =  -472.0038
Iteration 2:   log likelihood = -471.93849
Iteration 3:   log likelihood = -471.93843
Iteration 4:   log likelihood = -471.93843


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial lag model                                    Number of obs   =       312
                                                     Variance ratio  =     0.052
                                                     Squared corr.   =     0.077
Log likelihood = -471.93843                          Sigma           =      1.09
```

```
------------------------------------------------------------------------------
   nur1_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
nur1_mean    |
        frst |   .0000275   .0000473     0.58   0.561    -.0000653    .0001202
        recr |  -.0007437   .0008691    -0.86   0.392     -.002447    .0009596
        comm |   .0001972   .0004123     0.48   0.632    -.0006108    .0010052
        agri |   .0001377   .0001178     1.17   0.242    -.0000932    .0003686
        inds |   .0005025   .0003202     1.57   0.117    -.0001251    .0011301
        educ |   .0904945   .0390986     2.31   0.021     .0138627    .1671263
    employed |   .0182928   .1289974     0.14   0.887    -.2345373     .271123
   household |  -.0257393    .041427    -0.62   0.534    -.1069348    .0554562
  visit_freq |  -.0628463   .0511284    -1.23   0.219    -.1630562    .0373636
       _cons |    1.94704   .4130382     4.71   0.000       1.1375     2.75658
-------------+----------------------------------------------------------------
         rho |  -.5551659   .1783661    -3.11   0.002     -.904757   -.2055749
------------------------------------------------------------------------------
Wald test of rho=0:                    chi2(1) =   9.688 (0.002)
Likelihood ratio test of rho=0:        chi2(1) =   9.063 (0.003)
Lagrange multiplier test of rho=0:     chi2(1) =  10.055 (0.002)

Acceptable range for rho: -1.280 < rho < 1.000



. spatreg nur2_mean \$xlist1, weights(W) eigenval(E) model(lag)

initial:       log likelihood =   477.4689
rescale:       log likelihood =   477.4689
rescale eq:    log likelihood =   477.4689
Iteration 0:   log likelihood =   477.4689
Iteration 1:   log likelihood =  478.13022
Iteration 2:   log likelihood =  478.13106
Iteration 3:   log likelihood =  478.13106


Weights matrix
 Name: W
```

```
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial lag model                                  Number of obs   =        312
                                                   Variance ratio  =      0.009
                                                   Squared corr.   =      0.014
Log likelihood = 478.13106                         Sigma           =       0.05

------------------------------------------------------------------------------
   nur2_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
nur2_mean    |
        frst |   1.09e-06   2.26e-06     0.48   0.629    -3.34e-06    5.53e-06
        recr |   -.000038   .0000416    -0.91   0.361    -.0001195    .0000435
        comm |   .0000137   .0000198     0.70   0.487     -.000025    .0000525
        agri |   1.43e-06   5.63e-06     0.25   0.800    -9.61e-06    .0000125
        inds |  -1.52e-07   .0000153    -0.01   0.992    -.0000301    .0000298
        educ |   .0006438   .0018719     0.34   0.731     -.003025    .0043126
    employed |   .0032745    .006171     0.53   0.596    -.0088204    .0153693
   household |  -.0013099   .0019789    -0.66   0.508    -.0051885    .0025687
  visit_freq |   .0008907   .0024395     0.37   0.715    -.0038906     .005672
       _cons |   1.340623   .2401717     5.58   0.000      .869895    1.811351
-------------+----------------------------------------------------------------
         rho |  -.2383219    .220767    -1.08   0.280    -.6710173    .1943735
------------------------------------------------------------------------------
Wald test of rho=0:                   chi2(1) =   1.165 (0.280)
Likelihood ratio test of rho=0:       chi2(1) =   1.161 (0.281)
Lagrange multiplier test of rho=0:    chi2(1) =   0.958 (0.328)

Acceptable range for rho: -1.280 < rho < 1.000



. spatreg tree1_mean \$xlist1, weights(W) eigenval(E) model(lag)

initial:       log likelihood = -671.97874
```

```
rescale:       log likelihood = -671.97874
rescale eq:    log likelihood = -671.97874
Iteration 0:   log likelihood = -671.97874
Iteration 1:   log likelihood = -670.41569
Iteration 2:   log likelihood = -670.40051
Iteration 3:   log likelihood = -670.40051


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial lag model                                   Number of obs   =       312
                                                    Variance ratio  =     0.028
                                                    Squared corr.   =     0.038
Log likelihood = -670.40051                         Sigma           =      2.07

------------------------------------------------------------------------------
  tree1_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
tree1_mean   |
        frst |  -.0001367   .0000889    -1.54   0.124     -.000311    .0000376
        recr |   .0008024   .0016397     0.49   0.625    -.0024113    .0040161
        comm |  -.0012429   .0007956    -1.56   0.118    -.0028023    .0003164
        agri |   .0001624   .0002233     0.73   0.467    -.0002753    .0006002
        inds |   .0001468    .000605     0.24   0.808     -.001039    .0013327
        educ |  -.0025471   .0744656    -0.03   0.973    -.1484969    .1434028
    employed |   .0225795   .2444252     0.09   0.926     -.456485     .501644
   household |   .0500747   .0783795     0.64   0.523    -.1035463    .2036957
  visit_freq |  -.0502786   .0967592    -0.52   0.603    -.2399232     .139366
       _cons |    1.21439   .7075108     1.72   0.086    -.1723054    2.601086
-------------+----------------------------------------------------------------
         rho |   .3297325   .1852017     1.78   0.075    -.0332562    .6927212
------------------------------------------------------------------------------
Wald test of rho=0:                    chi2(1) =   3.170 (0.075)
```

```
Likelihood ratio test of rho=0:        chi2(1) =   2.993 (0.084)
Lagrange multiplier test of rho=0:     chi2(1) =   3.153 (0.076)

Acceptable range for rho: -1.280 < rho < 1.000



. spatreg tree2_mean \$xlist1, weights(W) eigenval(E) model(lag)

initial:       log likelihood =  384.40205
rescale:       log likelihood =  384.40205
rescale eq:    log likelihood =  384.40205
Iteration 0:   log likelihood =  384.40205
Iteration 1:   log likelihood =  384.70893
Iteration 2:   log likelihood =  384.70957
Iteration 3:   log likelihood =  384.70957


Weights matrix
 Name: W
 Type: Distance-based (inverse distance)
 Distance band: 0.0 < d <= 16.0
 Row-standardized: Yes


Spatial lag model                                  Number of obs   =       312
                                                   Variance ratio  =     0.031
                                                   Squared corr.   =     0.033
Log likelihood = 384.70957                         Sigma           =      0.07

------------------------------------------------------------------------------
  tree2_mean |      Coef.   Std. Err.      z    P>|z|     [95\% Conf. Interval]
-------------+----------------------------------------------------------------
tree2_mean   |
        frst |  -1.77e-07   3.04e-06    -0.06   0.954    -6.13e-06    5.78e-06
        recr |   .0000771   .0000561     1.37   0.169    -.0000328     .000187
        comm |   .0000313   .0000266     1.18   0.239    -.0000208    .0000835
        agri |   2.78e-07   7.61e-06     0.04   0.971    -.0000146    .0000152
```

```
        inds |   -.000016   .0000206    -0.77   0.439    -.0000564    .0000244
        educ |   .0010745    .002524     0.43   0.670    -.0038725    .0060214
    employed |   .0002988   .0083219     0.04   0.971    -.0160118    .0166094
   household |   .0035303   .0026739     1.32   0.187    -.0017104     .008771
 visit_freq |  -.0074958   .0032986    -2.27   0.023    -.0139609   -.0010307
       _cons |   1.541661   .3869576     3.98   0.000     .7832378    2.300084
-------------+----------------------------------------------------------------
         rho |   .1435051    .211858     0.68   0.498    -.2717289    .5587391
------------------------------------------------------------------------------
Wald test of rho=0:                      chi2(1) =   0.459 (0.498)
Likelihood ratio test of rho=0:          chi2(1) =   0.451 (0.502)
Lagrange multiplier test of rho=0:       chi2(1) =   0.391 (0.532)

Acceptable range for rho: -1.280 < rho < 1.000


.
end of do-file

.
```

Appendix I

MNL Estimation from Pilot Samples

Attributes	Coefficients	
alt2	-0.122	.
dis	-0.025	***
air1	1.344	***
air2	2.481	***
fac1	0.104	
fac2	0.258	.
noi1	0.230	
noi2	0.748	***
nur1	0.252	.
nur2	0.493	***
tree1	0.296	*
tree2	0.490	***
Sig: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’		
Log-Likelihood:	-625.38	
McFadden Rˆ2:	0.35722	
Likelihood ratio test:	chisq = 695.11 (p.value = 0.000)	

Table 72: MNL estimation from pilot samples

Bibliography

Adamowicz, W., Louviere, J., and Williams, M. (1994). Combining revealed and stated preference methods for valuing environmental amenities. Journal of Environmental Economics and Management, 26:271–292.

Adamowicz, W. L., Fletcher, J. J., and Graham-Tomasi, T. (1989). Functional form and the statistical properties of welfare measures. American Journal of Agricultural Economics, 71:414–421.

Adinolfi, C., Suárez-Cáceres, G. P., and Cariñanos, P. (2014). Relation between visitors' behaviour and characteristics of green spaces in the city of granada, south-eastern spain. Urban Forestry & Urban Greening, 13:534–542.

Aevermann, T. and Schmude, J. (2015). Quantification and monetary valuation of urban ecosystem services in munich, germany.

AirNow (2023). Air quality index (aqi) basics.

Albert, C., Aronson, J., Fürst, C., and Opdam, P. (2014). Integrating ecosystem services in landscape planning: requirements, approaches, and impacts. Landscape Ecology, 29:1277–1285.

Anderson, S. T. and West, S. E. (2006). Open space, residential property values, and spatial context. Regional Science and Urban Economics, 36:773–789.

Antrop, M. (2004). Landscape change and the urbanization process in europe. Landscape and Urban Planning, 67:9–26. Development of European Landscapes.

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. Journal of Environmental Planning and Management, 54:891–908.

Arnberger, A. and Eder, R. (2015). Are urban visitors' general preferences for green-spaces similar to their preferences when seeking stress relief? Urban Forestry & Urban Greening, 14(4):872–882.

Aronson, M. F., La Sorte, F. A., Nilon, C. H., Katti, M., Goddard, M. A., Lepczyk, C. A., ..., and Winter, M. (2014). A global analysis of the impacts of urbanization on bird and plant diversity reveals key anthropogenic drivers. Proceedings of the Royal Society B: Biological Sciences, 281(1780):20133330.

Aronson, M. F. J., Lepczyk, C. A., Evans, K. L., Goddard, M. A., Lerman, S. B., MacIvor, J. S., Nilon, C. H., and Vargo, T. (2017). Biodiversity in the city: key challenges for urban green space management. Frontiers in Ecology and the Environment, 15:189–196.

Badura, T., Ferrini, S., Burton, M., Binner, A., and Bateman, I. J. (2020). Using individualised choice maps to capture the spatial dimensions of value within choice experiments. Environmental and Resource Economics, 75:297–322.

Banfi, S., Filippini, M., and Horehájová, A. (2012). Using a choice experiment to estimate the benefits of a reduction of externalities in urban areas with special focus on electrosmog. Applied Economics, 44(3):387–397.

Barbosa, O., Tratalos, J. A., Armsworth, P. R., Davies, R. G., Fuller, R. A., Johnson, P., and Gaston, K. J. (2007). Who benefits from access to green space? a case study from sheffield, uk. Landscape and Urban Planning, 83:187–195.

Barnett, L. A., Kidman, G., and Howard, S. (2016). Children's drawing and writing: Insights for occupational therapy. Australian Occupational Therapy Journal, 63:316–324.

Bateman, I., Carson, R., Day, B., Hanemann, M., Hanley, N., Hett, T., Jones-Lee, M., and Loomes, G. (2002). Economic Valuation with Stated Preference Techniques. Edward Elgar Publishing.

Ben-Akiva, M. E. and Lerman, S. R. (1985). Discrete choice analysis: theory and application to travel demand, volume 9. MIT press.

Beninde, J., Veith, M., and Hochkirch, A. (2015). Biodiversity in cities needs space: a meta-analysis of factors determining intra-urban biodiversity variation. Ecology letters, 18 6:581–92.

Bertram, C. and Larondelle, N. (2017). Going to the woods is going home: Recreational benefits of a larger urban forest site — a travel cost analysis for berlin, germany. Ecological Economics, 132:255–263.

Bertram, C. and Rehdanz, K. (2015a). Preferences for cultural urban ecosystem services: Comparing attitudes, perception, and use. Ecosystem Services, 12:187–199.

Bertram, C. and Rehdanz, K. (2015b). The role of urban green space for human well-being. Ecological Economics, 120:139–152.

Bishop, J. T. (2002). Valuing forests a review of methods and applications in developing countries.

Blaug, M. (1992). The methodology of economics: Or, how economists explain. Cambridge University Press.

Bockstael, N. E., Strand, I. E., and Hanemann, W. M. (1987). Time and the recreational demand model. American Journal of Agricultural Economics, 69:293–302.

Bolund, P. and Hunhammar, S. (1999). Ecosystem services in urban areas. Ecological economics, 29(2):293–301.

Bonaiuto, M., Fornara, F., and Bonnes, M. (2003). Indexes of perceived residential environment quality and neighbourhood attachment in urban environments: a confirmation study on the city of rome. Landscape and Urban Planning, 65:41–52. Urban environmental quality and human wellbeing.

Bowler, D. E., Buyung-Ali, L. M., Knight, T. M., and Pullin, A. S. (2010). A systematic review of evidence for the added benefits to health of exposure to natural environments. BMC Public Health, 10:456.

Brainard, J. (1999). Integrating geographical information systems into travel cost analysis and benefit transfer. International Journal of Geographical Information Science, 13(3):227–246.

Brander, L. M. and Koetse, M. J. (2011). The value of urban open space: Meta-analyses of contingent valuation and hedonic pricing results. Journal of Environmental Management, 92:2763–2773.

Braun, V. and Clarke, V. (2006). Using thematic analysis in psychology. Qualitative Research in Psychology, 3:77–101.

Breitmeyer, B., Öğmen, H., and Chen, J. (2004). Unconscious priming by color and form: Different processes and levels. Consciousness and cognition, 13:138–157.

Breuste, J., Haase, D., and Elmqvist, T. (2013). Urban Landscapes and Ecosystem Services. John Wiley & Sons, Ltd.

Bronnmann, J., Liebelt, V., Marder, F., Meya, J., and Quaas, M. (2023). The value of naturalness of urban green spaces: Evidence from a discrete choice experiment. Land Economics.

Brouwer, R., Brander, L., Kuik, O., Papyrakis, E., and Bateman, I. (2013). A synthesis of approaches to assess and value ecosystem services in the eu in the context of teeb.

Buletin Mutiara (2019). Penang launches first urban archeological park in the country.

Buletin Mutiara (2021). Taman areca is a shining model for transforming idle public spaces.

Bullock, C. H. (2008). Valuing urban green space: Hypothetical alternatives and the status quo. Journal of Environmental Planning and Management, 51:15–35.

Calfapietra, C., Fares, S., and Loreto, F. (2009). Volatile organic compounds from italian vegetation and their interaction with ozone. Environmental Pollution, 157:1478–1486. Special Issue Section: Ozone and Mediterranean Ecology: Plants, People, Problems.

Cameron, A. C. and Trivedi, P. K. (1990). Regression-based tests for overdispersion in the poisson model. Journal of econometrics, 46(3):347–364.

Cameron, L., Maslen, R., Richardson, J., and Taylor, M. (2012). Words apart: Investigating language and affect in a forensic context using wordles. Discourse & Society, 23:377–400.

Cameron, T. A. (1988). A new paradigm for valuing non-market goods using referendum data: maximum likelihood estimation by censored logistic regression. Journal of environmental economics and management, 15(3):355–379.

Cameron, T. A. and James, M. D. (1987). Efficient estimation methods for" closed-ended" contingent valuation surveys. The review of economics and statistics, pages 269–276.

Campbell, D. (2007). Willingness to pay for rural landscape improvements: Combining mixed logit and random-effects models. Journal of Agricultural Economics, 58:467–483.

Campbell, D., Hutchinson, W. G., and Scarpa, R. (2009). Using choice experiments to explore the spatial distribution of willingness to pay for rural landscape improvements. Environment and Planning A: Economy and Space, 41:97–111.

Caparrós, B. A. E. and Schweppe-Kraft, B. (2020). Background paper: Exchange values and welfare values in the seea eea. seea eea revision, working group 5: Valuation and accounting treatments.

Caparrós, A., Campos, P., and Montero, G. (2003). An operative framework for total hicksian income measurement: application to a multiple use forest. Environmental and Resource Economics, 26:173–198.

Caparrós, A., Oviedo, J. L., Álvarez, A., and Campos, P. (2017). Simulated exchange values and ecosystem accounting: Theory and application to free access recreation. Ecological Economics, 139:140–149.

Carlsson, F. and Martinsson, P. (2003). Design techniques for stated preference methods in health economics. Health economics, 12(4):281–294.

Centers for Disease Control and Prevention (2023). What noises cause hearing loss?

Cesario, F. J. (1976). Value of time in recreation benefit studies. Land Economics, 52:32–41.

Chae, D.-R., Wattage, P., and Pascoe, S. (2012). Recreational benefits from a marine protected area: A travel cost analysis of lundy. Tourism Management, 33(4):971–977.

Champ, P. A., Boyle, K. J., Brown, T. C., and Peterson, L. G. (2003). A primer on nonmarket valuation (Vol. 3). Springer.

Chang, J., Qu, Z., Xu, R., Pan, K., Xu, B., Min, Y., Ren, Y., Yang, G., and Ge, Y. (2017). Assessing the ecosystem services provided by urban green spaces along urban center-edge gradients. Scientific Reports, 7:11226.

Cheshire, P. C. and Stabler, M. J. (1976). Joint consumption benefits in recreational site 'surplus'; an empirical estimate. Regional Studies, 10(3):343–351.

Chintantya, D. and Maryono, M. (2018). Comparing value of urban green space using contingent valuation and travel cost methods. E3S Web of Conferences, 31:7008.

ChoiceMetrics (2018). Ngene 1.2 user manual & reference guide.

Christie, M., Hanley, N., and Hynes, S. (2007). Valuing enhancements to forest recreation using choice experiment and contingent behaviour methods. Journal of Forest Economics, 13(2-3):75–102.

Cilliers, S., Cilliers, J., Lubbe, R., and Siebert, S. (2013). Ecosystem services of urban green spaces in african countries—perspectives and challenges. Urban Ecosystems, 16:681–702.

City Council of Penang Island (2022). Penang island voluntary local review.

Clarke, V. and Braun, V. (2017). Thematic analysis. The Journal of Positive Psychology, 12:297–298.

Cohen, P., Potchter, O., and Schnell, I. (2014). The impact of an urban park on air pollution and noise levels in the mediterranean city of tel-aviv, israel. Environmental Pollution, 195:73–83.

Comte, A., Campagne, C. S., Lange, S., Bruzón, A. G., Hein, L., Santos-Martín, F., and Levrel, H. (2022). Ecosystem accounting: Past scientific developments and future challenges. Ecosystem Services, 58.

Costanza, R., d'Arge, R., Groot, R. D., Farber, S., Grasso, M., Hannon, B., Limburg, K., Naeem, S., O'neill, R. V., Paruelo, J., et al. (1997). The value of the world's ecosystem services and natural capital. nature, 387:253–260.

Creel, M. D. and Loomis, J. B. (1990). Theoretical and empirical advantages of truncated count data estimators for analysis of deer hunting in california. American journal of agricultural economics, 72(2):434–441.

Czajkowski, M., Budziński, W., Campbell, D., Giergiczny, M., and Hanley, N. (2017). Spatial heterogeneity of willingness to pay for forest management. Environmental and Resource Economics, 68:705–727.

da Silva, B. F., Pena, J. C., Viana-Junior, A. B., Vergne, M., and Pizo, M. A. (2021). Noise and tree species richness modulate the bird community inhabiting small public urban green spaces of a neotropical city. Urban Ecosystems, 24:71–81.

de Magistris, T., Royo, A. G., and Aguado, L. M. A. (2014). Wine consumers preferences in spain: an analysis using the best-worst scaling approach. Spanish Journal of Agricultural Research, 3:529–541.

De Valck, J., Landuyt, D., Broekx, S., Liekens, I., De Nocker, L., and Vranken, L. (2017). Outdoor recreation in various landscapes: Which site characteristics really matter? Land Use Policy, 65:186–197.

del Saz Salazar, S. and Menéndez, L. G. (2007). Estimating the non-market benefits of an urban park: Does proximity matter? Land Use Policy, 24:296–305.

Department of Landscape Malaysia (2019). Dasar landskap negara: Malaysia negara taman terindah.

Department of Statistics Malaysia (2021). Salaries & wages survey report, malaysia, 2021.

Department of Statistics Malaysia (2023). The population of malaysia.

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

Dimoudi, A. and Nikolopoulou, M. (2003). Vegetation in the urban environment: microclimatic analysis and benefits. Energy and Buildings, 35:69–76. Special issue on urban research.

Domencich, T. A. and McFadden, D. (1975). Urban travel demand-a behavioral analysis.

Dwyer, J. F., Peterson, G. L., and Darragh, A. J. (1983). Estimating the value of urban forests using the travel cost method. Arboriculture & Urban Forestry.

Ecker, U. K., Lewandowsky, S., and Oberauer, K. (2015). Manipulating exposure time to create illusory correlations. Journal of Experimental Psychology: General, 144:1011–1025.

Englin, J. E., Holmes, T. P., and Sills, E. O. (2003). Estimating forest recreation demand using count data models. Forests in a market economy, pages 341–359.

Engström, G. and Gren, A. (2017). pros and cons of hedonic pricing. Ecology and Society, 22.

Erdem, S. and Rigby, D. (2013). Investigating heterogeneity in the characterization of risks using best worst scaling. Risk Analysis, 33(9):1728–1748.

Ezebilo, E. E. (2016). Economic value of a non-market ecosystem service: an application of the travel cost method to nature recreation in sweden. International Journal of Biodiversity Science, Ecosystem Services & Management, 12:314–327.

Faccioli, M., McVittie, A., Glenk, K., and Blackstock, K. (2016). Natural capital accounts: Review of available data and accounting options.

Fenichel, E. and Obst, C. (2019). Framework for the valuation of ecosystem assets (draft). working group 5: Valuation and accounting treatments. Technical report, Discussion Paper 5.2.

Fereday, J. and Muir-Cochrane, E. (2006). Demonstrating rigor using thematic analysis: A hybrid approach of inductive and deductive coding and theme development. International Journal of Qualitative Methods, 5:80–92.

Finn, A. and Louviere, J. J. (1992). Determining the appropriate response to evidence of public concern: the case of food safety. Journal of Public Policy & Marketing, 11(2):12–25.

Fischer, F. (1998). Beyond empiricism: policy inquiry in post-positivist perspective. Policy studies journal, 26(1):129–146.

Fischer, L. K., Honold, J., Botzat, A., Brinkmeyer, D., Cvejić, R., Delshammar, T., Elands, B., Haase, D., Kabisch, N., Karle, S. J., Lafortezza, R., Nastran, M., Nielsen, A. B., van der Jagt, A. P., Vierikko, K., and Kowarik, I. (2018). Recreational ecosystem services in european cities: Sociocultural and geographical contexts matter for park use. Ecosystem Services, 31:455–467. Assessment and Valuation of Recreational Ecosystem Services.

Fleming, C. M. and Cook, A. (2008). The recreational value of lake mckenzie, fraser island: An application of the travel cost method. Tourism Management, 29:1197–1205.

Franco, M. R., Howard, K., Sherrington, C., Ferreira, P. H., Rose, J., Gomes, J. L., and Ferreira, M. L. (2015). Eliciting older people's preferences for exercise programs: a best-worst scaling choice experiment. Journal of Physiotherapy, 61(1):34–41.

Freedman, D. A. (2006). On the so-called "huber sandwich estimator" and "robust standard errors". The American Statistician, 60(4):299–302.

Gidlöf-Gunnarsson, A. and Öhrström, E. (2007). Noise and well-being in urban residential environments: The potential role of perceived availability to nearby green areas. Landscape and Urban Planning, 83:115–126.

Giles-Corti, B., Broomhall, M. H., Knuiman, M., Collins, C., Douglas, K., Ng, K., Lange, A., and Donovan, R. J. (2005). Increasing walking: How important is distance to, attractiveness, and size of public open space? American Journal of Preventive Medicine, 28:169–176. Active Living Research.

Gillespie, R. (1997). Economic value and regional economic impact of minnamurra rainforest centre, budderoo national park.

Guest, G., MacQueen, K. M., and Namey, E. E. (2012). Applied thematic analysis. SAGE Publications.

Gómez-Baggethun, E. and Barton, D. N. (2013). Classifying and valuing ecosystem services for urban planning. Ecological Economics, 86:235–245. Sustainable Urbanisation: A resilient future.

Haase, D., Kabisch, S., Haase, A., Andersson, E., Banzhaf, E., Baró, F., Brenck, M., Fischer, L. K., Frantzeskaki, N., Kabisch, N., Krellenberg, K., Kremer, P., Kronenberg, J., Larondelle, N., Mathey, J., Pauleit, S., Ring, I., Rink, D., Schwarz, N., and Wolff, M. (2017). Greening cities – to be socially inclusive? about the alleged paradox of society and ecology in cities. Habitat International, 64:41–48.

Haines-Young, R. and Potschin, M. B. (2018). Common international classification of ecosystem services (cices) v5.1 and guidance on the application of the revised structure.

Hanauer, M. M. and Reid, J. (2017). Valuing urban open space using the travel-cost method and the implications of measurement error. Journal of environmental management, 198:50–65.

Hands, D. W. (1993). Popper and lakatos in economic methodology. In Rationality, institutions and economic methodology, pages 72–86. Routledge.

Hanemann, W. M. (1982). Applied welfare analysis with qualitative response models.

Hanley, N. D. (1989). Valuing rural recreation benefits: an empirical comparison of two approaches. Journal of Agricultural Economics, 40:361–374.

Harberger, A. C. (1971). Three basic postulates for applied welfare economics: an interpretive essay. Journal of Economic literature, 9(3):785–797.

Heal, G. and Kriström, B. (2005). National income and the environment. In Handbook of environmental economics, volume 3, pages 1147–1217. Elsevier.

Hellerstein, D. and Mendelsohn, R. (1993). A theoretical foundation for count data models. American journal of agricultural economics, 75(3):604–611.

Hendrey, G. R., Ellsworth, D. S., Lewin, K. F., and Nagy, J. (1999). A free-air enrichment system for exposing tall forest vegetation to elevated atmospheric co2. Global Change Biology, 5:293–309.

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

Heo, J.-N. and Kim, T.-G. (2013). Valuation and preference of urban agriculture park using choice experiment. Korean Journal of Organic Agriculture, 21(2):125–137.

Hess, S. and Palma, D. (2019). Apollo: A flexible, powerful and customisable freeware package for choice model estimation and application. Journal of Choice Modelling, 32:100170.

Hess, S. and Train, K. (2017). Correlation and scale in mixed logit models. Journal of choice modelling, 23:1–8.

Hole, A. R. (2007). Fitting mixed logit models by using maximum simulated likelihood. The Stata Journal, 7(3):388–401.

Howarth, R. B. and Farber, S. (2002). Accounting for the value of ecosystem services. Ecological economics, 41(3):421–429.

Hume, D. (1751). Philosophical essays concerning human understanding. M. Cooper.

Hynes, S., Hanley, N., and Scarpa, R. (2008). Effects on welfare measures of alternative means of accounting for preference heterogeneity in recreational demand models. American Journal of Agricultural Economics, 90:1011–1027.

Iamtrakul, P., Teknomo, K., and Hokao, K. (2005). Public park valuation using travel cost method. In Proceedings of the Eastern Asia Society for Transportation Studies.

James, P., Tzoulas, K., Adams, M. D., Barber, A., Box, J., Breuste, J., Elmqvist, T., Frith, M., Gordon, C., Greening, K. L., Handley, J., Haworth, S., Kazmierczak, A. E., Johnston, M., Korpela, K., Moretti, M., Niemelä, J., Pauleit, S., Roe, M. H., Sadler, J. P., and Thompson, C. W. (2009). Towards an integrated understanding of green space in the european built environment. Urban Forestry & Urban Greening, 8:65–75.

Janhäll, S. (2015). Review on urban vegetation and particle air pollution – deposition and dispersion. Atmospheric Environment, 105:130–137.

Jim, C. Y. and Chen, W. Y. (2006a). Perception and attitude of residents toward urban green spaces in guangzhou (china). Environmental Management, 38:338–349.

Jim, C. Y. and Chen, W. Y. (2006b). Recreation–amenity use and contingent valuation of urban greenspaces in guangzhou, china. Landscape and Urban Planning, 75:81–96.

Jim, C. Y. and Shan, X. (2013). Socioeconomic effect on perception of urban green spaces in guangzhou, china. Cities, 31:123–131.

Johnston, R. J., Jarvis, D., Wallmo, K., and Lew, D. K. (2015). Multiscale spatial pattern in nonuse willingness to pay: applications to threatened and endangered marine species. Land Economics, 91(4):739–761.

Kabisch, N., Qureshi, S., and Haase, D. (2015). Human–environment interactions in urban green spaces — a systematic review of contemporary issues and prospects for future research. Environmental Impact Assessment Review, 50:25–34.

Kaplan, R. and Kaplan, S. (1989). The experience of nature: A psychological perspective. Cambridge university press.

Karunarathne, L. and Gunawardena, P. (2020). Economic value of urban green space: A travel cost approach for viharamahadevi urban park, sri lanka. Journal of Tropical Forestry and Environment, 10.

Kerry, W. H. S. V. and Desvousges (1986). The Travel Cost Approach to Recreation Demand Modeling: An Introduction. Springer Netherlands.

Kinghorn, J., Snowball, J., Britz, P., and Weyl, O. (2014). A cross-model comparison of travel time inclusion techniques in recreational fishing demand analysis. Studies in Economics and Econometrics, 38(2):47–64.

Knapp, S., Aronson, M. F. J., Carpenter, E., Herrera-Montes, A., Jung, K., Kotze, D. J., Sorte, F. A. L., Lepczyk, C. A., MacGregor-Fors, I., MacIvor, J. S., Moretti, M., Nilon, C. H., Piana, M. R., Rega-Brodsky, C. C., Salisbury, A., Threlfall, C. G., Trisos, C., Williams, N. S. G., and Hahs, A. K. (2020). A research agenda for urban biodiversity in the global extinction crisis. BioScience, 71:268–279.

Lacey, A. and Luff, P. (2009). Telling the story: Foundational stylistic analysis for social scientists. International Journal of Social Research Methodology, 12:1–16.

Lafortezza, R., Carrus, G., Sanesi, G., and Davies, C. (2009). Benefits and well-being perceived by people visiting green spaces in periods of heat stress. Urban Forestry & Urban Greening, 8:97–108.

Lafortezza, R., Davies, C., Sanesi, G., and Konijnendijk, C. (2013). Green infrastructure as a tool to support spatial planning in european urban regions. iForest - Biogeosciences and Forestry, 6.

Lancsar, E., Louviere, J., Donaldson, C., Currie, G., and Burgess, L. (2013). Best worst discrete choice experiments in health: methods and an application. Social science & medicine, 76:74–82.

Lategan, L. G., Steynberg, Z., Cilliers, E. J., and Cilliers, S. S. (2022). Economic valuation of urban green spaces across a socioeconomic gradient: A south african case study. Land, 11.

Lee, J. A., Sneddon, J. N., Daly, T. M., Schwartz, S. H., Soutar, G. N., and Louviere, J. J. (2019). Testing and extending schwartz refined value theory using a best–worst scaling approach. Assessment, 26(2):166–180.

Lindholst, A. C., van den Bosch, C. C. K., Kjøller, C. P., Sullivan, S., Kristoffersson, A., Fors, H., and Nilsson, K. (2016). Urban green space qualities reframed toward a public value management paradigm: The case of the nordic green space award. Urban Forestry & Urban Greening, 17:166–176.

Lindsey, G. and Knaap, G. (1999). Willingness to pay for urban greenway projects. Journal of the American Planning Association, 65:297–313.

Lipovetsky, S. and Conklin, M. (2014). Best-worst scaling in analytical closed-form solution. Journal of choice modelling, 10:60–68.

Liquete, C., Cid, N., Lanzanova, D., Grizzetti, B., and Reynaud, A. (2016). Perspectives on the link between ecosystem services and biodiversity: The assessment of the nursery function. Ecological Indicators, 63:249–257.

Liu, W.-T., Chiou, C.-J., and Lin, J.-Y. (2014). Application of the travel cost method to measure the urban forest park's recreational value with first-visit satisfaction taken into account. Journal of Food, Agriculture & Environment, 12(2):1013–1024.

Liu, Z., Hanley, N., and Campbell, D. (2020). Linking urban air pollution with residents' willingness to pay for greenspace: A choice experiment study in beijing. Journal of Environmental Economics and Management, 104.

Lloyd-Smith, P., Abbott, J. K., Adamowicz, W., and Willard, D. (2019). Decoupling the value of leisure time from labor market returns in travel cost models. Journal of the Association of Environmental and Resource Economists, 6(2):215–242.

Lloyd-Smith, P., Abbott, J. K., Adamowicz, W., and Willard, D. (2020a). Intertemporal substitution in travel cost models with seasonal time constraints. Land Economics, 96:399–417.

Lloyd-Smith, P., Abbott, J. K., Adamowicz, W., and Willard, D. (2020b). Intertemporal substitution in travel cost models with seasonal time constraints. Land Economics, 96(3):399–417.

Lo, A. Y. H. and Jim, C. Y. (2012). Citizen attitude and expectation towards greenspace provision in compact urban milieu. Land Use Policy, 29:577–586.

Lockwood, M. and Tracy, K. (1995). Nonmarket economic valuation of an urban recreation park. Journal of Leisure Research, 27:155–167.

Loomis, J., Kent, P., Strange, L., Fausch, K., and Covich, A. (2018). Measuring the total economic value of restoring ecosystem services in an impaired river basin: results from a contingent valuation survey. Routledge.

Louviere, J., Lings, I., Islam, T., Gudergan, S., and Flynn, T. (2013). An introduction to the application of (case 1) best–worst scaling in marketing research. International journal of research in marketing, 30(3):292–303.

Luisetti, T., Turner, R. K., Bateman, I. J., Morse-Jones, S., Adams, C., and Fonseca, L. (2011). Coastal and marine ecosystem services valuation for policy and management: Managed realignment case studies in england. Ocean & Coastal Management, 54(3):212–224.

Macintyre, S., McKay, L., and Ellaway, A. (2019). Informing healthy urban planning in scotland: An overview and critique. Land Use Policy, 87:104049.

Madureira, H., Andresen, T., and Monteiro, A. (2011). Green structure and planning evolution in porto. Urban Forestry & Urban Greening, 10:141–149.

Madureira, H., Nunes, F., Oliveira, J., and Madureira, T. (2018). Preferences for urban green space characteristics: A comparative study in three portuguese cities. Environments, 5:23.

Mansfield, C., Pattanayak, S. K., McDow, W., McDonald, R., and Halpin, P. (2005). Shades of green: Measuring the value of urban forests in the housing market. Journal of Forest Economics, 11:177–199.

Marley, A., Islam, T., and Hawkins, G. (2016). A formal and empirical comparison of two score measures for best–worst scaling. Journal of choice modelling, 21:15–24.

Marley, A. and Pihlens, D. (2012). Models of best–worst choice and ranking among multiattribute options (profiles). Journal of Mathematical Psychology, 56(1):24–34.

Marley, A. A. and Louviere, J. J. (2005). Some probabilistic models of best, worst, and best–worst choices. Journal of mathematical psychology, 49(6):464–480.

Marschak, J. (1960). Binary choice constraints on random utility indications in: Arrow k, editor.

McConnell, K. E. (1992). On-site time in the demand for recreation. American Journal of Agricultural Economics, 74:918–925.

McConnell, K. E. (1995). Consumer surplus from discrete choice models. Journal of Environmental Economics and Management, 29:263–270.

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior.

McFadden, D. and Train, K. (2000). Mixed mnl models for discrete response. Journal of Applied Econometrics, 15(5):447–470.

McLeod, L. (2012). Textual analysis and the role of the researcher. Nurse Researcher, 19:7–18.

Mendes, I. and Proença, I. (2011). Measuring the social recreation per-day net benefit of the wildlife amenities of a national park: A count-data travel-cost approach. Environmental Management, 48(5):920–932.

Miller, R. W., Hauer, R. J., and Werner, L. P. (2015). Urban forestry: planning and managing urban greenspaces. Waveland press.

Mokas, I., Lizin, S., Witters, N., and Malina, R. (2019). Valuation of urban green spaces using virtual reality. 26/02/2019.

Moon, H. and Perron, B. (2008). Seemingly unrelated regressions. the new palgrave dictionary of economics. sn durlauf and le blume.

Nordhaus, W. D. and Kokkelenberg, E. C. (1999). Nature's numbers. Nat. Aca. Press, Washington, DC.

Obst, C., Hein, L., and Edens, B. (2016). National accounting and the valuation of ecosystem assets and their services. Environmental and Resource Economics, 64:1–23.

Obst, C. and Vardon, M. (2014). Recording environmental assets in the national accounts.

OpenStreetMaps (2023a). Location of malaysia.

OpenStreetMaps (2023b). Map of penang botanic gardens.

OpenStreetMaps (2023c). Map of penang youth park.

Ovaskainen, V., Neuvonen, M., and Pouta, E. (2012). Modelling recreation demand with respondent-reported driving cost and stated cost of travel time: A finnish case. Journal of Forest Economics, 18(4):303–317.

Palmquist, R. B., Phaneuf, D. J., and Smith, V. K. (2010). Short run constraints and the increasing marginal value of time in recreation. Environmental and Resource Economics, 46(1):19–41.

Pandey, A. K., Pandey, M., Mishra, A., Tiwary, S. M., and Tripathi, B. D. (2015). Air pollution tolerance index and anticipated performance index of some plant species for development of urban forest. Urban Forestry & Urban Greening, 14:866–871.

Panduro, T. E. and Veie, K. L. (2013). Classification and valuation of urban green spaces—a hedonic house price valuation. Landscape and Urban Planning, 120:119–128.

Parsons, G. R. (2003). The Travel Cost Model. Springer Netherlands.

Pathak, V., Tripathi, B. D., and Mishra, V. K. (2011). Evaluation of anticipated performance index of some tree species for green belt development to mitigate traffic generated noise. Urban Forestry & Urban Greening, 10:61–66.

Patterson, M. G., Cole, A. O., et al. (2013). Total economic value of new zealand's land-based ecosystems and their services. Ecosystem services in New Zealand–conditions and trends. Manaaki Whenua Press, Lincoln, New Zealand, pages 496–510.

Penang Global Tourism (2023a). Culture & heritage: Esplanade.

Penang Global Tourism (2023b). Nature & adventure: Sia boey.

Penang Green Council (2020a). Final report: Pegasus index 2019.

Penang Green Council (2020b). Penang green agenda 2030: Biodiversity and natural ecosystems.

Penang Green Council (2020c). Penang green agenda 2030: Green buildings and townships.

Penang Green Council (2020d). Penang green agenda 2030: Land use and planning.

Penang Institute (2020). Penang economic and development report 2019/2020.

Penang Institute (2023). Penang: Population and demographics.

Penang Town and Country Planning Department (2018). Rancangan struktur negeri pulau pinang 2030.

Peninsular Malaysia Town and Country Planning Department (2018). Rsnpp2030.

Peninsular Malaysia Town and Country Planning Department (2023). Integrated land use planning information system.

Phaneuf, D. J. and Requate, T. (2017). A course in environmental economics: Theory policy and practice.

Phaneuf, D. J. and Smith, V. K. (2005). Chapter 15 Recreation Demand Models, volume 2. Elsevier.

Popper, K. R. (1959). The logic of scientific discovery. Basic Books.

Rakotonarivo, O. S., Schaafsma, M., and Hockley, N. (2016). A systematic review of the reliability and validity of discrete choice experiments in valuing non-market environmental goods. Journal of environmental management, 183:98–109.

Renterghem, T. V. and Botteldooren, D. (2016). View on outdoor vegetation reduces noise annoyance for dwellers near busy roads. Landscape and Urban Planning, 148:203–215.

Rentziou, A., Gkritza, K., and Souleyrette, R. R. (2012). Vmt, energy consumption, and ghg emissions forecasting for passenger transportation. Transportation Research Part A: Policy and Practice, 46(3):487–500.

Revelt, D. and Train, K. (1998). Mixed logit with repeated choices: Households' choices of appliance efficiency level. Review of Economics and Statistics, 80(4):647–657.

Roberts, M., Glenk, K., and McVittie, A. (2022). Urban residents value multi-functional urban greenspaces. Urban Forestry & Urban Greening, 74:127681.

Roger, P. (2019). Natural resource and environmental economics. Pearson.

Samad, N. S. A., Abdul-Rahim, A. S., Yusof, M. J. M., and Tanaka, K. (2020). Assessing the economic value of urban green spaces in kuala lumpur. Environmental Science and Pollution Research, 27:10367–10390.

Sanesi, G., Padoa-Schioppa, E., Lorusso, L., Bottoni, L., and Lafortezza, R. (2009). Avian ecological diversity as an indicator of urban forest functionality. results from two case studies in northern and southern italy. Arboriculture & Urban Forestry (AUF), 35:80–86.

Scarpa, R. and Rose, J. (2008). Designs efficiency for non-market valuation with choice modelling: How to measure it, what to report and why. Australian Journal of Agricultural and Resource Economics, 52:253–282.

Scarpa, R., Thiene, M., and Hensher, D. A. (2012). Preferences for tap water attributes within couples: an exploration of alternative mixed logit parameterizations. Water Resources Research, 48.

Scarpa, R., Willis, K., and Acutt, M. (2005). Individual-specific welfare measures for public goods: A latent class approach to residential customers of yorkshire water.

Schaafsma, M. and Brouwer, R. (2020). Substitution effects in spatial discrete choice experiments. Environmental and Resource Economics, 75:323–349.

Schetke, S., Haase, D., and Kötter, T. (2012). Towards sustainable settlement growth: A new multi-criteria assessment for implementing environmental targets into strategic urban planning. Environmental Impact Assessment Review, 32:195–210.

Schipperijn, J., Ekholm, O., Stigsdotter, U. K., Toftager, M., Bentsen, P., Kamper-Jørgensen, F., and Randrup, T. B. (2010). Factors influencing the use of green space: Results from a danish national representative survey. Landscape and Urban Planning, 95:130–137.

Shaw, D. (1988). On-site samples' regression: Problems of non-negative integers, truncation, and endogenous stratification. Journal of econometrics, 37(2):211–223.

Shaw, W. D. and Jakus, P. M. (1996). Travel cost models of the demand for rock climbing. Agricultural and Resource Economics Review, 25:1–10.

Sheng, M. and Sharp, B. (2019). Aggregate road passenger travel demand in new zealand: A seemingly unrelated regression approach. Transportation research part A: policy and practice, 124:55–68.

Small, K. A. and Rosen, H. S. (1981). Applied welfare economics with discrete choice models. Econometrica: Journal of the Econometric Society, pages 105–130.

Smith, J. W., Siderelis, C., Moore, R. L., and Anderson, D. H. (2012). The effects of place meanings and social capital on desired forest management outcomes: A stated preference experiment. Landscape and Urban Planning, 106:207–218.

Smith, R. L. (1988). Extreme value theory for dependent sequences via the stein-chen method of poisson approximation. Stochastic processes and their applications, 30(2):317–327.

Smith, V. K., Desvousges, W. H., and McGivney, M. P. (1983). The opportunity cost of travel time in recreation demand models. Land Economics, 59(3):259–278.

Soe Zin, W., Suzuki, A., Peh, K. S.-H., and Gasparatos, A. (2019). Economic value of cultural ecosystem services from recreation in popa mountain national park, myanmar: a comparison of two rapid valuation techniques. Land, 8(12):194.

Swait, J. and Louviere, J. (1993). The role of the scale parameter in the estimation and comparison of multinomial logit models. Journal of marketing research, 30:305–314.

Sánchez, J. J., Baerenklau, K., and González-Cabán, A. (2016). Valuing hypothetical wildfire impacts with a kuhn–tucker model of recreation demand. Forest Policy and Economics, 71:63–70. Integrating Ecosystem Service Concepts into Valuation and Management Decisions.

Tapsuwan, S., Marcos-Martinez, R., Schandl, H., and Yu, Z. (2021). Valuing ecosystem services of urban forests and open spaces: application of the seea framework in australia. Australian Journal of Agricultural and Resource Economics, 65(1):37–65.

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

Thurstone, L. L. (1927). A law of comparative judgment. Psychological review, 34:273–286.

Train, K. and Weeks (2005). Discrete Choice Models in Preference Space and Willingness-to-Pay Space. Springer Netherlands.

Train, K. E. (1998). Recreation demand models with taste differences over people. Land economics, pages 230–239.

Train, K. E. (2003). Discrete Choice Methods with Simulation. Cambridge University Press.

Trojanek, R., Gluszak, M., and Tanas, J. (2018). The effect of urban green spaces on house prices in warsaw. International Journal of Strategic Property Management, 22:358+.

Twerefou, D. K. and Ababio, D. K. A. (2012). An economic valuation of the kakum national park: An individual travel cost approach. African Journal of Environmental Science and Technology, 6.

Tyrväinen, L., Pauleit, S., Seeland, K., and Vries, S. D. (2005). Benefits and uses of urban forests and trees. Urban forests and trees: A reference book, pages 81–114.

Tyrväinen, L. and Väänänen, H. (1998). The economic value of urban forest amenities: an application of the contingent valuation method. Landscape and Urban Planning, 43:105–118.

Tzoulas, K., Korpela, K., Venn, S., Yli-Pelkonen, V., Kaźmierczak, A., Niemela, J., and James, P. (2007). Promoting ecosystem and human health in urban areas using green infrastructure: A literature review. Landscape and Urban Planning, 81:167–178.

Umberger, W. J. and Mueller, S. C. (2010). Is presentation everything? using visual presentation of attributes in discrete choice experiments to measure the relative importance of intrinsic and extrinsic beef attributes. Technical report, Agricultural and Applied Economics Association.

United Nations et al. (2014a). System of environmental-economic accounting 2012: experimental ecosystem accounting. United Nations, New York.

United Nations et al. (2014b). System of Environmental-Economic Accounting 2012—Central Framework. United Nations, European Union, Food and Agriculture Organization of the United Nations, International Monetary Fund, Organisation for Economic Co-operation and Development, The World Bank.

United Nations et al. (2021). System of environmental-economic accounting—ecosystem accounting: Final draft. Technical report, DEPARTMENT OF ECONOMIC AND SOCIAL AFFAIRS, STATISTICS DIVISION, UNITED NATIONS.

Valck, J. D., Broekx, S., Liekens, I., Aertsens, J., and Vranken, L. (2017). Testing the influence of substitute sites in nature valuation by using spatial discounting factors. Environmental and Resource Economics, 66:17–43.

van den Berg, M., Wendel-Vos, W., van Poppel, M., Kemper, H., van Mechelen, W., and Maas, J. (2015). Health benefits of green spaces in the living environment: A systematic review of epidemiological studies. Urban Forestry & Urban Greening, 14:806–816.

Van Dongen, R. P. and Timmermans, H. J. (2019). Preference for different urban greenscape designs: A choice experiment using virtual environments. Urban Forestry & Urban Greening, 44:126435.

Villamagna, A. M., Angermeier, P. L., and Bennett, E. M. (2013). Capacity, pressure, demand, and flow: A conceptual framework for analyzing ecosystem service provision and delivery. Ecological Complexity, 15:114–121.

Voigt, A., Kabisch, N., Wurster, D., Haase, D., and Breuste, J. (2014). Structural diversity: A multi-dimensional approach to assess recreational services in urban parks. AMBIO, 43:480–491.

Vollmer, D., Ryffel, A. N., Djaja, K., and Grêt-Regamey, A. (2016). Examining demand for urban river rehabilitation in indonesia: Insights from a spatially explicit discrete choice experiment. Land Use Policy, 57:514–525.

Wan, C. and Shen, G. Q. (2015). Salient attributes of urban green spaces in high density cities: The case of hong kong. Habitat International, 49:92–99.

Wang, X. and Kockelman, K. M. (2007). Specification and estimation of a spatially and temporally autocorrelated seemingly unrelated regression model: application to crash rates in china. Transportation, 34:281–300.

Wanyu, L., ChyiRong, C., and JiunnCheng, L. (2014). Application of the travel cost method to measure the urban forest park's recreational value with first-visit satisfaction taken into account. Journal of Food Agriculture & Environment, 12:1013–1024.

Ward, F. A. and Loomis, J. B. (1986). The travel cost demand model as an environmental policy assessment tool: A review of literature. Western Journal of Agricultural Economics, 11:164–178.

Weitzman, M. L. (1976). On the welfare significance of national product in a dynamic economy.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. Econometrica: journal of the Econometric Society, pages 817–838.

White II, M. H. (2021). bwstools: An r package for case 1 best-worst scaling. Journal of choice modelling, 39:100289.

Wilkinson, A., Welch, C., and Young, C. E. (2015). Avoiding the pitfalls of word clouds: Social science, technology and scientific management in the 1930s. Business History, 57:1230–1244.

Willis, K. G. and Garrod, G. D. (1991). An individual travel-cost method of evaluating forest recreation. Journal of Agricultural Economics, 42:33–42.

Xu, C., Haase, D., and Pauleit, S. (2018). The impact of different urban dynamics on green space availability: A multiple scenario modeling approach for the region of munich, germany. Ecological Indicators, 93:1–12.

Yao, R. T., Scarpa, R., Turner, J. A., Barnard, T. D., Rose, J. M., Palma, J. H., and Harrison, D. R. (2014). Valuing biodiversity enhancement in new zealand's planted forests: Socioeconomic and spatial determinants of willingness-to-pay. Ecological economics, 98:90–101.

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

Zhang, F., Wang, X. H., Nunes, P. A. L. D., and Ma, C. (2015). The recreational value of gold coast beaches, australia: An application of the travel cost method. Ecosystem Services, 11:106–114. Marine Economics and Policy related to Ecosystem Services: Lessons from the World's Regional Seas.

Zhang, H., Chen, B., Sun, Z., and Bao, Z. (2013). Landscape perception and recreation needs in urban green space in fuyang, hangzhou, china. Urban Forestry & Urban Greening, 12:44–52.

Éva Terézia Vesely (2007). Green for green: The perceived value of a quantitative change in the urban tree estate of new zealand. Ecological Economics, 63:605–615. Ecological Economics of Coastal Disasters.

Özgüner, H. (2011). Cultural differences in attitudes towards urban parks and green spaces. Landscape Research, 36:599–620.