Summer 2015

Describing local community acceptance with discrete choice theory for enhanced community engagement

Sisi Que

Recommended Citation

DESCRIBING LOCAL COMMUNITY ACCEPTANCE WITH DISCRETE CHOICE THEORY FOR ENHANCED COMMUNITY ENGAGEMENT

by

SISI QUE

A DISSERTATION

Presented to the Faculty of the Graduate School of the

MISSOURI UNIVERSITY OF SCIENCE AND TECHNOLOGY

In Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

in

MINING ENGINEERING

2015

Approved by:

Kwame Awuah-Offei

Samuel Frimpong

Jason Baird

VA Samaranayake

Nathan W Weidner

ABSTRACT

This research sought to facilitate improved community (stakeholder) analysis by providing further insight on the determinants of local community acceptance using discrete choice theory. Specifically, the goals were to: (1) Identify, classify, and verify the important project characteristics and key demographic factors which affect local community acceptance of a mining project; (2) Account for the large number of relevant factors inherent in discrete choice experiments for mining community acceptance evaluation; and (3) Examine discrete choice models to select the most appropriate model for mining community consultation. The research will test the hypotheses that various discrete choice models can describe the local community's acceptance of mining projects.

Surveys were used to validate a classification of important mining project characteristics and demographic factors. Sixteen project characteristics and four demographic factors were identified as important for individual preferences for mining projects. A mixed style, blocking scheme, fractional factorial without interaction discrete choice experiment was proposed to overcome the challenge posed by the large number of relevant factors. The design was validated, revised, and implemented in Salt Lake City, UT to illustrate the usefulness of discrete choice theory in mining stakeholder analysis. Three candidate discrete choice models were evaluated to select the best model for mining stakeholder analysis. The results show that the conditional logit model, stratified by question, is the most suitable. The proposed approach has been demonstrated to answer three important questions for enhanced stakeholder analysis: (1) what are the factors that affect stakeholders' decision and how do these affect their preferences? (2) what is the effect of demographics on individual preferences? (3) what is the value of environmental and social impacts to individuals in the community?

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to all who helped and supported me through my PhD studies. First and foremost, I would like to express my appreciation for my graduate advisor, Dr. Kwame Awuah-Offei, for his constant support and guidance.

Furthermore, I would like to acknowledge my graduate committee members: Drs. Samuel Frimpong, Jason Baird, V. A. Samaranayake, and Nathan W. Weidner for their support and interest in my research.

Also, I appreciate Department of Mining & Nuclear Engineering, Missouri University of Science and Technology for giving me the opportunity to pursue my graduate studies. I also thank Ms. Tina Alobaidan, Ms. Barbara Robertson, Ms. Shirley Hall, Ms. Leanne Nuckolls, Ms. Diane Henke, and Ms. Judy Russell for the support and other administrative assistance.

Last, but not least, I would like to thank my husband for his unconditional love, my mother and father for their support and endless encouragement at different stages of my life.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

LIST OF TABLES

1. INTRODUCTION

1.1. BACKGROUND

Human development has been supported by metal and mineral products throughout human history and will continue in the future. Mining provides important products to meet society's needs, including raw materials for shelter, infrastructure, and manufacturing, and energy resources. World-wide, there are over 6,000 formal mining companies and 15 to 20 million artisanal and small scale miners operating in 30 countries (Ericsson & Löf, 2011; ICMM, 2012b). In the United States of America (USA), more than 14,000 mines mine for coal, metal ores and non-metallic minerals, according to 2012 data (National Mining Association, 2014).

The economic impact of US mining is summarized in Table 1-1. At the local level, mining provides a significant employment opportunity to the local community. In 2012, U.S. mines provided more than 634,000 jobs directly, and 1.27 million indirectly or induced. Thus, there is a total over 1.9 million full-time and part-time jobs created by US mining (National Mining Association, 2014). In addition, the direct labor income created by U.S. mining is over $46 billion with the total (direct, indirect and induced) exceeding $118 billion. At the national level, mining provides government revenues, foreign and domestic investment. According to National Mining Association (2014), mining activity (direct and indirect) generated total taxes of $46 billion. The contribution of US mining to the gross domestic product (GDP) is over than $225 billion in 2012.

Table 1-1. Economic contribution of U.S. Mining

Item	Direct	Indirect and Induced	Total
Employment	634,600	1,268,800	1,903,440
Labor Income (billions of dollars)	$46.2	$71.0	$118.2
Contribution to GDP (billions of dollars)	$102.1	$123.0	$225.1
Taxes Paid (billions of dollars)	$18.9	$26.9	$45.8

Source: Calculation based on IMPLAN modeling system (2012 database)

While there is no doubt that mineral and metal products make a significant contribution to development, the juxtaposed adverse impacts cannot be ignored. The negative environmental and social impacts of mining have attracted attention from governments, non-governmental organizations, the general public, and other stakeholders. In the past decade, concerns over sustainable development[1] have increased over the world (Dechant, Altman, Downing, & Keeney, 1994; Epstein & Roy, 2003; Freeman & Gilbert, 1988; Friedman & Miles, 2001; Gao & Zhang, 2006; Mathews, 1997; Rotheroe, Keenlyside, & Coates, 2003; Rowe & Enticott, 1998; Schaefer, 2004; Shrivastava, 1995). Mining companies cannot proceed with mining as in the past since global expectations have changed the role of business. How to contribute to sustainable development has become a key challenge for mining.

The industry has moved from environmental compliance (and associated standards like ISO 14001), to corporate social responsibility (CSR) programs, to social license to operate, and now to sustainability reporting with standards like the Global Reporting Initiative (GRI) (Brown, de Jong, & Lessidrenska, 2009; Browne, Stehlik, & Buckley, 2011; Hedberg & Malmborg, 2003; Thomson & Boutilier, 2011; A. Willis, 2003; Wood, 1991, 2010). Currently, most of the major mining houses produce audited annual sustainability reports that document their sustainability impacts (Fonseca, 2010). Additionally, there are calls for mines and mining businesses, like their counterparts in other sectors, to operate in a way that creates shared value for all stakeholders (Porter & Kramer, 2011). All these show that mines and mining businesses have a role to play in sustainable development of their host communities and the world at large. However, this can only be done with a concerted effort to develop and operate mines sustainably.

While the whole world benefits from mining's contributions, most of the resulting detrimental impacts on the environment and society fall on the local community or communities. Compared to other stakeholders, protecting the local communities' interests has been a key element of sustainable development efforts in mining (R Hamann, Patel, & Pressend, 2002; Ralph Hamann, 2003). Hence, community engagement is the key to sustainable development in mining, and also the main challenge for mines.

---

[1] The ability of current generations to meet their needs without compromising the ability of future generations to meet their own needs (Brundtland, 1987)

There are numerous examples of mining projects that have been postponed, interrupted, and even shut down due to poor community engagement (Browne et al., 2011; Davis & Franks, 2011; Moffat & Zhang, 2014; Prno & Scott Slocombe, 2012; Thomson & Boutilier, 2011). Stakeholder-related risk has been shown to be one of the major non-technical risks responsible for these delays (Ruggie, 2010; Davis and Franks 2011). Davis and Franks (2011) estimates the delay cost to be approximately US$ 10,000/day, during the exploration stage of a new mine. These costs are even higher during production when the costs of labor, equipment ownership, and deferred production are much higher. From a company's standpoint, community engagement is the best way to mitigate these community-related risks and achieve sustainable development.

1.2. STATEMENT OF PROBLEM

A key part of community engagement is community consultation, which includes three main parts: stakeholder identification, stakeholder analysis and iterative consultation (ICMM, 2012a; IFC, 2007). Stakeholder analysis is one of the key challenges since misunderstanding stakeholders will misguide the whole community consultation effort.

Discrete choice theory, based on the Nobel winning work by McFadden (1974), has transformed the world of market research. Discrete choice theory analyses an individual decision marker's preferences in discrete choices. Discrete choice theory, has been successfully used in econometrics and other disciplines to understand behavior in choice situations (Dimitropoulos & Kontoleon, 2009; Walekhwa, Mugisha, & Drake, 2009; K. Willis, Scarpa, Gilroy, & Hamza, 2011; Winslott Hiselius, 2005). IBM used discrete choice theory to study of the demand for laptop computers and reconfigure their product line to target various country-specific market segments. AT&T wireless used this modeling framework to assess demand for proposed wireless communication services (StatWizards LLC, n.d.). Also, choice theory has been used to evaluate community acceptance of renewable energy projects (K. Willis et al., 2011) and assess people's preferences for railway transportation of hazardous materials (Winslott Hiselius, 2005).

In mining, as far as the author is aware, only Ivanova, et al. have used discrete choice theory to understand the decision-making process of local communities regarding preferred mineral project development options (Ivanova, Rolfe, Lockie, & Timmer, 2007; Ivanova & Rolfe, 2011). The author hypothesizes that discrete choice theory can provide a framework for describing community acceptance of mining projects. This will provide additional information, hitherto unavailable, for stakeholder analysis and issue identification. With the increasing conflicts between mining companies and host communities (Hodge, 2014), it is crucial to develop methods to provide more insights for community engagement. Research is required to provide a general framework for including discrete choice theory (discrete choice experiments and modeling) into improved community consultation in mining. This will provide further insight on the determinants of local community acceptance and the relationships between those determinants.

Successful application of discrete choice theory will allow mining companies to better understand what kind of mining project the community prefers and which demographic factors are crucial in dividing opinions. Discrete choice experiments can be used to obtain, from respondents, preferred mine developments from several choice sets. By identifying patterns in these choices, discrete choice models how different individuals respond to different mine development options. Discrete choice modeling allows a mining company to examine the effect of each mining project attribute (or characteristic) on individual and community preferences. Compared to traditional stakeholder analysis methods, a mining company will have quantitative tool for planning, designing, and managing a mining project. This data driven community consultation could facilitate better community engagement, enhance social license to operate, and, hopefully, lead to reduced conflicts between mines and the host communities.

The three main challenges of a discrete choice theory framework for stakeholder (community) analysis are: (1) How do you identify, classify, and verify the important factors (attributes of the mining project) that may affect local community acceptance of a mining project? (2) How do you design effective discrete choice experiments with large number of relevant factors, without overloading respondents? (3) How do you select the

most appropriate discrete choice model to describe the local community's acceptance of mining projects?

The first challenge can affect the success of the whole discrete choice experiment and model. The factors considered need to be broad enough to cover the key issues that might be important to different respondents, and easy enough to provide useful feedback. Ivanova et al. (2007) and Ivanova & Rolfe (2011) tracked five and seven mining project characteristics, respectively (Ivanova et al., 2007; Ivanova & Rolfe, 2011). Further work, with emphasis on identifying the key mining project characteristics from the plethora of candidate characteristics, is required to improve the reliability of discrete choice models and further refine how this approach can be used in community analysis. Pursuant to this challenge, three further questions have to be answered: (1) How do you identify the important mining project characteristics for discrete choice experiments? (2) How do you find the key demographic factors, which are significant vis-à-vis people's perception of the importance of the mining characteristics? (3) Is there a difference between attitudes of people who live in mining and non-mining communities (i.e. people with and without significant mining experience)? Without answers to these three important questions, discrete choice experiments and modeling would not be efficient and effective, nor produce valid models to help with community analysis.

The second huddle of incorporating discrete choice modeling into mining community analysis is how to design good discrete choice experiments (DCEs) for mining community consultation. For effective and efficient discrete choice experiment design, there are three important questions that cannot be ignored: (1) What is the optimum number of factors to consider in one choice set? (2) How do you design discrete choice experiments for mining community consultation? (3) How do you validate the discrete choice experiment design? Without answers to these questions, discrete choice experiment design would not yield useful data to help with community analysis.

The final challenge is how to select the most appropriate discrete choice model to describe local community's acceptance of mining projects. This task involves: (1) conduct a comprehensive literature review of discrete choice models; (2) identify the candidate discrete choice models for mining community acceptance modeling; and (3)

evaluate the goodness-of-fit of the candidate discrete choice models to select the most suitable discrete choice model for mining community acceptance.

This PhD study sought to overcome the above mentioned technical challenges of applying discrete choice theory to community consultation for mining projects. This work will be a significant contribution to knowledge and the literature on community analysis in mining. The research provides a framework for effective and efficient discrete choice experiments and modeling.

1.3. RESEARCH OBJECTIVES AND SCOPE

The goal of this PhD research is to facilitate improved community (stakeholder) analysis by providing further insight on the determinants of local community acceptance using discrete choice theory. Pursuant to the overall goal of this study, the specific objectives are to:

(1) Identify, classify, and verify the important mine characteristics and key demographic factors which affect local community acceptance of a mining project;

(2) Account for the large number of relevant factors inherent in discrete choice experiments for mining community acceptance evaluation; and

(3) Examine discrete choice models to select the most appropriate model for mining community consultation.

The research will test the hypotheses that various discrete choice models can describe the local community's acceptance of mining projects.

The research has two main limitations that need to be clarified. First, this research provides a general framework for including discrete choice theory into improved community consultation in mining. In the case study, discrete choice experiment (survey) is designed for a specific mining community (Salt Lake City, Utah, USA) to illustrate how to conduct such experiments. Thus, the resulting model applies to the target mining community only. However, the general framework and research approach can be used for other mining communities and even other fields. Secondly, the discrete choice modeling advocated in this framework treats all participants, equally. Thus, this model can be applied for the groups in which individuals have equal rights to support or reject a project.

If other researchers want to employ this framework for multi-stakeholders (employees, customers, affected communities and the general public) at the same time, a possible approach is to establish one discrete choice model (DCM) for each group, and combine the results.

1.4. RESEARCH METHODOLOGY

Figure 1-1 presents the research framework adopted in this work.
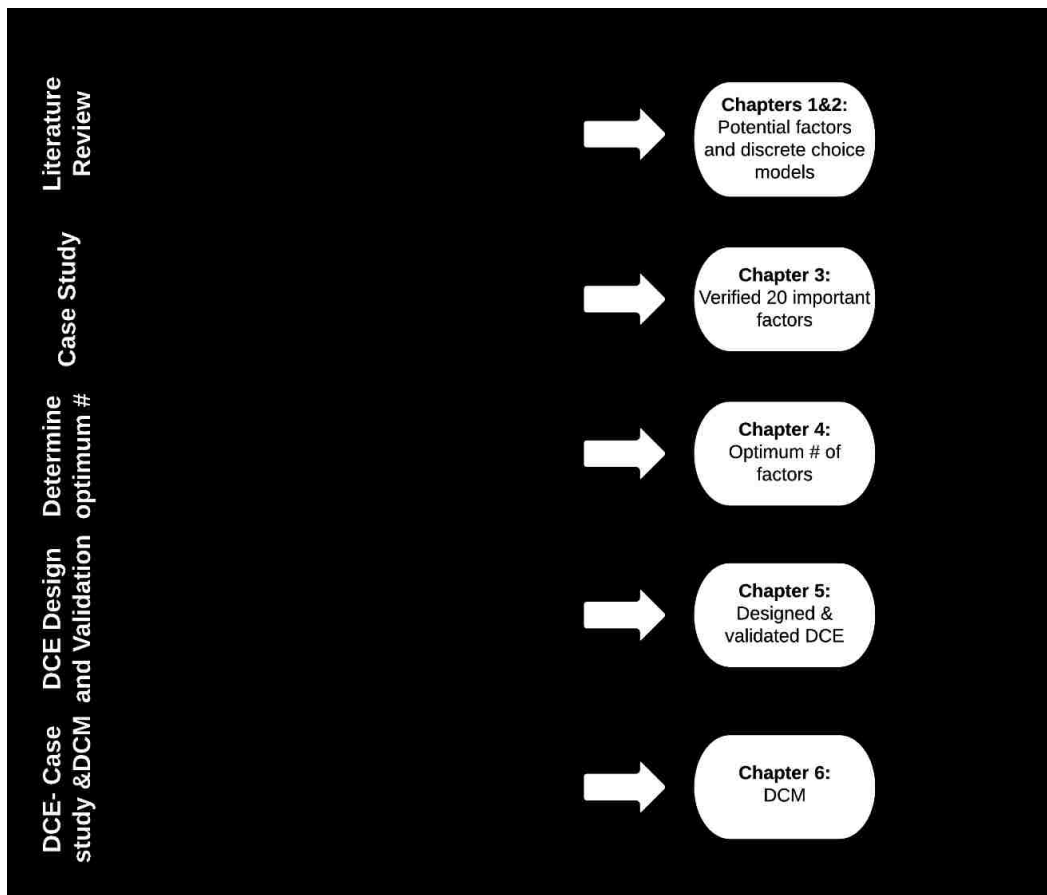
Figure 1-1. Activities/tasks in this research

To achieve objective (1), critical literature review was used to identify and classify the potential important mining characteristics and demographic factors that affect community acceptance of mining project. Online surveys were used to capture respondents' perception of the level of importance of the identified factors in their decision to support a mining project. Relevant statistical analysis is used to determine the

important demographic factors and verify the classification of the mining characteristics. The research also explored differences in the level of importance data from respondents from mining and control populations.

For objective (2), surveys were designed with choice sets that had the number of factors varied from three to six. An online survey was used to identify the optimum number of factors to be included in each choice set (set of alternatives) for respondents in mining community consultation. Then, a "blocking scheme" discrete choice experiment (DCE) was designed using the determined optimum number factors and the verified classification of factors from the results of objective (1). To address DCE validation, which is a major challenge, a focus group study was used to capture respondents' perception of the difficulty, clarity, reliability and validity of the DCE. Statistical analysis was used to analyze the level of difficulty and clarity from the focus group results.

Finally, the main research hypothesis was tested with data from the discrete choice experiments to achieve objective (3). The survey was conducted in Salt Lake City with more than 600 participants. The Statistical Analysis System (SAS) procedures were used to fit the DCE data to the conditional logit, strata conditional logit, and mixed logit models. Then, the results of different models were compared to determine the most suitable discrete choice model for mining community consultation.

1.5. STRUCTURE OF THE DISSERTATION

This dissertation contains seven sections. The rest of the dissertation is structured as follows. Section 2 presents a review of relevant literature. Section 3 discusses research on how to identify and classify critical factors for discrete choice experiment. Section 4 presents research on how to determine the optimum number of factors for mining community consultation using discrete choice experiments. An approach for discrete choice experiment design for mining community acceptance is provided in Section 5. A case study of discrete choice experiment and discrete choice modeling is presented in Section 6. Section 7 provides the conclusions of this study and recommendations for future work.

2. LITERATURE REVIEW

2.1. SUSTAINABILITY AND COMMUNITY ENGAGEMENT

Sustainable development is defined as the ability of current generations to meet their needs without compromising the ability of future generations to meet their own needs (Brundtland, 1987). Since sustainable development includes social, economic and environmental impacts, these have been referred to widely as the triple bottom line (Munashinge, M. and Shearer, 1995). Also, sustainable development has been defined in relation to social, natural, human, physical, and financial capital (the five capitals) (Goodwin, 2003). Sustainable development has been said to be 'an ambitious new project intended to act as the focus of human endeavor in the twenty-first century' (Meadowcraft, 2000).

In the past decade, concerns about corporate sustainability have increased over the world (Dechant et al., 1994; Epstein & Roy, 2003; Freeman & Gilbert, 1988; Friedman & Miles, 2001; Gao & Zhang, 2006; Mathews, 1997; Rotheroe et al., 2003; Rowe & Enticott, 1998; Schaefer, 2004; Shrivastava, 1995). Besides the bad publicity from environmental misadventures and the resulting stricter government legislation and public pressures, poor sustainability performance affects the triple-bottom line and long-term profitability of a business. Thus, businesses have both an interest and a responsibility to incorporate sustainable development into their long-term business strategy (Elkington, 1997; Gao & Zhang, 2006; Grant, 1997; Johnson & Scholes, 1993; Russo & Fouts, 1997).

Sustainable development can only be given real meaning by investigating the ideas through a multi-stakeholder approach (Rotheroe et al., 2003). A stakeholder is any group or individual who can affect or is affected by the achievement of the organization's objectives (Freeman, 1984). The Institute of Social and Ethical Account Ability (ISEA, 1999) defines stakeholder engagement as "the process of seeking stakeholder views on their relationship with an organization in a way that may realistically be expected to elicit

them." A mining project and its stakeholders are interdependent. This relationship is confirmed by Rotheroe (2003), who indicates that industry has to engage stakeholders in the decision-making process and throughout the whole project to achieve sustainable development (Cheney & Christensen, 2001).

In recent years, mining has witnessed an increasing demand for sustainable development from the public and regulators, as well as internal advocates who cite the sector's own long-term benefit (Hodge, 2014). Many mining companies realize the important role of other stakeholders and emphasize stakeholder engagement in the process of mine planning and design, operation, and closure. For the mining sector, ICMM (2012) defines stakeholders as a comprehensive list of people and groups who may be affected by, can affect, or have an interest in a project. Examples include the local and indigenous groups, employees and contractors, labor unions, suppliers, governments and regulators, media, non-governmental organizations, and investors (BHP Billiton, 2014; Rio Tinto, 2012). In mining industry terms, the community is generally defined as the inhabitants of the immediate and surrounding areas who are affected by a company's activities (MCMPR, 2005). Actually, local communities are the first stakeholder on the International Council on Mining and Metals (ICMM) Checklist of possible stakeholders (ICMM 2012).

It is increasingly evident that mining community engagement is important for successful of mining operations (indeed, for all industrial activity). The examples of mining projects that have been disrupted due to lack of community support, cited earlier are proof of this (Browne et al., 2011; Davis & Franks, 2011; Moffat & Zhang, 2014; Prno & Scott Slocombe, 2012; Thomson & Boutilier, 2011). Community engagement is critical for obtaining permits prior to commencing mining. Actually, community

acceptance is a requirement for the permitting process in some jurisdictions (e.g. Peru[2]).
In the USA, the local community's acceptance is not necessarily a requirement for
granting a permit. However, public participation is required during environmental impact
assessment (EPA, 1998).

This concept of community approval of mining operations and its relationship to
socio-political risk has been formalized as the social license to operate, in the last decade
(Thomson & Boutilier, 2011). The social license to operate (SLO) is defined as a
community's perceptions of the acceptability of a company and its local operations
(Thomson & Boutilier, 2011). SLO is inversely proportional to the level of socio-political
risk faced by a mining operation. For instance, it has been shown that the time it takes for
the major international oil companies to bring a project online nearly doubled in the
decade preceding 2008, with the delay adding significant extra costs to projects (Davis &
Franks, 2011). Community-related risk has been shown to be one of the major non-
technical risks responsible for these delays (Davis & Franks, 2011). For a mining project,
the cost of delays can be equally significant. As stated earlier, Davis and Franks (2011)
estimates the delay cost to be approximately US$ 10,000/day, during the exploration
stage of a new mine. Good community engagement is the best way to mitigate these
community-related risks.

Currently, some mines are gradually coming to understand the special importance
of the host community, and are attempting to address this issue by referring to local
communities as 'primary' or 'key' stakeholders. However, even with increased effort
the mines and mining businesses still struggle to avoid community conflict. In fact, there
appears to be a rise in conflict in the face of increased community engagement from
mines (Hodge, 2014).

---

[2] Peru passed a Law on the Right of Consultation of Indigenous Peoples in 2011 in
accordance with various international conventions they had ratified.

2.2. COMMUNITY ENGAGEMENT IN MINING

The decision-making process is greatly impacted by the characteristics of stakeholders, whether individuals, groups or organizations. Stakeholder analysis is the tool to analyze this impact and has gained increasing popularity in the last decade. Stakeholder analysis approaches have a long history in business management applications starting from the 1930s (Clarkson MBE, 1995). However, not until the 1990s, the techniques were not considered useful for analyzing the policy making process (Anonymous, 1996). In the last few decades, the usefulness of stakeholder analysis has been investigated by a number of researchers (Brugha & Varvasovszky, 2000; Clarkson MBE, 1995; Gregory & Keeney, 1994; Hill & Jones, 1992; Thomas & Palfrey, 1996; Thompson, 1996).

Stakeholder analysis is the process of understanding the behavior and interests of a group of targeted stakeholders, who have the potential to influence an organization, project, or policy direction, through surveys and data analysis (B Crosby, 1992; R. Mason & Mitroff, 1981; Walt, 1994). The results are used to manage stakeholders by knowing and satisfying their preferences and facilitating the decision making processes. Stakeholder analysis is also helpful for policy makers or managers to better understand stakeholders as a basis for formulating better policies or management strategies.

The basic analysis technique is described by Bryson (1995). It offers a quick and useful way of: identifying stakeholders and their interests, clarifying stakeholders' views of a local organization, identifying some key strategic issues and beginning the process of identifying coalitions of support and opposition. Bryson describes how this technique was used to bring about major change in a state department of natural resources in the United States, because it showed participants how existing strategies ignored important stakeholders – who refused to be ignored – as well as what might be done to satisfy the stakeholders. The technique involved nine steps, starting with brainstorming to find the

list of potential stakeholders and ending with identifying and recording longer-term issues with individual stakeholders and with stakeholders as a group (Bryson 1995).

Currently, the most accepted stakeholder analysis method was published by Reed et al. (2009), which has been cited 165 times in the literature. The method, shown in Figure 2-1, has three main steps: (i) identifying stakeholders; (ii) differentiating between and categorizing stakeholders; and (iii) investigating relationships between stakeholders.
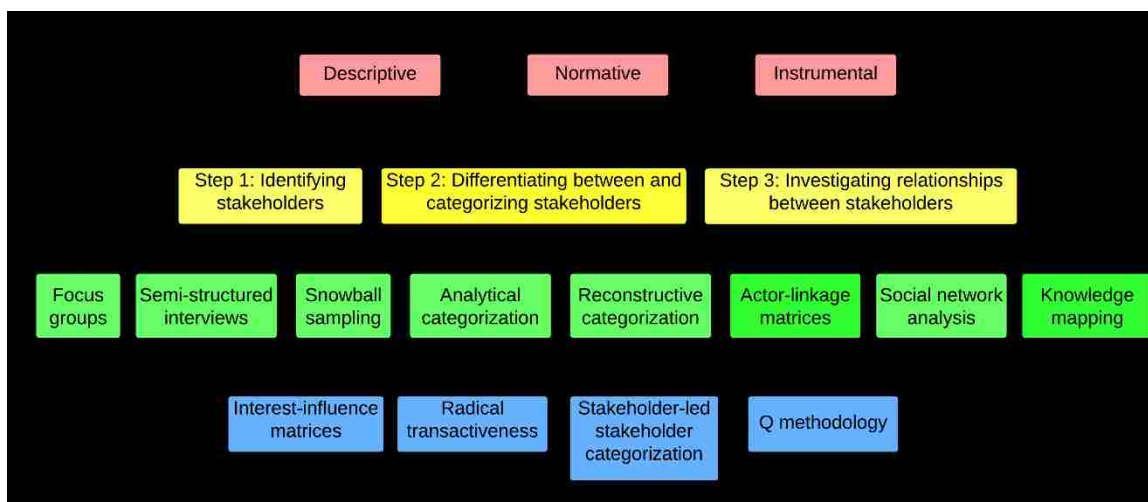
Figure 2-1. Schematic representation of rationale, typology and methods for stakeholder analysis (Reed et al., 2009)

In mining, organizations like the International Finance Corporation (IFC) and International Council on Mining & Metals (ICMM) have discussed stakeholder engagement in varying degrees (ICMM, ICRC, IFC, 2011; ICMM, 2008, 2009, 2010, 2012a; IFC, 1998, 2007, 2009, 2010a, 2010b). The literature contains many contributions in this area (Azapagic, 2004; Davis & Franks, 2011; Gunningham & Sinclair, 2009; Jenkins & Yakovleva, 2006; Kempe, 1983; Moffat & Zhang, 2014; O'Faircheallaigh, 2012; Thomson & Boutilier, 2011). There is a burgeoning method that has developed for stakeholder engagement in the mining industry, which includes three main parts, as

recommended by Reed et al. (2009): stakeholder identification, stakeholder analysis and iterative consultation (ICMM, 2012a; IFC, 2007).

This stakeholder analysis procedure, currently in practice, is likely to remain the key evaluation process through which stakeholder opinions are assessed in a mining project. The most widely used method for stakeholder analysis is suggested by the International Council on Mining & Metals (ICMM, 2012a). This method requires the analyst(s) to evaluate each stakeholder's view of the project (positive, neutral, negative), how influential they are (high, medium, low) and how greatly they will be impacted by the project (high, medium, low). Stakeholders' information is filled in a stakeholder analysis matrix (Table 2-1), and then classified into three groups: highly influential supporter of the project, neutral about the project, and highly influential opponent of the project. The result of stakeholder analysis is critical and provides the key to evaluating stakeholder opinions of a mining project during the iterative consultation process. Thus, stakeholder analysis affects the whole stakeholder engagement process.

Table 2-1. Stakeholder analysis matrix (ICMM, 2012a)

Name/group of stakeholders	View of project			Influence			Impact		
	Pos	Neutral	Neg	H	M	L	H	M	L
X	✓			✓					✓
Y		✓			✓				✓
Z			✓	✓			✓		

While the local communities are listed first in ICMM (2012) checklist of possible stakeholders, their special status does not lead to any special attention in the stakeholder analysis procedure. Compared to other stakeholders (such as government, internal company stakeholders like employees and unions, and regulators), the local community is

the most voiceless group but, often, has the most diverse opinions and diversity in demands. This is heightened in cases where mining occurs on land belonging to indigenous people (Native Americans), and poor and disadvantaged communities. This particularity makes community engagement in mining difficult, requiring special attention and unique methods for stakeholder analysis (IFC, 2007).

Current stakeholder analysis processes for engaging local communities (ICMM, 2012a) are mainly qualitative, using public forums, surveys, analysis of comments to public announcements of permit application and others. The goal for the stakeholder analysis is to understand the local community by classifying community into three groups: highly influential supporter, neutral, and highly influential opponent of the project (Table 2-1). This is not enough to ensure the success of the whole consultation process. The goals of community analysis should include: (1) what are the factors that affect stakeholders' decision and how these affect the decision? (2) what is the effect of demographics on individual preferences? (3) what is the value of environmental and social impacts to individuals in the community?

Current qualitative community analysis methods alone may not provide enough insight into the community's needs, concerns, and level of acceptance to achieve the goals of community analysis process. Additional methods (qualitative or quantitative) that provide unique insight can be helpful in providing information that is not currently available. This will ensure mines and mining businesses target the right people in the community and focus on the right issues, in their community engagement. There is a need for some quantitative methods, including computer modeling, to augment the current qualitative methods. Results obtained by such analysis should, however, be complemented by the insights gained through other methods of analyzing communities' preferences.

Discrete choice theory, based on the Nobel winning work by McFadden (1974) has transformed the world of market research. As a statistical analysis method, discrete choice theory aims at analyzing individual decision marker's preferences. Discrete choice modeling can help us understand what kind of mining project individuals in a community prefer by comparing different hypothetical options. By identifying patterns in these choices, discrete choice models will provide insight into how different individuals respond to different mining options. DCM will allow mining companies to examine the significance of different mining impacts (including social, economic, and environmental) and other aspects of a project on the preferences of different groups of in the local communities. Compared to traditional stakeholder analysis methods, the mining company will have a quantitative tool for planning, designing, operating, and managing the mining project to facilitate better community engagement.

As far as this author is aware, only Ivanova, et al. have used discrete choice theory to understand the decision-making process of local communities regarding preferred mineral project development options (Ivanova et al., 2007; Ivanova & Rolfe, 2011). Discrete choice theory can provide a framework for successfully describing community acceptance of mining projects, which can be incorporated into community engagement activities in mining. This area of research is under-explored and further research is required to formulate such a framework.

2.3. FACTORS THAT AFFECT COMMUNITY ACCEPTANCE

There are many factors that can affect an individual's perception of a mining project, which in turn affects whether he/she supports the mine or not. There is a lot in the literature on this subject. Generally, the factors that affect community acceptance are the impacts of the mine on the environment and host community, the mine owner (the corporate reputation etc.) and governance issues, and demographics of the community.

2.3.1. Mining Impacts. The mining impacts have positive impacts, negative impacts, and other impacts.

2.3.1.1. Positive impacts. Mining operations can result in three obvious positive impacts: job opportunities, income increase, and infrastructure improvement.

The impact of job opportunities and related economic impacts (income increases) were introduced in Section 1 and summarized in Table 1-1. In 2012, U.S. mines provided more than 634,000 jobs directly, and 1.27 million indirectly or induced (National Mining Association, 2014). ICMM (2012a) describes job opportunities as the first issue and most often asked question by members of local communities is, "how many jobs will go to their community members", when they hear that a mine may be developed in their community. Income increases due to higher paying jobs and/or the unemployed joining the mine's supply chain is another important impact of mining (ICMM, 2012a; Petkova, Lockie, Rolfe, & Ivanova, 2009). The direct labor income created by U.S. mining is over $46 billion with the total (direct, indirect and induced) exceeding $118 billion in 2012 (National Mining Association, 2014). Petkova et al.(2009) indicate that the relatively high incomes of people working in the mining and allied industry were seen, by the local community, to generate positive impacts on all towns.

Infrastructure improvement is another obvious positive impact of mining, and it includes educational institutions, health services, power and water supply, sewerage and sanitation, transport infrastructure including roads, rail, air and sea transport and the accessibility of services ICMM (2012a). Some of this investment in infrastructure is for business purposes (for instance, a quarry needs to improve roads so their product can be transported efficiently to market). However, a significant portion also comes through corporate social responsibility programs that invest in the host community. For example, in BHP Billiton's 2014 sustainability report, the company reports that its commitment to

invest 1% of pre-tax profits in community programs resulting in $241.7 million invested in community programs (BHP Billiton, 2014).

 2.3.1.2. Negative impacts. However, mining also has juxtaposed adverse impacts, including environmental pollution, increases in housing costs, labor shortages for other businesses, traffic and crime increase. The environmental issue is the main issue of the anti-mining movement and the first reason for rejecting mining. The environmental impacts include water use and pollution, air, land, and noise pollution.

 The United States Geological Survey (USGS) estimates a drop of 300 meters in the water table of the areas surrounding open-pit mines in Nevada, due to the mining water demand (Rockwell, 2000). The Betze-Post mine alone pumps out 380,000 cubic meters (100 million gallons) of groundwater per day (Solley, Pierce, & Perlman, 1999). Acid mine drainage at the Summitville gold mine in Colorado alone destroyed all the biological life within seventeen miles of the Alamosa River. The place was designated a Federal Superfund site and the Environmental Protection Agency (EPA) spent $30,000 a day in treating the drainage (Earthworks and Oxfam America, 2004). Opponents of mining are concerned about potential environmental impacts, in particular, possible water contamination (ICMM, 2010).

 The contaminated water will contaminate the land, resulting in significant impacts to terrestrial ecosystems, including accumulation of toxic elements in soil, soil acidification, damage to soil biota, loss of soil fertility, plant contamination, plant toxicity, and food chain contamination (Dudka & Adriano, 1997). Solid waste is another big issue, since mining products are, mostly, a small fraction of total excavated mass. In gold mining, one ton ore may be refined into only 1 gram gold, with the rest being waste. In addition, several tons of barren rock may be mined to expose the ore or valuable material. The amount of solid waste tends to increase with time since improved mining technology makes it possible to exploit low-grade deposits with time.

Air pollution is another important impact. The major area of concern is dust –
from excavation and transportation, causing air quality degradation (ICMM, ICRC, IFC,
2011). In addition, the processing (including refining) of material produces pollutants
(e.g. oxides of nitrogen and sulfur) that pollute the air. Worldwide, smelters add 142
million tons of sulfur dioxide to the atmosphere every year—13 percent of global
emissions (Earthworks and Oxfam America, 2004).

Noise pollution results from traffic, blasting and operating heavy machinery
(ICMM, ICRC, IFC, 2011). Noise pollution is the single largest type of community
complaint (ICMM, 2009). BHP Billiton reports that out of 536 complaints in 2008, 200
were related to noise (BHP, 2008). Ivanova & Rolfe (2011) also identified noise impacts,
together with vibration and dust, as a significant factor (90% confidence) in explaining
community members' preferences for mining developments.

Beside the environmental issues, increases in housing costs and labor market
shortages are juxtaposed negative impacts of mining projects. Petkova et al. (2009) did a
qualitative social impact assessment of post-development impacts of mining on six
communities in the Bowen Basin in Queensland, Australia, following the boom in coal
prices between 2003 and 2008. The result of accommodation and staff shortages are
shown in Table 2-2. The ten years growth rate of median weekly rents from 1998 to 2008
were all at least 160% for the five studied communities with reported data. The found
accommodation in short supply and expensive in all six surveyed communities. Also,
Ivanova and Rolfe (2011) found 'housing and rental prices' were significant at 5% level
for explaining preferences for mine development options. Mining can lead to labor
shortages, especially for other businesses in the local community that cannot compete
with large mines for talent. Labor shortage for other business is listed as '*staff shortages'*
in Petkova et al., (2009), which occurred at five from six assessed mining communities.

Table 2-2. Accommodation and staff shortages (Petkova et al., 2009)

Median weekly rents	1998	2003	2008	Growth rate 1998-2008	Accommodation in short supply and expensive	Staff shortages
Moranbah	137	235	680	+393%	X	X
Nebo	117	220	450	+283%	X	X
Rolleston	80	85	220	+175%	X	X
Blackwater	145	140	380	+162%	X	X
Springsure	100	137	260	+160%	X	X
Coppabella	N/A	N/A	N/A	N/A	X	N/A

The tendency of traffic and crime increase in mining regions should also be encapsulated in the analysis. Two social impact assessment (SIA) studies of Central Queensland's Coppabella coal mine were undertaken in 2002–2003 and 2006–2007 to provide a reference point for predictive assessments of proposed resource extration projects (Lockie, Franettovich, Petkova-Timmer, Rolfe, & Ivanova, 2009). The study reports that residents have the perception of increased crime risk and believe that crimes against property and general anti-social behaviour were accelerating in the community. Although, the police reported that any increase in criminal activity was proportional to population growth from 2003 to 2006, it still represent an absolute increase in the criminal activity. This criminal activity increase is supported by other research results. Hajkowicz et al. (2011) suggests that the indicators on crime, domestic violence, and alcohol abuse reflect serious social problems in mining communities. The norms for acceptable levels of alcohol consumption are higher within the mining workforce, for example (Midford et al., 1997).

Traffic increase has also been observed in the two social impact assessment (SIA) studies (Lockie et al., 2009). Residents believed that traffic volumes and accidents have increased, including the large trailers and mining equipment. Road use statistsics indicate that traffic volumes did increase with the bulk of additional traffic associated with miners traveling between their places of employment and residence area. In addition, the

increased road traffic and incidence of drivers travelling home while fatigued following end of shift were documented in the mining communities by comparing the impacts identified in independent studies of Coppabella Coal Mine and eight other EIA studies in the Bowen Basin (Ivanova et al., 2007; Lockie et al., 2009).

     2.3.1.3. Other impacts. Besides the obvious positive and negative impacts, mining projects can also cause population increases and culture impacts. Also, two additional attributes of the mine affect the communities perception of the intensity and duration of impacts: mine buffer (how far the mine is from the community) and life (duration of mining operation).

     A consequence of a boom in mining is the associated population growth, especially in small community without enough skilled labors (Lockie, Franetovich, Sharma, & Rolfe, 2008). Resource exploitation can be directly linked to local population changes as there is often population growth from migrants looking for job opportunities. This is shown by the population census of four mining communities in the Bowen Basin, Queensland, Australia in Table 2-3 (Petkova et al., 2009). The mining boom started at 2001 and the population growth is apparent in four of six studied communities. The six years population growth from 2001 to 2006 are varied from +2.4% to 18.5%.

Table 2-3. Description of case study communities

No. permanent residents	Blackwater	Moranbah	Nebo	Springsure
2001*	4,913	6,124	238	770
2006**	5,031	7,133	282	829
Growth rate 2001-2006	+2.4%	+16.5%	+18.5%	+7.7%

*ABS (2001); ** ABS (2006)

     A new mining company and migrants looking for job opportunities have impacts on the community's way of life, culture and traditions. Indigenous populations have been particularly affected and their traditional ways of life changed, sometimes, without their consent. The diverse cultural backgrounds of the mining communities and management

styles of the mining companies are a factor in determining the extent of this impact (Sassoon, 1998). ICCM specifically identifies cultural (heritage) impacts as a factor in community engagement (ICMM, 2012a). Cultural impacts include any effects on the cultural norms and practices, which include effects on intangible and tangible cultural heritage, and access to and vibrancy of cultural facilities. This will be of critical importance when indigenous peoples are present within the area of impact for the mining project.

Beside the above two factors, mine buffer and life affect the mining community's perceptions of the impacts of the mine. Community opposition to a mining operation is an all too familiar picture, and this phenomenon has been called the 'not in my backyard syndrome' or NIMBY-ism (MPE, 2011). The key to NIMBY opposition is the location of the proposed construction. The 'backyard' has grown so vastly that, today, NIMBY-ism affects companies all over the world. Ivanova and Rolfe (2011) found 'buffer for mine impacts' to be a significant (at 5% level) factor that explains community preferences for mine developments. The mine life is a measure of the persistence of all impacts (positive and negative). So it determines how long the job opportunities and noise impacts, for example, will last. It is a measure of the 'length contract' of contract, which has been found to be a significant factor (at 1% level) for explaining local acceptability of renewable energy adoption in an ageing population (K. Willis et al., 2011).

2.3.2. Mine Owner and Governance. While indigenous peoples are custodians of the land, the mine owner and government regulators decide how to design, plan, process and manage the mine. The way these decisions are made have a significant impact on the communities perception of the mine's owners and government. Hence, a key factor is governance, including the mechanism for making permit decisions and availability of transparent information.

The decision making mechanism and availability of independent and transparent information are complement each other. Local communities need to have the right to be engaged in the decision making mechanism first, then independent and transparent information is meaningful for them. And available independent and transparent information is the foundation for them to make meaningful decisions in the engagement. The decision making mechanism describes how decisions are made when disagreements arise on the impacts (positive and negative) of mining. The information refers to all information relevant to the decision to permit a mine including reports mining impacts and baseline studies as will be contained in an environmental impact assessment (EIA), for example. These decision making mechanisms vary from the purely legal (i.e. the mining company meets the regulatory requirements) to those that take cognisance of the SLO and seeks legitimacy (Muradian, Martinez-Alier, & Correa, 2003). The information is provided by the mining company and/or government currently. The local community often does not trust the available information on the potential impacts from both sources (ICMM, 2012a). The information should be independent and transparent, and should be provided by multiple groups with technical expertise but no commercial stake in the industry. The information should cover both the broad industry and also relating to specific proposals, which can facilitate local community participation in the decision-making and help the community develop.

2.3.3. Community Demographics. Compared to the mining project characteristics, there is much less in the literature that discusses demographic factors that affect an individual's likelihood to support a (proposed) mining project in their community. Four demographic factors (age, gender, income and number of children) are used in the only previous choice experiment in a mining community (Ivanova & Rolfe, 2011). These four demographic factors were identified as significant at 1% or 5% level, and the coefficients are shown in Table 2-4. The positive coefficients of female (gender), number of children, and age mean that the individuals who are female, older, or have more children are more likely to support the mining project than individuals who are male, younger, or have fewer children.

Dimitropoulos & Kontoleon (2009) showed that the level of education was significant for local acceptability of wind-farm investment at 5% level. This author hypothesizes that the level of education will be important for mining decisions, as well. The negative coefficient of education, shown in Table 2-4, means that there is a higher probability that people with higher education level will be opponents of the mining project than the people with lower education level.

Table 2-4. Demographic factors (Dimitropoulos & Kontoleon, 2009; Ivanova & Rolfe, 2011)

Factors	Coefficient	Standard error
Female[1]	1.243**	0.259
Number of children[1]	0.261**	0.098
Income[1]	0.000*	0.000
Age[1]	0.037*	0.015
Education[2]	- 0.422*	-2.293

**significant at the 1% level
 *significant at the 5% level
[1]Ivanova & Rolfe 2011
[2]Dimitropoulos & Kontoleon 2009

A survey was done to understand the local community after a massive demonstration and violent conflict (Muradian et al., 2003). Gender, level of education, and age were used as important background characteristics. Beside these three, the main economic activity (job field) is also included as background information at Muradian et al.(2003). In addition, job field as a factor is confirmed by Mason, Paxton, Parr, & Boughen (2010) in their study of community perceptions of seafloor exploration. In the surveyed community, differences in opinions were observed based on how closely the respondent's job was related to seafloor exploration and mining.

2.4. DISCRETE CHOICE THEORY AND MODELS

Discrete choice analysis can be employed to describe the influence of the characteristics of decision makers (demographics) and the attributes of alternatives and choices they are presented with. Discrete choice models take many forms, including: binary logit, binary probit, multinomial logit (MNL), conditional logit (CL), nested logit(NL), generalized extreme value(GEV), multinomial probit (MNP), mixed logit (ML) models (Train, 2002).

In this section, the author discusses the two most popular discrete models: the multinomial logit and conditional logit models. In addition, a special case of the CL model in which the data is stratified by question or choice set (referred to in this work as the "conditional logit model stratified by question") as well as the multinomial probit and mixed logit models are discussed.

2.4.1. Discrete Choice Theory. The basic theory of discrete choice modeling is random utility maximization (Marschak, 1959). The individual decision maker's overall preference of a choice alternative is a function of the utility, which the alternative holds for the individual. This individual's utility (U_{ni}) for an alternative is separable into two components, as shown in Equation (2-1): (i) the component which can be explained by the observed (by a researcher) variables; and (ii) the component, which can be explained by unobserved variables − often, deemed random.

$$U_{ni} = V_{ni} + \varepsilon_{ni} \tag{2-1}$$

U_{ni}: utility of alternative i to individual n

V_{ni}: observed component measured for alternative i of individual n

ε_{ni}: unobserved random component for alternative i of individual n

It is postulated that an individual will prefer the choice alternative perceived to have the greatest utility. The probability that individual n prefers the mining project or plan i of choice set J, is shown in Equation (2-2).

$$\begin{aligned} P_{ni} &= \text{Prob}\left(U_{ni} \geq U_{nj}, \ \forall i \neq j, \ i \text{ and } j \in J\right) \\ &= \text{Prob}\left(V_{ni} + \varepsilon_{ni} \geq V_{nj} + \varepsilon_{nj}, \ \forall i \neq j, \ i \text{ and } j \in J\right) \\ &= \text{Prob}\left(\varepsilon_{ni} - \varepsilon_{nj} \geq V_{nj} - V_{ni}, \ \forall i \neq j, \ i \text{ and } j \in J\right) \end{aligned} \tag{2-2}$$

j: alternatives (other than i)

J: the total number of alternatives.

U_{nj}: utility of alternative j to individual n

V_{nj}: observed component measured for alternative j of individual n

ε_{nj}: unobserved random component for alternative j of individual n

2.4.2. Multinomial Logit Model. In the multinomial logit (MNL) model, also called multinomial logistic regression, the observed utility of each alterantive V_{ni} is a linear function of X_n and the random component (ε_{ni}). The utility and probability are shown in Equations (2-3) and (2-4). The utility for each alternative depends on the same variables, $\boldsymbol{X_n}$, but the coefficients are different for different alternatives. X_n is a vector of characteristics specific to the n-th individual and the variables contain only individual characteristics. $\boldsymbol{\beta_i}$ is a vector of coefficients specific to the i-th alternative. Thus, this model involves choice-specific coefficients and only individual specific repressors. The error terms, $\boldsymbol{\varepsilon_{ni}}$, are assumed to be independently and identically distributed (iid) with a type 1 extreme value distribution.

$$U_{ni} = V_{ni} + \varepsilon_{ni} = \beta_i X_n + \varepsilon_{ni} \tag{2-3}$$

β_i: a vector of coefficients specific to the ith alternative

X_n: characteristics specific to the nth individual

ε_{ni}: iid extreme value

The probability of choice i to individual n is:

$$P_{ni} = \frac{\exp(\beta_i X_n)}{\sum_{j=1}^{J} \exp(\beta_j X_n)} \tag{2-4}$$

2.4.3. Conditional Logit Model. The conditional logit model (CL), sometimes also called the multinomial logit model, was first formulated by McFadden in the 1970s (Daniel McFadden, 1974). In this model, the observed utility of each alternative, V_{ni}, is a linear function of X_{ni} and the random component (ε_{ni}). The error terms, $\boldsymbol{\varepsilon_{ni}}$, are assumed to be independently and identically distributed (iid) with type 1 extreme value distribution. $\boldsymbol{X_{ni}}$ is a vector of attributes specific to the ith alternative as perceived by the nth individual. The utility and probability are shown in Equations (2-5) and (2-6).

$$U_{ni} = V_{ni} + \varepsilon_{ni} = \beta X_{ni} + \varepsilon_{ni} \tag{2-5}$$

β: a coefficient vector for X_{ni}

X_{ni}: a vector of attributes specific to the ith alternative as perceived by the nth individual

The probability of choice i to individual n is:

$$P_{ni} = \frac{\exp\left(\beta X_{ni}\right)}{\sum_{j=1}^{J} \exp\left(\beta X_{nj}\right)} \tag{2-6}$$

The Equation (2-5) is quite similar to Equation (2-3) for the MNL model. However, the explanatory variables X_{ni} do not only include characteristics specific to the nth individual, but also describing the relationship between the chooser (nth individual) and the option (ith option). It is an important feature that distinguishes the conditional logit model from the MNL model. In addition, the MNL model has separate coefficient vectors, β_i, for each of the possible outcomes. Compared to the MNL model, there is only one coefficient vector but different X vectors, for each outcome in the conditional logit model.

As a result of these two characteristics, the conditional logit model has an important advantage over the MNL model. The model has significantly fewer parameters than the MNL model. While each factor of the CL model has one coefficient, that of MNL model has the number of coefficients equal to the number of its levels minus one.

2.4.4. Conditional Logit Model Stratified by Question. The conditional logit model stratified by question (SCQ), sometimes also called stratified logistic model, is a special instance of the CL model.

In this instance of the CL model, at least one variable must be specified to invoke a stratified analysis. In SAS, the variable can be either character or numeric, but the procedure treats them as categorical variables. The STRATA statement partitions the input data set into non-overlapping subgroups (SAS, 2014). The stratified logistic model has the form shown in Equation (2-7).

$$\text{logit}\left(\pi_{hi}\right) = \alpha_h + x'_{hi}\beta \tag{2-7}$$

Where π_{hi} is the event probability for i^{th} observation in stratum h having covariates x'_{hi}, and where the stratum-specific intercepts α_h are the nuisance parameters that are to be conditioned out.

The SCQ model does not have an intercept, as can be understood from Equation (2-6). An intercept can be included by transforming Equation (2-6) to Equation (2-8). From Equation (2-8), we can see that the standard CL model can be transformed into a special case of the conditional logit model by appropriate coding of the explanatory variables.

$$P_{ni} = \frac{\exp\left(\beta X_{ni}\right)}{\sum_{j=1}^{J}\exp\left(\beta X_{nj}\right)} = \frac{\exp\left(\beta_0 + \beta X_{ni}\right)}{\sum_{j=1}^{J}\exp\left(\beta_0 + \beta X_{nj}\right)} = \frac{\exp\left(\beta_0\right)\cdot\exp\left(\beta X_{ni}\right)}{\exp\left(\beta_0\right)\cdot\sum_{j=1}^{J}\exp\left(\beta X_{nj}\right)} \tag{2-8}$$

Compared to the CL model, the stratified analysis in the CLQ model makes it possible to compare the options in each choice set of the DCE. In the following discrete choice experimental design, all the possible combination of these 16 mining characteristics will be divided into non-overlapping subgroups. In the CL model, the fitting algorithm can only analyze the local mining communities' preference by comparing all possible combinations at the same time. However, in the real DCE, participants were answering questions one by one, and each question has limited options. The stratified analysis instructs the algorithm to consider the data by choice set, which makes it more practically applicable. It better represents the way respondents considered the choices. Based on this, the CLQ model appears more suitable for mining stakeholder analysis.

Despite these apparent differences, the multinomial logit and conditional logit models have the same three shortcomings. First, the coefficient vector, β, is fixed in the MNL and CL models. This means different individuals with the same surveyed characteristics will make the same choice given the same choice set. In reality, individuals with the same characteristics might make different choices. Thus, the fixed coefficient, β, is not reasonable.

Second, the MNL and conditional logit models have the independence of an irrelevant alternatives (iia) property, since the error terms, ε_{ni}, are assumed to be independently and identically distributed (Train, 2002). The probability ratio of alternatives i and k only depend on alternatives i and k in the MNL and conditional logit models and does not depend on the other alternatives (see Equation 2-9, which is based on Equations 2-4 and 2-6).

$$\frac{P_{ni}}{P_{nk}} = \frac{\exp(V_{ni})/\sum_{j=1}^{J}\exp(V_{nj})}{\exp(V_{nk})/\sum_{j=1}^{J}\exp(V_{nj})} = \exp(V_{ni} - V_{nk}) \tag{2-9}$$

The iia property means that there is no cross elasticity among the alternatives. If an attribute of one alternative *j* is changed, the changes in the other alternatives' probabilities are not dependent on the changed alternative *j*. Yet, this is not true in some choice situations. For example, assume there are three kinds of vehicles in a market: large gasoline cars, small gasoline cars and small electric cars. Their current market shares are 66%, 33% and 1%, respectively. Also, assume that a government subsidy increases the market share of the small electric car from 1% to 10 %. Using the MNL and conditional logit model, the market share of the other two cars would be predicted to drop while still maintaining the same ratio. The market share of large gasoline cars would drop from 66 % to 60%, and that of small gasoline cars would drop from 33% to 30% (maintaining the 2:1 ratio). The ratio of the market share of these two vehicles have to be 2:1 since their utility rate is 2:1, and is not dependent on any other alternatives. However, this prediction is unrealistic. Since the electric car is small, subsidizing it can be expected to draw more from small gas cars than from large gasoline cars.

Thirdly, the MNL and conditional logit models have the potential to capture dynamics of repeated choice. However, the repeated choice has to be independent over time since the error terms, ε_{ni}, are assumed to be independently and identically distributed (iid) in the MNL model. Thus, the MNL and conditional logit models cannot handle repeated choice situations if the choices are correlated over time.

2.4.5. Multinomial Probit Model. While the multinomial probit model (MNP) is not a popular model, it is an important model in the history of discrete choice model development. The first binary probit model was derived by Thurstone (1927). Hausman and Wise (1978) and Daganzo (1979) employed and developed it for choice behavior (Daganzo, 1979; Hausman & D.Wise, 1978).

The utility equation is the same as the conditional logit model (Equation 2-5), but the ε_{ni} are assumed to be normally distributed with mean of zero and covariance matrix,

Ω. The probability density function (pdf) of ε_n and probability of choice are shown in Equations (2-10) and (2-11), respectively.

The probability density function (pdf) of ε_n is:

$$\Phi\left(\varepsilon_n\right)=\frac{1}{\left(2\pi\right)^{\frac{1}{2}}\left|\Omega\right|^{\frac{1}{2}}}\exp(-\frac{1}{2}\varepsilon_n^{'}\Omega^{-1}\varepsilon_n) \tag{2-10}$$

$$\varepsilon_n \equiv \left(\varepsilon_{n1},.......,\varepsilon_{nJ}\right) \sim N\left(0,\Omega\right)$$

$\varepsilon_n^{'}$: a vector of ε_n, and $\varepsilon_n^{'}=\left(\varepsilon_{n1};\varepsilon_{n2};.......;\varepsilon_{nJ}\right)$

For individual n, the probability of choice i, is:

$$\begin{aligned} P_{ni} &= \mathrm{Prob}\left(V_{ni}+\varepsilon_{ni}\geq V_{nj}+\varepsilon_{nj},\ \forall i\neq j\right) \\ &= \int I\left(V_{ni}+\varepsilon_{ni}\geq V_{nj}+\varepsilon_{nj},\ \forall i\neq j\right)\Phi\left(\varepsilon_n\right)d\varepsilon_n \end{aligned} \tag{2-11}$$

Where I($\bullet$) is an indicator function: it equals 1 when the expression inside the parenthesis is real and 0, otherwise.

Compelling progress to the MNP was made by Haaijer et al. (1998). They accounted for random variation in the coefficients β over decision-makers, instead of having it be fixed as before. The coefficients β were assumed to be normally distributed in the population with mean b and covariance W. And the parameters b and W can be estimated by the MNP model.

Thus, the three limitations of the MNL model are all relaxed in the MNP model. Firstly, two people who have the same surveyed characteristics can make different choices since there is a covariance W between the normally distributed coefficients, β_n. Secondly, MNP does not have the iia property and can represent any substitution pattern, because ε_n are assumed to be normally distributed with mean 0 and covariance matrix Ω. In the previous example, the large gasoline and small gasoline cars' market shares would not have to maintain the 2:1 ratio after the small electric car's market share changes.

Their market share will be relative to the change in small electric car's market share.

Finally, the MNP model can handle repeated choice situation where choices are

correlated over time by expanding the covariance matrix Ω of the errors ε_n. The details

are not explained here since the current research does not include dynamic choice

modeling of mining local community acceptance, the details can be found at Train (2002).

   2.4.6. Mixed Logit Model. The mixed logit (ML) model, also called random

parameters logit model, was proposed by Mcfadden and Train (2000). In the ML model,

the distribution of coefficents, $f(\beta)$, is not limited to the normal distribution like in the

MNP model. The ML model can utilize any distribution for the random coefficients. The

most popular distributions of the random parameters are uniform, triangular, normal and

lognormal distributions. The probability of choice and logit probability are shown in

Equations (2-12) and (2-13). Mixed logit probabilities are the integral of standard logit

probabilities over the coefficents distribution function, $f(\beta)$.

$$P_{ni} = \int L_{ni}(\beta) f(\beta) d\beta \qquad (2\text{-}12)$$

$L_{ni}(\beta)$: the logit probability evaluated at parameters β

$$L_{ni}(\beta) = \frac{\exp\left[V_{ni}(\beta)\right]}{\sum_{j=1}^{J} \exp\left[V_{nj}(\beta)\right]} \qquad (2\text{-}13)$$

$f(\beta)$: any distribution of parameters β

   McFadden and Train (2000) show that any choice model can be approximated by

the ML model with appropriate specification of the observed variables and distribution of

coefficients (Mcfadden & Train, 2000). The MNP is a special case of the ML model

where the coefficent distribution function, $f(\beta)$, is a normal distribution.

While the ML model is the most advanced discrete choice model available, its practical application is challenging. First of all, the modeling algorithm, MDC PROC, in the general statistics software, SAS, cannot estimate the coefficients of demographic factors in the ML model. This is because MDC PROC can only compare the responses one question or choice set at a time. Since for each choice set, the demographic factors are the same (i.e. the same individual chose one choice and did not choose the others), the algorithm cannot help predict the influence of demographics factors on the choice. Train (2002) provides advanced methods to estimate the coefficients of demographic factors. However, this method is computationally expensive and not implemented in widely used statistical packages, like SAS. This makes it difficult to apply this model in mining stakeholder analysis.

What is more important, the ML model is more suitable for factors with continuous levels since the coefficients will be estimated as distributions. However, the continuous levels are difficult to include in choice experiments. The literature review of factors that affect community acceptance in Section 2.3, show that factors such as "job opportunities", "income increase" and "mine life" could have continuous levels. However, in discrete choice experimental design, their levels will have to be selected as several representative levels, or there will be too many combinations in the DCE. Once the levels of these factors are limited to representative levels for meaningful solicitation of information, it is difficult to estimate the coefficient distributions in an ML model at any significant level. As anexample, K. Willis et al. (2011) studied four factors ("capital cost", "energy bill per month", "maintenance cost", and "contract length") that have continuous levels in monetary units and years. The authors designed the discrete choice experiment with four levels each for these fours factors, and were able to estimate the coefficient of only one factor as a distribution at the 1% significance level (K. Willis et al., 2011).

2.4.7. Model Discussion. The conditional logit model is the most popular discrete choice model and has been used to understand the decision-making process of local communities regarding preferred mineral project development choice (Ivanova et al., 2007; Ivanova & Rolfe, 2011)[3].

The conditional logit model stratified by question is a special instance of the CL model with a stratified conditional logistic regression to compare the options in each choice set. As discussed in Section 2.4.4, it better represents the way respondents consider the choices. Based on this, the CLQ model appears more suitable for mining stakeholder analysis.

The multinomial logit model is not a suitable model for mining stakeholder analysis due to the fact that it has too many parameters. There are potential 16 mining characteristics and six demographic factors affecting local mining communities' acceptance of mining project (see Section 3). If each factor has only three levels, the number of required coefficients will be 44. The huge number of coefficients makes it difficult to use the discrete choice model result for reasonable inferences in mining stakeholder analysis.

Both of the MNL and CL models have two limitations: fixed taste coefficients β_n and the iia property. The fixed taste coefficient may restrict application for mining community acceptance modeling, since the models do not allow for uncertainty modeling around the coefficients, say with distributions like the ML model. This means individuals with the same surveyed demographic factors will always be modeled to have the same preferences. Yet, the MNL and CL models restrict one attribute to one fixed coefficient, which is a limitation for this application.

---

[3] Ivanova et al. (2007) and Ivanova & Rolfe (2011) refer to their models as MNL models. However, it is apparent from Train (2002) that these models are indeed CL models.

Additionally, the iia property is not true for some mining local acceptance choice situations. Consider the choice set presented in Table 2-5, for example. Assume that 66%, 33% and 1%, respectively, choose options 1, 2, and 3. Assume also that after public education, more people have been convinced of the ability of the mining company to implement the proposed 3:1 wetland compensation plan (i.e. 3 acres of wetlands will be built elsewhere for every acre of wetlands impacted), leading to an increase in the percentage of people in favor of Option 3 to 10%. The MNL model (with the iia property) will predict proportional decreases in Options 1 and 2 to 60% and 30%, respectively (similar to the previous example with cars). However, it is likely that more of those in favor of Option 1 (those influenced mostly by wetland concerns) will change their mind with this change than those who were in favor of Option 2 (wetlands were not an issue for those).

Table 2-5. Sample choice set

Choice options	Acres of wetlands impacted by mine	New jobs created
Option 1	0	500
Option 2	1,000	500
Option 3	1,000	1,000

These two constraints (constant β_n and iia property) are relaxed in the mixed multinomial logit model. Also, the ML model allows each random coefficient to follow any distribution (instead of being restricted to the normal distribution as in the MNP model). However, practical application of the ML model is challenging for the reasons stated in Section 2.4.6. First of all, model fitting is computationally expensive and not easily available in commercial statistical software. Secondly, the ML model is more appropriate for factors with continuous levels and, even then, it is difficult to estimate the coefficient distributions at any significant level (Revelt & Train, 1998; K. Willis et al., 2011). Thus, while the two constraints (constant β_n and iia property) that plague the MNL

and CL models are relaxed in the advanced ML model, the cost of relaxing these two limitations is big.

In this study, the author would like to evaluate the performance of three candidate discrete choice models (CL, CLQ and ML models) to determine the most appropriate model for mining stakeholder analysis

2.5. SUMMARY OF SECTION TWO

From the above, discussion, the following main points summarize the discussions in this section.

1. Community engagement is important for sustainable development in mining

2. The literature review shows that there are many factors that affect community acceptance. These factors include the impact of the mine, the mine owner's track record and governance issues (local, regional and national), and community demographics.

3. Discrete choice modeling shows significant potential to improve stakeholder analysis, which is an important part of community engagement

4. The candidate discrete choice models for mining stakeholder analysis are conditional logit model, conditional logit model stratified by question, and mixed logit models.

3.   CLASSIFYING CRITICAL FACTORS THAT INFLUENCE COMMUNITY ACCEPTANCE OF MINING PROJECTS FOR DISCRETE CHOICE EXPERIMENTS IN THE UNITED STATES

3.1. INTRODUCTION

The first challenge of the discrete choice theory framework for stakeholder (community) analysis is, how do you identify, classify, and verify the important factors (attributes of the mining project) that may affect local community acceptance of a mining project? This challenge can affect the success of the whole discrete choice experiment and model. The factors considered need to be broad enough to cover the key issues that might be important to different respondents, and easy enough to provide useful feedback. Ivanova et al. (2007) and Ivanova & Rolfe (2011) tracked five and seven mining project characteristics, respectively. Section 2 provides a discussion on the important factors that affect a community's (or individual's) acceptance of a mining project. Further work, with emphasis on classifying and verifying the key mining project characteristics from the plethora of candidate characteristics, is required to improve the reliability of discrete choice models and further refine how this approach can be used in community analysis. Pursuant to this challenge, three further questions have to be answered: (1) How do you classify and verify the important mining project characteristics for discrete choice experiments? (2) How do you find the key demographic factors, which are significant vis-à-vis people's perception of the importance of the mining characteristics? (3) Is there a difference between attitudes of people who live in mining and non-mining communities (i.e. people with and without significant mining experience)? Without answers to these three important questions, discrete choice experiments and modeling would not be efficient and effective, nor produce valid models to help with community analysis.

To bridge this gap, this section describes a qualitative data collection process, with the aim of facilitating better choice experiment (survey) design for discrete choice modeling. Among qualitative methods, online surveys are useful in an initial exploratory

or hypothesis-generating phase (Tey et al., 2012). This work used an online survey to validate a classification of the important factors in an individual's choice to accept a mining project. The objectives of this online survey were to: (1) validated the author's classification of mining project characteristics, which affect people's decision to support a proposed mining project; (2) identify the key demographic factors that will affect people's evaluation of project characteristics; and (3) test whether there are significant differences between attitudes of respondents who live in mining communities[4] and non-mining communities[5]. The author conducted a literature review (Section 2) to identify six demographic factors and classify mining project characteristics into 16 independent factors that would affect community acceptance. Although the list of project characteristics that affect an individual's choice to support a mine or not can be long, the author chose a classification system that balances environmental, social, and economic impacts, with a view on balanced choice experiments. The survey of residents of mining and non-mining communities was used to test the research hypotheses and evaluate the differences between the results of respondents living in mining and non-mining communities.

This work will be a significant contribution to knowledge and the literature on community acceptance in mining. The research provides preliminary results for effective and efficient discrete choice experiments and modeling.

3.2. DETERMINANTS OF COMMUNITY SUPPORT OF MINING

Obviously, there are many ways to classify the factors that affect community support and any classification is subjective. In this work, the author attempted to classify

---

[4] Communities where there is significant mining and life is affected by mining activity
[5] Communities where there is no significant mining and no significant impact of mining on life

the factors equally into four main groups: environmental, economic, social, and governance and miscellaneous others. The factors themselves are framed in order to easily facilitate the design of choice experiments (i.e. it is easy to set different levels of each factor). These choices are subjective and are not put forth as universally correct but the most suitable for preliminary choice experiments. Based on a critical review of the literature, the author hypothesized the classification of important mining project characteristics[6] and list of key demographic factors in Table 3-1 and Table 3-2 as the preliminary list of factors that influence community support (Que & Awuah-Offei, 2014). This list does not include all possible factors but contains the common factors that most people will consider in making a decision to support a mining project or not. Thus, the author thinks this is a good start for general discrete choice experiments. However, the list of factors for discrete choice experiments might vary from one context to another depending on the unique characteristics of the project. Professionals involved in community consultation using discrete choice experiments should select and validate factors, as appropriate, to ensure valid stakeholder input.

3.2.1. Mining Project Characteristics. Table 3-1 shows the list of 16 project characteristics from the four categories discussed above (Dudka & Adriano, 1997; ICMM, ICRC, IFC, 2011; ICMM, 2010, 2012a; IFC, 2009; Ivanova & Rolfe, 2011; Lockie et al., 2009; Muradian et al., 2003; Petkova et al., 2009; Schooten, Vanclay, & Slootweg, 2003; K. Willis et al., 2011). In the interest of brevity, the author cites three references for each factor, and provides a brief explanation of the factors in each category below.

---

[6] Mining project (or mine) characteristic here refers to attributes of the development option that could affect an individual's choice to accept a project or not. Factors beyond the mine itself are included (see Table 3-1).

Table 3-1. Classified characteristics of mining projects that are hypothesized to be
determinants of community acceptance

Determinant	References
Social	
Population changes	Lockie et al. 2009; ICMM 2012
Infrastructure improvement	ICMM 2012; Petkova et al., 2009
(e.g transportation, education, human services, communications and IT, hospitals, and shopping)	
Cultural impacts	ICMM 2012; Schooten et al. 2003
(e.g. impacts on archaeological and historical sites, native American artifacts, historical burial sites, arts and culture)	
Traffic and crime increase	Lockie et al., 2009; ICMM 2011
Economic	
Job opportunities	ICMM 2012; IFC 2009
Income increase	Petkova et al. 2009; Ivanova and Rolfe 2011
Cost of housing or housing shortage	Ivanova and Rolfe 2011; Petkova et al. 2009
Labor shortage for other business	Petkova et al. 2009
Environmental	
Noise pollution	ICMM 2011; Petkova et al., 2009
Water shortage or pollution	Ivanova and Rolfe 2011, ICMM 2010
Air pollution	ICMM 2011, Dudka & Adriano 1997
Land pollution	Dudka & Adriano 1997, Lockie et al. 2009
Governance and others	
Decision making mechanism on the mine's permits	Muradian et al. 2003
(e.g. decisions are based solely on what is legal; or decision makers consider input from local communities)	
Independent and transparent information	Muradian et al. 2003; ICMM 2012
(e.g. the availability of information on impacts from independent and trusted sources in addition to or including the mining company, government or non-governmental organizations)	
Mine buffer	Ivanova and Rolfe, 2011
(distance of respondent's residence from mine)	
Mine life (how long the mine will last)	Willis et al. 2011

3.2.1.1 Social aspects. Resource exploitation can be directly linked to local population changes as there is often population growth from migrants looking for job opportunities. In addition, many mining towns have significant non-resident workers living in temporary accommodation or company provided mining camps (Lockie et al. 2009). Mine development often includes expansion and/or improvement of local infrastructure to facilitate the mining activities. Sometimes, these improvements are not to the direct benefit of the mine but are done as part of CSR programs. ICCM specifically identifies cultural (heritage) impacts as a factor in community engagement (ICMM 2012). Cultural impacts include any effects on the cultural norms and practices, which include effects on intangible and tangible cultural heritage, and access to and vibrancy of cultural facilities (e.g. community meeting places). Many mine developments result in increases in traffic and crime (Lockie et al. 2009).

3.2.1.2 Economic aspects. The economic impacts of mining activities are well documented and include job opportunities (both direct and indirect) and income increases due to higher paying jobs and/or the unemployed joining the supply chain (Petkova et al., 2009; ICMM, 2012).

However, mining can also lead to increases in housing costs and labor shortages, especially for other businesses in the local community that cannot compete with large mines for talent (Petkova et al., 2009; Ivanova and Rolfe, 2011). For instance, in five out of six communities studied by Petkova (2009) scarcity of labor for other businesses was identified as an issue.

3.2.1.3 Environmental aspects. There are many environmental impacts of mining. The author chose to categorize these impacts into four broad impacts that are easy for respondents to understand and make stated preference choices (in a discrete choice survey) feasible without overwhelming respondents with information. The four broad impacts (water shortage or pollution, air pollution, land pollution/impacts and noise pollution) were selected with due regard to the significance of noise (ICMM 2011), which was added to the three traditional categories.

3.2.1.4 Governance and others. As with the other categories, there are many factors relating to governance and decision making during the legal permitting process. The author chose to include two important governance factors: the 'decision making mechanism' and the availability of 'independent and transparent information' (Muradian et al. 2003). The decision making mechanism describes how decisions are made when disagreements arise on the impacts (positive and negative) of mining. These mechanisms vary from the purely legal (i.e. the mining company meets the regulatory requirements) to those that take cognisance of the SLO and seeks legitimacy. Often, during conflicts around mining impacts, most of the information on impacts and baseline studies is provided by the mining company and/or government. The local community often does not trust the available information on the potential impacts from both sources. SLO or acceptance (note that SLO is not acceptance; acceptance is a lower level of SLO) is easier to achieve when there is independent and transparent information.

Two additional factors were included in this section because of their significance. Mining projects differ significantly in their mine lives. This has been shown to have significant impacts on community acceptance (K. Willis et al., 2011). The role of the not-in-my-back-yard (NIMBY) phenomenon is well documented in community engagement. This is included in as the mine buffer, which is the buffer between the respondent and mining impacts (Ivanova & Rolfe, 2011).

3.2.2. Demographic Factors. Compared to the mining project characteristics, there is much less in the literature that discusses demographic factors that affect an individual's likelihood to support a (proposed) mining project in their community. Four demographic factors (age, gender, income and number of children) are used in the only previous choice experiment in a mining community (Ivanova & Rolfe, 2011). The author hypothesized that the level of education will be important as well. Dimitropoulos & Kontoleon (2009) showed that the level of education was significant for local acceptability of wind-farm investment. Muradian et al. (2003) identified the 'job field' as an important demographic factor, with differences in opinions based on how closely the respondent's job was related to seafloor exploration and mining.

The six potential demographic factors are shown in Table 3-2 (Dimitropoulos & Kontoleon, 2009; Ivanova & Rolfe, 2011; Muradian et al., 2003). The author hypothesized that the factors that are correlated to the likelihood to support a project will be correlated to the ranking of the importance of the mine characteristics.

Table 3-2. Demographic factors that are hypothesized to be determinants of community acceptance

Demographic factors	References
Age	Ivanova & Rolfe 2011
Gender	Ivanova & Rolfe 2011
Income	Ivanova & Rolfe 2011
Education	Ivanova & Rolfe 2011
Job field	Dimitropoulos & Kontoleon 2009
Number of children	Muradian et al. 2003

3.3. EXPERIMENTAL METHOD

The experimental and analytical approach includes 10 steps, which are shown in Figure 3-1. Key aspects of this method are discussed in the following subsections.
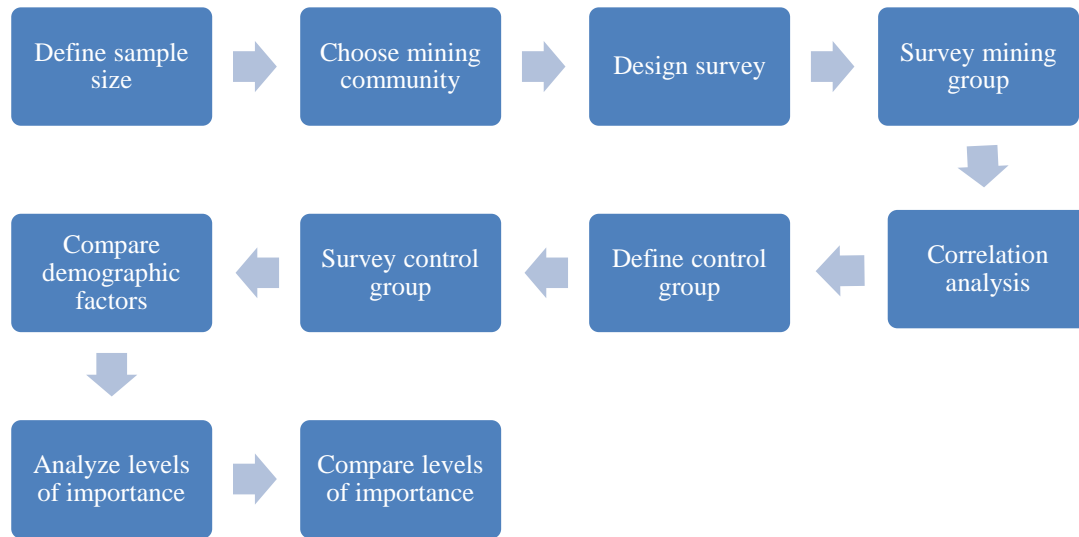
Figure 3-1. Flowchart of experimental process

3.3.1. Sample Size Determination. The sample size is an important feature of any survey, where the goal is to make inferences about a large population from a sample. The sample size should be determined based on data collection cost and acceptable sampling error (Robert E. Odeh, 1975). The sample size estimation can be targeted toward determining the correlation among the demographic attributes and attitudes towards project characteristics. In this study, we wish to detect true underlying correlations of 0.2 or higher as statistically significant. The value 0.2 was chosen because any value below this would be of little practical importance. The testing method for significance can be based on the *t*-statistic in Equation (3-1):

$$t = r\sqrt{\frac{n-2}{1-r^2}} \tag{3-1}$$

where *r* is the estimated Spearman correlation and *n* is the sample size.

Under the null hypothesis of zero correlation, this statistic has an approximate Student's *t* distribution with $n-2$ degrees of freedom. For a sample size of 100, the statistic takes the value $t = 2.0207$. The critical value of *t* with 98 degrees of freedom is 1.984, so the null hypothesis will be rejected here. Based on this, a sample size of 100

was chosen so that a correlation of 0.2 or higher between the demographic attributes and attitudes towards the project characteristics will be deemed significant. In addition, Fisher's exact test is used, in one instance, to determine if there is association between gender (which is binary) and the respondents' preferences on the project characteristics. Using the SAS sample size calculation procedure (PROC POWER), it is estimated that a sample size of 100 is sufficient to detect a difference of 20% or more between respondent's choice preferences 80% of the time (M.G.Kendall & A.Stuart, 1973) .

3.3.2. Sampling Comparable Respondents from Non-mining Communities.

Individual preferences for a new mine may be influenced by past experience with mining. In order to test whether mining experience of a respondent (living in a community with significant mining is used as a proxy for experience with mining) affects his or her preferences, 100 individuals living in mining communities and 100 living in non-mining communities were recruited to complete the online survey. These two surveys were conducted in June and October, 2013, respectively.

The first 100 individuals were randomly selected from 20 mining communities across the whole USA (see Appendix A for the full list of communities). After identifying the demographic factors significantly associated with attitudes towards rankings of the mining project characteristics using correlation analysis, another 100 individuals from 20 non-mining communities (the list of communities is in Appendix B) were selected such that the new sample matched the distribution of important (defined based on results of the first survey) demographic factors in the sample from of the mining communities. The surveys were computer assisted personal interviews, administered by Qualtrics, a well-known market research firm. Respondents were tracked by their zip code using the IP address they used to access the survey.

Among the respondents from mining communities, 48 out of the 100 stated that they live near a mine. Of these, 36 reported living within 30 miles of a mine: 24 live

within 10 miles of a mine; seven live between 11 to 20 miles away from a mine; and another five live between 21 to 30 miles from a mine. Eleven of the 100 participants self-declared to have *experience* with mining (e.g. working for a mine, familiarity with mining activities, studying about mining etc.). Sample answers to Question 7 include:

"Family worked in mining"; "dad was a miner"; "spouse employed by mine"; "study of mines and mining"; "grandfather worked in mines"; "work at the mine"; "my father used to work for the mine so I have seen what they do there and how it's processed [etc.]"; and "I had uncles who worked in coal mines in Kentucky".

3.3.3. Online Survey Design. The online survey was conducted with a three part questionnaire. These topical issues were reviewed by colleagues of the author, for their relevance, clarity, and efficiency, beforehand. The full survey is shown in the Appendix C.

The first part of the survey contained background questions regarding the respondent's socioeconomic status, as well as their zip code and past experience with mining. The demographic questions include age, gender, income, education, job field, and number of children.

The second part of the questionnaire involved attitudinal questions. Participants were asked to rank the importance of each project characteristic, by selecting a number from 1 to 7 ("not at all important" to "extremely important"), in their decision to support a mine, if a new mine were to be opened in their hometown. A short description was also given for each characteristic. The third part contained an open ended question about what other characteristic is important to the participant.

Two quality control questions were inserted in the survey. If a participant did not 'pass' the quality control questions, their data was deleted (less than 10% of participants gave invalid answers and their data was deleted and not counted towards the 100). In

addition, data was regarded as invalid if the participant completed the survey in less than one third of the average survey time (150 seconds).

3.4. DATA ANALYSIS AND RESULTS

3.4.1. Determining Significant Demographic Factors of the Mining Group.

The first sample of 100 comprised a stratified random sample of participants living in one of 20 mining communities. Then correlation and Fisher's exact tests were performed to identify important demographic factors. The correlation analysis was used for age, education, income, job field, and number of children, since these demographics are ordinal alternatives in the survey. The Fisher's exact test was used to determine if the level of importance attributed to mining characteristics is independent of gender. A demographic factor was regarded as important if it was significantly (at the 0.05 significance level) correlated to at least one of the 16 mine characteristics. The important demographic factors were matched in the control group survey.

Correlation analysis was done using the SAS CORR procedure to estimate the Spearman rank correlation coefficient and test the null hypothesis of zero correlation against the alternative hypothesis that the coefficient is non-zero (SAS, 2007a, 2007i, 2007j). If the p-value is less than the significance level α (= 0.05), the null hypothesis is rejected, implying that there is significant correlation. Fisher's exact test analysis was done using the SAS FREQ procedure (Exact Fisher/mc) to identify whether there is a significant difference between genders (SAS, 2007b, 2007c; Stokes, Davis, & Koch, 2012). The results of correlation analysis and Fisher's exact test are shown in Table 3-3. All statistically significant coefficients are shown in bold font.

Table 3-3. Correlation coefficients (and p-values in parenthesis) of mining group ranking and demographic factors. Gender results are based on the fisher's exact test (p-values). Statistically significant correlation coefficients or fisher's test results are shown in bold font

	Age	Education	Income	Job field	Number of children	Gender
Population changes	-0.240 (0.0171)	0.024 (0.813)	-0.066 (0.538)	-0.083 (0.414)	0.058 (0.275)	(0.255)
Infrastructure improvement	-0.124 (0.225)	0.071 (0.490)	0.120 (0.256)	-0.191 (0.059)	-0.097 (0.347)	(0.122)
Cultural impact	-0.394 (<.0001)	0.136 (0.184)	-0.071 (0.503)	-0.016 (0.879)	0.014 (0.895)	(0.130)
Traffic and crime increase	-0.094 (0.356)	-0.051 (0.623)	-0.230 (0.030)	0.041 (0.690)	0.016 (0.879)	(0.037)
Job opportunities	-0.137 (0.175)	-0.212 (0.035)	0.028 (0.792)	-0.134 (0.183)	-0.037 (0.715)	(0.046)
Income increase	-0.025 (0.805)	-0.236 (0.019)	-0.026 (0.808)	-0.051 (0.616)	-0.073 (0.478)	(0.064)
Cost of housing or housing shortage	-0.173 (0.090)	-0.169 (0.099)	-0.093 (0.385)	0.109 (0.288)	0.010 (0.922)	(0.023)
Labor shortage for other businesses	-0.140 (0.170)	0.046 (0.653)	-0.093 (0.388)	0.166 (0.103)	-0.017 (0.871)	(0.878)
Noise pollution	-0.196 (0.053)	0.091 (0.376)	-0.005 (0.960)	-0.119 (0.244)	-0.048 (0.640)	(0.519)
Water shortage or pollution	-0.173 (0.084)	0.012 (0.907)	-0.320 (0.002)	0.076 (0.448)	0.143 (0.157)	(0.015)
Air pollution	-0.185 (0.065)	0.025 (0.803)	-0.260 (0.012)	0.069 (0.494)	0.052 (0.608)	(0.229)
Land pollution	-0.169 (0.093)	0.040 (0.697)	-0.265 (0.011)	0.153 (0.128)	0.068 (0.502)	(0.005)
Decision making mechanism*	0.146 (0.163)	-0.055 (0.601)	-0.121 (0.271)	-0.029 (0.783)	-0.147 (0.161)	(0.048)
Information available**	0.102 (0.332)	0.153 (0.145)	-0.010 (0.927)	-0.063 (0.551)	0.044 (0.676)	(0.302)
Mine buffer	-0.286 (0.004)	0.100 (0.330)	0.000 (0.998)	-0.077 (0.447)	0.062 (0.542)	(0.027)
Mine life	0.095 (0.355)	-0.057 (0.582)	-0.108 (0.317)	0.027 (0.794)	-0.069 (0.506)	(0.049)

* Decision making mechanism: Decision making mechanism on the mine's permits
** Information available: Whether or not there is independent and transparent information available

As shown in Table 3-3, there is a significant negative correlation between income and the possible negative impacts, which include increases in traffic and crime, water shortage or pollution, air pollution, and land pollution. This means participants with higher incomes ranked traffic, crime, and pollution issues lower than those with lower incomes (a lower ranking means the respondent did not think the particular factor is as important in his/her decision to support or not support a mining project).

Also, Table 3-3 shows a negative correlation between education and job opportunities and income increase. This means respondents with higher education are less concerned about new job opportunities and potential income increases associated with the mining operation. This negative correlation may be because people with higher education have lower need to change jobs and work for the new mine or consider mining-related jobs to be less desirable.

Age is observed to be negatively correlated to population changes, cultural impacts, and mine buffer. These three factors would highly affect lifestyle. It appears, from the results, that younger people in mining communities care more about these lifestyle impacts.

From the Fisher's exact test results, there is a significant difference between female and male rankings of eight mining characteristics. They are traffic and crime increase, job opportunities, cost of housing or housing shortage, water shortage or pollution, land pollution, decision making mechanism, mine buffer, and mine life. It is important to note that the Fisher's exact test (nor this author) does not seek to determine which group (male or female) rank a particular characteristic higher/lower. It only seeks to determine whether there is a significant difference in the distribution of responses from the two groups. The goal of this research was to determine whether gender matters and should, therefore, be included in a discrete choice experiment.

From the above (Table 3-3), the author concluded that age, gender, income, and education are the important demographic factors for the respondents from mining communities. Number of children and job field were observed not to be significantly correlated with respondents' choices.

3.4.2. Comparing the Two Samples. A second set of 100 participants were recruited from 20 non-mining communities. Based on the results of correlation analysis of the initial responses from the mining group, the four important demographic factors (age, gender, income, and education) of this sample were intended to match those of the first set of 100 from mining communities. The distributions of these four socio-economic variables of respondents are summarized and compared in Figure 3-2(a-d). The distributions were compared using the SAS FREQ procedure (Exact Fisher test) (Stokes et al., 2012). The results are shown in Table 3-3 and Table 3-4. The null hypothesis of no significant difference between cumulative distributions of each demographic variable for the mining and control groups was tested using the Exact Fisher test. All p-values are greater than 0.05. The results mean there is not enough evidence to reject the null hypothesis at $\alpha = 0.05$. The author concluded that the two samples are similar (with respect to the age, gender, income, and education) and any significant differences in opinions will not be due to differences in these demographic factors.
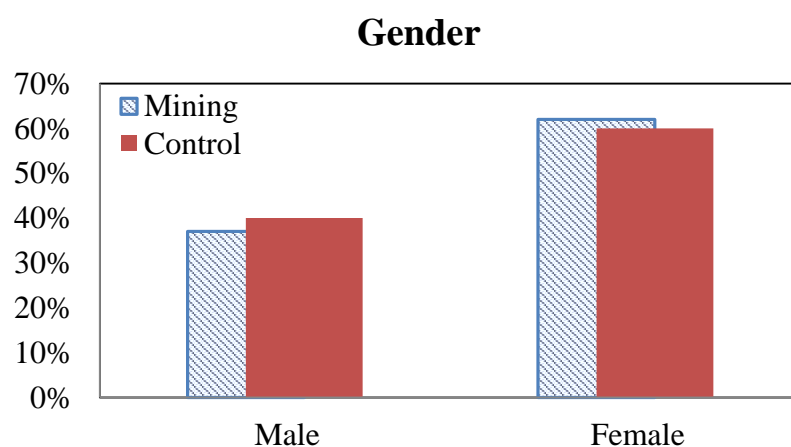
Figure 3-2 Distribution summary of the demographic factors (a) Gender comparison

Age
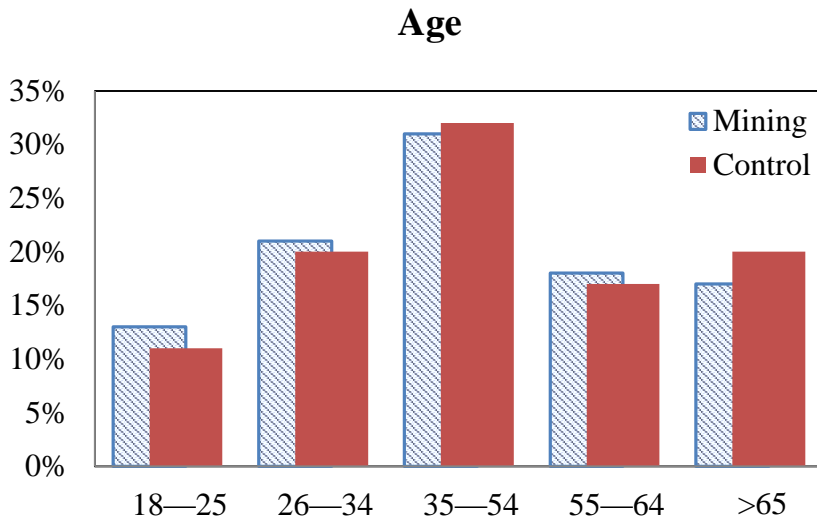
Figure 3-2 Distribution summary of the demographic factors (cont.) (b) Age comparison
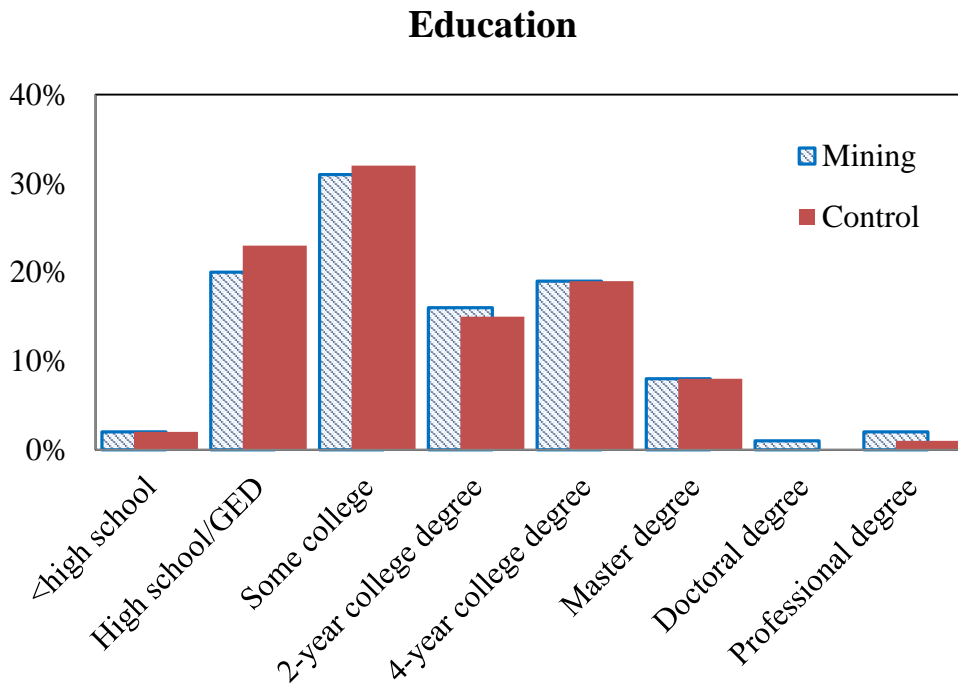
Education

Figure 3-2 Distribution summary of the demographic factors (cont.) (c) Education comparison
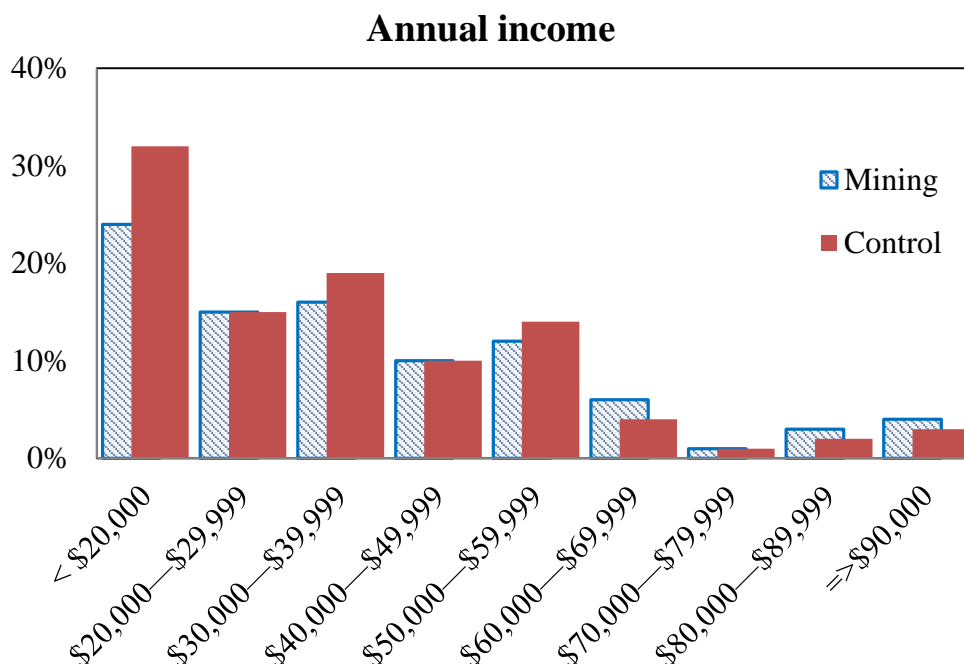
Annual income

Figure 3-2 Distribution summary of the demographic factors (cont.) (d) Annual income comparison

Table 3-4. Exact fisher test results comparing demographics of mining and control groups

Demographic factors	P-value
Gender	0.771
Age	0.976
Highest education	0.999
Annual income	0.984

3.4.3. Determining Significant Demographic Factors of the 'Control Group'.

The correlation analysis and Fisher's exact test carried out for the mining group was also done for the second set of 100 participants who live in non-mining communities using the same procedure. The results of correlation analysis and Fisher's exact test are shown in Table 3-5. All statistically significant ($\alpha = 0.05$) coefficients are shown in bold font.

Table 3-5. Correlation coefficients (and p-values in parenthesis) of control group ranking and demographic factors. Gender results are based on the fisher's exact test (p-values). Statistically significant correlation coefficients and fisher's test results are shown in bold font

	Age	Education	Income	Gender
Population changes	-0.072 (0.481)	0.182 (0.074)	-0.056 (0.586)	(0.526)
Infrastructure improvement	**-0.241** **(0.016)**	0.073 (0.470)	0.110 (0.277)	(0.128)
Cultural impact	-0.065 (0.526)	0.183 (0.070)	-0.001 (0.990)	(0.953)
Traffic and crime increase	-0.078 (0.447)	-0.076 (0.460)	0.004 (0.968)	(0.240)
Job opportunities	-0.017 (0.871)	**-0.228** **(0.023)**	**0.207** **(0.040)**	(0.533)
Income increase	-0.183 (0.073)	-0.266 (0.008)	0.061 (0.554)	(0.631)
Cost of housing or housing shortage	-0.055 (0.588)	-0.133 (0.193)	-0.027 (0.795)	**(0.013)**
Labor shortage for other businesses	**-0.290** **(0.004)**	-0.087 (0.395)	**0.201** **(0.049)**	(0.499)
Noise pollution	0.084 (0.408)	**0.326** **(0.001)**	-0.035 (0.734)	(0.350)
Water shortage or pollution	0.014 (0.889)	-0.001 (0.989)	-0.021 (0.840)	(0.518)
Air pollution	0.011 (0.913)	0.083 (0.411)	-0.047 (0.640)	(0.341)
Land pollution	-0.008 (0.934)	0.138 (0.172)	0.059 (0.557)	(0.816)
Decision making mechanism*	0.178 (0.087)	0.128 (0.221)	0.153 (0.144)	(0.764)
Information available**	**0.274** **(0.007)**	0.142 (0.168)	0.017 (0.870)	(0.286)
Mine buffer	0.065 (0.526)	0.176 (0.082)	-0.039 (0.706)	(0.398)
Mine life	-0.094 (0.366)	0.133 (0.198)	0.030 (0.771)	(0.336)

* Decision making mechanism: Decision making mechanism on the mine's permits
* *Information available: Whether or not there is independent and transparent information available

As with the mining group, all four demographic factors are correlated to at least one of the 16 mine characteristics. This confirms these demographic factors as important to the decision to accept a mine. However, in terms of which mine characteristics and the sign (positive/negative) of the correlation coefficients, the results of the control group are different from that of the mining group.

As shown in Table 3-5, income is no longer negatively ($p > 0.05$) correlated with potential negative impacts (traffic & crime increase, water shortage or pollution, air pollution, and land pollution). There is a significant positive correlation between income and job opportunities and labor shortage for other businesses. This means respondents with higher incomes ranked job opportunities and labor shortages for other businesses, which are indicators of economic opportunity, higher than those with lower incomes. This was not observed with those respondents living in mining communities. Since the two samples have similar demographic distributions, these differences may be attributable to the different attitudes the two groups may have towards mining project attributes. Such diversity could arise from the different exposures the two groups have had to mining; with the mining group having a more intimate and real-life experience while the control group's attitudes may have formed through indirect knowledge.

Level of education is negatively correlated to job opportunities (similar to the mining group) and positively correlated to noise pollution (the mining group had a negligible correlation coefficient – 0.091/ $p = 0.376$). The positive correlation means respondents with higher education ranked noise pollution higher.

As shown in Table 3-3, age is negatively correlated to infrastructure improvement and labor shortage for other businesses. This means younger people ranked infrastructure improvements and labor shortage for other businesses higher than those who are older. There is a significant positive correlation between age and information available, which means older people in non-mining communities care more about whether or not there is

independent and transparent information available. In contrast to the mining group, females and males significantly differed only on their ranking of the cost of housing or housing shortage instead of eight mine characteristics.

For the non-mining group, labor shortage for other businesses is observed to be significantly (positive or negative) correlated to age and income. This was not so for the mining group. This result is reasonable since participants in the control group are all employed by 'other businesses.'

3.4.4. Evaluating the Importance of Project Characteristics. The SAS UNIVARIATE procedure was used to analyze the level of importance data from both mining and control groups (SAS, 2007l, 2007m). Table 3-6 shows the 95% confidence bounds of the median level of importance, rounded to the nearest integer, computed without assuming any specific distribution. The detailed comparison of the importance scores of the mining and control groups are shown in Figure 3-3(a—q). For all 16 mining project characteristics, the distributions of responses on the level of importance are skewed to the right (i.e. most respondents ranked higher than 4 – "neither important nor unimportant").

3.4.4.1. Social impacts. With respect to the social impacts (Figure 3-3a-d and Table 3-6), respondents from mining and non-mining communities have similar opinions of the importance of population changes, cultural impact, and traffic & crime increase. The median levels of importance for these factors are 5—"somewhat important", (5, 6) — above "somewhat important" but below "very important", and 6 — "very important", respectively. However, the two groups differ slightly in their opinion of the importance of infrastructure improvement. While respondents who do not live in mining communities view infrastructure improvement as very important, those who live in mining communities think it falls within "somewhat important" to "very important."

Table 3-6. Level of importance of mining project characteristics

Mining characteristics	Mining group	Control group
Population changes	5	5
Infrastructure improvement	(5,6)	6
Cultural impact	(5,6)	(5,6)
Traffic and crime increase	6	6
Job opportunities	(6,7)	(6,7)
Income increase	6	6
Cost of housing or housing shortage	(5,6)	(5,6)
Labor shortage for other businesses	(4,5)	5
Noise pollution	5	(5,6)
Water shortage or pollution	(6,7)	(6,7)
Air pollution	(6,7)	(6,7)
Land pollution	(6,7)	(6,7)
Decision making mechanism on the mine's permits	(5,6)	(5,6)
Whether or not there is independent and transparent information available	(5,6)	(5,6)
Mine buffer	6	6
Mine life	6	(5,6)

1 Not at all Important

2 Very Unimportant

3 Somewhat Unimportant

4 Neither Important nor Unimportant

5 Somewhat Important

6 Very Important

7 Extremely Important
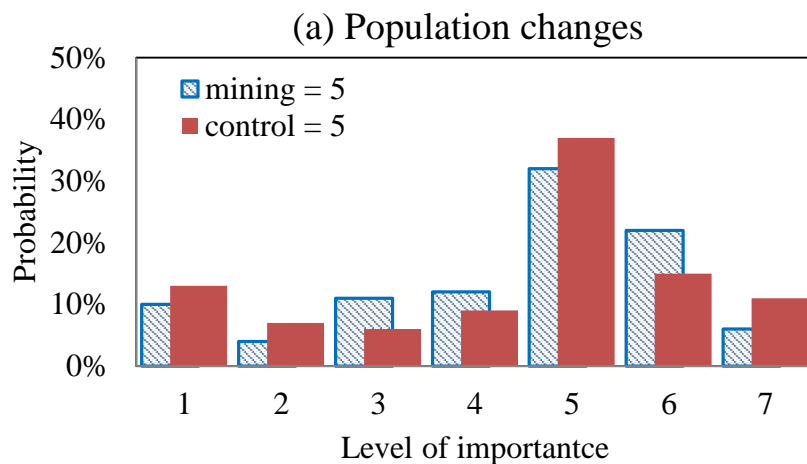
Figure 3-3. Level of importance
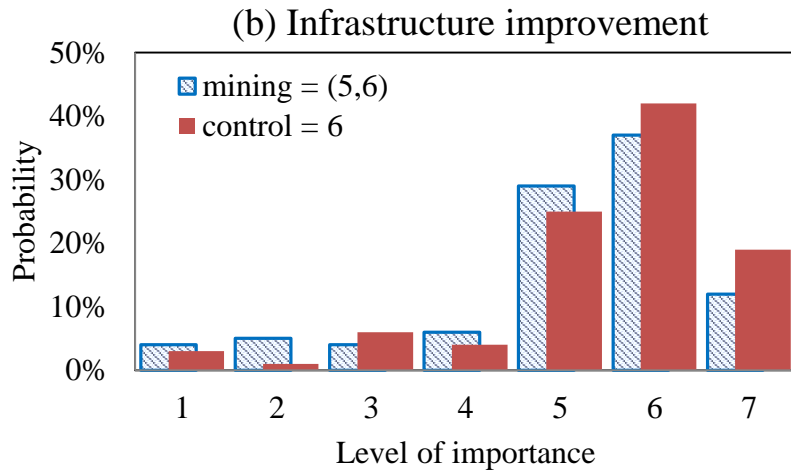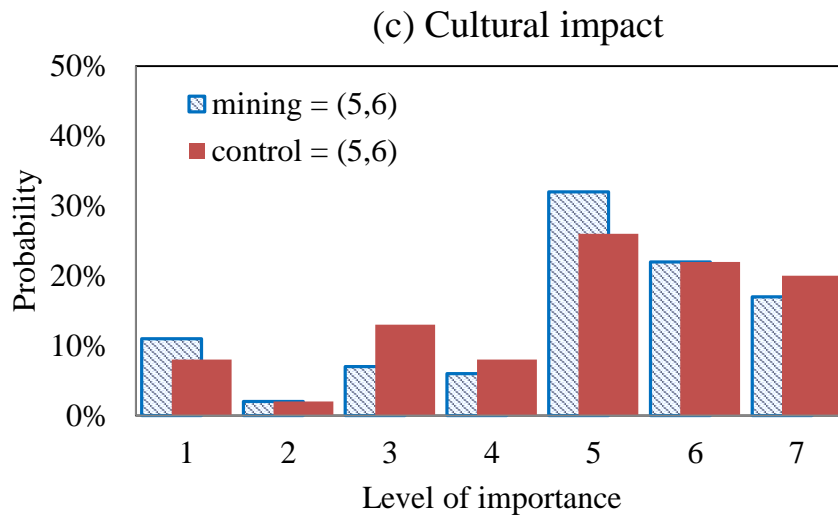
Figure 3-3. Level of importance (cont.)

Figure 3-3. Level of importance (cont.)

(d) Traffic and crime increase
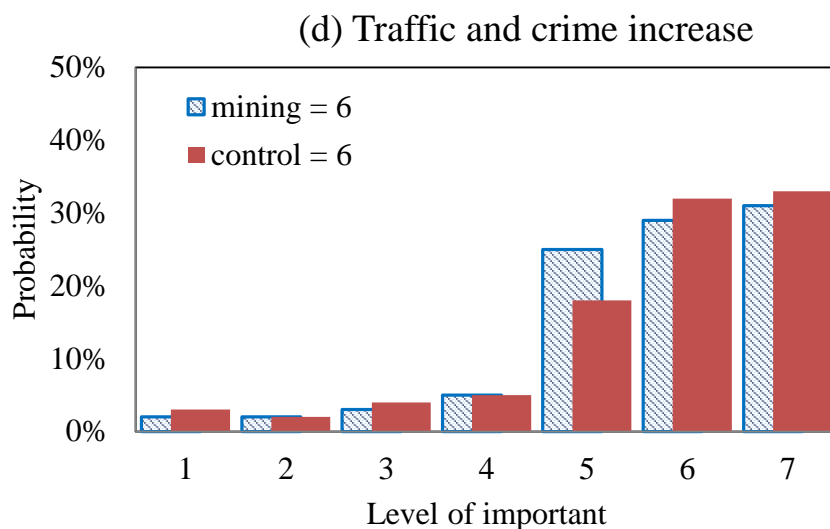
Figure 3-3. Level of importance (cont.)

For population changes, infrastructure improvement and cultural impact (Figure 3-3a, b and c), around or less than 10 % of the participants, in both groups, selected each level of importance from 1 to 4 (from 'Not at all important' to 'Neither Important nor Unimportant'). Less than 40% in both groups thought these factors are not important in their decision to support or not support a mining project in their community. Most of the remaining participants selected a level of importance of 6 (very important) for infrastructure improvement. For population changes and cultural impact, most participants selected 5 (somewhat important). While significantly fewer respondents selected 6 or 7, the drop in number of respondents selecting 6 or 7, with regards to cultural impact is moderate compared to population changes.

For traffic and crime increase, the distribution of responses is very skewed with an increasing number of respondents selecting higher levels of importance, in both groups. The mode of the distribution for both groups is level 6 (Very Important).While relatively more people from the mining group think its level of importance is 5

(Somewhat Important), more people from the control group view its level of important as 6 or 7 ('Very Important' or 'Extremely Important').

3.4.4.2. Economic impacts. Figure 3-3e-h and Table 3-6 show the mining and non-mining groups have similar opinions of the importance of job opportunities, income increase, and cost of housing or housing shortage in their decision. The median level of importance of these factors are (6, 7) — above "very important" but below "extremely important", 6 —"very important, and (5, 6) — above "somewhat important" but below "very important", respectively, for both groups. However, the two groups differ slightly in their opinion of the importance of labor shortage for other businesses. While the control group views labor shortage for other businesses as somewhat important, the mining group thinks it falls within "neither important or unimportant" to "somewhat important."
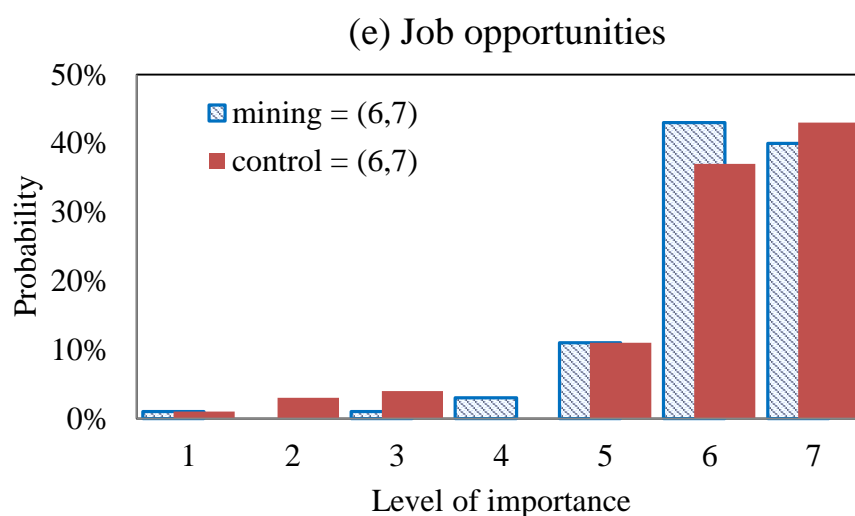
Figure 3-3. Level of importance (cont.)
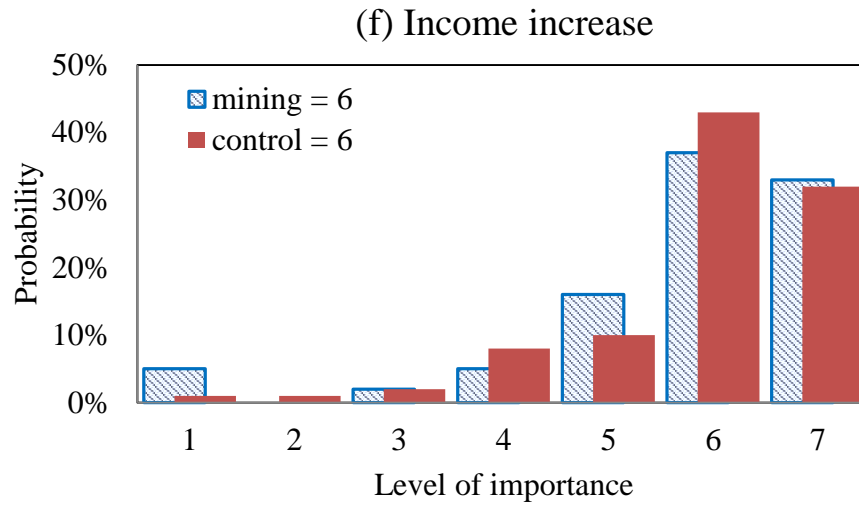
(f) Income increase

Figure 3-3. Level of importance (cont.)

(g) Cost of housing or housing shortage

Figure 3-3. Level of importance (cont.)

(h) Labor shortage for other business
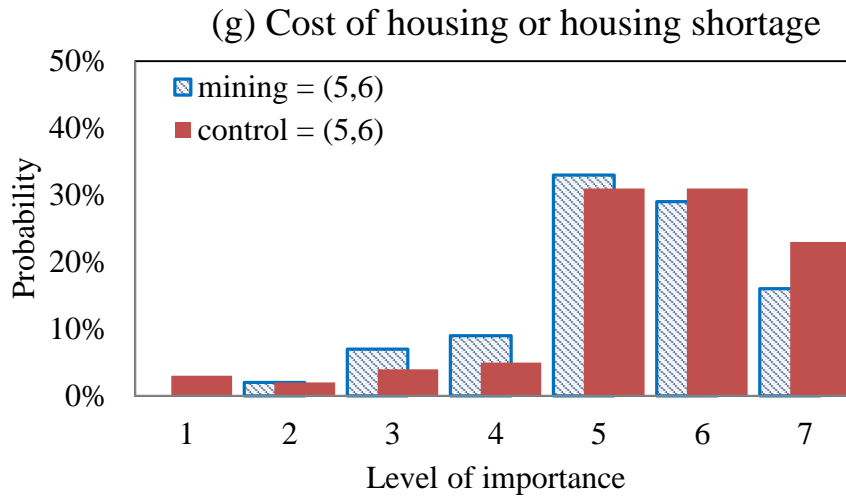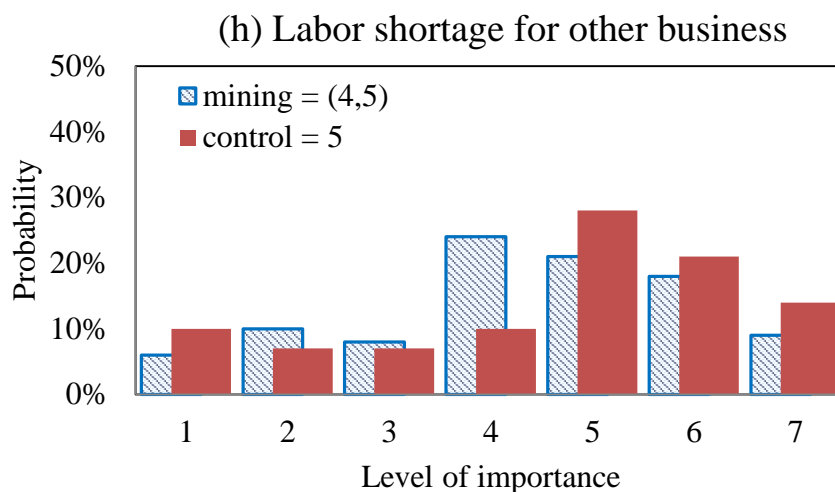
Figure 3-3. Level of importance (cont.)

For job opportunities and income increases (Figure 3e and f), around or less than 5 % of the participants, in both groups, selected each level of importance from 1 to 4 (from 'Not at all important' to 'Neither Important nor Unimportant'). Most of the remaining participants split evenly between level of importance of 6 and 7 (very important and extremely important).

For cost of housing or housing shortage and labor shortage for other business, most participants selected 5 (somewhat important) for the level of importance, with fewer people selecting 6 and 7.

3.4.4.3. Environmental impacts. From Figure 3-3i-l and Table 3-6, the mining and non-mining groups have similar opinions of water, air, and land pollution.

All their median levels of importance for these factors are (6, 7) — above "very important" but below "extremely important". However, the two groups differ slightly in their opinion of the importance of noise pollution. While the mining group views noise pollution as somewhat important, the control group thinks it is above "somewhat important" and below "very important".
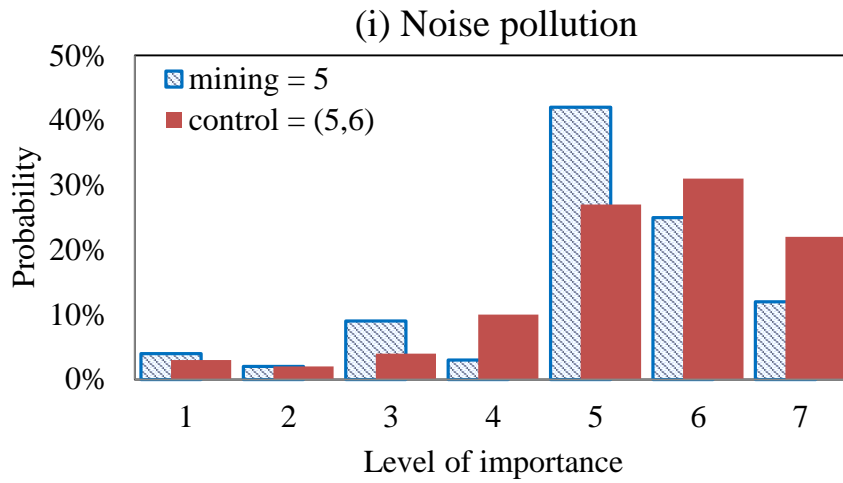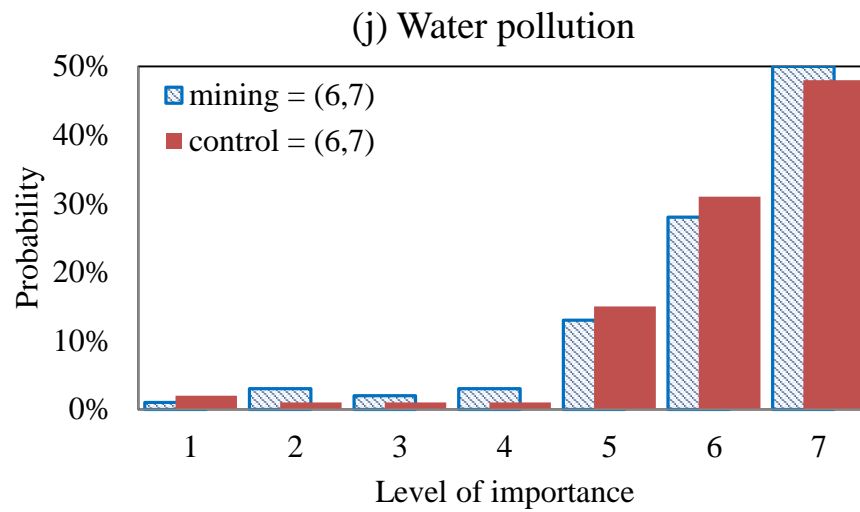
(i) Noise pollution

Figure 3-3. Level of importance (cont.)

(j) Water pollution

Figure 3-3. Level of importance (cont.)

(k) Air pollution
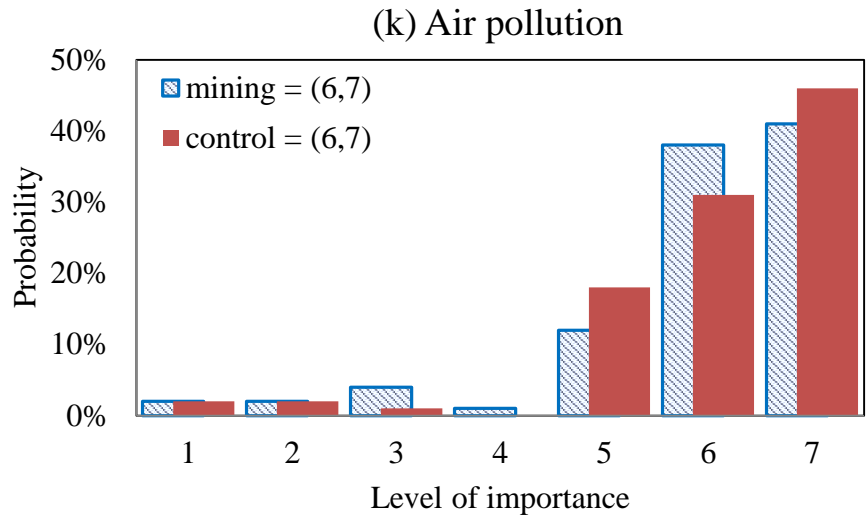
Figure 3-3. Level of importance (cont.)
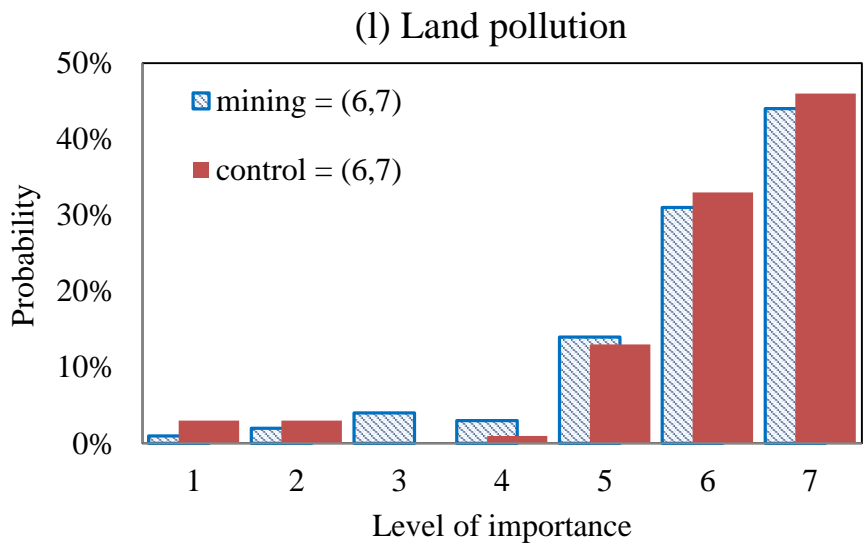
(l) Land pollution

Figure 3-3. Level of importance (cont.)

For water, air, and land pollution, the distribution of responses is very skewed with an increasing number of respondents selecting higher levels of importance, in both

groups. More than 70% of respondents rank the importance of these three impacts as 6 or 7 ('Very Important' or 'Extremely Important').

        3.4.4.4. Governance and other aspects. Figure 3-3m-p and Table 3-6 show that the respondents living in mining and non-mining communities have similar opinions of decision making mechanism, independent and transparent information, and mine buffer. The median level of importance for the first two are (5, 6) — above "somewhat important" but below "very important", while mine buffer's median level of importance is 6 —"very important". However, the two groups differ slightly in their opinion of the importance of mine life. While the control group views mine life as very important, the mining group thinks it falls within "somewhat important" to "very important."
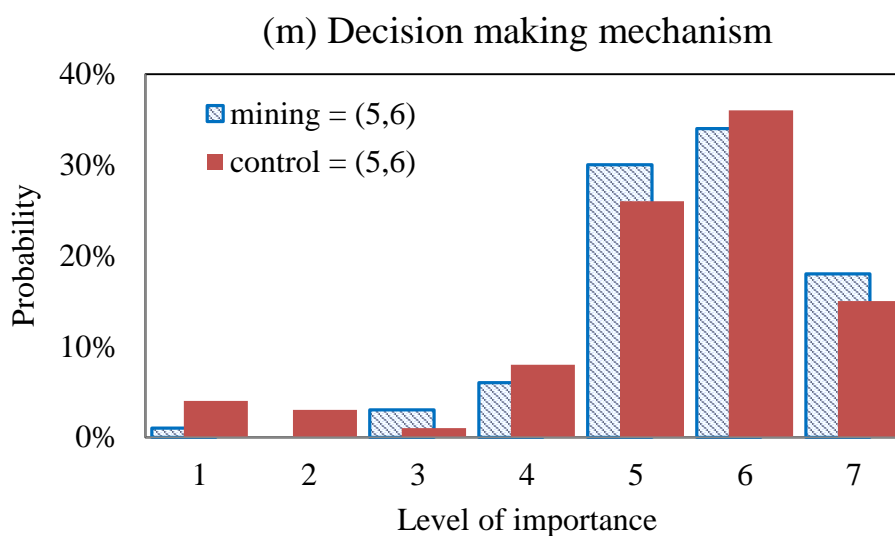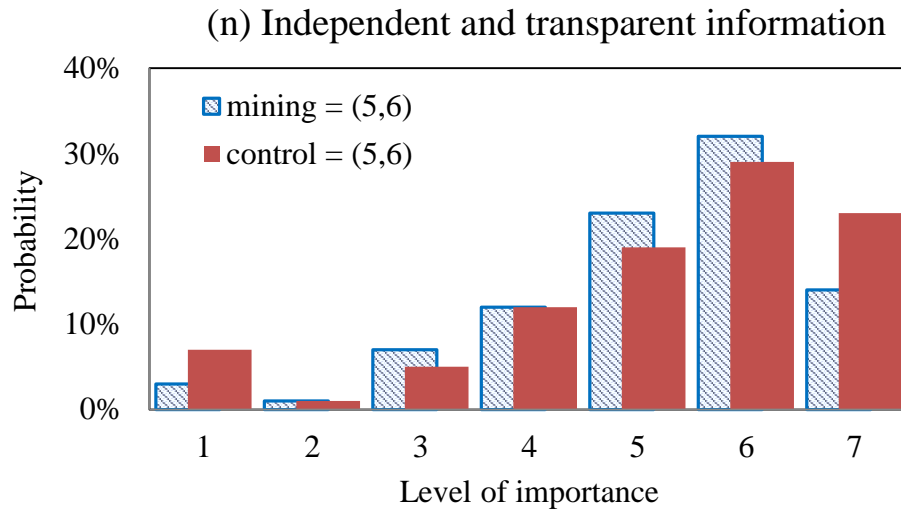
(m) Decision making mechanism

Figure 3-3. Level of importance (cont.)

(n) Independent and transparent information

Figure 3-3. Level of importance (cont.)

(o) Mine buffer

Figure 3-3. Level of importance (cont.)

(p) Mine life
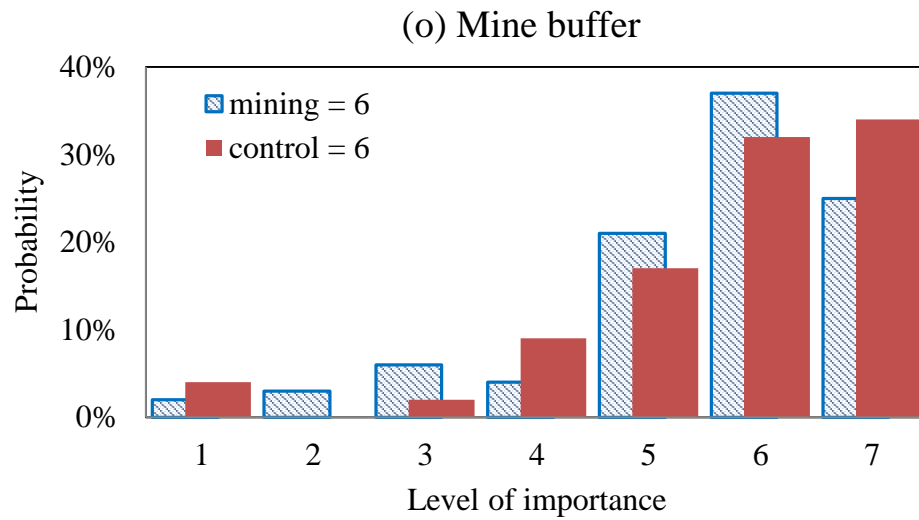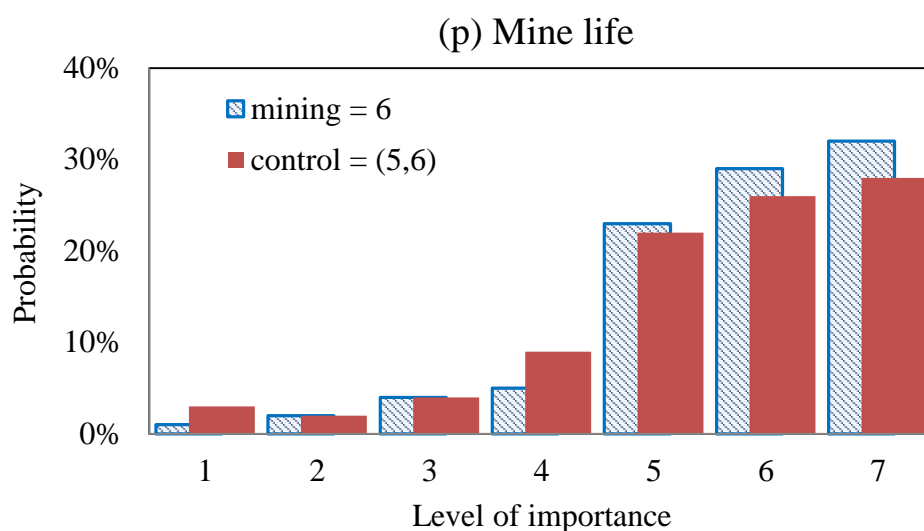
Figure 3-3. Level of importance (cont.)

For decision making mechanism, and independent & transparent information (Figure 3-3m and n), less than 20% in both groups thought these factors are not important in their decision to support or not support a mining project in their community. The probability of selection increases as the level of importance increases until level 6 (very important). Most of the participants, in both groups, view the level of importance of these two project characteristics as 6 (very important).

For mine buffer (the distance between the mine and the respondent's residence), most participants decided the level of importance is 6 or 7 (very important or extremely important). While relatively more people from the mining group think its level of importance is 6 (very important), more participants from the control group view its level of importance as 7 (Extremely Important). For mine life, the probability of selection increases as the level of importance increases. Most of the participants, in both groups, view mine life as extremely important (level 7).

3.4.4.5. Summary. In summary, as shown in Table 3-6, all 16 mining project characteristics are verified as important factors for both groups of respondents. Both groups consider job opportunities, water shortage or pollution, air pollution and land pollution very important, with their median level of importance between 6 and 7, which means above "very important" and below "extremely important." The median level of importance assigned to traffic and crime increase, income increase, and mine buffer is equal to 6 (very important) in both mining and control groups.

Population changes, cultural impact, cost of housing or housing shortage, decision making mechanism, and mine buffer were also assigned the same median level of importance by both groups. The median level of importance for population changes is 5 (somewhat important). The median level of importance for the other four factors is between 5 and 6, which means above "somewhat important" and below "very important."

The two groups' opinions appear to differ, very slightly (by less than 1 point on a 7 point scale), on the remaining four mine characteristics: infrastructure improvement, labor shortage for other businesses, noise pollution, and mine life. For the first three factors, respondents living in mining communities appear to view them as less important than those who live in non-mining communities. But mine life is more important for the mining group than the control group.

The respondents provided many suggestions for additional factors in the open ended question at the end of the survey. 48 out of 200 indicated explicitly that the provided list was adequate with answers like, *"none", "n/a", "I don't know" or "the survey captures it all."* Most of the suggestions in the other answers were not consistently recurring, except for land subsidence which was mentioned in some form by 6 out of the 200 participants. The author is of the opinion that land pollution alone (as defined in this survey) did not seem to capture fully the respondents' perception of subsidence, which appears to be an important factor for these respondents. The author

suggests either combining land subsidence explicitly with land pollution in the project characteristics list or as a separate factor.

3.5. DISCUSSIONS

This Section is an attempt to provide a proper preliminary study for discrete choice experiments for mining community engagement. The work attempts to provide a systematic process to: (1) identify key demographic factors; (2) determine the important mining project characteristics, which affect people's acceptance of a mining project (in most cases); and (3) evaluate whether there are differences between individuals living in mining and non-mining communities.

As shown in Table 3-3 and Table 3-5, four, out of six, demographic factors have been identified as significantly correlated to respondents' opinion of the level of importance of mine characteristics (at $\alpha = 0.05$). The author postulates that this observed correlation means these demographic factors will be important explanatory variables of the decision to accept or reject a mining project. Hence, these demographic factors (at least in the USA context) should be included in any discrete choice experiments aimed at facilitating mining community engagement.

This result partly confirms three (age, gender and income) of the four demographic factors used by Ivanova & Rolfe (2011) in their discrete choice experiment in Australia as useful in the USA context too. However, 'number of children' was not observed to be significantly correlated to the respondents' answers in this research. Level of 'Education', which has been found significant for local acceptability of wind-farm investment (Dimitropoulos & Kontoleon, 2009), was found to be significant in this work. 'Job field' was not identified as important in this research, although the author hypothesized it will be important, based on Muradian (2003).

Also, the specific nature of the correlation (positive or negative) should provide some insights into which demographic groups are more or less likely to accept certain mining projects. Of particular interest is the observed negative correlation between income and traffic and crime increase, water shortage or pollution, air pollution, and land pollution, for the respondents from mining communities. For these same respondents, a negative correlation was observed between education and job opportunities and income increase. It will appear then that higher income earners and more educated individuals perceived these six characteristics as less important than lower income and less educated individuals. Further research is required to determine the underlying reasons for this. Stated preference surveys (discrete choice experiments) can reveal how these same demographic groups will make hypothesized choices, once they are presented with alternatives that require trade-offs between these. However, anecdotally, this result will seem to challenge the perception that poor and uneducated individuals do not care about these negative impacts.

From Table 3-5, it is apparent that all sixteen mining project characteristics were deemed important by both groups of respondents. This result confirms what is known in the literature (Dudka & Adriano, 1997; ICMM, ICRC, IFC, 2011; ICMM, 2010, 2012a; IFC, 2009; Ivanova & Rolfe, 2011; Lockie et al., 2009; Muradian et al., 2003; Petkova et al., 2009; Schooten et al., 2003; K. Willis et al., 2011). The primary goal of this work was to validate the classification of mining impacts, by the author, for discrete choice experiments. The list seems to have widely captured all the factors respondents will consider in their decision to support a mining project, as no persistent independent themes emerged from the open ended question. Only land subsidence appeared consistently in their responses. This will suggest that discrete choice surveys with this classification of factors will need to find a way to incorporate land subsidence. It is important to note that these insights are limited to the USA context and further work will

be required to extend this to other contexts. Even in the USA, for a given mine, some of these factors may not be as relevant and a specific issue (even though captured by the classification – e.g. land subsidence and land pollution in this work) may have to be included in the discrete choice experiment, explicitly.

The survey results will seem to suggest that the most important mining project characteristics are job opportunities, water shortage or pollution, air pollution and land pollution with median of level of importance above "very important" and below "extremely important." The least important project characteristics are population changes and labor shortage for other businesses, which was found to be above "neutral" and below or equal "somewhat important." These differences may be marginal (1- to 2-points on a 7 point scale) but are statistically significant (i.e. low probability that the observation is by chance). Further work is required to confirm them. However, they provide some insight into how mine managers and community engagement professionals may approach mine design and community engagement.

The author hypothesizes that the observed difference (however, slight) in opinions between respondents living in non-mining and mining communities is due to knowledge and experience gaps. The mental representation of a problem – the perception of a situation – is central to the decision-making process. This perception is based on what you know about a problem and all judgments made are based on this perception (Chris Horn 2006). Respondents from non-mining communities know mining from their acquired knowledge, through information from the media and other sources. However, respondents from mining communities know mining from practical experience as well as the information available to respondents from non-mining communities. As shown in Table 3-4, the respondents from mining communities view infrastructure improvement, labor shortage for other businesses, and noise pollution as less important, but think mine life is more important, than those from non-mining communities, based on their practical

experience with mining. Although the author is not sure what the specific reasons for this observation are, the difference in experience and knowledge is, most likely, at the core. For example, very few members of the general public know how big the differences in mine life (which can exceed 100 years) can be. Respondents who live in mining communities and know how broad the spectrum of mine lives are, are more likely to rank mine life as a more important charactersitic.

In this study, 16 mining project charactersitics (positive and negaitive impacts) and six demogrphic factors were evaluated based on a literature review. The author carefully selected impacts that were independent and covered a broad spectrum of impacts. The choice of these 16 characteristics are validated by the open-ended question that did not identify any significantly different characteristic (land subsidence has the potential to be highly correlated to land pollution). The results of this study suggest that these 16 mine characteristics (or other classification that covers the same spectrum) and four of the six demographic factors should be included in any discrete choice experiments for mining community engagement in the USA. Ignoring any of these will result in a large intercept (constant) in the discrete choice model and unreliable model predictions.

Mining industry professionals can use these research results in three specific ways: (1) As shown in Table 3-6, job opportunities appear more important than higher incomes. Thus, during community consultation a company should highlight the increased job opportunities rather than higher salaries. Mine design alternatives that improve job opportunities (i.e. higher labor needs) are more likely to be supported by individuals in the local community than those that reduce labor needs but provide higher paying technical jobs. This is particularly pertinent given the shift towards autonomous mining systems. Also, noise pollution (somewhat important) is much less important than air, water and land pollution (above "very important" and below "extremely important."). More effort should be dedicated to air, water, and land pollution issues, during the

consultation process, than on noise pollution control. (2) Respondents from mining and non-mining communities have slightly different opinions of four mine characteristics: infrastructure improvement, labor shortage for other businesses, noise pollution, and mine life. Although, this result still needs further work to fully confirm (the differences in this section may be just marginal), the result might mean some adjustment in community engagement is necessary during permitting in a community with no prior experience with mining. (3) Gender, income, age, and education are important predictors of an individual's decision to accept or reject a proposed mining project. It is important, therefore, to engage different sectors of the community differently since their priorities are different. The "listening" part of community consultation is important to identify the different priority issues for different demographic sectors.

The exact nature of the interactions between these 16 mining project characteristics and four demographic factors in stated preference surveys are explored in Section 6. The author conducted discrete choice modeling of mining community acceptance in a selected community. This methodology developed and the example choice model will help facilitate better community engagement by providing a unique perspective, which is not yet widely used in mining community acceptance despite wide application in other disciplines.

3.6. SUMMARY OF SECTION THREE

All sixteen project characteristics, identified and classified through a literature review, were confirmed as important to the decision to accept or not accept a mining project. The most important mining project characteristics are job opportunities, water shortage or pollution, air pollution and land pollution. Respondents living in mining and non-mining communities have similar opinions of 12 mine characteristics and appear to differ on four (infrastructure improvement, labor shortage for other businesses, noise

pollution, and mine life). Four of the six demographic factors were confirmed to be significantly ($p < 0.05$) correlated with respondents' opinion of the importance of the mine characteristics.

During design and management of mines and community consultation for mining projects, these results can be used as a guide. The results will facilitate better choice experiment (survey) design for discrete choice modeling. Such discrete choice models can provide a viable framework for quantitative data-driven community engagement and sustainable mine management.

4. DETERMINING THE OPTIMUM NUMBER OF FACTORS FOR MINING DISCRETE CHOICE EXPERIMENTS

4.1. INTRODUCTION

The second huddle of incorporating discrete choice modeling into mining community analysis is how to design good discrete choice experiments (DCEs) for mining community consultation. According to the research results presented in Sections 2 and 3, surveyed respondents drawn from mining and non-mining communities consider at least 16 important characteristics in their decision to support a mining project (Que, Awuah-Offei, & Samaranayake, 2015). However, most choice experiments use fewer than 10 attributes, with the average being around 5 or 6 (Ryan & Gerard, 2003). Ivanova and Rolfe (2011) considered only five characteristics of the mine development options in order to keep it "simple and concise" so that respondents can complete the survey with ease (i.e. reasonable cognitive burden). However, using only five attributes led to a high alternate specific constant, which indicates that the selected attributes do not fully explain respondent's preferences (Train, 2002).

Discrete choice experiments with large number of factors result in complicated choices, which require significant cognitive effort by respondents (Caussade, Ortúzar, Rizzi, & Hensher, 2005; Hoyos, 2010). This can lead to a gap between the cognitive ability of respondents and the cognitive burden of the decision they are asked to make. Using 16 attributes in a discrete choice experiment will, most likely, lead to a higher than bearable cognitive burden for respondents. The challenge then is how to incorporate enough attributes, while ensuring the choice experiment will lead to reasonable cognitive burden and, consequently, valid results.

In this section, the author proposes an approach to overcome this challenge for mining community engagement, where there are many attributes to the choice (as many as 16, based on the results from Section 3). The approach is based on incorporating all

important characteristics (16 in this work) into discrete choice experiments by using block scheme designs, in which factors are split into several discrete choice experiments.

4.2. BLOCK SCHEME DESIGN

Block scheme design is a method that can be used when there are many attributes because of respondent burden and/or sample size considerations. In the block scheme design, the attributes are split into several separate discrete choice experiments. It is an important new method to consider since it is not always possible or desirable to reduce the number of attributes used.

There are two reasons for keeping the number of attributes relatively small (Ryan & Gerard, 2003). Firstly, with a large number of attributes, individuals may use other decision heuristics or lexicographic decision rules instead of make trade-offs. This outcome violates a key assumption of the economic choice theory that rules out the interpretation of the data as utilities (Scott, 2002). By contrast, a smaller number of attributes reduces the task complexity for respondents who are more likely to enable compensatory decision rules. The choice sets are constructed with reasonable factors, which reduces the cognitive burden placed on respondents. A cognitive demanding set may cause respondents to randomly select a choice rather than making a rational choice.

The second reason for having a smaller number of attributes is more pragmatic; the fewer the permutations of attributes and levels, the smaller the number of choice sets that needs to be presented. This will reduce the need to block the choice sets across different versions of the questionnaire, and therefore may reduce the sample size required to complete a given number of choice sets, which, in turn, reduces the cognitive burden of the choice tasks. The full combination of 16 factors, each with three levels, is $3^{16} = 43,046,721$. If the factors are split into four blocks of four attributes each, the full

combination is $3^4 \times 4 = 324$. This means, not only will researchers save huge amounts on survey cost, but also each participant will deal with much less choice sets.

Blocking is usually used to reduce the number of choice sets that each respondent has to answer, whereas we use it to reduce the number of attributes. For example, the block scheme design has been used in Witt, Scott, & Osborne (2009). In that study, a choice experiment in the situation where there are 11 attributes, of which 10 have four levels, and one has three levels, is carried out. Instead of presenting each respondent with all 11 attributes, the attributes are 'blocked' into three experimental designs.

However, before using a block scheme design to account for the large number factors in discrete choice experiments for mining community acceptance evaluation, there is a need to investigate the optimal number of factors for one choice set. Without this, the block scheme may still be to tasking to the respondent (still too many attributes in each choice set) or inefficient (too few factors in each choice set), leading to higher than necessary costs.

To bridge this gap, this section introduces block scheme design for discrete choice experiments and uses an online survey to determine the optimal number of attributes per choice set. The main research objective is to determine whether there is an optimal number of factors to include in choice experiments or not. Discrete choice experiments were designed with different number of factors, from three to six. Respondents' were asked to rate the required effort and difficulty of each choice set. The effort and difficulty level ratings were tracked for each design to find the optimal number of factors. This research provides preliminary results for effective and efficient block schemed discrete choice experiment design.

4.3. EXPERIMENTAL METHOD

4.3.1. Sample Size Determination. The margin of error is a statistic expressing the amount of random sampling error in a survey's results. Margin of error occurs whenever a population is incompletely sampled. The larger the margin of error, the less confidence the results represent the whole population.

$$E = z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$$ (4-1)

E: margin of error

α: level of confidence

$z_{\alpha/2}$: critical value

σ: standard deviation of the population

n: sample size

The Equation (4-1) is used to calculate the margin of error of a sample, assuming the population is normally distributed (Taylor, 2013). To balance data collection costs and the need for acceptable margin of error, the authors chose a sample size of 200, which corresponds to 5 % margin of error at 95% level of confidence.

4.3.2. Survey Design. An online survey was conducted with a questionnaire, consisting of three parts, to evaluate the optimum number of factors that should be included in one choice set. The first part of the survey contained background questions regarding the respondent's socioeconomic status. The demographic questions included age, gender, income, education, job field, and number of children.

The second part contained four discrete choice experimental designs with number of factors varied from three to six, in increments of one. Each design has nine choice sets with the same factors but different combinations of the factor levels. Factors were selected from the 16 factors validated in Section 3 (Table 3-1) by including one factor from each of the four categories (economic, social, environmental, and governance), and

balancing the positive, negative and neutral effects. The factors included in each design

and their levels are shown in Table 4-1. A choice set sample is shown in Table 4-2.

Table 4-1. Factors of each case design

Design	No. of factors	Factors
1	3	Income increase (Positive, Economic) (1)+ $100 per month (2)+ $300 per month (3)+ $500 per month Population increase **(Neutral, Social)** (1) A reduced rate of population growth (only 2%) (2) Continued population growth (average rate 4%) (3) An increased population growth (6%) Mine life (Neutral, Governance) (1)10 years (2)20 years (3)30 years
2	4	Air pollution (Negative, Environmental) (1)No increase in pollution (2)A slight increase in pollution (3)A moderate increase in pollution
3	5	Job opportunities (Positive, Economic) (1)300 people employed directly by the mine (2)600 people employed directly by the mine (3)900 people employed directly by the mine
4	6	Permit approval decision making mechanism (Neutral, Governance) (1)Final decision solely by Government agency (2)Final decision by Government agency after significant public input (3)Final decision by Government agency after negotiating with local representatives

Participants were asked to select one of three mine options, if a new mine were to

be opened in their community. Each choice set has four alternatives: Option 4 is "Too

Complex to Decide." Following each question, participants were asked to rank their level

of perceived mental effort (i.e., 36 mental effort ratings), by selecting a number from 1 to

7 ("very easy" to "very difficult"). After completing the questions in each discrete choice

experiment design, participates were asked to assess the perceived difficulty of the choice experiments with three, four, five, and six factors (i.e., four difficulty ratings). The same seven levels (from very easy to very difficult) were used in the difficulty rating. The time it took for each participant to complete each design was tracked by the system.

Table 4-2. Example of choice sets

	Income increase	Population increase	Mine life	Air pollution	I would choose
Option 1	+$300 per month	Continued population growth (average rate 4%)	20 years	A slight increase in pollution	☐
Option 2	+$100 per month	Continued population growth (average rate 4%)	30 years	No increase in pollution	☐
Option 3	+$500 per month	A reduced rate of population growth (only 2%)	20 years	A moderate increase in pollution	☐
Option 4	Too complex to decide				☐

This survey was conducted in October, 2013. Based on the sample size calculation, 210 individuals were recruited to complete the online survey. This survey was supposed to be a 30 minute online survey, and the data was automatically eliminated from anyone who completed the survey in less than 10 minutes. The data of eight participants were deleted because they did not follow instructions. The data from the remaining 202 participants was used in the following data analysis.

Among the respondents, 52% participants were female and the other 48% were male. The ages ranged from 19 to 78 years old, with a mean age of 52.84. This survey was computer-assisted personal interviews, administered by Qualtrics. The system tracked respondents by their zip code using the IP address they used to access the survey.

4.4. DATA ANALYSIS

4.4.1. Correlation Analyses. The frequency with which respondents chose Option 4, "Too Complex to Decide", for each design is shown in Table 4-3. There are 1,818 total responses for each design (=9×202, nine choice sets for each case and 202 respondents for each choice set). The design with three factors has the lowest frequency of Option 4 answers: 84/1818, which is less than 5%. The design with four factors has a higher frequency at 143 responses, and the designs with five and six have even higher frequency at 152 and 158 responses, respectively.

Table 4-3. Frequencies for Option 4 selections

Design	Number of factors	Observed frequency of Option 4 "Too complex to choose"	Percent
1	3	84/1818	4.62%
2	4	143/1818	7.86%
3	5	152/1818	8.36%
4	6	158/1818	8.69%

Correlation analysis was performed to examine whether there is a significant relationship between the frequency of choosing Option 4 and the number of factors in one choice set (Rodriguez, 2006). The correlation coefficient is a measure of the relationship strength, and varies between -1 and +1. The bigger the absolute value of the correlation coefficient, the stronger the relationship. Positive correlation coefficient means the two variables are proportional while negative correlation coefficient indicates the two variables are inversely proportional to each other (Rodriguez, 2006).

Hypothesis tests are used in correlation analysis to evaluate whether the observed correlation coefficient is due to random sampling or not. For correlation analysis, the typical null hypothesis is that the correlation coefficient is 0, and the alternative hypothesis is that the correlation coefficient is different from 0. The p-value (p) is the probability that the null hypothesis is true. If $p < \alpha$ (the level of significance), then the

probability of the correlation coefficient being 0 is smaller than α (typically 0.05). Thus, there is enough evidence to reject the null hypothesis, and conclude that the correlation coefficient is significantly different from 0 at significant level α. On the other hand, if $p >$ α, then the test fails to reject the null hypothesis, and we conclude that there is not enough evidence to counter the claim that the correlation coefficient is significantly different from 0 at significant level α.

There are three popular correlation coefficients and their corresponding tests: Pearson, Spearman and Kendall correlation coefficients (Gibbons, 2006; Pirie, 2006; Rodriguez, 2006; Stuart, 2006). Figure 4-1 shows a selection logic for selecting the best test. Pearson and Spearman test can be used when the variables are interval variables. The Kendall test can be used if the variables are ordinals or can be ranked as ordinals. The Chi-square test can be used if the variables cannot be ranked as ordinals (Jeffrey, 2006; Koch & Bhapkar, 2006; Stuart, 2006). In this case study, the variables are ordinals from levels 1 to 7. Thus, the Kendall test was used to compute p-values for Kendall's tau-b to detect the correlation coefficient.

The Kendall correlation analysis was done using the SAS FREQ Procedure, which computes the Kendall tau-b correlation coefficient and uses the estimate in hypothesis testing ($H_0 : r = 0$). The result of Kendall test is shown in Table 4-4. The null hypothesis cannot be rejected at $\alpha = 0.05$ with p-value <0.0001, and we can conclude that the Kendall's tau-b is significantly different from 0 at significant level 0.05. The Kendall's Tau-b coefficient is 0.0503. This is small and raises some doubt on whether the frequency of selecting "too complex to choose" is significantly correlated to the number of factors in the choice set. Further evidence is needed to support this claim.
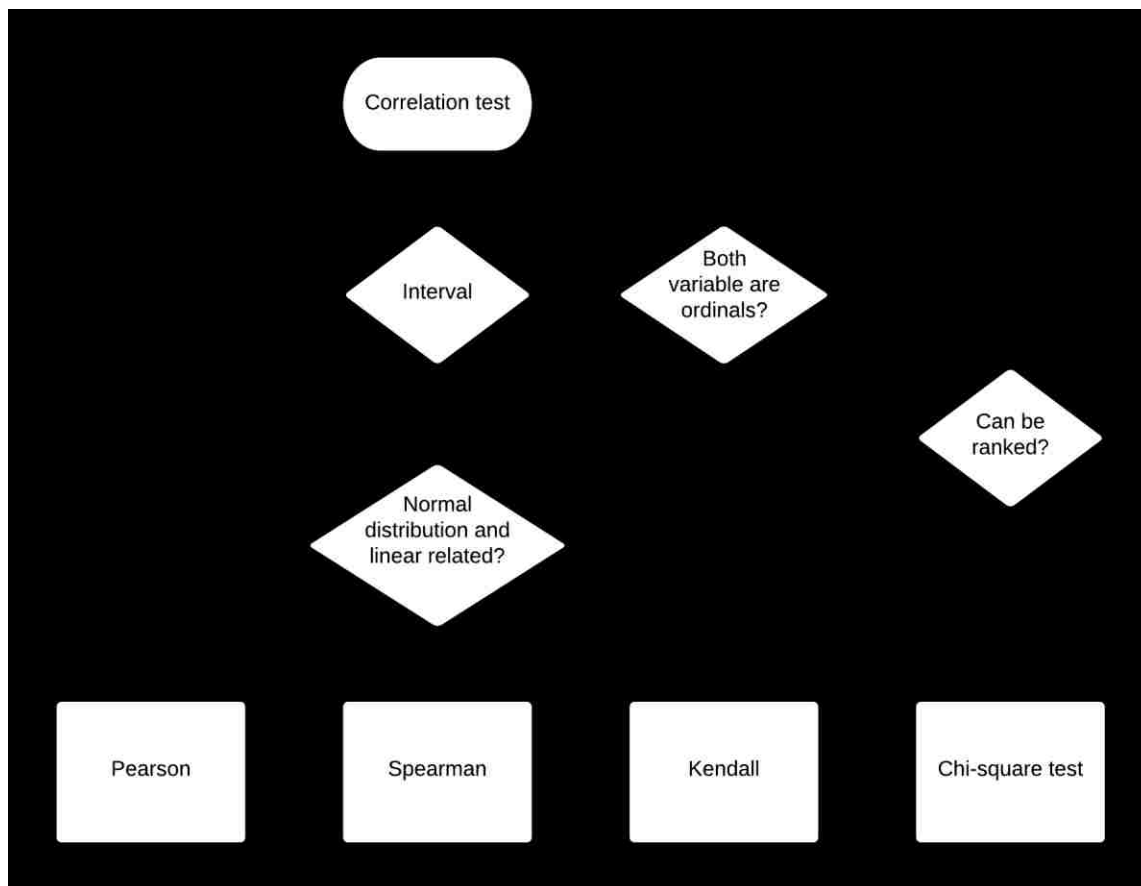
Figure 4-1. Flowchart of correlation test approach

Table 4-4. Result of Kendall's tau-b correlation test

The FREQ Procedure	
Statistics for Table of Exposure by Response	
Kendall's Tau-b	0.0503
ASE	0.0101
95% Lower Conf Limit	0.0304
95% Upper Conf Limit	0.0702
Test of H_0: Tau-b = 0	
ASE under H_0	0.0103
Z	4.8921
One-sided Pr > |Z|	< 0.0001
Two-sided Pr > |Z|	< 0.0001

4.4.2. Analysis of Effort and Difficulty Ratings. In this study, the effort and difficulty level ratings were acquired for questions from all four designs. Statistical analysis was used to evaluate whether or not there is a significant difference between the level of effort and difficulty ratings for these four designs. The data on effort and difficulty level ratings from the fours designs were treated as independent groups.

There are three data analysis methods for more than two independent groups: ANOVA, Welch ANOVA, and Kruskal-Wallis (Schlotzhauer, 2009). Figure 4-2 shows the logic for selecting the proper test. ANOVA and Welch ANOVA tests can be used when the variables are interval variables. In this case study, the data on the level of effort and difficulty were treated as ordinal. Thus, the Kruskal-Wallis test is the most appropriate test.

The null hypothesis is that 'there is no significant difference between the effort/difficulty level distributions of these four designs'. The p-value (p) is the probability that the null hypothesis is true. If $p < \alpha$ (the level of significance), then we have enough evidence to reject the null hypothesis, and conclude there is at least one significant difference between the effort/difficulty level ratings of these four designs at significant level α (typically 0.05). On the other hand, if $p > \alpha$, then the test fails to reject the null hypothesis at significant level α.

Nemenyi and Dunn's multiple comparison are multiple comparison tests based on significant Kruskal–Wallis results. In this case, the Dunn's multiple comparison test is used since the data are classified as tied rank.
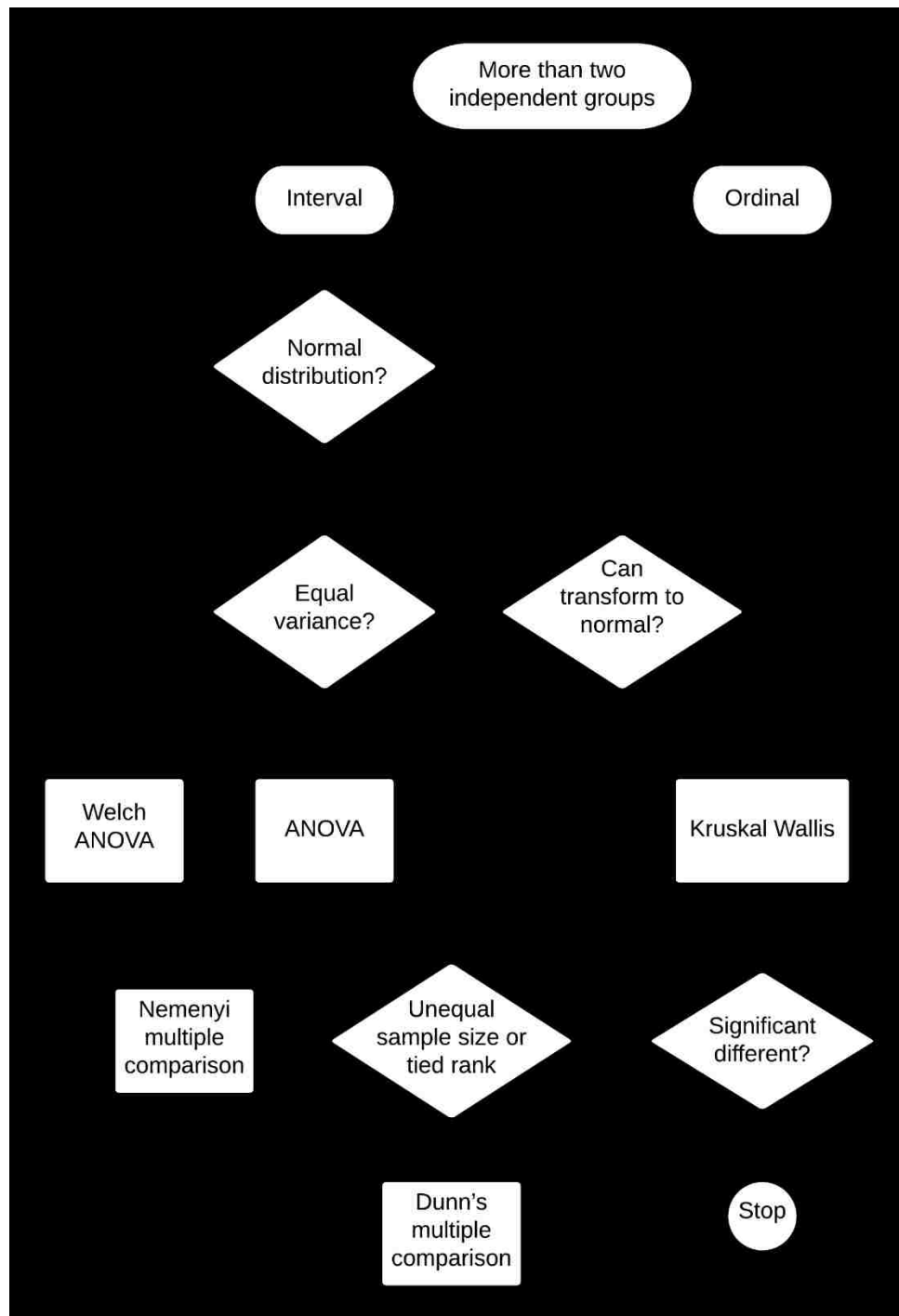
Figure 4-2. Logic for selecting data analysis method for more than two independent groups

An important part of Dunn's test is an estimate of the standard error (SE). For this test, the SE is calculated using Equation (4-2) (Dunn, 1964). Then, a Q_{AB} statistic and a Q critical value are calculated using Equations (4-3) and (4-4), respectively (Zar, 2010).

$$SE_{AB} = \sqrt{\left(\frac{N(N+1)}{12} - \frac{\Sigma t}{12(N-1)} \right)\left(\frac{1}{n_A} + \frac{1}{n_B} \right)} \tag{4-2}$$

Where

$$\Sigma t = \sum_{i=1}^{K}\left(t_i^3 - t_i \right)$$

K is the number of comparison groups

N is the total sample size

Σt is the total count of tied rank

$$Q_{AB} = \frac{\overline{R_B} - \overline{R_A}}{SE} \tag{4-3}$$

Where

$\overline{R_i}$ is the rank for group i

$$Q = 1 - \left(\frac{Alpha}{K \times (K-1)} \right) \tag{4-4}$$

The null hypothesis for each pairwise comparison is that there is no significant difference between these two groups. If $Q_{AB} > Q$, then we have enough evidence to reject the null hypothesis, and conclude that there is a significant difference between the difficulty/confusion level ratings of these two cases at significant level α (typically 0.05). On the other hand, if $Q_{AB} < Q$, then the test fails to reject the null hypothesis at significant level α.

4.4.2.1. Effort level analysis. The effort level rating distributions are shown as histograms and box plots in Figure 4-3 and Figure 4-4, respectively. From Figure 4-3, all the distributions are relatively asymmetric and skewed to the left. Most of the participants selected level 2 (Easy) or 3 (Somewhat easy) for effort ratings of all four designs with number of factors from three to six. As the number of factors increase from three to six, the percentage of respondents selecting level 4 (Neutral) increased from 15.7% to 18.2%. Similarly, the percentage of respondents selecting levels 5, 6 and 7 increase from 6.9 % to 11.6%, 2.4% to 4.3%, and 0.7% to 1.2%. Figure 4-4 shows that the median effort level ratings increase as the number of factors increase, even though the mean effort level ratings of each design are very similar.
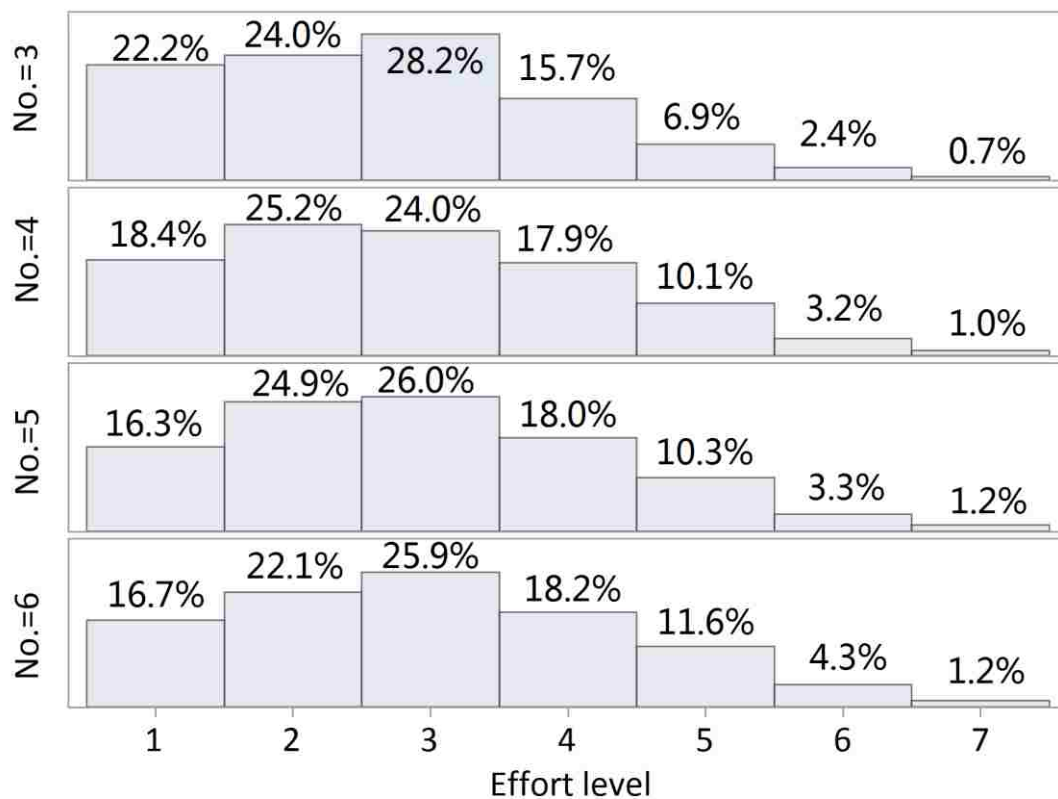
Figure 4-3. Effort level rating histograms

Level 1 Vary easy
Level 2 Easy
Level 3 Somewhat easy
Level 4 Neutral
Level 5 Somewhat difficult
Level 6 Difficult
Level 7 Very difficult
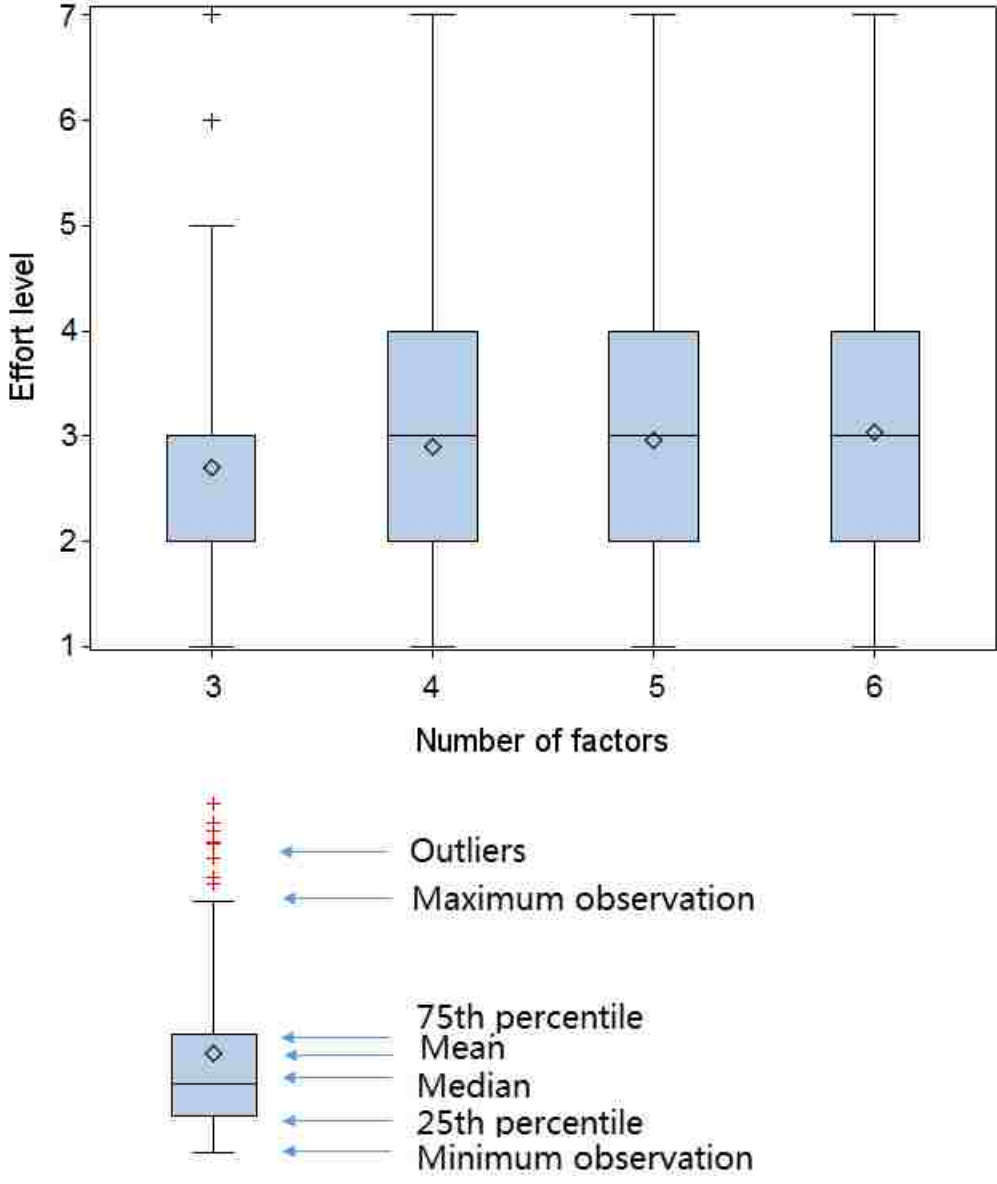
Figure 4-4. Effort level rating boxplot

The Kruskal-Wallis test was done, using the SAS PROC NPAR1WAY WILCOXON procedure (SAS, 2007k), to evaluate whether there is a significant difference between these four designs. The p values were estimated as <0.0001, which means the test rejected the null hypothesis. It can be concluded that there is at least one significant difference among the effort level ratings of these four designs, at significant level 0.05.

Dunn's multiple comparison was used to find the detail differences between each pair of designs (Elliott & Hynan, 2011; Schlotzhauer, 2009). As shown in Table 4-5, the Q critical value is 2.638 at significant level $\alpha=0.05$. The effort level ratings of the three-factor design is significantly different from that of the other three designs. The calculated Q_{AB} statistic (3.85, 5.21, and 6.75) exceeds the Q critical value. The effort level ratings of the four-factors design is significantly different from the six-factors design ($Q_{AB} = 2.90$). However, the effort level ratings of the four-factors design is not significantly different from that of the five-factors design ($Q_{AB} = 1.36$), and effort level ratings of the five-factors design is not significantly different from that of the six-factors design ($Q_{AB} = 1.54$).

Table 4-5. Results of the Dunn's multiple comparisons test for effort level ratings

Comparison group=Number of factors					
Compare	Diff	SE	Q_{AB}	Q(0.05)	Conclude
6 vs 3	459.81	68.08	6.75	2.638	Reject
6 vs 4	197.71	68.08	2.90	2.638	Reject
6 vs 5	104.80	68.08	1.54	2.638	Do not reject
5 vs 3	355.01	68.08	5.21	2.638	Reject
5 vs 4	92.91	68.08	1.36	2.638	Do not reject
4 vs 3	262.10	68.08	3.85	2.638	Reject

4.4.2.2. Difficulty level analysis. The difficulty level rating distributions and boxplots are shown in Figure 4-5 and Figure 4-6, respectively. From Figure 4-5, the mode of the distribution for the three- and four-factors design is level 2 at 26.8% and 25.9%; for the five-factors design is level 4 at 21.0%; and the six-factors design is level 5 at 19.5%. For the three-factors design, the distribution is relatively asymmetric and skewed to the left. However, as the number of factors increase, the distribution is less and less skewed to the left. The distribution of the six-factors design is nearly symmetric. Note from Figure 4-6 that there is a jump between the difficulty ratings of three- and four-factors designs and five- and six-factors designs. The median of the difficulty level ratings increase as the number of factors increase.
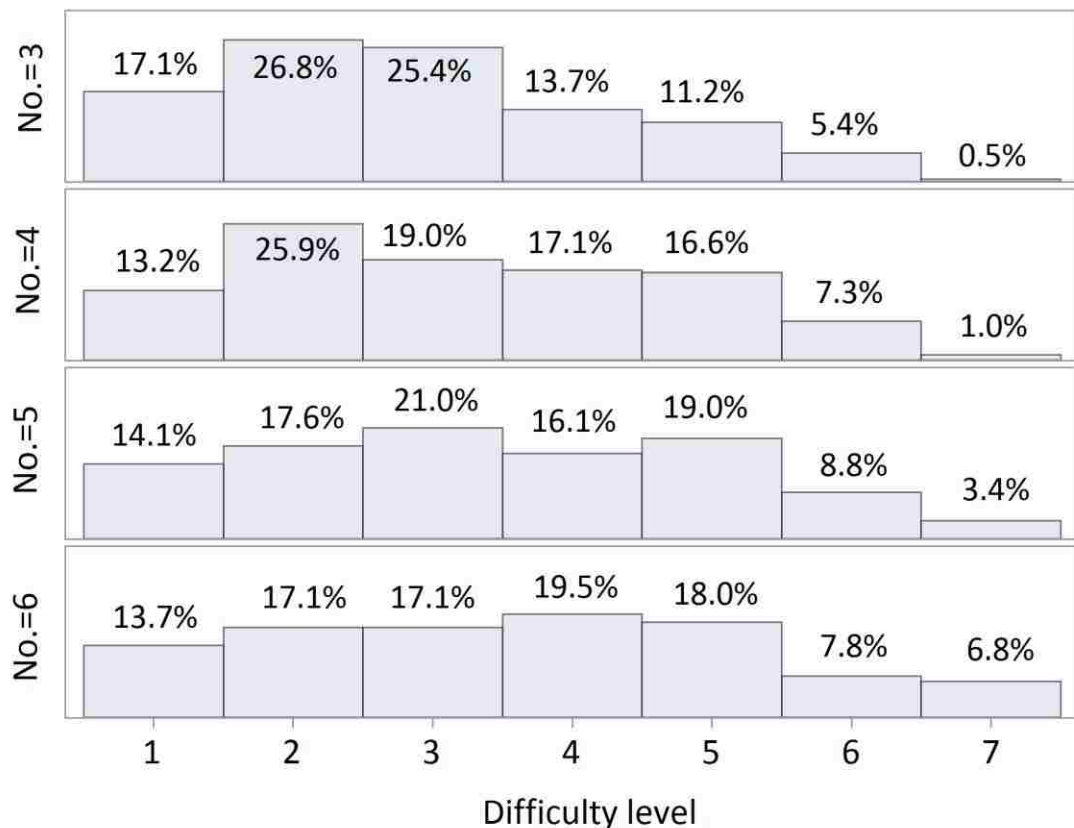

Figure 4-5. Difficulty level rating histograms

Level 1 Very easy
Level 2 Easy
Level 3 Somewhat easy
Level 4 Neutral
Level 5 Somewhat difficult
Level 6 Difficult
Level 7 Very difficult
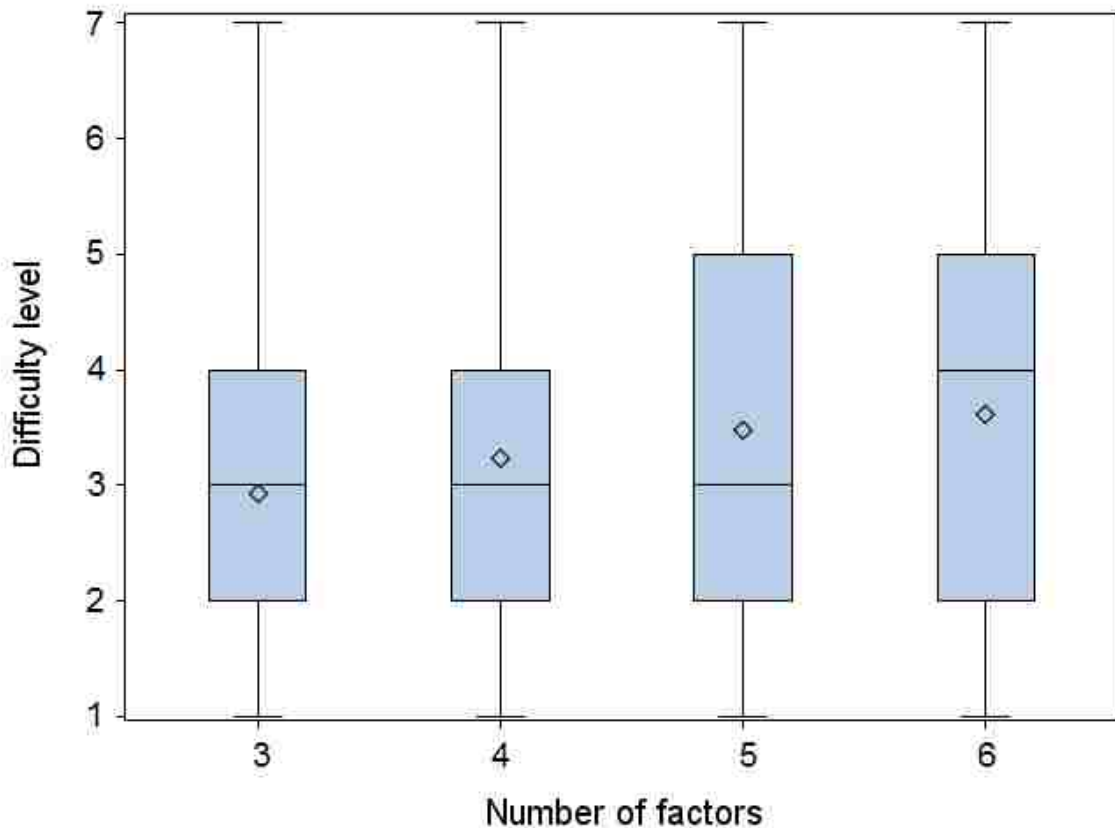
Figure 4-6. Difficulty level rating boxplot

The same Kruskal-Wallis test and Dunn's multiple comparison were used to analyze whether or not the number of factors included in one choice set had a significant effect on the difficulty level ratings. The null hypothesis of the Kruskal-Wallis test was rejected at $\alpha = 0.05$ ($p = 0.0003$). The result means there is significant difference between the difficulty level ratings of these four designs. The result of Dunn's multiple comparisons is shown in Table 4-6. The Q critical value is 2.638 at significant level 0.05.

While the difficulty level ratings of the three-factors design is not significantly different from that of the four-factors design ($Q_{AB} = 1.93$), it is significantly different from the designs with five and six factors. The calculated Q_{AB} statistic (3.36 and 4.03, respectively) exceeds the Q critical value. The difficulty level ratings of four-factors design is not significantly different from that of the six-factors design ($Q_{AB} = 2.10$). In addition, the pair wise comparison of difficulty level ratings of five- vs four and six- vs five-factors designs show no significant differences.

Table 4-6. Results of the Dunn's multiple comparisons test for difficulty level ratings

Comparison group=Number of factors					
Compare	Diff	SE	Q_{AB}	Q(0.05)	Conclude
6 vs 3	92.87	23.02	4.03	2.638	Reject
6 vs 4	48.40	23.02	2.10	2.638	Do not reject
6 vs 5	Do not reject (within non-sig. comparison)				
5 vs 3	77.25	23.02	3.36	2.638	Reject
5 vs 4	Do not reject (within non-sig. comparison)				
4 vs 3	44.47	23.02	1.93	2.638	Do not reject

4.4.3. Analysis Based on Duration of Survey. The time it took for each participant to complete the survey section with each design was tracked, and the comparative box plot is shown in Figure 4-7.

The mean duration is higher than the median in all four designs. The mean and median of time duration of the design with three factors are 30 and 22 seconds, respectively. For the design with four factors, both mean and median decrease slightly to 29 and 21 seconds, respectively. After that, the mean and median increase to the highest values for the design with five factors at 33 and 27seconds, respectively. Then, both mean and median drop sharply to the lowest values at 25 and 20 seconds, respectively, for the design with six factors.
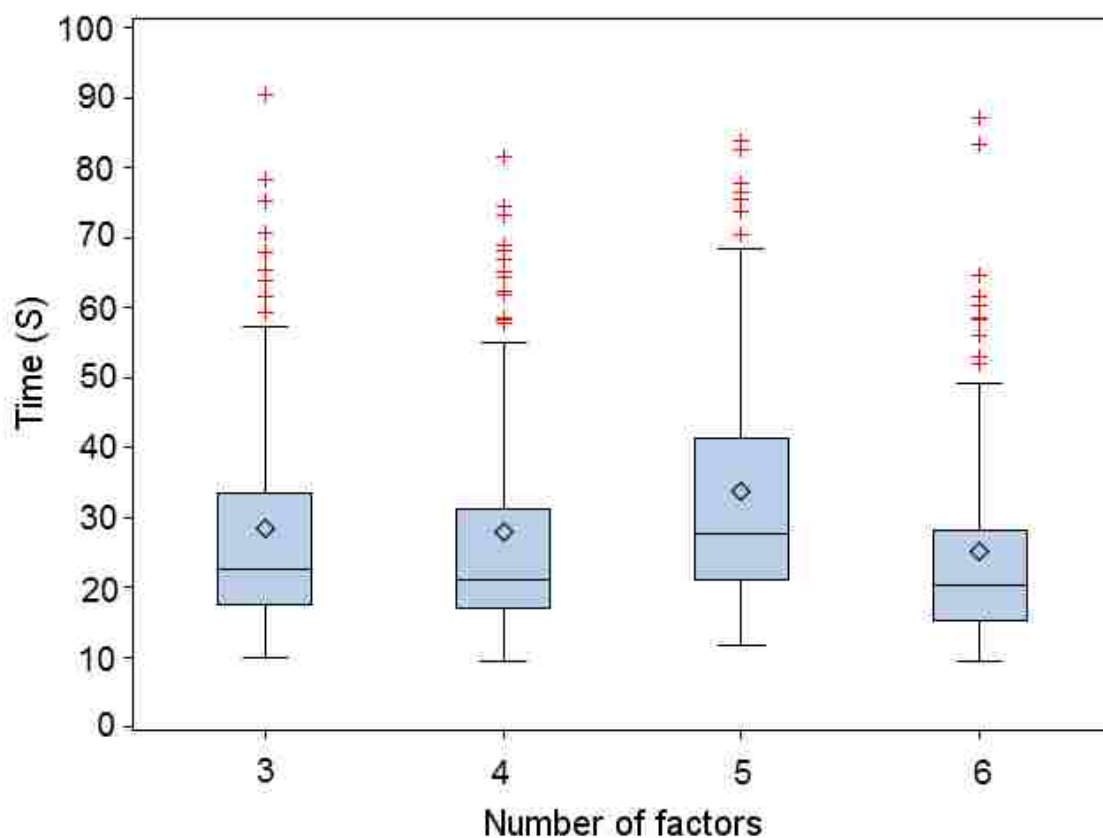
Figure 4-7. Comparative box plots for time consumed of the four designs

The data of survey duration are interval variables. Based on the logic for selecting the proper test shown in Figure 4-2, ANOVA and Welch ANOVA tests can be used to compare the duration data. However, the prerequisite of normality needs to be tested, first.

Common normality tests include Shapiro-Wilk (W) test (Shapiro, Wilk, & Chen, 1968), Kolmogorov-Smirnov (KS) test, Anderson-Darling (AD) test (Anderson & Darling, 1952), and Cramer-vol Mises (CM) test (Anderson, 1961). These four normality tests are all used in this study, and the null hypothesis in all tests are that the data follows a normal distribution. SAS PROC UNIVAREATE PROCEDURE was used to perform these tests on the duration data. The results are shown in Table 4-7 (Survey duration). The results of the tests show that the null hypothesis in all tests (data follows normal distribution) is

rejected and none of the survey duration of four designs follows normal distribution (all p-values are less than 0.01).

One approach to handle the violation of the normality assumption is to transform the data by using a natural log transformation (Zhou, Gao, & Hui, 1997). The same statistical tests were performed on the log transformed data and the results are shown in Table 4-7 (Log-Survey duration). Again the results show that the log-transformed data does not follow the normal distribution (all p-values are less than 0.01).

Table 4-7. Normality test of survey duration and log-survey duration

Survey duration								
	Design 1		Design 2		Design 3		Design 4	
Test	Statistic	P-value	Statistic	P-value	Statistic	P-value	Statistic	P-value
Shapiro-Wilk	0.580	<0.0001	0.484	<0.0001	0.805	<0.0001	0.646	<0.0001
Kolmogorov-Smirnov	0.211	<0.0100	0.249	<0.0100	0.151	<0.0100	0.195	<0.0100
Cramer-von Mises	3.480	<0.0050	4.748	<0.0050	1.750	<0.0050	2.964	<0.0050
Anderson-Darling	19.009	<0.0050	25.022	<0.0050	10.054	<0.0050	16.592	<0.0050
Log-Survey duration								
	Design 1		Design 2		Design 3		Design 4	
Test	Statistic	P-value	Statistic	P-value	Statistic	P-value	Statistic	P-value
Shapiro-Wilk	0.944	<0.0001	0.921	<0.0001	0.970	0.0003	0.943	<0.0001
Kolmogorov-Smirnov	0.093	<0.0100	0.115	<0.0100	0.083	<0.0100	0.089	<0.0100
Cramer-von Mises	0.397	<0.0050	0.620	<0.0050	0.301	<0.0050	0.342	<0.0050
Anderson-Darling	2.377	<0.0050	3.623	<0.0050	1.742	<0.0050	2.188	<0.0050

Thus, Kruskal-Wallis test and Dunn's multiple comparison were used to analyze whether or not the number of factors included in one choice set had a significant effect on

the survey duration. The null hypothesis of Kruskal-Wallis test was rejected at $\alpha = 0.05$

($p < 0.0001$). The result means there is significant difference between the survey duration

of these four designs. The result of Dunn's multiple comparisons is shown in Table 4-8.

Table 4-8. Results of the Dunn's multiple comparisons test for survey duration

Comparison group=Number of factors					
Compare	Diff	SE	Q_{AB}	Q(0.05)	Conclude
5 vs 3	87.27	23.22	3.76	2.638	Reject
5 vs 4	111.33	23.22	4.79	2.638	Reject
4 vs 3	24.06	23.22	1.04	2.638	Do not reject

From Table 4-8, it is shown that there is no significant difference between the

survey duration of designs with three and four factors ($Q_{AB} = 1.04$). This finding is

confirmed by the result that there is no significant difference between the difficult level

ratings of these two designs. In Figure 4-7, it appears the duration of design with four

factors is slightly less than that of the design with three factors. It is possible that the

respondents are learning as they proceed through the survey. Caussade et al. (2005)

observed that there is a learning effect in choice situations. Though the four-factors

design has one additional factor, the questions, framework, and style of choice situations

are the same.

In addition, the survey duration of the design with five factors is significantly

different from those with three and four factors. The calculated Q_{AB} statistic (3.76 and

4.79, respectively) exceeds the Q critical value (2.638). The design with six factors is

special in the survey duration analysis. From Figure 4-7, the survey duration drops

sharply to the lowest at 25 and 20 seconds. Most likely, the design with six factors is too

complicated, and participants are randomly selecting the choices. Thus, the duration data

from this design is treated as invalid and not used in the Kruskal-Wallis test and Dunn's

multiple comparison.

4.5. DISCUSSION

First of all, participants indicated that designs with the lower numbers of factors required less mental effort, and the choices were easier to make. As the number of factors increased, the perceived level of required mental effort and the perceived difficulty increased, as well. The result of Kendall correlation analysis appears to support the fact that the number of factors and the frequency of "too complex to choose" are positively correlated, with correlation coefficient of 0.0503. Although the magnitude of the correlation coefficient will seem to be small, the existence of such a correlation appears to be supported by the effort and difficulty ratings data. These findings support our hypothesis that a higher number of factors will lead to greater amounts of cognitive load (e.g. higher effort and difficulty ratings).

In addition, as shown by the results of the Kruskal-Wallis test of effort ratings, difficulty level ratings, and survey duration, there are significant differences between the participants' effort ratings, difficulty ratings, and survey duration for these four designs at significant level 0.05. The results of Dunn's multiple comparison tests, which were used for pairwise comparison of the designs, are summarized in Figure 4-8, Figure 4-9 and Figure 4-10.
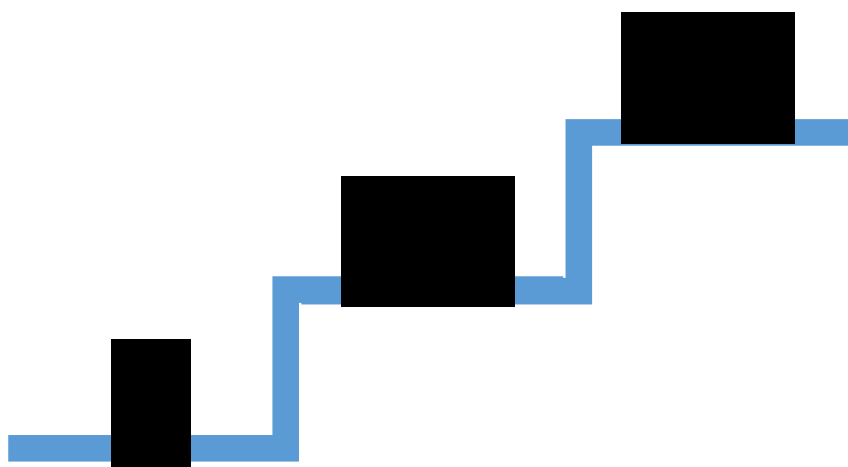
Figure 4-8. Summary plot of effort level ratings comparisons
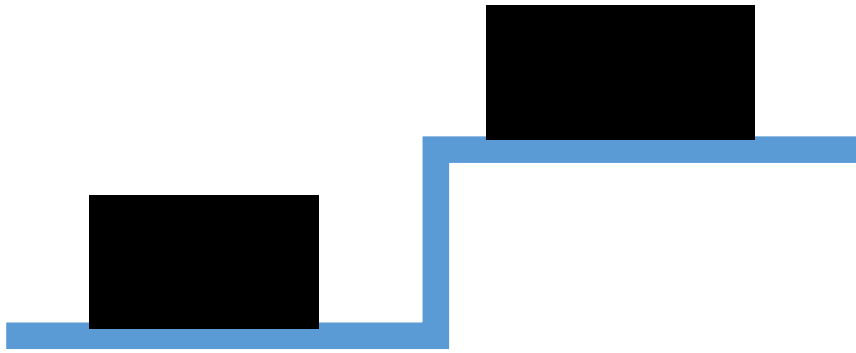
Figure 4-9. Summary plot of difficulty level ratings comparisons

Figure 4-10. Summary plot of survey duration comparisons

From Figure 4-8 and Figure 4-9, there is no significant difference between designs with five and six factors for both effort and difficulty level ratings. From Table 4-3 and Figure 4-5, the design with six factors is not easy as compared to the other three designs. Almost 10% (specifically, 8.69%) of the time, respondents indicated that choice sets with six factors are too complex to make a decision. The difficulty level rating mode for this design is level 4 (neutral), and 52.1% of the time, respondents indicated that choice sets with six factors are "neutral" (neither difficult or easy) difficulty or difficult. These results indicate that the designs with five and six factors are not good options for the block scheme experimental design.

The design with three factors is easiest among these four designs. Only 4.62% of the responses indicated the questions with three factors are 'too complex to decide.' 74.4%

and 69.3% of the time, respondents indicated that the effort and difficulty level, respectively, of these choice sets were equal or less than level 3 (somewhat easy). While the design with four factors needs, relatively, more effort and is ranked as more difficult than the design with three factors, it is still easy enough for the block scheme experimental design. 67.6% and 58.1% of the time, respondents indicated that the effort and difficulty level ratings, respectively, were equal or less than level 3 (somewhat easy). In addition, there is no significant difference between the designs with three and four factors for both the difficulty level and survey duration, as shown in Figure 4-9 and Figure 4-10. Based on these results, both of the designs with three and four factors are good options for the block scheme experimental design. The lower the number of factors included in one choice set, the more blocks are need in the block scheme experimental design, which will increase the cost of the overall survey. Thus, the design with four factors is the optimal choice to reduce cognitive burden and reduce costs.

4.6. COMPARISON WITH KLEIN ET AL. (2015) ANALYSIS

This case study has been published by Klein et al. (2015), who also identified the optimum number of factors in one choice set as four. This author conducted independent analysis of the data used by Klein et al. (2015) in this dissertation to ensure completeness. The main difference between this section and Klein et al. (2015) is the data analysis method. As shown by the method selection logic shown in Figure 4-2, the author chose the Kruskal-Wallis test and Dunn's multiple comparisons test since the effort and difficulty level were treated in this work as ordinal data. However, Klein et al. (2015) used the ANOVA test since they assumed the data to be interval data, an assumption that is usually accepted in social science journals for Likert scales. Most of the time (as long as the distribution is not too skewed), an ANOVA and a Kruskal-Wallis produce the similar results (Schlotzhauer, 2009). The main divergence (between this work and Klein

et al. (2015)) is whether to treat the Likert scales as ordinal or interval data in different fields of study. In engineering analysis, data like Likert scales are usually considered ordinal (Schlotzhauer, 2009).

4.7. SUMMARY OF SECTION FOUR

From the above, discussion, the following main points summarize the discussions in this section.

1. This section attempts to find a way to include all (16) factors in the discrete choice experiments while ensuring reasonable cognitive burden. This is to be done through a block scheme choice experimental design. Experiments were conducted with designs, which differed in the number of factors used to develop the choice sets. The number of factors varied from three to six factors.

2. There is no significant difference between effort and difficulty level ratings of the design with five and six factors. These two designs are not 'easy' enough for participants, and not good options for the block scheme experimental design.

3. The designs with three and four factors are 'easy' enough for participants. The design with four factors is the optimal choice for the block scheme experimental design since it balances cognitive burden and survey cost.

5   DISCRETE CHOICE EXPERIMENTAL DESIGN FOR MINING COMMUNITY ACCEPTANCE

5.1. INTRODUCTION

The second hurdle for incorporating discrete choice modeling into mining community analysis is how to design good discrete choice experiments (DCEs) for mining community consultation. For effective and efficient discrete choice experiment design, there are three important questions which need to be answered: (1) What is the optimum number of factors to consider in one choice set? (2) How do you design discrete choice experiment for mining community consultation? (3) How do you validate the discrete choice experiment design to ensure the data collected with by the survey is useful? Without answers to these questions, discrete choice experiment design would not yield useful data to help with community analysis.

In Section 4, the author addressed the first question and found the optimal number of factors in one choice set is four. Consequently, the other two questions are addressed in this section. This section presents an approach for designing DCEs which is validated with a pilot study. A focus group was used to examine the clarity of instructions and difficulty of the survey questions. The objectives of the work in this section are to: (1) formulate a general approach to discrete choice experimental design for mining community engagement; and (2) validate the proposed DCE design, using the 16 factors from Section 3.

5.2. RELEVANT MINING PROJECT CHARACTERISTICS FOR COMMUNITY CONSULTATION

Table 5-1 is the list of 16 project characteristics from four categories used in the discrete choice experiment in this work.

Table 5-1. Mining project characteristics used as attributes in choice experiment design (after Que et al. (2014))

Determinant
Social
Population increase
Infrastructure improvement
Traffic increase
Crime increase
Economic
Job opportunities
Income increase (for all local residents)
Increase in housing costs
Labor shortage for other business
Environmental
Noise pollution
Water shortage or pollution
Air pollution
Land pollution and subsidence
Management and others
Permit approval decision making mechanism
Availability of independent and transparent information on potential impacts of mine
Mine buffer (Home distance from mine)
Mine life

Compared to Section 3[7], these project characteristics were modified for clarity, effectiveness, and validity. First of all, "cultural impact" was deleted since there is no clear definition of cultural impact, which makes it invalid in a survey instrument. Second, "traffic and crime increase" were separated into "traffic increase" and "crime increase" since they are really two different items. In addition, "cost of housing or housing shortage" was revised to "increase in housing costs", "decision making mechanism on the mine's permits" has been revised to "permit approval decision making mechanism" , and "independent and transparent information" has been revised to "availability of independent and transparent information on potential impacts of mine" in an effort to provide clarity and conciseness. This list of mining project characteristics (impacts and

---

[7] the work in Section 3 has been published by Que et al. (2015)

other attributes) are then used as attributes, which are varied to generate alternatives in discrete choice experiments.

5.3. DISCRETE CHOICE EXPERIMENTAL DESIGN

Two important concepts for discrete choice experimental design are presented in this section. These are the concepts of D-error as well as the attributes and their levels. This is then followed by discussions on discrete choice experimental design for mining community consultation.

5.3.1. The Concept of D-error. D-error is used as an indicator of effectiveness of discrete choice experiments (Kuhfeld, 2010). As shown in Equations 5-1 and 5-2, D-error is the geometric mean of the eigenvalues and D-efficiency is the inverse of D-error. The D-efficiency is greater than 0 (i.e. no upper bound). If the effectiveness of a discrete choice experiment design is evaluated by the D-efficiency, it is difficult to know how good the experimental design is. The relative D-efficiency resolves this shortcoming of D-efficiency. Relative D-efficiency (Equation 5-3) ranges from 0 to 100%. The relative D-efficiency is used in this research.

$$\text{D-error} = \left| \Sigma \right|^{1/K}$$

(5-1)

$$\text{D-efficiency} = 1 / \left| \Sigma \right|^{1/K}$$

(5-2)

$$\text{Relative D-efficiency} = \frac{1}{\left| \Sigma \right|^{1/K} \times \text{number of choice sets}} \times 100\%$$

(5-3)

where

Σ ----the covariance matrix of choice sets design

K --- the number of parameters

5.3.2. Attributes and Levels. The number of attributes and number of levels per attribute is a key part of designing a choice experiment (Caussade et al., 2005; Hoyos, 2010). Firstly, the selected attributes and levels for each attribute should be important and relevant to the choice and potential participants. In this case study, the attributes and levels for each attribute are shown in Appendix D. The detailed discussion of how the author arrived at these attributes is contained in Section 3.

Secondly, the attributes and levels need to be realistic and framed appropriately. Take the levels of the attributes "job opportunities" and "income increase", as examples. The potential increases in number of jobs and incomes due to a mine varies depending on mine and community size. These levels should be selected carefully to account for different mines and associated local communities. In this study, all levels were determined bearing in mind a mine close to Salt Lake City, since the final survey was conducted there.

Thirdly, the attribute levels should easily be understood by the average respondent while providing useful information (Bateman et al., 2002; Bergmann, Hanley, & Wright, 2006). For example, one can add an explanation to the attribute "mine buffer" to explain that it is 'the distance of the respondent's home from the mine.' Examples such as transportation, education, human services, and internet can be provided for the attribute, infrastructure improvement.

5.3.3. Experimental Design Considerations. The experimental design considerations include four main aspects.

5.3.3.1. Stated preference vs. revealed preference. Revealed preference (RP) and stated preference (SP) are the primary discrete choice experiment methods. RP is a conventional method, which refers to situations where the choice is actually made in real market situations. In contrast to RP, SP refers to situations where a choice is made by considering hypothetical situations. The choice options of SP are similar to that of RP except that the choices in RP are limited to reality. This feature results in a major advantage of SP.

Currently, the most popular method is to have both RP and SP choice data (Ivanova & Rolfe, 2011; Winslott Hiselius, 2005). In the case study presented in this research, the status quo option shows the average value of each attribute/characteristic in the real world. The other two alternatives, in the choice set, use the same attributes but different combination of attribute levels to generate hypothetical situations. The design of this study is, consequently, a mixed design as recommended by Louviere, Hensher, and Swait (2003). Thus, we can show participants various real and possible hypothetical mining scenarios (Hensher, Rose, & Greene, 2005).

5.3.3.2. Block design. Discrete choice experiments with large number of factors result in complicated choices, which require significant cognitive effort by respondents (Caussade et al., 2005; Hoyos, 2010).

This can lead to a gap between the cognitive ability of respondents and the cognitive burden of the decision they are asked to make. Ivanova and Rolfe (2011) considered only five characteristics/attributes in developing mine development options, in order to keep it "simple and concise" so that respondents can complete the survey with reasonable cognitive burden. However, using only five attributes led to a high alternate

specific constant, which indicates the selected attributes do not fully explain respondents'

preferences.

In this case study, there are 16 important attributes or characteristics of the mining

project (Que et al., 2015). The challenge is to find a way to include all 16 attributes while

designing choice experiments with reasonable cognitive burden. Previous research (Klein

et al., 2015) has addressed the cognitive load for participants while completing discrete

choice experiments with hypothetical mining project choice sets. The optimal number of

attributes in one question was determined to be four (Klein et al., 2015). Thus, the 16

mining characteristics were divided into four blocks. Each block includes one factor from

each of the four categories (Table 5-1), and the factors are chosen to balance the positive

and negative effects (Table 5-2). There are three positive impacts included in these 16

characteristics, thus Block 4 does not have one. However, participants will understand

that the main objective is to make trade-offs, assuming all other factors are at status quo

levels in the previous blocks.

Table 5-2. Attribute blocks for DCE ([*] Positive attribute is shown in bold font)

Block	Attributes[*]
1	Job opportunities Water pollution and shortage Permit approval decision making mechanism Population increase
2	**Income increase** (for all local residents) Air pollution Availability of independent and transparent information on potential impacts of mine Crime increase
3	Increase in housing costs Noise pollution **Infrastructure improvement** (transportation, education, human serves, internet) Mine buffer (Home distance from mine)
4	Labor shortage for other business Land pollution and subsidence Traffic increase Mine life

5.3.3.3. Fractional factorial design. Fractional factorial designs refer to survey designs, which use only a fraction of the total number of treatment combinations. In the case study, each block has four factors with three levels and the full set of combinations is $81=3^4$. This means we will have more than 40 choice sets (each choice set includes two hypothetical alternatives plus the status quo option) for each block. This number of questions (each choice set will be presented as a question) is too high. Thus, fractional factorial design is used in this case study (Hensher et al., 2005; Louviere et al., 2003).

5.3.3.4. With or without interaction. A discrete choice experiment can examine both main factors as well as interactions. In this case study, the design only includes main factors because interactions rarely account for much of the choice. As suggested by Dawes and Corrigan (1974), main effects typically account for 70 to 90 % of the explained variance, two-way interactions typically account for 5 to 15%, and Higher-order interactions account for the remaining explained variance (Louviere et al., 2003).

5.3.4. Generating Experiments. As discussed in the previous section, the discrete choice experiment will be designed as a mix style, blocking scheme, fractional factorial without interaction experiment. This design has five main steps: experimental size determination, candidate design construction, efficient experiment design, duplicate check, and labeling. The functions and respective SAS Macros are shown in Table 5-3.

Table 5-3. Steps in generating experiments

Step	SAS Macro	Function
1. Experimental size determination	%MktRuns (SAS, 2007h)	Suggests sizes for balanced factional factorial experiment designs.
2. Candidate design construction	%MktEx (SAS, 2007f)	Creates efficient factorial designs with the selected size
3. Efficient experiment design	%ChoicEff (SAS, 2007d)	Finds optimal experimental designs for choice experiments and evaluates choice designs
4. Duplicate check	%MktDups (SAS, 2007e)	Detects duplicate choice sets and alternatives within generic choice sets
5. Labeling	%MktLab (SAS, 2007g)	Labels factors and their levels for each block

5.3.4.1. Experimental size determination. The size of the experiment needs to be found to achieve perfect balance and orthogonality or, at least, to minimize violations of orthogonality and balance in the following experimental design. A design is orthogonal when all of the parameter estimates are uncorrelated. A design is balanced when all levels of each factor occurs equally often. As discussed in Section 5.3.1, the relative D-efficiency is an indicator of effectiveness (orthogonality and balance) of discrete choice experiments. The design size is selected to achieve the maximum possible relative D-efficiency.

The SAS %MktRuns macro was used to determine the design size with four factors, each with three levels (SAS, 2007h). The possible sizes and corresponding relative D-efficiency are shown in Table 5-4.

Table 5-4. Experimental size suggestion output

Saturated	= 9	
Full Factorial	= 81	
Observation	Reasonable design size	
1	9 [*S]	
2	18 [*]	
3	12	
4	15	
5	10	
6	11	
7	13	
8	14	
9	16	
10	17	

[*] - 100% efficient design.

[S] - Saturated Design - The smallest design that can be made.

The saturated design (smallest design that can be made) is nine, and the full design size is 81. Designs with 100% relative D-efficiency can be achieved with size 9 and 18. Both satisfy the desire for a reasonable sample size and to maximize the relative

D-efficiency. In this case study, each question (choice set) includes the status quo

alternative and two hypothetical alternatives. Thus, the author chose the design size of 18

since 18 hypothetical alternatives can be divided into nine choice sets of two alternatives

each.

      5.3.4.2. Candidate design construction. In this step, a researcher needs to

construct a candidate design with size 18 and 100% D-efficiency, which can be used to

find the most efficient design quickly, in the next step. The coordinate-exchange

algorithm (CoordX) of Meyer and Nachtsheim (1995) provides a way to search the

candidate designs by initializing the design with an orthogonal and random design

(Kuhfeld, 2010). The CoordX algorithm stops if, at any time it finds a perfect, 100%

efficient, orthogonal, and balanced design. The CoordX algorithm, as implemented in the

SAS %ChoicEx macro, was used to find a candidate design with 100% D-efficiency

(SAS, 2007f). The solution (candidate design) is shown in Table 5-5.

Table 5-5. CoordX algorithm output (100% efficient)

Observation	x1	x2	x3	x4
1	1	1	1	1
2	1	1	2	2
3	1	2	1	3
4	1	2	3	1
5	1	3	2	3
6	1	3	3	2
7	2	1	1	3
8	2	1	3	1
9	2	2	2	2
10	2	2	3	3
11	2	3	1	2
12	2	3	2	1
13	3	1	2	3
14	3	1	3	2
15	3	2	1	2
16	3	2	2	1
17	3	3	1	1
18	3	3	3	3

5.3.4.3. Efficient experiment design. In this step, the 18 hypothetical alternatives (Table 5-5) need to be divided into nine choice sets of two alternatives each. A modified Fedorov (MFed) candidate-set-search algorithm was used in this step (Cook & Nachtsheim, 1980; Fedorov, 1972; Kuhfeld, 2010). A random initial design was constructed from the candidate design. This is then evaluated by exchanging alternatives/sets until increase the D-efficiency stabilizes at a local maximum. This process is repeated with different initial designs to find the best design for all possible initial designs.

MFed algorithm is implemented in the SAS %ChoicEff macro, which was used to find the most efficient random scheme to pair the 18 alternatives into nine choice sets (SAS, 2007d). The candidate-set-search output is shown in Table 5-6, which shows two runs with both converging in four iterations. The first run returns the highest local maximum D-efficiency. The process was repeated with 145 initial designs (two iterations are shown in Table 5-6 as examples). The first run in Table 5-6 returns the maximum D-efficiency and corresponds to the pairing shown in Table 5-7.

Table 5-6. Sample candidate-set-search results

Design	Iteration	D-Efficiency	D-Error	Relative D-Efficiency
1	0	0	.	.
	1	5.970687	0.167485	66%
	2	6.142330	0.162805	68%
	3	6.500560	0.153833	72%
	4	6.500560	0.153833	72%
2	0	1.837117	0.544331	20%
	1	5.759409	0.173629	64%
	2	5.845501	0.171072	65%
	3	6.167149	0.162149	69%
	4	6.167149	0.162149	69%

The relative D-efficiency is 72%. This step changes the relative D-efficiency (from 100%) since it depends on the number of choice sets (Equation 5-3). In this study, the author deemed this adequate. If the relative D-efficiency is too low, the designer can change the number of choice sets or the size of candidate alternatives to improve the design. The number of attributes and number of levels for each attribute can also affect the relative D-efficiency. However, these are difficult or unrealistic to change in real cases.

Table 5-7. Efficient experimental design result

Observation (Table 5-5)	Set
4	1
13	1
8	2
15	2
9	3
18	3
16	4
11	4
12	5
3	5
5	6
14	6
17	7
2	7
1	8
10	8
6	9
7	9

5.3.4.4. Duplicate check and labeling. At this step, the design needs to be checked for duplicate choice sets and alternatives. The experimental design needs to be labeled with the full description of factors and their levels for each block.

The SAS %MktDups macro is used to check for duplicates (SAS, 2007e). In this case, there are no duplicate choice sets and no duplicate alternatives within the choice sets. The SAS %MktLab macro is used to label the experiments with factors and levels

(SAS, 2007g). The status quo alternative was added to each question (choice set) manually. Finally, there are four discrete choice experiments in this work: one for each of the four blocks. Each survey has nine choice sets: one status quo option plus two hypothetical alternatives. This results in 36 choice sets (questions).

5.4. FOCUS GROUP STUDY

An online survey was conducted with a two part questionnaire to validate the design. The first part of the survey contained background questions regarding demographics, the respondent's socioeconomic status, and past experience with mining. The demographic questions included age, gender, income, education, job field, and number of children.

The second part would have contained 36 choice sets from four blocks. This is too much for one respondent. To prevent fatigue, the plan for the final survey was to give each respondent three choice set from each block. This means each respondent will see 12 choice sets, and three respondents are required to answer all 36 choice sets. This approach has been used in several applications where there are too many choice sets (Witt et al., 2009). In this focus group study, the author used only one of the surveys (i.e. only 12 choice sets) to examine the clarity of instructions and difficulty of the survey questions.

In the second part of the survey, participants were asked to select one of three mine options, if a new mine were to be opened in their community (12 such questions were asked, three from each block). Following the questions from each block, participants were asked to rank the level of difficulty and confusion of each block's questions, by selecting a number from 1 to 5 ("not difficult at all" to "very difficult" and "not confusing at all" to "very confusing"). Additional open questions 'what made the

questions difficult/ confusing?' were shown to them if they chose level 4 (somewhat difficult/ confusing) or level 5 (very difficult/ confusing).

This survey was conducted in October, 2014. Twenty-five people participated in this survey, and 22 of them completed it (i.e. answered all question to the end). The participants were recruited from residents in Rolla, MO. Two quality control questions were inserted in the survey. Data was regarded as invalid if the participant completed the survey in less than the minimum expected survey time (150 seconds). Two participants did not 'pass' the quality control questions, and data from these two participants were deleted. Thus, data from 20 participants was used for the data analysis.

Among the 20 respondents, nine were male and eleven were female. Ten people stated that they live near a mine. Of these, 6 reported living within 10 miles of a mine and the remaining four live more than 30 miles of a mine. The Missouri University of Science and Technology Experimental Mine and Capital Quarries' limestone quarry are located in Rolla, MO. Seven of the 20 participants self-declared to have experience with mining (e.g. working for a mine, familiarity with mining activities, studying about mining etc.). Sample answers to this Question include:

"I visited the experimental mine at MST", "Mining Engineer, work with Mining Companies and for Mining Companies" and "intern at underground coal mine and a surface diamond mine".

5.5. DATA ANALYSIS

  5.5.1. Analysis of Difficulty and Confusion Ratings. In the focus group study, data on the level of difficulty and confusion was acquired for questions and instructions from all four blocks. The main objective of the data analysis is evaluate whether or not there is a significant difference between the level of difficulty and confusion rating for these four blocks. Thus, the data on difficulty and confusion ratings from the different blocks were treated as independent groups.

  There are three data analysis methods for more than two independent groups: ANOVA, Welch ANOVA, and Kruskal-Wallis (Schlotzhauer, 2009). Figure 4-2 shows the logic for selecting the proper test. ANOVA and Welch ANOVA tests can be used when the variables are interval variables. In this case study, the data on the level of difficulty and confusion are treated as ordinal. Thus, the Kruskal-Wallis test is the most appropriate test.

  The null hypothesis is that 'there is no significant difference between the difficulty/confusion level distributions of these four blocks.' The p-value (p) is the probability that the null hypothesis is not true. If $p < \alpha$ (the level of significance), then we have enough evidence to reject the null hypothesis, and conclude there is at least one significant difference among the difficulty/confusion level ratings of these four blocks at significant level α (typically 0.05). On the other hand, if $p > \alpha$, then the test fails to reject the null hypothesis at significant level α.

  Kruskal-Wallis test was done using the SAS PROC NPAR1WAY WILCOXON PROCEDURE (SAS, 2007k). The p values were estimated as 0.9536 and 0.8469 for difficulty and confusion level comparisons, respectively. The tests fail to reject the null hypothesis, and we conclude that there is not enough evidence to prove there is at least one significant difference among the difficulty and confusion level ratings of these four blocks at significant level 0.05.

To examine the hypothesis test more closely, the author estimated 95% confidence bounds of the medians using the SAS UNIVARIATE procedure for both data sets (SAS, 2007l, 2007m). Table 5-8 shows the results, rounded to the nearest integer, computed without assuming any specific distribution. It shows the median level of difficulty and confusion is similar for all four blocks, although, Block 1 has slightly higher medians.

Table 5-8. Level of difficult/confusing scale for each block

	Difficulty level	Confusing level
Block 1	(2,4)	(2,4)
Block 2	(2,3)	(2,3)
Block 3	(2,3)	(2,3)
Block 4	(2,3)	(2,3)

1 Not difficult/confusing at all
2 Not very difficult/confusing
3 Acceptable
4 Somewhat difficult/confusing
5 Very difficult/confusing

From Figure 5-1 and Figure 5-2, most of the participants selected level 3 (acceptable) for both difficulty and confusion ratings of all four blocks. All the distributions are relatively symmetric, although a few more selected levels 1 and 2 than levels 4 and 5, in most cases. For Block 1, 30% of participants selected levels 4 (somewhat difficult/confusing) and 5 (very difficult/confusing) for both difficulty and confusion rating. However, significantly fewer respondents ($\leq$20%) selected levels 4 or 5 in the other three blocks. It is the main reason why the median level of both difficulty and confusion rating for Block 1 are higher than the other three blocks.
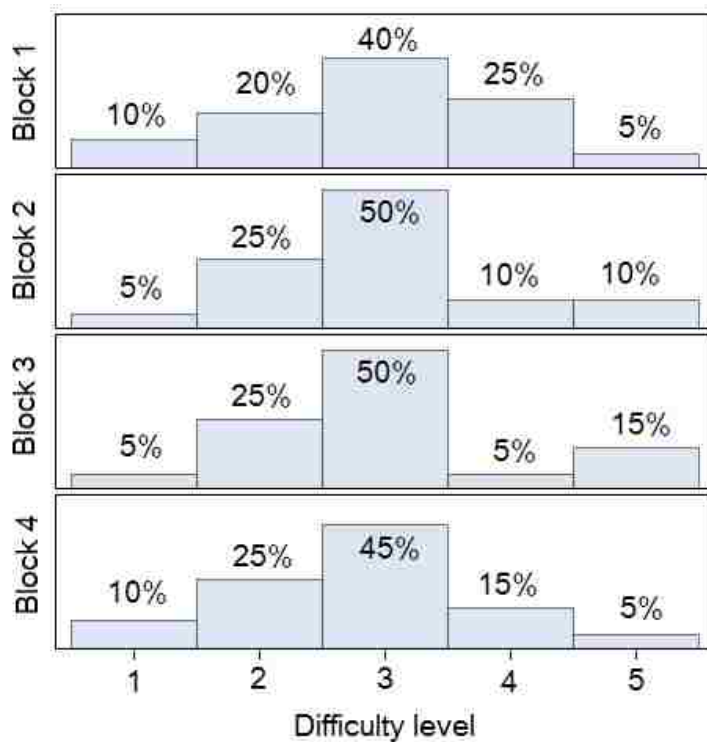
Figure 5-1. Difficulty levels
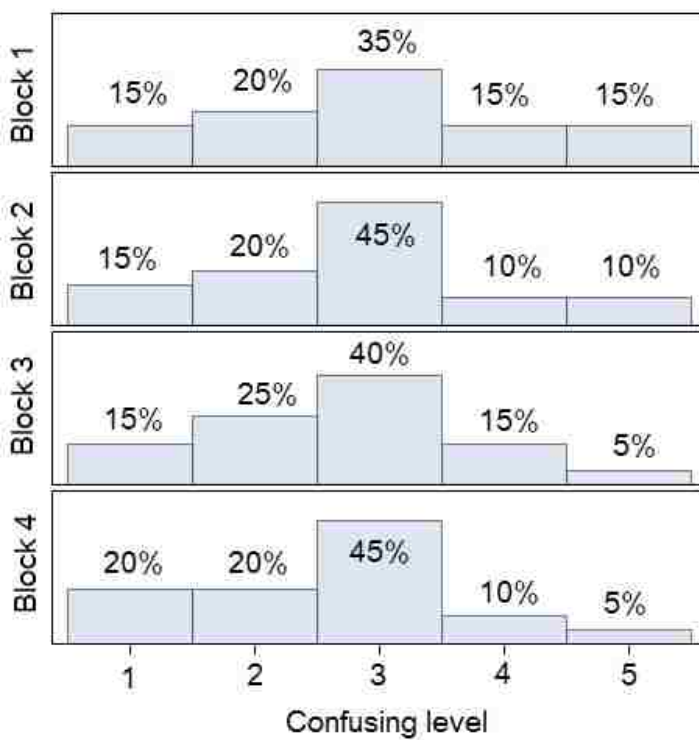
Figure 5-2. Confusing levels

5.5.2. Analysis Based on Duration of Survey. The time duration it took for each participant to complete each block was tracked, and the comparative box plot is shown in Figure 5-3.

The same Kruskal-Wallis test was used to test the null hypothesis that there is no significant difference between the duration of these four blocks (SAS, 2007k). The p value was estimated as 0.6373, and fail to reject the null hypothesis. This result means there is not enough evidence to prove that there is at least one significant difference among the duration of the survey for respondents on these four blocks, at a significant level of 0.05.
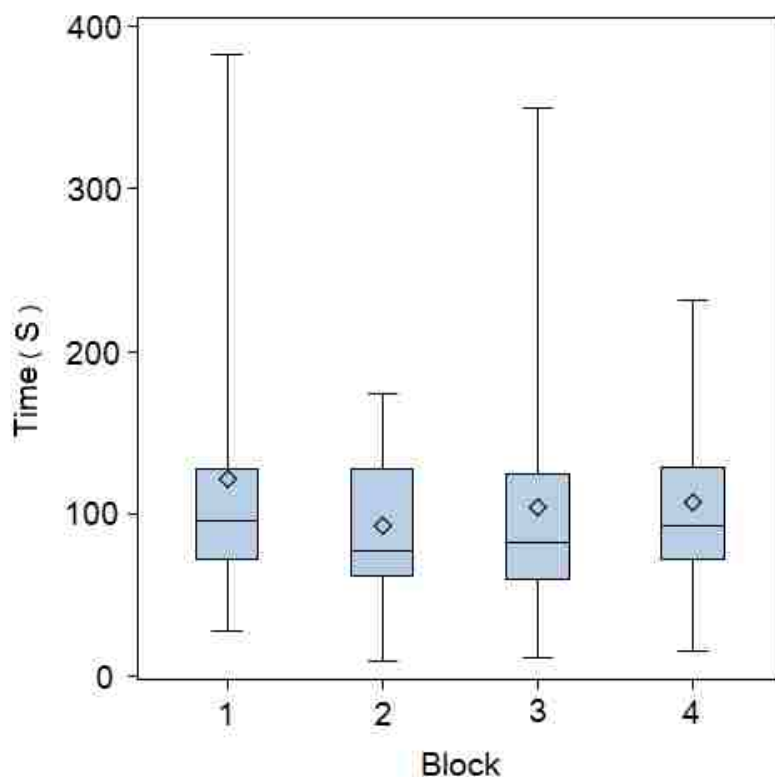
Figure 5-3. Time for completing each block

Figure 5-3 shows that the mean of duration is higher than the median in all blocks and the variation for all blocks is similar. The mean and median of time duration is

highest for Block 1, which is 121 and 96 seconds, respectively. For Block 2, both mean and median drop sharply to the lowest at 93 and 78 seconds, respectively. After that, the mean and median stays steady for Blocks 3 and 4 at 104 and 83 seconds, and 107 and 93 seconds, respectively.

5.5.3. Open Question Comments. The respondents provided 32 comments in response to the open questions "what made the questions difficult/confusing?"

There were five comments for each question (difficulty/clarity) from Block 1. The other three blocks have fewer comments: three to four comments for each question. The answers are quite similar for these two questions, thus, author did not separate them. These comments can be generally classified into three types: clarity, information, and levels and factors.

Some participants did not understand the main purpose of this survey (i.e. the instructions on the purpose was not clear to them). This is indicated by comments such as "Do I choose these according to what I want this mine to be?" In addition, the block introduction, which introduces the factors and their levels before each block seems to have confused some respondents. This is shown by comments like "Mostly just the way it started -- I was trying to select options on first page".

It appears some participants wanted to more details and background information about the survey. Respondents provided comments like "Not enough information or detail" and "No explanation of how factors relate to question". What is more, this survey is designed with a block scheme to ensure reasonable cognitive burden. However, it appears this unsettled some participants who wanted information on the hidden/unknown variables. They seem to have missed the instruction that "all other factors are the same for all alternatives."

It appears most of the levels are not clear enough. First of all, labor shortage for other businesses, infrastructure improvement, and all environmental factors have similar

levels "negligible", "no increase", "slight", "moderate", which appears to confuse some respondents. Respondents did not understand the relative degrees, and provided comments like "Is slight>moderate?" and "I don't understand if slight<negligible<moderate?" Secondly, the levels of factors population increase, traffic increase, increase in housing costs and crime increase seem to be confusing (e.g. "a reduced rate of increase" and "2% reduction in increase …"). There were comments such as "What is reduction in increase? Is reduction in increase= decrease?" Thirdly, the percent levels were also difficult to understand since participants did seem to have a concept of the current rate.

There were also comments like "I can't say I really liked any of the options so I had to weigh in what I thought best" and "no obvious option". These comments are acceptable since the objective of choice experiments is evaluate how respondents make trade-offs.

5.6. RESULTS AND DISCUSSION

5.6.1. Design Evaluation. As discussed in the design section, the discrete choice experiment is a mix style, blocking scheme, fractional factorial without interaction design. The 16 mining project characteristics are divided into four blocks as shown in Table 5-2.

Every block has nine questions with three alternatives each. The first alternative is always the status quo option and shows the average impact/value of each attribute in the survey area. The other two alternatives use the same attributes but different levels to generate hypothetical alternatives. The relative D-efficiency of the discrete choice experiment is 72%. Each survey (for each respondent) contained 12 choice sets, three from each of the blocks, in order to prevent fatigue and to ensure each participant provides input on each block's attributes. Hence, there will be three surveys in the real discrete choice experiment for collecting data for discrete choice modeling. In this focus

group study, the author used one such survey to examine the clarity of instructions and difficulty of the survey questions.

As shown by the results of the Kruskal-Wallis test, there is no significant difference between the participants' difficulty and confusion level ratings for these four blocks, at significant level 0.05. In addition, the median level of difficulty and confusion rating for all blocks are (2, 3) — above "not very difficult/confusing" but below "acceptable" -- and (2, 4), — above "not very confusing" but below "somewhat confusing" -- respectively. This means the discrete choice experiment design achieves one of the main goals for questions in all blocks, which is a survey that is not unduly difficult (i.e. reasonable cognitive burden).

While there is no significant difference among the difficulty and confusion ratings of the four blocks, it appears Block 1 has a slightly higher rating than the other three blocks. This finding is confirmed by the result of time duration analysis: both the median and mean peaked at Block 1, although there is no significant difference in statistics. This slight difference may be partly because Block 1 is slightly more difficult and confusing than the other blocks. However, it is also possible that respondents are learning as they proceed through the survey. Caussade et al. (2005) observed that there is a learning effect in choice situations. Though the factors in each block are different, the question, framework, and style of choice situations are the same. Hence, these four blocks should be random ordered in the actual discrete choice experiment. This way, one can prevent unnecessary variance in the discrete choice model.

5.6.2. Survey Revision. Although, the respondents rated the difficulty and confusion levels as acceptable, some revision is necessary based on feedback from the open-ended question. The main purpose of this survey needs to be stated more clearly to help participants understand the survey and provide effective data for the discrete choice model.

With respect to clarity of the survey and instructions, an introduction will be added to provide background information to respondents so they can appreciate why and how a mining project can affect their lives. A sample problem will be given to help participants understand the survey questions. Also, the instructions will highlight that the each choice set includes only four attributes and assumes the remaining attributes are all at the status quo level. The revised survey introduction is shown in Appendix E. A video will be inserted in the induction to help participants understand the instructions better. In addition, the block introduction, which introduces the attributes and their levels before each block, was revised to prevent participants from trying to select answers at this stage. The whole survey including the video introduction, and discrete choice questions is available in Appendix F.

The author has revised the levels and/or provided more explanation to ensure respondents understand each level and the relative scale. First, the confusing levels "a reduced rate of increase" and "2% reduction in increase …" have instead been revised to percent increase/decrease or by comparing the level with the current rate. Take the levels of traffic increase as an example, the new levels are compared to the current traffic rate to yield "Lower than current rate", "Same as current rate" and "Higher than current rate". All revised levels are shown in Appendix D in italic. Second，to prevent participants from confusing the relative degrees of the levels, the author highlighted the middle level as yellow, 'worse' level as red and 'better' level as green. The text highlighting provides a quick reference for respondents, who can now easily appreciate the relative scale of the

levels. Third, the factor "land pollution and subsidence" is revised as "land pollution" since the proposed study site is Salt Lake City, Utah, USA, which has mainly surface mining activities. Land subsidence is not a significant problem. Also, it will be difficult to define clear levels that will be useful for "pollution" and "subsidence" at the same time. In addition, the factor "water pollution and shortage" is revised to "water pollution" for the same reason. In another context, subsidence or shortage may be more important than pollution or both will be equally important. The choice of whether to include one, the other, or both should be made on the basis of relevance to the community consultation process.

5.7. SUMMARY OF SECTION FIVE

This Section is an attempt to design and validate a discrete choice experiment for mining community engagement. The work attempts to (1) show the general approach to discrete choice experimental design for mining community engagement; and (2) provide a research note on validation data analysis with a case study.

The discrete choice experiment is designed as mixed style, blocking scheme, factional factorial without interaction experiment. The relative D-efficiency of the discrete choice experiment was 72%. Based on the focus group results, the discrete choice experiment design achieved acceptable difficulty and clarity for questions in all blocks.

In addition, on the basis of the validation result, the four blocks should be random ordered in the discrete choice experiment to avoid the learning effect affecting only Block 1. Revision is also done to address some of the concerns raised with respect to clarity, lack of information, and definition of attribute levels.

6   DICRETE CHOICE EXPERIMENT—A CASE STUDY

6.1. INTRODUCTION

The final challenge of this dissertation work is to illustrate the usefulness of the discrete choice experimental design suggested in Section 5 for stakeholder analysis in mining. This is done by conducting the discrete choice experiment in Salt Lake City, UT and analyzing the results to make useful inferences. A major technical challenge in this step is how to select the most appropriate discrete choice model to describe the local community's acceptance of mining projects. This task involves: (1) conducting a comprehensive literature review of discrete choice models to identify candidate discrete choice models that are most appropriate for modeling mining community acceptance (Section 2.4); and (2) evaluating the candidate discrete choice models to select the most suitable discrete choice model for mining community acceptance. The comprehensive literature review of discrete choice models was done in Section 2, in which candidate models were identified. Those candidate models are the conditional logit (CL), conditional logit stratified by questions (CLQ), and mixed logit (ML) models. The second task is addressed in this Section, which examines various discrete choice models to select the most appropriate model for mining community consultation based on the case study presented in section five.

When produced from a properly selected DCM method, the discrete choice modeling results can provide valuable information for mining companies during design, planning and management of mining projects. The results can support the whole consultation process by answering three important questions: (1) what are the factors that affect stakeholders' decision and how do these affect the decision-making? (2) what is the effect of demographics on individual preferences? (3) what is the value of environmental and social impacts to individuals in the community?

6.2. DISCRETE CHOICE EXPERIMENT (SURVEY)

6.2.1 Sample Size Determination. As discussed in Section 5, this survey targeted 600 respondents (three blocks with 200 respondents per block). The sample size (200 per block) was based on the largest proportion of respondents that were likely to choose a particular option, which was estimated to be 0.5 (based on the largest proportion in the preliminary survey, where 50% of respondents chose water pollution as very important (Que et al., 2015). This analysis indicated that a sample size of 183 would be adequate. The sample size was rounded up to a more conservative 200 respondents per block for convenience in budgeting.

6.2.2. Sampling Respondents from Mining Communities. The survey (designed and validated in Section 5) was conducted in Salt Lake City for three reasons: (1) it has a population of 186,440, which makes it relatively easy to find enough participants to complete the discrete choice experiment via online interviews; (2) Rio Tinto Kennecott's mine is very visible in the community and has an effect on locals; and (3) Rio Tinto has a comprehensive sustainable development report available online, which provides relevant data for designing discrete choice experiments. . In a real stakeholder analysis situation, a mine seeking to engage in consultation should dedicate enough resources to design the DCE and acquire the data, since these particular characteristics (which made it easy to acquire responses via online surveys) of this case study will not be relevant for determining ease of application.

The survey was conducted with a two part questionnaire. The first part contained demographic questions, including respondent's age, gender, income, and education. This means that three separate respondents were required in order to receive a complete set of answers to all 36 choice sets. This approach has been used in several applications where there are too many choice sets (Witt et al., 2009). For each choice set, participants were asked to select one of three mine options, if a new mine were to be opened in Salt Lake City. The full survey can be found in Appendix F.

The survey was estimated to take 15 minutes to complete. Two quality control questions were inserted in the survey (Appendix F). Data from a participant was regarded as invalid if the participant completed the survey in less than seven minutes, which was estimated to be the minimum expected survey time. Data from participants who did not 'pass' the quality control questions or survey duration control, was deleted from the data set. The survey was computer assisted personal interviews, administered by Qualtrics, a well-known market research firm. The survey was conducted from January to March 2015, and respondents were tracked by their zip code using the IP address they used to access the survey.

6.2.3. Demographic Distribution of Respondents. A total of more than 1,810 people were invited to participate in this survey, as shown in Table 6-1. Of these, 1,062 responded, and 882 of them completed the survey (i.e. answered all questions). Forty-four participants were excluded for failing to answer the quality control question correctly or completing the survey in less than seven minutes.

Table 6-1. Survey participant statistics

No. of participants	Group 1	Group 2	Group 3	Total
Invited	755	669	>386	>1,810
Started	485	316	386	1,062
Completed (i.e. answered all question to the end)	300	261	261	882
Terminated by quality control question or survey duration	12	10	22	44
Excluded due to demographic factors	74	40	36	150
Final qualified	214	211	203	628

Based on the results of correlation analysis discussed in Section 3, the goal was to match the four important demographic factors (age, gender, income, and education) of the intended participants to those of the Salt Lake City (SLC). In all, 150 people were

terminated due demographic factors (i.e. accepting them would have unduly biased the pool). While the gender and age of the respondents matched those of SLC, the average education and annual income did not achieve this goal. This was partly due to the difficulty of recruiting a representative sample using online surveys. The limitations of the survey method are discussed in Section 6.5.4. A total of 628 qualified participants, recruited from Salt Lake City, were included in the final pool of participants. The statistics of the four important demographic variables of respondents are summarized and compared in Table 6-2.

Table 6-2 Demographic distribution of participants

Demographics	Group 1	Group 2	Group 3	Sum	SLC*
Gender					
Male	50%	49%	44%	47%	50%
Female	50%	51%	56%	53%	50%
Age					
18—25	4%	11%	12%	9%	18%
26—34	26%	24%	34%	28%	26%
35—54	34%	38%	30%	34%	31%
55—64	19%	16%	12%	16%	12%
>65	17%	10%	12%	13%	13%
Highest education					
<high school	0%	0%	0%	0%	14%
High school/GED	7%	9%	13%	10%	18%
Some college, vocational, or 2 year college degree	33%	36%	31%	34%	27%
Bachelor's degree and higher	59%	55%	55%	56%	41%
Annual income					
< $20,000	4%	6%	10%	7%	22%
$20,000—$39,999	22%	21%	19%	21%	23%
$40,000—$59,999	22%	24%	21%	22%	18%
=>$60,000	52%	49%	51%	51%	37%

*United States Census Bureau, 2010 Census

6.3. DATA PROCESSING

The survey data was collected by groups. Each group contained 12 choice sets, three from each of the four blocks. Raw data for the first 10 observations of group one is shown in Figure 6-1 as an illustration. The data consist of a subject (participant) number followed by 16 integers, which represent the participant's responses. The first four integers represent the demographic factors: gender, age, education, and income.


Figure 6-1. Raw data for the first 10 observations of Group 1

These demographic factors were coded as integers as shown in Table 6-3, in which the "prefer not to answer", option was coded as a blank entry. The next 12 integers are in the range 1 to 3 and represent the answers for each choice set. For example, the first subject chose alternative 3 in choice set 1 of block 1, alternative 3 in choice set 2 of block 1, and so on.

Table 6-3. Demographic factor codes

Demographic factors and levels		Level code
Gender	Male	1
	Female	2
	Prefer not to answer	.
Age	18-24	1
	25-34	2
	35-54	3
	55-64	4
	65 or over	5
	Prefer not to answer	.
Highest level education	Less than high school	1
	High school/GED	2
	Some college, vocational, or 2 year college degree	3
	Bachelor's degree and higher	4
	Prefer not to answer	.
Combined annual household income	below $20,000	1
	$20,000-$39,000	2
	$40,000-$59,000	3
	$60,000 or more	4
	Prefer not to answer	.

For modeling purposes, the raw data needed to be converted to binary coded (1=chose, 0=did not choose) data. A sample binary coded data sample for answers of the first three respondents in Figure 6-1 are shown in Figure 6-2.

The data starts with subject number and the four demographic factors. Next are the attributes X1, X2, X3, and X4, which are the factors used in the choice sets in Block 1 (job opportunities, water pollution, permit approval decision making mechanism, and population increase). The relative integers, in the range 1 to 3, show the level of each factor in the choice set. The next number describes the decision (i.e. whether the option was chosen or not).The final number in the sequence, describes the choice set.

After coding the data as shown in Figure 6-2, the data sets from the four blocks and three groups are combined to form over 21,600 observations: 4 blocks $\times 3$ groups $\times$ 200 subjects (minimum) $\times 3$ choice sets $\times 3$ alternatives. This data was then used to fit the three candidate discrete choice models to the respondents' preferences.

```
title group1block1;

data block1;
input subject gender age education income X1 X2 X3 X4 decision set;

datalines;
1  1  3  3  4  2  2  2  2  0  1
1  1  3  3  4  1  2  3  1  0  1
1  1  3  3  4  3  1  2  3  1  1
1  1  3  3  4  2  2  2  2  0  2
1  1  3  3  4  2  1  3  1  0  2
1  1  3  3  4  3  2  1  2  1  2
1  1  3  3  4  2  2  2  2  1  3
1  1  3  3  4  2  2  2  2  0  3
1  1  3  3  4  3  3  3  3  0  3
2  2  3  3  4  2  2  2  2  0  1
2  2  3  3  4  1  2  3  1  0  1
2  2  3  3  4  3  1  2  3  1  1
2  2  3  3  4  2  2  2  2  0  2
2  2  3  3  4  2  1  3  1  0  2
2  2  3  3  4  3  2  1  2  1  2
2  2  3  3  4  2  2  2  2  1  3
2  2  3  3  4  2  2  2  2  0  3
2  2  3  3  4  3  3  3  3  0  3
3  1  3  3  4  2  2  2  2  1  1
3  1  3  3  4  1  2  3  1  0  1
3  1  3  3  4  3  1  2  3  0  1
3  1  3  3  4  2  2  2  2  1  2
3  1  3  3  4  2  1  3  1  0  2
3  1  3  3  4  3  2  1  2  0  2
3  1  3  3  4  2  2  2  2  1  3
3  1  3  3  4  2  2  2  2  0  3
3  1  3  3  4  3  3  3  3  0  3
```

Figure 6-2. Sample coded data for the first three subjects of Group 1 Block 1

6.4. LOG-LIKLIHOOD RATIO INDEX AND ALGORITHMS

6.4.1. The Log-likelihood Ratio Index. The likelihood ratio index (LRI, also called pseudo- R^2) is frequently used in discrete choice modeling to measure how well the models fit the data. Specifically, the statistic measures how well the model, with its estimated parameters, performs compared with a model in which all the parameters are zero (which is usually equivalent to having no model at all).

Breffle and Rowe (2002) report that an LRI of 0.12 is typical for cross sectional data[8]. As a rule of thumb, well fitted models have an LRI greater than 0.2, and it is rare to find cases with LRI greater than 0.4 (Hoyos, 2010). LRI comparison is conducted on the basis of the log likelihood function, evaluated at both the estimated parameters and at zero for all parameters. The likelihood ratio index is defined by Equation (6-1).

$$\rho = 1 - \frac{LL(\boldsymbol{\beta})}{LL(\mathbf{0})} \tag{6-1}$$

where $LL(\boldsymbol{\beta})$ is the value of the log likelihood function at the estimated parameters while $LL(\mathbf{0})$ is its value when all parameters are set equal to zero. If the estimated parameters are not better than the zero parameters in terms of the likelihood function, then $LL(\boldsymbol{\beta}) = LL(\mathbf{0})$ and $\rho = 0$. This is the lowest value of ρ, since $LL(\boldsymbol{\beta}) \geq LL(\mathbf{0})$ if β would be the maximum likelihood estimate. At the other extreme, suppose the estimated model was so good that each sampled decision-maker's choice can be predicted perfectly, then the likelihood function at the estimated parameters would be one. This is because the probability of observing the choices that were actually made is one. And, since the log of one is zero, the log likelihood function would be zero at the estimated parameters, making $LL(\boldsymbol{\beta}) = 0$ and $\rho = 1$ in this scenario. This is the highest value that ρ can take. In summary, the likelihood ratio index ranges from zero, when the estimated parameters are no better than zero parameters, to one, when the estimated parameters perfectly predict the choices of the sampled decision-makers. The log-likelihood function has the form shown in Equation (6-2).

$$LL(\boldsymbol{\beta}) = \sum_{n=1}^{N} \ln P_n(\boldsymbol{\beta}) / N \tag{6-2}$$

---

[8] Cross-sectional data, in statistics and econometrics, is a type of data collected by observing many subjects (such as individuals, firms, countries, or regions) at the same point in time, or without regard to differences in time.

where $P_n(\boldsymbol{\beta})$ is the probability of the observed outcome for decision maker n, N is the

sample size, and $\boldsymbol{\beta}$ is a $K \times 1$ vector of parameters. If the log-likelihood function was

divided by N, LL would be the average log-likelihood in the sample. Doing so does not

change the location of the maximum, since N is fixed for a given sample and yet

facilitates the interpretation of some of the procedures. All the procedures operate the

same way regardless of whether the log-likelihood is or is not divided by N.

6.4.2. The Newton-Raphson Algorithm. This section describes the Newton-

Raphson algorithm that is used to maximize a likelihood function.

This algorithm is used to find the optimum value of $\boldsymbol{\beta}$ that maximizes $LL(\boldsymbol{\beta})$

Referring to Figure 6-3, the goal is to locate $\boldsymbol{\beta}$ which occurs at a maximum in $LL(\boldsymbol{\beta})$.

Note in the figure that $LL(\boldsymbol{\beta})$ is always negative, since the likelihood is a probability

between 0 and 1 and the log of any number between 0 and 1 is negative. The maximum

of $LL(\boldsymbol{\beta})$ can be found by "walking up" the likelihood function until no further

increase can be found (within numerical tolerance). A starting value β_0 can be specified

as 0 at the beginning, and the subscribe "t" of $\boldsymbol{\beta}_t$ refers to the number of steps that $\boldsymbol{\beta}_t$ has

moved from $\boldsymbol{\beta}_0$. Each iteration moves, to a new value of the parameters at which $LL(\boldsymbol{\beta})$

is higher than at the previous value, if the algorithm is convergent. The question is how to

find the best value for $\boldsymbol{\beta}_{t+1}$. There are two related questions: (1) in what direction, within

the search space, should the algorithm proceed to search for the optimum? (2) once the

search direction is determined, what step size should the algorithm use to find the next

candidate solution?

The gradient $\mathbf{g}_t$ at $\boldsymbol{\beta}_t$ is the vector of first derivatives of the function $LL(\boldsymbol{\beta})$

evaluated at $\boldsymbol{\beta}_t$ (Equation 6-3). The gradient has dimension K $\times$ 1, where K is the number

of parameters to be estimated for the model. As shown in Figure 6-4 for the 1-

dimensional case, this vector indicates the direction of increase in the likelihood function

(i.e. $\boldsymbol{\beta}_t$ moves LL towards maximum in the positive direction of $\mathbf{g}_t$ and $\boldsymbol{\beta}_t$ moves LL

away from maximum in the negative direction of $\mathbf{g}_t$).
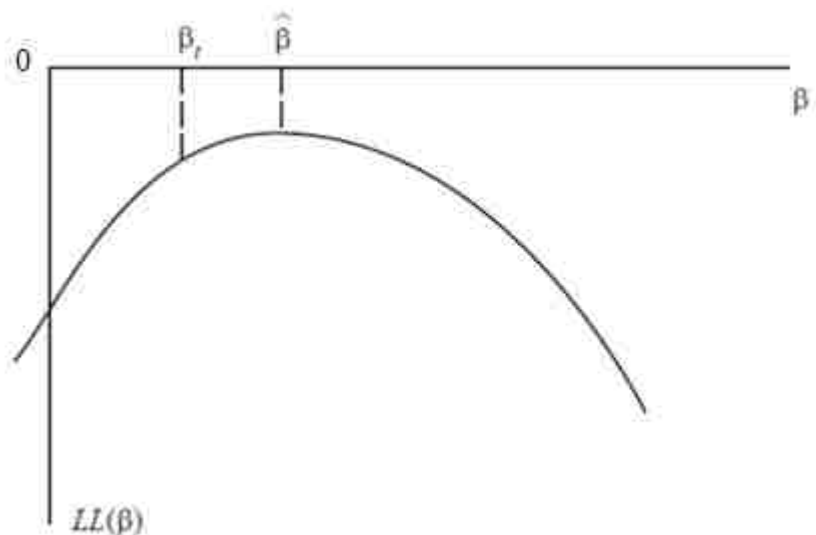
Figure 6-3. Maximum likelihood estimate (Train, 2002)

The Hessian, H_t, is the K $\times$ K matrix of second derivatives of $LL(\beta)$, (Equation

6-4). The Hessian can be used to determine how large a step should be made based on the

direction suggested by the gradient. If the determinant of the Hessian is high, it implies

that the slope changes quickly, as in Figure 6-5(a), and the maximum is likely to be close.

Hence, the algorithm takes small step sizes. Conversely, if the determinant of the

Hessian is small, then that the slope is not changing much and the maximum is likely to

be further away. Hence, the algorithm takes larger step sizes. After calculating g_t and H_t,

an iterative algorithm to find the optimum can be developed using Equation (6-5), where the inverse of the Hessian matrix is used as the step size. The algorithm terminates when the maximum $LL(\boldsymbol{\beta})$ has been found.

$$g_t = \left(\frac{\partial LL(\beta)}{\partial \beta} \right)_{\beta_t} \tag{6-3}$$

$$H_t = \left(\frac{\partial g_t}{\partial \beta'} \right)_{\beta_t} = \left(\frac{\partial^2 LL(\beta)}{\partial \beta \partial \beta'} \right)_{\beta_t} \tag{6-4}$$

$$\beta_{t+1} = \beta_t + \left(-H_t^{-1} \right) g_t \tag{6-5}$$

Positive slope $\Rightarrow$ move forward   Negative slope $\Rightarrow$ move backward

$LL(\beta)$ (a) $LL(\beta)$ (b)
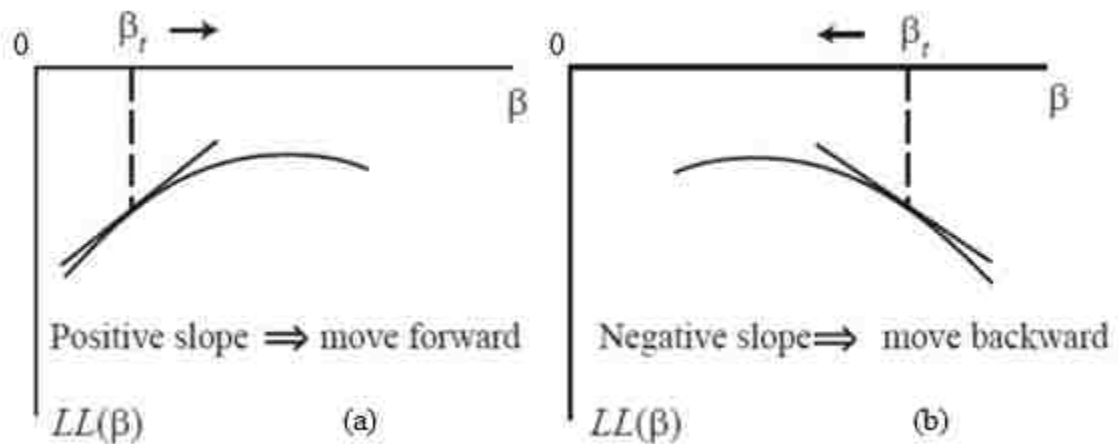
Figure 6-4. Direction of step follows the slope (Train, 2002)
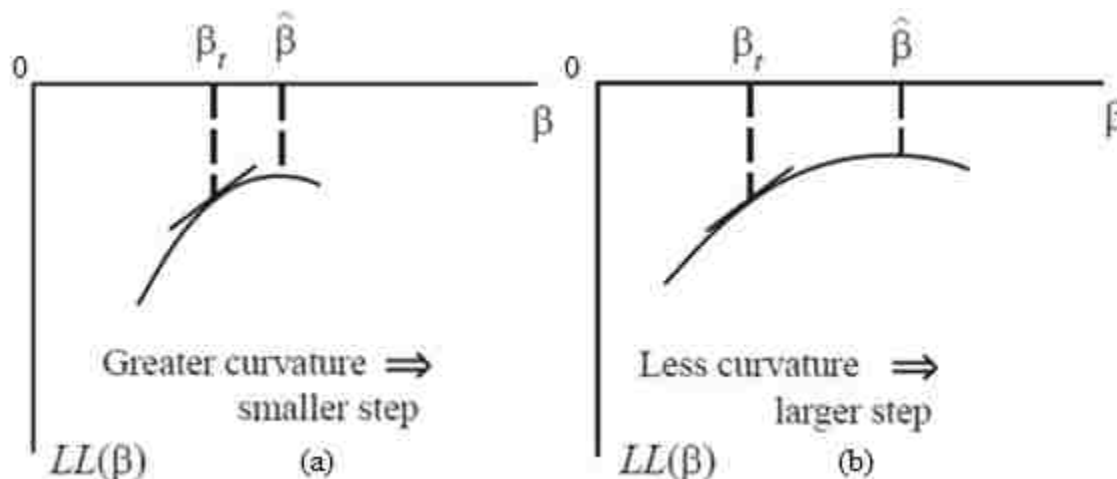
Figure 6-5. Step size is inversely related to curvature (Train, 2002)

6.5. DISCRETE CHOICE MODELING RESULTS

In this study, the author selected the most appropriate model by fitting the discrete choice experiment data to the three candidate discrete choice models: conditional logit model (CL), conditional logit model stratified by question (CLQ), and mixed multinomial logit model (ML).

6.5.1. Conditional Logit Model. Modeling with the conditional logit model (CL) was done using the SAS LOGISTIC procedure. Based on the CL model result (Table 6-4), it can be concluded that the estimated values agreed with the discrete choice experimental data reasonably well.

The goodness-of-fit of the CL model is quite good, with an LRI (pseudo-R^2) of 0.2687. As mentioned above, well-fitting models have a LRI greater than 0.2 (Hoyos, 2010). The percent concordant[9] of the CL model is 73.3%, the percent discordant[10]

---

[9] Percent Concordant: Percentage of pairs where the observation with the desired outcome (event) has a higher predicted probability than the observation without the outcome (non-event).

equals 23.5%, and the percent tied[11] is 3.2%. In general, higher percentages of concordant

pairs and lower percentages of discordant and tied pairs indicate a more desirable model.

Table 6-4. Conditional logit model result

Parameter	Coefficient	Standard Error	Odds Ratio	WTP $/month (Standard error)
Intercept	0.8931**	0.3632		
Economic				
Job opportunities	1.1259***	0.0481	3.083	341(15)
Income increase	0.6600***	0.0518	1.935	200(16)
Increase in housing costs	-1.0416***	0.0506	0.353	-316(15)
Labor shortage for other business	-0.0924**	0.0433	0.912	-28(13)
Environmental				
Noise pollution	-0.9580***	0.0507	0.384	-290(15)
Water pollution	-0.1956***	0.0479	0.822	-59(15)
Air pollution	-1.0952***	0.0552	0.334	-332(17)
Land pollution	-0.2485***	0.0451	0.780	-75(14)
Social				
Population increase	-0.0709	0.0465	0.932	N/A
Infrastructure improvement	0.6527***	0.0475	1.921	198(14)
Crime increase	-1.1753***	0.0548	0.309	-356(17)
Traffic increase	-0.1938***	0.0431	0.824	-59(13)
Governance and others				
Decision making mechanism	0.1634***	0.0464	1.178	50(14)
Information available	0.8460***	0.0532	2.330	256(16)
Mine buffer	0.6684***	0.0479	1.951	203(15)
Mine life	0.1181***	0.0431	1.125	36(13)
Demographic factor				
Age	0.0100**	0.0074		
Gender	-0.0200*	0.0118		
Household income	0.0043*	0.0017		
Education	0.0013*	0.0008		

***1% significance level, **5% significance level, *10% significance level.

[10] Percent Discordant: Percentage of pairs where the observation with the desired outcome (event) has a lower predicted probability than the observation without the outcome (non-event).

[11] Percent Tied: Percentage of pairs where the observation with the desired outcome (event) has same predicted probability than the observation without the outcome (non-event).

Of the 16 mining characteristics and four demographic factors included in the model (Table 6-4), the Wald χ^2 test results indicate that 15 of the 16 mining characteristics, (population increase being the exception), and all four demographic factors that have statistically significant influence on the participants' choices. The null hypothesis of the Wald χ^2 test is that the coefficient of a factor is equal to zero. If the p-value is less than the significance level α, the null hypothesis is rejected, implying that the factor has significant influence on the choice to accept or reject a particular mining project by the mining community. In this model, the author used three significance levels (i.e. 0.01, 0.05, and 0.1).

6.5.1.1. Positive mining characteristics. Since the levels of each factor are coded as 1, 2 and 3, the coefficients f the factors represent the relative degrees of influence of the factors (i.e. bigger coefficient means larger influence). As illustrated by the results in Table 6-4, job opportunities, income increase, infrastructure improvement, decision making mechanism, information available, mine buffer, and mine life have positive impacts on the preference for particular mining projects at the 1% significance level. Increasing their levels will increase the probability of acceptance of a mining project in the communities.

Contrary to the author's expectation, mine life is estimated as a positive factor at the 1% significance level. In the discrete choice experimental design (Section 5), the author colored the first level (20 years) of mine life as green representing the "best" level, the second level (30 years) as yellow indicating the intermediate level, and the third level (40 years) as red representing the "worse" level. However, participants appear not to have been biased by the experimental design. This finding is important for mining project design and planning since it will suggests—if true for other communities—that mining communities prefer longer mine lives (at least in the range of 20 to 40 years). This result may be because individuals in mining communities know that most of the mining's

positive impacts only last as long as the mining project, but the negative impacts may last much longer than the mine life. Hence, such individuals are likely to opt for longer mine lives. The reasons behind this result and whether it can be generalized should be explored in future studies.

The coefficient results from Table 6-4 show that the degree of influence of the seven positive mining project characteristics are job opportunities (6, 7) > information available (5, 6) > mine buffer (6) > income increase (6) > infrastructure improvement (5, 6) > decision making mechanism (5, 6) > mine life (6). The numbers in the parentheses represent the mining group result from Table 3-6 (level of importance of mining project characteristics). Comparing these two results, the order of the factors based on the discrete choice model coefficients are different from the order that would be expected based on the results in Section 3. For example, "availability of independent and transparent information on potential impacts of mine" has a level of importance (5, 6), above somewhat important and less than very important. However, its degree of influence in participants choices is larger than mine buffer and mine life, both of which have a level of importance (6) -- very important. Also, while the mining project characteristic "mine life" has a level of importance (6) -- very important, its degree of influence is the lowest among these seven positive mining characteristics and is lower than "availability of independent and transparent information on potential impacts of mine", "infrastructure improvements" and "permit approval decision making mechanism", all of which have level of importance (5, 6) – above somewhat important and less than very important.

The differences show that just ranking the factors independently cannot give us the full picture. Section 3 measures each factor independently by asking respondents to rank the level of importance of each factor. In discrete choice experiments, respondents are presented with real and hypothetical options and asked to choose one option. In doing so, respondents are forced to make trade-offs to choose one option over the other. The

discrete choice model estimates the relative importance of these 16 mining characteristics to respondents when making choices about what they prefer. This is why the discrete choice model result can give us a better sense for what people really value.

Factors similar to the mining project characteristics "mine buffer" and "job opportunities" were studied by Ivanova and Rolfe (2011) and Ivanova et al. (2007) as "buffer for mine impacts" and "jobs for partners/children". In their conditional logit (CL) model, the coefficients were estimated as 0.248 at significance level 5%, and 0.278 at significance level 1%. In this CL model, the coefficient of mine buffer is 0.6684 at significance level 1%, and that of job opportunities is 1.1259 at significance level 1%. The coefficients are difficult to compare be compared directly due to the different levels and unit definition. However, the two results agree that both mining project characteristics are positive, at the 1% or 5% significance level. Also, it is shown that "job opportunities" has greater influence than "mine buffer" in both models.

The odds ratio results are shown in the fourth column of Table 6-4. The odds ratio is the exponentiated values of the coefficients, so these can be interpreted as odds ratios between levels (Allison, 2012). For example, the coefficient of job opportunities is 1.1259, and the odds ratio is calculated as $e^{1.1259} = 3.083$. This means that the odds of choosing job opportunities at the second level (600 people employed directly by the mine) is three times the odds of choosing job opportunities at the first level (300 people employed directly by the mine).

And the odds of choosing job opportunities at the third level (900 people employed directly by the mine) is three times the odds of choosing job opportunities at the second level (600 people employed directly by the mine). The odds ratios of the mining characteristics "income increase", "infrastructure improvement", and "availability of independent and transparent information" on potential impacts of mine and mine buffer are almost equal to two, which means there is twice the chance that an individual

will choose the level 2 over level 1, or level 3 over level 2 of these four mining

characteristics. The other two factors, "decision making mechanism" and "mine life"

have an odds ratio close to one, which means that the odds are the same that an individual

will choose one level over the other.

Willingness-to-pay (WTP) is a measure designed to determine the amount of

money that individuals are willing to forfeit in order to obtain some benefit from the

undertaking of some specific action or task. WTP is calculated as the ratio of two

parameter coefficients in MNL and CL models. It requires at least one attribute that is

measured in monetary units in the discrete choice experimental design, and the

coefficient of the factor with monetary unit will provide a financial indicator for all other

factors.

In this work, the mining characteristic "income increase (for all local residents)"

is the factor with a monetary unit (US$), and its coefficient has been estimated as 0.6600

at the 1% significance level. In calculating a measure of WTP, it is important that both

factors to be used in the calculation are found to be statistically significant, otherwise no

meaningful WTP measure can be established. The WTP of job opportunities from the CL

model can be calculated as Equation (6-6).

$$\text{WTP}_{job} = \frac{\beta_{job}}{\beta_{income}} \times 200 = \frac{1.1259}{0.6600} \times 200 = 341\$ / month \tag{6-6}$$

The WTP is multiplied by 200 since the level differential for income increase is

200$/month, but all coefficients were estimated with factors coded as 1, 2, and 3 (so the

differential is 1). This multiplication converts the WTP of job opportunities into dollars

per month. The WTP of $341/month for job opportunities means that directly employing

300 less people (300 is the level differential) at the mine will be acceptable to local

residents only if their incomes increase by $341/month. All WTP results are shown in

Table 6-4, with the standard error shown in the parentheses. The WTP measures are especially important for governance factors which are difficult to value in monetary terms. The WTP of these four governance factors are \$50, \$256, \$203 and \$36 per month for each level decrease, respectively.

6.5.1.2. Negative mining characteristics. As the model results in Table 6-4 show, the following factors were considered by participants to be negative impacts: increase in housing costs, labor shortage for other businesses, noise pollution, water pollution, air pollution, land pollution, crime increase and traffic increase. Labor shortage for other business is significant at the 5% significance level while all the other seven factors are significant at the 1% significance level. Increasing the levels of these eight factors will decrease the probability of the acceptance of mining project. While the coefficient of population increase is estimated as negative, it is not found to be a statistically significant influence (i.e. p-value > 0.1). A similar finding was reported by Ivanova et al. (2007), where population increase is uncorrelated with the decision to adopt a mining project.

As explained before, the coefficients of factors can represent the relative degree of influence of the factors in explaining individual choices. Thus, the absolute value of the coefficients can be used to rank the degree of influence of the eight negative mining characteristics as crime increase (6) > air pollution (6, 7) > increase in housing costs (5, 6) > noise pollution (5) > land pollution (6, 7) > water pollution (6, 7) > traffic increase (6) > labor shortage (4, 5). Similar to the positive factors, the ranking is different from what one would expect using the results from Section 3.

In Table 3-6, the factors "crime increase" and "traffic increase" were combined together and the importance level is found to be (6) -- very important. In the DCE, it was thought to be better to separate these two items. The results validate this choice. They are different items, and have very different degrees of influence. The results indicate crime increase is much more important to residents than traffic increase.

Also, although the respondents in the earlier survey (Table 3-6) ranked increase in housing cost relatively lower at (5, 6) -- above somewhat important and less than very important, in the DCE it has much higher degree of influence than the three pollutions items. The differences show why the DCE is superior to the ranking survey. The participants of the raking survey may have thought pollution is important to them because they never had to make trade-offs. When they were required to make trade-offs with housing costs, they chose lower housing costs. The discrete choice model can compare the immediate interest of participants and the factors which are difficult to value monetary terms.

In addition, the importance level of noise pollution was much less than the other three pollution impacts in the earlier survey. However, in the DCE, its influence is greater than land and water pollution. This result agrees with our literature review result in Section 2 that noise pollution is the single largest type of community complaint (ICMM, 2009). BHP Billiton reports that out of 536 complaints in 2008, 200 were related to noise (BHP, 2008).

"Increase in housing costs" and "water pollution" were studied by Ivanova and Rolfe (2011) as "housing and rental prices" and "water restrictions". In their CL model, the coefficients were estimated as 0.284 at the 5% significance level, and 0.218 at the 10% significance level. In this CL model, the coefficient of the factor, "increase in housing costs" is -1.0416 at significance level 1%, and that of "water pollution" is -0.1956 at significance level 1%. While the sign of these coefficients seems different, it is only because of how levels were defined in the different discrete choice experiments. In Ivanova and Rolfe (2011), the levels of "Housing and rental prices" and "water restriction" were defined from "best" to "worse" whereas the author defined the levels in this work from "worse" to "best."

The absolute value of the coefficients of "water pollution" and "water restrictions" are close to 0.2, which means their relative importance is similar in the different discrete choice experiments. The levels of water pollution are from "less than similar mine in the area" to "more than similar mine in the area", and that of water restrictions are from "none of households, town parks and gardens are drier than now" to "none of households, town parks and gardens are greener than now".

The absolute value of the coefficient of "increase in housing costs" in this work is only half of the coefficient of "housing and rental prices", in Ivanova and Rolfe (2011). It may be because the levels of the increase in housing costs in this work only varied in the range from "3% increase every year in 10 years" to "7% increase every year in 10 years", and that of the housing and rental prices used by Ivanova and Rolfe varied in the much bigger range of "25% increase" to "25% decrease".

As expected, all negative factors have odds ratio less than one (Table 6-4). For "increase in housing costs", the odds of choosing the second level (5% increase every year for 10 years) are only one third of the odds of choosing the first level (3% increase every year for 10 years), and the odds of choosing the third level (7% increase every year for 10 years) are also only one third of the odds of choosing the second level (5% increase every year for 10 years). Noise pollution, air pollution, and crime increase have similar odds ratio as increase in housing costs. Labor shortage for other businesses, water pollution, and land pollution all have odds ratios close to one.

The WTP of population increase was not estimated since population increase is not statistically significant, and the WTP is not meaningful. The WTP of increase in housing costs, noise pollution, air pollution, and crime increase is around -$300/month. For example, if the mining project causes housing costs to increase from first level (3% increase every year in 10 years) to second level (5% increase every year in 10 years), this 2% difference, to the respondents, is equivalent to losing $316 per month in income for

all local residents. The WTP for labor shortage for other businesses, water pollution, land pollution, and traffic increase are much less.

WTP measures are important for environmental-economic studies in which a common objective is the valuation of non-monetary attributes. In this study, it has been found that individuals in Salt Lake City treat noise pollution and air pollution as much more important than the water pollution and land pollution. Further work is necessary to determine whether this result can be generalized in some form to all mining communities.

6.5.1.3. Demographic factors. As the model result in Table 6-4 shows, the estimated coefficients (and p-values) of demographic factors revealed that age (5% level), gender (10% level), household income (10% level), and education (10% level) have significant effects on individual preferences. That is, individuals who differ in these demographic factors are likely to have different opinions.

The coefficient of gender is estimated as negative at 10% significance level. Since the first level of gender is male, and the second one is female, this shows a negative relationship between female and the probability of choosing a particular preference. Compared with the coefficients of mining characteristics, it is obvious that the influence of demographic factors is secondary to the mining characteristics. While individuals with variable backgrounds in Salt Lake City may have different preferences with regards to a mining project, the characteristics or impacts of the mining project itself are much more important. In fact, the influence of the demographics itself is, most likely, the result of how individuals of different demographics perceive the mining characteristics.

The results of demographic factors of the CL model confirm the correlation results (Fisher's exact test results) in the earlier survey reported in Table 3-3 (summary in Table 6-5). As shown in Table 6-5, age is observed to be negatively correlated to population changes, cultural impacts, and mine buffer. These three factors would highly affect lifestyle. It appears, from the results, that younger people in mining communities

care more about these lifestyle impacts. It may be the main reason why individuals with a higher age are more likely to prefer any particular mine option than individuals with a lower age.

Table 6-5 also shows a significant negative correlation between income and the possible negative impacts, which include traffic and crime increase, water shortage or pollution, air pollution, and land pollution. This means participants with higher incomes ranked traffic, crime, and pollution issues lower than those with lower incomes, and it can explain why people with higher income are more likely to prefer any particular mine option than people with a lower income.

Table 6-5. Summary significant results from Table 3-3

Demographic factors	Correlated characteristics of mining projects	Correlation coefficients (p-values)
Age	Population changes (M)	-0.240 (0.0171)
	Cultural impact(M)	-0.394 (<.0001)
	Mine buffer(M)	-0.286 (0.004)
Income	Traffic and crime increase (M)	-0.230 (0.030)
	Water shortage or pollution (M)	-0.320 (0.002)
	Air pollution (M)	-0.260 (0.012)
	Land pollution (M)	-0.265 (0.011)
Education	Job opportunities (M)	-0.212 (0.035)
	Income increase (M)	-0.236 (0.019)
Gender	Traffic and crime increase (M)	(0.037)
	Job opportunities (M)	(0.046)
	Cost of housing or housing shortage (M)	(0.023)
	Water shortage or pollution (M)	(0.015)
	Land pollution (M)	(0.005)
	Decision making mechanism (M)	(0.048)
	Mine buffer (M)	(0.027)
	Mine life (M)	(0.049)

These differences in the preferences of individuals with different demographics are most likely due to differences in how individuals of different demographics perceive the mining attributes (Que et al. 2015). Future work needs to be done to establish the exact differences in preferences of males and females which was found to be the most

significant demographic factor. The same discrete choice experiment could administered to males and females. The difference between the two discrete choice models can show the exact differences in preferences of males and females.

There is also a negative correlation between education and job opportunities and income increase. This means respondents with higher education are less concerned about new job opportunities and potential income increases associated with the mining operation. This negative correlation may be because people with higher education have less of a need to change jobs and work for the new mine or consider mining-related jobs to be less desirable. However, people with higher education are still more likely to prefer any particular mine option than people with lower education.

From the Fisher's exact test results of gender, there is a significant difference between female and male rankings of eight mining characteristics. They are traffic and crime increase, job opportunities, cost of housing or housing shortage, water shortage or pollution, land pollution, decision making mechanism, mine buffer, and mine life. The Fisher's exact test only seeks to determine whether there is a significant difference between the distribution of responses from the male and female groups, but does not determine which group (male or female) rank a particular characteristic higher/lower. However, the tendency is revealed in the CL model result. Females have a significant negative relationship with the probability of mining project adoption. That is, females are less likely to prefer any particular mine option that males.

"Age", "income" and "gender" were studied by Ivanova and Rolfe (2011) and Ivanova et al. (2007). In their MNL model, the coefficients were estimated as 0.037 (age) at significance level 5%, 0.000 (income) at significance level 5%, and 1.243 (gender) at significance level 1%.

In this CL model, the coefficient of age is 0.037 at the 5% significance level, and that of income is 0.0043 at the 10% significance level. While the coefficients are difficult

to compare due to differences in levels, unit definition, and modeling parameters, these results confirm that both of these demographic factors are positive at 5% or 10% significance level. Also, the degree of influence of "age" is greater than that of "income" in both models.

The coefficient of gender is difficult to compare since the definition of levels of gender are not shown in Ivanova and Rolfe (2011). If Ivanova and Rolfe (2011) defined the first level as female and second level as male, then the results of this work are confirmed by theirs. Otherwise, they are different. However, regardless of whether gender is negatively or positively related, the absolute value of the gender coefficient is the largest of all demographic factors, which means gender has the largest influence among all demographic factors in the decision of mining project adoption.

6.5.2. Conditional logit model stratified by question. Modeling with the conditional logit model stratified by question (CLQ) was done using the SAS LOGISTIC procedure and the STRATA statement. With the STRATA statement, the LOGISTIC algorithm has the ability to do a stratified analysis. The CLQ model results are shown in Table 6-6.

In the CL model, the LOGISTIC procedure, by itself, analyzes the local mining communities' preference by comparing only the decision codes, 0 and 1, of all questions together (Figure 6-2, tenth column). However, in the real case, participants were answering questions one by one. And in each question, they only had three options. The STRATA statement instructs the algorithm to consider the data by choice set (Figure 6-2, eleventh column). The CLQ model fit the observed data reasonably well. The goodness-of-fit of the CLQ model is slightly better than the CL model. The LRI (pseudo-R^2) of the CLQ model is 0.2696 compared to 0.2687 for the CL model. The percent concordant of the CLQ model increased to 78.5 and the percent discordant and percent tied are decreased to 18.7 and 2.8, respectively.

Table 6-6. Conditional logit model stratified by question result

Parameter	Coefficient	Standard Error	Odds Ratio	WTP $/month (error)
Economic				
Job opportunities	1.3886***	0.0562	4.009	221(9)
Income increase	1.2541***	0.0697	3.505	200(11)
Increase in housing costs	-1.7527***	0.0706	0.173	-280(11)
Labor shortage for other business	-0.1117**	0.0463	0.894	-18(7)
Environmental				
Noise pollution	-1.6794***	0.0713	0.186	-268(11)
Water pollution	-0.3471***	0.0566	0.707	-55(9)
Air pollution	-1.8216***	0.0735	0.162	-291(12)
Land pollution	-0.2707***	0.0488	0.763	-43(8)
Social				
Population increase	-0.2570***	0.0532	0.773	-41(8)
Infrastructure improvement	1.1575***	0.0601	3.182	185(10)
Crime increase	-1.6939***	0.0703	0.184	-270(11)
Traffic increase	-0.1742***	0.0453	0.840	-28(7)
Governance and others				
Decision making mechanism	0.2028***	0.0499	1.225	32(8)
Information available	1.2606***	0.0649	3.528	201(10)
Mine buffer	1.2141***	0.0620	3.367	194(10)
Mine life	0.1402***	0.0460	1.150	22(7)
Demographic factor				
Age	0.0028*	0.0015		
Gender	-0.0093*	0.0033		
Income	0.0021*	0.0017		
Education	0.0017*	0.0009		

*** 1% significance level, **5% significance level, *10% significance level.

6.5.2.1. Comparing taste coefficient results. Compared to the CL model result, the coefficient of population increase is estimated as -0.2570 at the 1% significance level in the CLQ model. It is the main difference between the CLQ and CL models. In the CL model, population increase was estimated as a non-significant factor.

The same seven mining project characteristics are estimated as positive impacts at the same 1% significance level. These are job opportunities, income increase, infrastructure improvement, decision making mechanism, information available, mine

buffer, and mine life . Increasing the levels of these factors will increase the probability of acceptance of a mining project.

However, all positive coefficients are bigger than those of the CL model. The influences of these seven mining project characteristics are estimated to be higher than those of the CL model. The coefficient of job opportunities increased from 1.1259 to 1.3886. The coefficients of income increase, infrastructure improvement and mine buffer almost doubled to 1.2541, 1.1675 and 1.2141, respectively. The coefficients of decision making mechanism, information available, and mine life increased slightly.

The negative project characteristics in the conditional logit model stratified by question have the same tendencies as the previous results. All coefficients of these eight characteristics are still negative in the CLQ model, at the same significance level (1% or 5%). These are increase in housing costs, labor shortage for other businesses, noise pollution, water pollution, air pollution, land pollution, crime increase and traffic increase. Increasing levels of these factors will decrease the mining project adoption probability. Also, most of the absolute values of the negative coefficient (seven of eight, with traffic increase as the exception) are bigger than those of the CL model. Their influences are estimated to be higher than in the CL model. The coefficients of increase in housing costs, noise pollution and air pollution are almost doubled to -1.7527, -1.6794 and -1.8216, respectively. The coefficients of water pollution, land pollution, crime increase and labor shortage decreased much more gradually.

As the model results in Table 6-6 show, the coefficients of demographic factors of CLQ model are similar but less than that that of CL model. All four demographic factors are estimated at the 10% significance level. This result of CLQ model confirms that the influence of demographic factors is less than the mining characteristics. This tendency is more predominant when the decisions of mining communities have been stratified by the choice sets.

6.5.2.2. Comparing odds ratio. The odds ratios of positive factors in CLQ model are bigger than those of the CL model.

For example, the odds ratio of job opportunities was 3.0883 in the CL model, and it increases to 4.009 in the CLQ model. It means the odds of choosing job opportunities at the second level (600 people employed directly by the mine) is four times the odds of choosing job opportunities at the first level (300 people employed directly by the mine), and the odds of choosing job opportunities at the third level (900 people employed directly by the mine) is four times the odds of choosing job opportunities at the second level (600 people employed directly by the mine). The odds ratios of income increase, infrastructure improvement, information available, and mine buffer were estimated to be around 2 in the CL model. In the CLQ model, their odds ratios are increased to 3.505, 3.182, 3.528, and 3.367, respectively. The other two factors, decision making mechanism and mine life, which have odds ratio close to one in the CL model, have slightly increased odds ratios.

As expected, the odds ratios of the seven of eight negative mining factors (except traffic increase) plus the new significance of a negative mining factor (i.e. population increase) are decreased. The odds ratios of increase in housing costs, noise pollution, air pollution, and crime increase in the CL model are almost one third of that in the CLQ model. In the CLQ model, these odds ratios were decreased by a half, which means one-sixth of people are likely to choose the level 2 compared to level 1, or level 3 compared to level 2 for these four mining project characteristics. The odds ratios of labor shortage for other business, water pollution, land pollution and population increase were slightly lower.

6.5.2.3. Comparing WTP results. Comparing the WTPs of the CL and CLQ models, the WTPs of all positive mining characteristics are lower in the CLQ model. As shown in Equation (6-6), the WTP of job opportunities is proportional to the ratio between the coefficients of job opportunities and income increase. While the coefficient of job opportunities increased from 1.1259 to 1.3886, that of income increase doubled, in the CLQ model. Thus, the WTP of job opportunities decreased from $341 to $221/month. Similarly, the WTPs of infrastructure improvement, decision making mechanism, information available, mine buffer and mine life decreased to $185, $32, $201, $194 and $22 per month for each level increase, respectively.

The WTP of population increase can now be estimated in the CLQ model since the factor was found to be statistically significant at the 1% significance level. As shown in Table 6-6, the WTP of population increase is -$41/month. This means if population increases from the first (2% annually) to second level (4% annually) due to a mining project increases, this additional 2% per annum is equivalent to an additional $41 per month income for all local residents. The absolute value of WTPs of all the eight negative mining characteristics decreased, meaning the factors are of less value for the local mining residents.

6.5.3. Mixed Logit Model. The mixed logit model was done using the MDC procedure in SAS. This procedure supports three distributions: normal, lognormal and uniform. The coefficient of each mining project characteristic was tested with a model of these three distributions. Only two factors were found to fit distributions at a significance level. The coefficient of labor shortage for other businesses was found to be lognormally distributed with mean -0.0746 and standard deviation 0.0188 at 10% significance level. The coefficient of mine buffer is normally distributed with mean 0.7519 (1% level) and standard deviation 0.0165 (10% level). The results are shown in Table 6-7.

The LRI of the ML model, which was 0.3127, was bigger than the CL and CLQ models. The coefficients of the other factors (those with no distributions) were similar to those of the CL and CLQ models. However, as discussed in Section 2.4, there is a big cost to using this advanced model to relax the iia limitation and use random parameter distributions. Thus, whether the ML model is suitable for a particular case of mining stakeholder analysis depends on whether relaxing the restrictions (iia and fixed taste coefficients) is important to the intended use of the model.

First of all, the coefficient of the demographic factors could not be estimated in ML model with the MDC PROC algorithm in the general software SAS. From both the CL and CLQ results, the coefficients of all four demographic factors are much smaller than that of mining characteristics, and most of them are significant only at the 10% level. This means that the influence of demographic factors is far less than the mining characteristics, in this case. Therefore, in this case, a model that does not include demographic factors is acceptable.

Second, the ML model is more proper for factors with continuous levels since the coefficients will be estimated as distributions. In this case, only two mining characteristics have been estimated as distributions at any significance level. Thus, this limitation restricts the application of ML model in mining stakeholder analysis a lot. Also,

the odds ratio and WTPs cannot be estimated in the model since the factors are no longer

assumed to have the independence of irrelevant alternatives (iia) property. The

importance of this limitation depends on whether or not this information is necessary for

achieving the goals of the community engagement.

Table 6-7. ML model result

Parameter	Coefficient	Standard Error
Economic		
Job opportunities	0.9139***	0.0518
Income increase	0.7502***	0.0613
Increase in housing costs	-1.0828***	0.0543
Labor shortage for other business_M	-0.0746*	0.0434
Labor shortage for other business_S	0.0188*	0.0036
Environmental		
Noise pollution	-1.0362***	0.0544
Water pollution	-0.2218***	0.0514
Air pollution	-1.1147***	0.0598
Land pollution	-0.1803***	0.0504
Social		
Population increase	-0.1727***	0.0516
Infrastructure improvement	0.7195***	0.0459
Crime increase	-1.0269***	0.0615
Traffic increase	-0.1145***	0.0400
Governance and others		
Decision making mechanism	0.1262***	0.0470
Information available	0.7458***	0.0564
Mine buffer_M	0.7519***	0.0483
Mine buffer_S	0.0165*	0.0039
Mine life	0.0930**	0.0397

M: mean, S: standard deviation

6.5.4. Discussion. From the case study results, the CLQ model appears more suitable for mining stakeholder analysis. This is not only because of the bigger LRI and better percent concordant in this particular case study. The main reason is that the CLQ model can do stratified analysis, which makes it more practically applicable. It better represents the way respondents considered the choices.

While the mixed logit model is the most advanced model with the biggest LRI, its disadvantages have never been discussed well in other studies. First, the coefficients of demographic factors could not be estimated in the ML model with the most widely used advanced statistical software SAS (with its MDC PROC). Second, not all the factors of the ML model can be estimated as distributions at any significant level. Based on the author's experience, the ML model is more appropriate for factors with continuous levels (e.g. time). Third, the odds ratio and WTPs cannot be estimated with the ML model since the independence of irrelevant alternatives (iia) property is relaxed in this model.

As shown by the CLQ model result (Table 6-6), the goal to use discrete choice theory for mining stakeholder analysis has been achieved by answering three important questions:

(1) What are the factors that affect individual's decision and how do these affect the decision?

In Salt Lake City, there are 16 mining project characteristics that affect the communities' acceptance of a mining project, at 1% and 5% significance levels (Table 6-6). Of these, seven factors are positive (i.e. they positively correlate to the likelihood of individuals accepting a project) and the remaining are negative. The degree of influence of the positive mining project characteristics are job opportunities > information available > income increase > mine buffer > infrastructure improvement > decision making mechanism > mine life. The degree of influence of the negative mining project characteristics are air pollution > increase in housing costs > crime increase > noise

pollution > water pollution > land pollution > traffic increase > labor shortage. These results are much more realistic than those obtained by soliciting such results from respondents independently, as was done in the survey in Section 3, for example. The discrete choice theory results are based on hypothetical choices that force respondents to make real trade-offs based on their perceived importance.

(2) What is the effect of demographics on individual preferences?

Age, household income, education and gender were found to significantly affect individual preferences at 5% and 10% significance levels. The result reveals that, in Salt Lake City, older males with higher household incomes and more education are more likely to prefer any particular mine option than younger females with lower household income and less education.

(3) What is the value of environmental and social impacts to individuals in the community?

There are eight negative mining project characteristics affecting individual acceptance of a mining project at 1% and 5% significance levels (Table 6-6). The results can be used to estimate the average Salt Lake City resident's WTPs (or value) for the undesirable effects of these factors. The ranking of predicted value of the negative effects, in additional income, is air pollution ($291/month) > increase in housing costs ($280/month) > crime increase ($270/month) > noise pollution ($268/month) > water pollution ($55/month) > land pollution ($43/month) > traffic increase ($28/month) > labor shortage ($18/month).

The number of internet surveys has increased dramatically in the last 10 years. Online surveys have a number of advantages over traditional survey modes. First, online surveys are not limited by the space and time of respondents. Second, online surveys allow researchers to use multimedia elements. Take this case study as example, the author inserted two videos (a respondent was not allowed to skip this) to help the

respondents understand the survey background information and the survey questions. Third, the survey duration of each question can be recorded to allow researchers to track each participant. These functions are not available to other survey modes.

However, online surveys have their drawbacks. First of all, online surveys rely on the internet and can be subject to significant biases resulting from under-coverage and nonresponse. Not everyone in the mining community has access to the Internet. Hence, the demographic distribution of respondents may be significantly different from the mining community's. This drawback affected this case study's sampling. While 150 participants were excluded due to demographic factors, the distributions of income and education still could not be completely matched to the CLQ population, very well. As shown by the demographic distribution of participants in Table 6-2, the main problem is that not enough respondents with lower incomes and education were reached by our market research partner, Qualtrics. While the population of SLC has 14% and 22% of people with education less than high school and annual income less than $20,000, respectively, less than 1% and 7% of the participants in this survey had that level of education or income, respectively. This tendency is confirmed by research by the Pew Research Center (2015), which found that people with lower incomes, less education, living in rural areas or ages 65 and older are underrepresented among internet users.

In addition, volunteer bias is present in any survey. A voluntary sample is made up of people who self-select into the survey. Often, these people have a strong interest in the main topic of the survey. The sample is chosen by the viewers, not by the survey administrator. Thus, the sample used in this work would be a voluntary sample, not a random sample. Therefore, the resulting sample and the following result tends to over represent individuals who have strong opinions. This was mitigated to an extent in this survey because Qualtrics recruited participants from its large network of users who have

signed up to take surveys of all kinds, limiting the possibility of someone only taking the survey because they have a strong bias for or against mining generally.

Third, the factors and their level were colored-coded to be clear and easy for the participants to understand, based on the feedback from the focus group survey in Section 5. However, these color codes may have biased some of the results by priming the respondents to think in a particular way. Take mine life as an example. In the discrete choice experimental design (Section 5), the author colored the first level (20 years) of mine life green representing the "best" level; the second level (30 years) yellow, indicating the intermediate level; and the third level (40 years) red, representing the "worst" level. Contrary to the author's expectation, mine life is estimated as a positive factor with coefficient 0.1402 (CLQ model result, Table 6-6) at the 1% significance level. The coefficient of mine life is low compared to other positive factors. The coefficient may have been bigger if the factor had been color coded differently.

Based on these discussions and experiences, the online survey alone is not the best option for mining stakeholder analysis using discrete choice modeling. Discrete choice experiments for mining stakeholder analysis does not really benefit from the "no space limitation" of online surveys since the target is the several local communities around the mining project. A mixed-mode survey is recommended for the local mining stakeholder analysis using DCE. For example, mining companies could use face-to-face interviews with the support of multimedia elements to overcome the limitation of paper surveys. The participants still can get background information with video and graphics.

6.7. SUMMARY OF SECTION SIX

This Section illustrates the usefulness of discrete choice theory for stakeholder analysis in mining by conducting a discrete choice experiment in Salt Lake City, UT and analyzing the results to make useful inferences. A major technical challenge was an

attempt to select the most appropriate discrete choice model to describe the local community's acceptance of mining projects. The discrete choice experiment designed in Section 5 was conducted in Salt Lake City, UT. The conditional logit (CL), strata conditional logit (CLQ) and mixed logit (ML) models were evaluated using the log-likelihood ratio index as a measure of goodness-of-fit.

After balancing the advantages and disadvantages of each model, the CLQ model is recommended as the most appropriate discrete choice model for mining stakeholder analysis. The CLQ model had the second highest LRI; the ML model had the highest LRI. However, the disadvantages of the ML model restrict its application in mining stakeholder analysis.

More importantly, all three questions posed have been successfully answered by the selected (CLQ) model result. This achieves the goal of illustrating that discrete choice theory can be used for stakeholder analysis in mining. Discrete choice modeling results can be a guideline for the mining company during mining project design, planning and management. Discrete choice theory can support successful community consultation.

7 CONCLUSIONS, RECOMMENDTATIONS &FUTURE WORK

7.1. SUMMARY AND CONCLUSIONS

A key part of community engagement is community consultation, which includes three main parts: stakeholder identification, stakeholder analysis and iterative consultation (ICMM, 2012a; IFC, 2007). Stakeholder analysis is one of the key challenges in community consultation since misunderstanding stakeholders will misguide the whole community consultation effort. Current stakeholder analysis processes (ICMM, 2012a) are mainly qualitative, and classify stakeholders into three groups: highly influential supporter of the project, neutral about the project, and highly influential opponent of the project.

This kind of stakeholder analysis alone is not enough to support the success of the whole consultation process. The main goals of community analysis should include answering the following questions: (1) what are the factors that affect stakeholders' decisions and how do these factors affect their decisions/preferences? (2) what is the effect of demographics on individual preferences? (3) what is the value of environmental and social impacts to individuals in the community?

Discrete choice theory, based on the Nobel winning work by McFadden (1974) has transformed the world of market research. As a statistical analysis method, discrete choice theory aims at analyzing individual decision marker's preferences. Discrete choice modeling can help us understand what kind of mining project individuals in a community prefer by comparing different hypothetical options. By identifying patterns in these choices, discrete choice models will provide insight into how different individuals respond to different mining options. DCM will allow mining companies to examine the significance of different mining impacts (including social, economic, and environmental) and other aspects of a project on the preferences of different groups of in the local communities. Compared to traditional stakeholder analysis methods, the mining company

will have a quantitative tool for planning, designing, operating, and managing their mining project in order to facilitate better community engagement.

The goal of this PhD research is to facilitate improved community (stakeholder) analysis by providing further insight on the determinants of local community acceptance using discrete choice theory. Pursuant to the overall goal of this study, the specific objectives are to:

1. Identify, classify, and verify the important mine characteristics and key demographic factors that affect local community acceptance of a mining project;

2. Account for the large number of relevant factors inherent in discrete choice experiments for mining community acceptance evaluation; and

3. Examine discrete choice models to select the most appropriate model for mining community consultation. The research will test the hypotheses that various discrete choice models can describe the local community's acceptance of mining projects.

Pursuant to the aims of this study, online surveys were done in 20 mining communities and 20 non-mining communities to validate a classification of important mining project characteristics developed from a comprehensive literature review. A discrete choice experiment was then designed, based on the validated list of important factors, for Salt Lake City, UT, a select mining community. Three candidate discrete choice models were applied for the discrete choice experiment data.

Based on the work in this dissertation, the following conclusions can be drawn:

1. On research objective one:

(1) All sixteen project characteristics, identified and classified through a literature review, were confirmed as important to the decision to accept or not accept a mining project. The most important mining project characteristics are job opportunities, water shortage or pollution, air pollution and land pollution. This list of characteristics is not put forth as universally true. This is found to be

generally true in the United States. Depending on the particular context of a specific project, however, the list may change as appropriate.

(2) Respondents living in mining and non-mining communities have similar opinions of 12 mine characteristics and appear to differ on four (infrastructure improvement, labor shortage for other businesses, noise pollution, and mine life). The candidate hypothesizes that this is due to differences in the experience of these two groups. Hence, stakeholder analysis in a community with or without prior mining should be approached differently.

(3) Four of the six selected demographic factors were confirmed to be significantly (p < 0.05) correlated with respondents' opinion of the importance of the mine characteristics. Gender, income, age, and education are important predictors of an individual's decision to accept or reject a proposed mining project.

2. On research objective two:

(1) A mixed style, blocking scheme, factional factorial discrete choice experiment without interaction is proposed as a solution to overcome the large number of relevant factors in mining community analysis. The relative D-efficiency of the discrete choice experiment was 72%.

(2) A design with four factors in each choice set is optimal for the block scheme experimental design. Using four factors in each choice set balances the survey cost with reasonable cognitive burden.

(3) A focus group study was used to validate the experimental design. The discrete choice experiment design achieved acceptable difficulty and clarity for questions in all blocks.

3. On research objective three:

(1) In the case study, all three candidate discrete choice models showed acceptable goodness-of-fit.

(2) The conditional logit model stratified by question was found to be better than the conditional logit model in the case study, since the CLQ model has the ability to do a stratified analysis by choice set.

(3) While the mixed logit model is the most advanced discrete choice model, its disadvantages restrict its application for mining stakeholder analysis.

Also, this work has successfully demonstrated that discrete choice theory can be used in mining community consultation for stakeholder analysis. Three important questions posed in support of community consultation can by answered with discrete choice theory.

7.2. CONTRIBUTION OF THE PHD RESEARCH

1. Contribution to knowledge on factor selection, identification, and verification

This dissertation is the first attempt to provide research on classifying and verifying the key mining project characteristics from the plethora of candidate characteristics for discrete choice experimentation. Section 3, and the statistical analysis methods in it, will be helpful for researchers who would employ discrete choice theory in any kind of economic or project development on: (1) How to classify and verify the important project characteristics for discrete choice experiments? (2) How to find key demographic factors, which are significant vis-à-vis people's perception of the importance of the project characteristics? (3) Is there a difference between attitudes of local and non-local communities? This contribution serves as a starting point for efficient choice experiment (survey) design and effective discrete choice modeling.

2. Contribution to knowledge on discrete choice experiment design.

This dissertation is the first attempt, to the best of the author's knowledge, to design discrete choice experiments (DCEs) with such a large number (16) of factors and

also to attempt to address the challenges associated with the clarity and difficulty of questions. The experience of DCE design and validation can be borrowed by other discrete choice theory researcher. The most important design contribution is that DCEs: (1) can be designed as mixed style, including both a status quo option and hypothetical situations; (2) can be designed as a block scheme, using the optimal number of factors, determined by balancing the survey cost and reasonable cognitive burden; (3) need to be validated using a focus group, and then revised based on the feedback. This contribution is the foundation of effective and efficient discrete choice experiment design, even with the large number of factors.

3. *Contribution to knowledge on the discrete choice modeling of mining stakeholders*

This dissertation is the first attempt at comprehensive discrete choice modeling for mining stakeholders. This work includes 20 factors (16 mining project characteristics and four demographic factors) in a discrete choice model to analyze the mining stakeholders. The only other examples of discrete choice modeling in mining contain five or seven factors (Ivanova et al., 2007; Ivanova & Rolfe, 2011). In this study, the author designed the DCE as a block scheme, using the optimal number of factors.

4. *Contributions to knowledge on the most suitable discrete choice model for mining stakeholder analysis*

This dissertation is the first attempt to critically evaluate models in order to select the most suitable for mining stakeholder analysis. Ivanova and Rolfe (2011) and Ivanova et al. (2007) used the MNL model without any discussion of whether it was the most suitable model for mining stakeholder analysis. This work is the first research to highlight the relevance of the conditional logit model stratified by question and discuss the disadvantages of the mixed logit model. The knowledge gained from applying the CLQ and ML models to stakeholder analysis are helpful for researchers who would employ these two models similar applications: (1) The CLQ model is more suitable for

stakeholder analysis than the popular CL model since the CLQ model can do stratified analysis, which makes it more practically applicable. (2) While the ML model is the most advanced model, its limitations affects its usefulness in mining stakeholder analysis.

7.3. RECOMMENDATIONS FOR FUTURE WORK

The following are recommendations for future research that will improve on the present work and further our understanding of local community acceptance of mining:

1.   *Evaluation of cultural impact*

In Section 3, the importance level of "cultural impact" was ranked at (5, 6) — above "somewhat important" but below "very important" – by respondents. However, this factor was deleted in the discrete choice experimental design since there is no clear definition of cultural impact, which makes it invalid in a survey instrument. This factor can be separated into several clearly defined mining project characteristics, which are suitable for further discrete choice experimental design and discrete choice modeling.

2.   *Inclusion of the effect of females with and without children*

In Section 3, the demographic factor "number of children" was observed not to be significantly correlated with respondents' choices. That is why the number of children was not included in the following discrete choice model. It may be possible that whether the respondent has any children at all is affects the respondents choices even if the total number does not. This can be tested by comparing the distribution of respondents with and without children and even further splitting the respondents into males and females, with and without children. If it is observed that whether a respondent has children or not has an effect on the distribution of rankings, then it is possible that the discrete choice model may be improved by including this factor. The nested logit model is a possible approach to achieve this.

3.   Extending the observations on the classification of important factors and discrete choice modeling to other mining contexts

Currently, most mining projects are moving to developing countries, such as China. However, the benefits and costs to local mining communities have not been studied thoroughly. The discrete choice model result can be used to get a better understanding of the determinants of community acceptance in those contexts. The selection and classification of the mining characteristics and demographic factors may vary between different communities and countries. Thus, the author suggests the whole methodology of this dissertation should be applied to select the important factors for a given target mining communities' acceptance of a mining projects, design the discrete choice experiments, and conduct discrete choice modeling with the data. Then, the discrete choice model result would help the mining community to plan, design, process, and manage mining projects better.

4.   Extending the discrete choice modeling to non-mining communities

The case study in this dissertation focused on a community with a large mining project located nearby. Thus, the discrete choice model result shows the opinions of people who are used to a mining project. It will be interesting if this study can be duplicated in non-mining communities. The differences between the mining and non-mining communities will be important information for mining companies. The result may be similar to the importance level ranking result from mining and non-mining communities in Section 3. However, there may be some significant differences as the results in Section 3 also indicates.

APPENDIX A:

ONLINE SURVEY MINING COMMUNITIES LIST

APPENDIX A: ONLINE SURVEY MINING COMMUNITIES LIST

1. Eureka, Nevada   68859

2. Elko, Nevada 89801 89802 89803

3. Gillette, WY 82716 82717   82718 82731   82732

4. Butte, Montana 59701   59702   59703 59707 59750

5. Lemhi County, Idaho   83465

6. Carthage, Missouri 64836

7. Centralia, in Lewis County, Washington 98531

8. Fairbanks mining district of Alaska 99701 99702 99705 99706 99707 99708
   99709   99710 99711 99712   99714 99716 99767 99775 99790

9. Lead, South Dakota 57754

10. Boron, California 93516   93596

11. Bunker, Missouri 63629

12. Carbondale, IL   62901   62902 62903

13. Terre Haute, IN 47801   47802 47803 47804   47805 47807 47808 47809

14. Cutler, IL 62238

15. Percy, IL 62272

16. Steeleville, IL 62288

17. Somerset, PA 15501 15510

18. Jenners, PA 15546

19. Jennerstown, PA 15547

20. Meyersdale, PA 15552

APPENDIX B:

ONLINE SURVEY NON-MINING COMMUNITIES LIST

APPENDIX B: ONLINE SURVEY NON-MINING COMMUNITIES LIST

1. Kly, Nevada 89301

2. Carson city, Nevada 89701 89702 89703 89704

3. 89705 89706 89711 89712 89713 89714 89721

4. Douglas, WY 82633

5. Idaho falls, Idaho 83401 83402 83403 83404 83405 83406 83415

6. Hamilton County, Illinois 62817 62828 62829 62860 62859

7. Winfield, Kansas 67156

8. Ellensburg, Washington 98926 98950

9. Wasilla, Alaska 99623 99629 99652 99654 99687

10. Chadron, Nebraska 69337

11. Lone pine, California 93545

12. Murrayville, Illinois 62668

13. Owensboro, Kentucky 42301 42302 42303 42304

14. Muncie, IN 47302 47303 47304 47305 47306 47307 47308

15. Morganfield, Kentucky 42437

16. Ortonville, Michigan 48462

17. Chaffee, MO 63740

18. Clearfield, PA 16830

19. Rigby city, Idaho 83442

20. Kingsley city, Iowa 51028

21. Arcola city, Illinois 61910

APPENDIX C

ONLINE SURVEY

APPENDIX C: ONLINE SURVEY

Q1. Do you live or have you ever lived near a mine?
Yes
No

Q2. If yes, how far is/was it?
< or =10 miles
11-20 miles
21-30 miles

Q3. What is your zip code?

Q4. Do you have any experience with mining (e.g. working for a mine, familiarity with mining activities, studying about mining)? If yes, what is it?
Yes
No

Q5. What is your gender?
Male
Female
Prefer not to answer

Q6. How old are you?
18-25
26-34
35-54
55-64
65 or over
Prefer not to answer

Q7. What is the highest level of education you have completed?
Less than High School
High School / GED
Some College
2-year College Degree
4-year College Degree
Masters Degree
Doctoral Degree
Professional Degree (JD,
MD)
Prefer not to answer

Q8. What is your annual income?

Below $20,000
$20,000 - $29,999
$30,000 - $39,999
$40,000 - $49,999
$50,000 - $59,999
$60,000 - $69,999
$70,000 - $79,999
$80,000 - $89,999
$90,000 or more
Prefer not to answer

Q9. In which industry are you employed?
Forestry, fishing, hunting or agriculture
support
Mining
Utilities
Construction
Manufacturing
Wholesale trade
Retail trade
Transportation or warehousing
Information
Finance or insurance
Real estate or rental and leasing
Professional, scientific or technical
services
Management of companies or enterprises
Admin, support, waste management or
remediation services
Educational services
Health care or social assistance
Arts, entertainment or recreation
Accommodation or food services
Other services (except public
administration)
Unclassified establishments

Q10. How many children (under the age of 18) do you have?
0
1
2
3
4
5 +
Prefer not to answer

Q11-26: If a new mine is to open in your area, carefully consider the following factors, and rank the importance of each factor in your decision to support or not support the mine.

1 Not at all Important
2 Very Unimportant
3 Somewhat Unimportant
4 Neither Important nor Unimportant
5 Somewhat Important
6 Very Important
7 Extremely Important
8 Do not know
9 Prefer not to answer

Mining impacts
Population changes
Infrastructure improvement (e.g transportation, education, human service, Internet, hospital, and shopping)
Cultural impact (e.g. impacts on archaeological and historical sites, native American artifacts, historical burial sites, arts and culture)
Traffic and crime increase
Job opportunities
Income increase
Cost of housing or housing shortage
Labor shortage for other businesses
Noise pollution
Water shortage or pollution
Air pollution
Land pollution
Decision making mechanism on the mine's permits (e.g. decisions are based only on what is legal; or decision makers consider input from local communities)
Whether or not there is independent and transparent information available
Mine buffer (distance of your residence from mine)
Mine life (how long the mine will last)

Q27: What other factor(s) (characteristics of the mining operation) is important for you?

APPENDIX D:

FACTORS AND LEVELS

APPENDIX D: FACTORS AND LEVELS

Economic:

1. Job opportunities
(1)300 people employed directly by the mine
(2)600 people employed directly by the mine
(3)900 people employed directly by the mine

2. Income increase (for all local residents)
(1)+ $100 per month
(2)+ $300 per month
(3)+ $500 per month

3. Increase in housing costs
(1)2% reduction in increase every year for 10 years – *3% increase every year in 10 years*
(2)0 % additional increase every year for 10 years – *5% increase every year in 10 years*
(3)2 % additional increase every year for 10 years – *7% increase every year in 10 years*

4. Labor shortage for other business
(1)Negligible (no noticeable effect on other businesses)
(2)Slight (other businesses take longer to fill vacancies but don't have to pay more)
(3)Moderate (other businesses take longer to fill vacancies and have to offer higher wages)

Environmental:

1. Noise pollution
(1)No increase in pollution – Less than similar mine in the area
(2)A slight increase in pollution – *Same as similar mine in the area*
(3)A moderate increase in pollution – *More than similar mine in the area*

2. Water pollution and shortage – *Water pollution*
(1)No increase in pollution – Less than similar mine in the area
(2)A slight increase in pollution – *Same as similar mine in the area*
(3)A moderate increase in pollution – *More than similar mine in the area*

3. Air pollution
(1)No increase in pollution – Less than similar mine in the area
(2)A slight increase in pollution – *Same as similar mine in the area*
(3)A moderate increase in pollution – *More than similar mine in the area*

4. Land pollution and subsidence – *Land pollution*

(1)No increase in pollution – Less than similar mine in the area

(2)A slight increase in pollution – *Same as similar mine in the area*

(3)A moderate increase in pollution – *More than similar mine in the area*

Social:

1. Population increase

(1) A reduced rate of population growth (only 2%) – *2% annually*

(2) Continued population growth (average rate 4%) – *4% annually*

(3) An increased population growth (6%) – *6% annually*

2. Infrastructure improvement (transportation, education, human serves, internet)

1) Slight improvement

2) Moderate improvement

3) Considerable improvements

3. Traffic increase

(1) A reduced rate of traffic increase – *Lower than current rate*

(2) Continued average rate of traffic increase – *Same as current rate*

(3) An increased rate of traffic increase – *Higher than current rate*

4. Crime increase

(1) A reduced rate of crime increase – *Lower than current rate*

(2) Continued average rate of crime increase – *Same as current rate*

(3) An increased rate of crime increase – *Higher than current rate*

Management and other:

1. Permit approval decision making mechanism

(1)Final decision solely by Government agency

(2)Final decision by Government agency after significant public input

(3)Final decision by Government agency after negotiating with local representatives

2. Availability of independent and transparent information on potential impacts of mine

(1)Information reported by mining company only

(2)Information reported/verified by government agency

(3)Information reported/verified by third party (e.g. non-profit or independent expert)

3. Mine buffer (Home distance from mine)

(1)5 mile
(2)10 mile
(3) >20 mile

4. Mine life
(1)20 years
(2)30 years
(3)40 years

APPENDIX E:

SURVEY INSTRUCTIONS AND SAMPLE PROBLEM

APPENDIX E: SURVEY INSTRUCTIONS AND SAMPLE PROBLEM

Instructions: If a new mining project was to start near Salt Lake City, it would likely affect you and the community in many different ways.

The objective of this survey is to understand what kind of mining project you prefer by varying the impacts (better or worse) of all project characteristics.

In the survey, you will be presented with several possible combinations of impacts and mine characteristics. You will then be asked to choose your preferred alternative from the three options available. Impacts will be color coded to help you quickly understand how they might affect the community. Impacts colored yellow represent current conditions in Salt Lake or the impact of a similar mine in the area. Impacts colored red are worse than the current conditions and those colored green are better than current conditions.

Shown below are 16 project impacts or mine characteristics classified into four major categories. For each set of options, assume all the other characteristics have the same impacts for all choices. This activity will involve making choices such as whether an increase in one dimension (ex: jobs) is worth an increase in another (ex: crime).

Economic:
1. Job opportunities
2. Income increase (for all local residents)
3. Increase in housing costs
4. Labor shortage for other business

Environmental:
1. Noise pollution
2. Water pollution
3. Air pollution
4. Land pollution and subsidence

Social:
1. Population increase
2. Infrastructure improvement (transportation, education, human serves, internet)
3. Traffic increase
4. Crime increase

Management and other:
1. Permit approval decision making mechanism
2. Availability of independent and transparent information on potential impacts of mine
3. Mine buffer (Home distance from mine)
4. Mine life

This page contains a sample problem.

The following pages will contain questions regarding a hypothetical mining project coming to your town/city of residence. Each page will contain a table with several different options you can choose for this hypothetical mine. For example, look at this table:

Column Definitions:
Job Opportunities - New jobs created in the area as a result of the mining operation
Water Pollution - water pollution from mining operations
Permit Approval Decision Making Mechanism – The process government workers will use to approve the mine's permit application
Population Increase – How the population will grow each year for ten years after the mine opens

A new mine will be opened near Salt Lake. Carefully consider each of the following options. Suppose all other unknown conditions/characteristics are the same, which option would you choose?

	Job Opportunities	Water Pollution	Permit Approval Decision Making Mechanism	Population Increase
Option A	600 people employed directly by the mine	Same as similar mine in the area	Final decision by Government agency after significant public input	4% annually
Option B	300 people employed directly by the mine	Same as similar mine in the area	Final decision by Government agency after negotiating with local representatives	2% annually
Option C	900 people employed directly by the mine	Less than similar mine in the area	Final decision by Government agency after significant public input	6% annually

○ Option A
○ Option B
● Option C

So, for this example, you are asked to consider that a hypothetical mine is coming to your community. There are three possible options listed above that describe possible changes to your community when the mining project begins. For example, "Option 1" in the table above would offer 600 jobs for community, same water pollution as a similar mine in the area, the mine permit will approved by government agency after significant public input, and increase the population of the community by 4% annually. You are being asked to read through all four options and select which option you would prefer.

Please set aside a block of time (approximately 15 minutes) to complete this task without disruption or distraction. Please also take this survey in an area where you will not be distracted.

APPENDIX F:

VIDEO INTRODUCTION AND SURVEY

APPENDIX F: VIDEO INTRODUCTION AND SURVEY

1. INTRODUCTION

The whole survey including the video introduction, and discrete choice questions are available online at http://web.mst.edu/~kabp3/jem2015supplement.htm

BIBLIOGRAPHY

ABS. (2001). *Australian Bureau of Statistics. Census Data. Canberra: Australian Government.*

ABS. (2006). *Australian Bureau of Statistics. Census Data. Canberra: Australian Government.*

Allison, P. D. (2012). *Logistic Regression Using SAS: Theory and Application* (Second Edi.). SAS Institute Inc.

Anderson, T. W. (1961). On the distribution of the two-sample Cramer-von Mises critetion. *Annals of Mathematical Statisics*.

Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain "goodness of fit" criteria based on stochastic processes. *Annals of Mathematical Statisitcs*, *23*(2).

Anonymous. (1996). Blair raises the stake: Labour's leader says he wants to turn Britain into a "stakeholder economy". What might this mean? *The Economist Newspaper*.

Azapagic, A. (2004). Developing a framework for sustainable development indicators for the mining and minerals industry. *Journal of Cleaner Production*, *12*(6), 639–662.

B Crosby. (1992). *Stakeholder analysis: a vital tool for strategic managers*. Washington DC: USAID.

Bateman, I. J., Carson, R. T., Day, B., Hanemann, M., Hanley, N., Hett, T., … Swanson, J. (2002). *Economic Valuation with Stated Preference Techniques: A Manual.* Cheltenham: Edward Elgar, Ltd.

Bergmann, A., Hanley, N., & Wright, R. (2006). Valuing the attributes of renewable energy investments. *Energy Policy*, *34*(9), 1004–1014.

BHP. (2008). *Resourcing the future: Sustaiability report 2008*.

BHP Billiton. (2014). *Sustainability Reporting Navigator*.

Brown, H. S., de Jong, M., & Lessidrenska, T. (2009). The rise of the Global Reporting Initiative: a case of institutional entrepreneurship. *Environmental Politics*, *18*(2), 182–200. doi:10.1080/09644010802682551

Browne, A. L., Stehlik, D., & Buckley, A. (2011). Social licences to operate: For better not for worse; for richer not for poorer? The impacts of unplanned mining closure for "fence line" residential communities. *Local Environ*, *16*, 707–725.

Brugha, R., & Varvasovszky, Z. (2000). Stakeholder analysis : a review. *Health Policy and Planning*, *15*(3), 239–246.

Brundtland. (1987). *World Commission on Environment and Development (WCED)* (p. 43).

Bryson, J. (1995). *Strategic Planning for Public and Nonprofit Organizations*. San Francisco CA: Jossey- Bass.

Caussade, S., Ortúzar, J. D. D., Rizzi, L. I., & Hensher, D. a. (2005). Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B: Methodological*, *39*(7), 621–640. doi:10.1016/j.trb.2004.07.006

Cheney, G., & Christensen, L. T. (2001). Organisational communication: linkages between internal and external communication. In *The New Handbook of Organisational Communication: Advances in Theory, Research, and Methods* (pp. 231–69). Sage, Thousand Oaks, CA.

Clarkson MBE. (1995). A stakeholder framework for analyzing and evaluating corporate social performance. *Academy of Management Review*, *20*(1), 92–117.

Cook, R. D., & Nachtsheim, C. J. (1980). A Comparison of Algorithms for Constructing Exact D- optimal Designs. *Technometrics*, *22*, 315–324.

Daganzo, C. (1979). *Multinomial Probit: The Theory and Its Application to Demand Forecasting*. New York: Academic Press.

Davis, R., & Franks, D. M. (2011). The costs of conflict with local communities in the extractive industry. In *First international seminar on social responsibility in mininig* (pp. 1–13). Sanitiaco, CHILE.

Dechant, K., Altman, B., Downing, R. M., & Keeney, T. (1994). Environmental leadership: from compliance to competitive advantage. *Academy of Management Executive*, *8*(3), 7–28.

Dimitropoulos, A., & Kontoleon, A. (2009). Assessing the determinants of local acceptability of wind-farm investment: A choice experiment in the Greek Aegean Islands. *Energy Policy*, *37*(5), 1842–1854. doi:10.1016/j.enpol.2009.01.002

Dudka, S., & Adriano, D. C. (1997). Environmental Impacts of Metal Ore Mining and Processing: A Review. *Journal of Environment Quality*, *26*(3), 590. doi:10.2134/jeq1997.00472425002600030003x

Dunn, O. J. (1964). Multiple contrasts using rank sums. *Technometrics*, *6*, 241–252.

Earthworks and Oxfam America. (2004). *Dirty Metals: Mining , Communities and the Environment*.

Elkington, J. (1997). *Cannibals with Forks: the Triple Bottom Line of 21st Century Business*. Capstone Publishing, Oxford.

Elliott, A. C., & Hynan, L. S. (2011). A SAS macro implementation of a multiple comparison post hoc test for a Kruskal-Wallis analysis. *Computer Methods and Programs in Biomedicine*, *102*(1), 75–80.

EPA. (1998). *Principles of Environmental Impact Assessment* (p. 196).

Epstein, M. J., & Roy, M.-J. (2003). Making the business case for sustainability: linking social and environmental actions to financial performance. *Journal of Corporate Citizenship*, *9*, 79–96.

Ericsson, M., & Löf, F. (2011). *Overview of state ownership in the global minerals industry, Long term trends and future*.

Fedorov, V. (1972). *Theory of Optimal Experiments, translated*. (W. J. Studden & E. M. Klimko, Eds.). New York: Academic Press.

Fonseca, A. (2010). How Credible are Mining Corporations' Sustainability Reports? A Critical Analysis of External Assurance under the Requirements of the International Council on Mining and Metals. *Corporate Social Responsibility and Environmental Management*, *17*, 355–370.

Freeman, R. E. (1984). *Strategic Management: A Stakeholder Approach*. Prentice-Hall, Englewood Cliffs, NJ.

Freeman, R. E., & Gilbert, D. R. (1988). *Corporate Strategy and the Search for Ethics*. Prentice Hall, Englewood Cliffs, NJ.

Friedman, A. L., & Miles, S. (2001). Social responsible investment and corporate social and environmental reporting in the UK: an exploratory study. *British Accounting Review*, *33*(4), 523–48.

Gao, S. S., & Zhang, J. J. (2006). Stakeholder engagement, social auditing and corporate sustainability. *Business Process Management Journal*, *12*(6), 722–740. doi:10.1108/14637150610710891

Gibbons, D. (2006). Rank tests. In *Encyclopedia of Statistical Sciences*.

Goodwin, N. (2003). Five kinds of capital : useful concepts for sustainable. In *the American Association of Legal Scholars annual meeting*.

Grant, R. W. (1997). *Contemporary Strategy Analysis: Concepts, Techniques, Applications*. Blackwell Business Books, Oxford.

Gregory, R., & Keeney, R. (1994). Creating policy alternatives using stakeholder values. *Management Science*, *40*(8), 1035–1048.

Gunningham, N., & Sinclair, D. (2009). Regulation and the Role of Trust: Reflections from the Mining Industry. *Journal of Law and Society*, *36*(2), 167–194. doi:10.1111/j.1467-6478.2009.00462.x

Haaijer, M., Wedel, M., Vriens, M., & Wansbeek, T. (1998). Utility covariances and context effects in conjoint mnp models. *Marketing Science*, *17*, 236–252.

Hajkowicz, S. a., Heyenga, S., & Moffat, K. (2011). The relationship between mining and socio-economic well being in Australia's regions. *Resources Policy*, *36*(1), 30–38. doi:10.1016/j.resourpol.2010.08.007

Hamann, R. (2003). Mining companies ' role in sustainable development : the " why " and " how " of corporate social responsibility from a business perspective, *20*(2). doi:10.1080/0376835032000085910

Hamann, R., Patel, Z., & Pressend, M. (2002). Competing visions and conflicting strategies: a southern African perspective on the World Summit. *Environment*, *44*(6), 8–21.

Hausman, J., & D.Wise. (1978). A conditional probit model for qual- itative choice: Discrete decisions recognizing interdependence and heterogeneous preferences. *Econometrica*, *48*, 403–429.

Hedberg, C., & Malmborg, F. (2003). The global reporting initiative and corporate sustainability reporting in Swedish companies. *Corporate Social Responsibility and Environmental Management*, *10*, 153–164.

Hensher, D., Rose, J., & Greene, W. (2005). *Applied choice analysis: a primer* (pp. 1–742). Cambridge University.

Hill, C., & Jones, T. (1992). Stakeholder-agency theory. *Journal of Management Studies*, *29*(2), 131–154.

Hodge, R. A. (2014). Mining company performance and community conflict: moving beyond a seeming paradox. *Journal of Cleaner Production*.

Hoyos, D. (2010). The State of the Art of Environmental Valuation With Discrete Choice Experiments. *Ecological Economics*, *69*(8), 1595–1603.

ICMM. (2008). *Planning for integrated mine closure : Toolkit*. Retrieved from www.icmm.com/page/9568/planning-for-integrated-mine-closure-toolkit

ICMM. (2009). *Human rights in the mining & metals inductry: Handling and resolving local level concerns & grievances*. Retrieved from www.icmm.com

ICMM. (2010). *Good practice guide: Indigenous peoples and mining*. Retrieved from www.icmm.com/library/indigenouspeoplesguide

ICMM. (2012a). *Community development toolkit* (pp. 1–222). Retrieved from www.icmm.com

ICMM. (2012b). *Mining ' s contribution to sustainable development – an overview Mining ' s contribution to sustainable development*.

ICMM, ICRC, IFC, I. 2011. I. (2011). *Voluntary principles on security and human rights: Implementation guidance tools*. Retrieved from www.voluntaryprinciples.org/resources

IFC. (1998). *Doing better business through effective public consultation and disclosure* (pp. 1–79). Washington DC.

IFC. (2007). *Stakeholder engagement : A good practice handbook for companies doing business in emerging markets*. Washington DC.

IFC. (2009). *Good practice note: Addressing grievances from project-affected communities*. Washington DC. Retrieved from www.ifc.org/ifcext/sustainability.nsf/content/publications_resettlement

IFC. (2010a). *Guide to human rights impact assessment and management (HRIAM)*. Washington DC. Retrieved from www.ifc.org/ifcext/sustainability.nsf/Content/Publications_Handbook_HRIA

IFC. (2010b). *Strategic Community Investment: A good practice handbook for companies doing business in emerging markets*. Washington DC. Retrieved from www.ifc.org/ifcext/sustainability.nsf/content/publications_resettlement

ISEA. (1999). *ISEA AccountAbility 1000 (AA1000): Standard, Guidelines and Professional Qualification*. Institute of Social and Ethical AccountAbility, London.

Ivanova, G., & Rolfe, J. (2011). Assessing development options in mining communities using stated preference techniques. *Resources Policy*, *36*(3), 255–264. doi:10.1016/j.resourpol.2011.03.001

Ivanova, G., Rolfe, J., Lockie, S., & Timmer, V. (2007). Assessing social and economic impacts associated with changes in the coal mining industry in the Bowen Basin,

Queensland, Australia. *Management of Environmental Quality: An International Journal*, *18*(2), 211–228. doi:10.1108/14777830710725867

Jeffrey, L. (2006). Kendall's Tau II. In *Encyclopedia of Statistical Sciences*.

Jenkins, H., & Yakovleva, N. (2006). Corporate social responsibility in the mining industry: Exploring trends in social and environmental disclosure. *Journal of Cleaner Production*, *14*(3-4), 271–284.

Johnson, G., & Scholes, K. (1993). *Exploring Corporate Strategy: Text and Cases,*. Prentice-Hall International, Hemel Hempstead.

Kempe, J. (1983). Review of water pollution problems and control strategies in the South African mining industry. *Water Science & Technology*, *15*, 27–58. Retrieved from http://www.iwaponline.com/wst/01502/wst015020027.htm

Klein, B., Awuah-Offei, K., Abel, B., Schnell, D., Stancil, R., & Verslues, J. (2015). Optimal Number of Factors for Mining Community Consultation using Choice Experiments. *Journal of Environmental Management*, *Under revi*.

Koch, G., & Bhapkar, V. (2006). Chi-square test — I. In *Encyclopedia of Statistical Sciences*.

Kuhfeld, W. (2010). *Marketing research methods in SAS. ... Graphical Techniques. Cary, NC, SAS-Institute TS-722* (SAS 9.2., pp. 1–1309). SAS Institute Inc.

Lockie, S., Franetovich, M., Sharma, S., & Rolfe, J. (2008). Democratisation versus engagement? Social and economic impact assessment and community participation in the coal mining industry of the Bowen Basin, Australia. *Impact Assessment and Project Appraisal*, *26*(3), 177–187. doi:10.3152/146155108X357257

Lockie, S., Franettovich, M., Petkova-Timmer, V., Rolfe, J., & Ivanova, G. (2009). Coal mining and the resource community cycle: A longitudinal assessment of the social impacts of the Coppabella coal mine. *Environmental Impact Assessment Review*, *29*(5), 330–339. doi:10.1016/j.eiar.2009.01.008

Louviere, J. J., Hensher, D. A., & Swait, J. D. (2003). *Stated Choice Methods Analysis and applications* (pp. 1–418). Cambridge University Press.

M.G.Kendall, & A.Stuart. (1973). *The Advanced Theory of Statistics, Volume 2: Inference and Relationship* (p. Sections 31.19, 31.21). Griffin.

Marschak, J. (1959). Binary choice contraints on random utility indications. *Mathematical Methods in the Social Sciences, Stanford Univeristy Press*, 312–329.

Mason, C., Paxton, G., Parr, J., & Boughen, N. (2010). Charting the territory: Exploring stakeholder reactions to the prospect of seafloor exploration and mining in Australia. *Marine Policy*, *34*(6), 1374–1380. doi:10.1016/j.marpol.2010.06.012

Mason, R., & Mitroff, I. (1981). *Challenging strategic planning assumptions*. New York: John Wiley & Sons.

Mathews, M. R. (1997). Twenty-five years of social and environmental accounting research: is there a silver jubilee to celebrate? *Accounting, Auditing & Accountability Journal*, *10*(4), 481–531.

McFadden, D. (1974). Coditional logit analysis of qualitative choice behavior. In *Frontiers in Econometrics* (P. Zarembk., pp. 105–142). New York: Academic Press.

McFadden, D. (1974). *Conditional Logit Analysis of Qualitative Choice Behavior* (pp. 105–142). New York: Frontiers in Econometrics.

Mcfadden, D., & Train, K. (2000). Mixed MNL models for discrete response. *Joural of Applied Econometrics*, *15*(5), 447–470.

MCMPR. (2005). *Principles for Engagement with communities and stakeholders*.

Meadowcraft, J. (2000). Sustainable development: a new(ish) idea for a new century? *Political Studies*, *48*, 370–87.

Meyer, R. K., & Nachtsheim, C. J. (1995). The Coordinate-Exchange Algorithm for Constructing Exact Optimal Experimental Designs. *Technometrics*, *37*(1), 60–69.

Midford, R., Marsden, A., & Phillips, M. (1997). Workforce alcohol consumption patterns at two Pilbara mining-related worksites. *J. Occup. Health Saf.*, *13*(3), 267–274.

Moffat, K., & Zhang, A. (2014). The paths to social licence to operate: An integrative model explaining community acceptance of mining. *Resources Policy*, *39*, 61–70. doi:10.1016/j.resourpol.2013.11.003

MPE. (2011). Community engagement: Not in my backyard. *Mining, People and the Environment*, (July), 26–27.

Munashinge, M. and Shearer, W. (1995). *Defining and Measuring Sustainability* (pp. 1–482). The World Bank, Washington.

Muradian, R., Martinez-Alier, J., & Correa, H. (2003). International capital versus local population: The environmental conflict of the Tambogrande mining project, Peru. *Society &Natural …*, *16*(November 2001), 775–792. doi:10.1080/08941920390227176

National Mining Association. (2014). *The Economic Contributions of U.S. Mining*.

O'Faircheallaigh, C. (2012). Community Development Agreements in the Mining Industry: An Emerging Global Phenomena. *Community Development*. Retrieved from Available at <http://dx.doi.org/10.1080/15575330.2012.705872>

Petkova, V., Lockie, S., Rolfe, J., & Ivanova, G. (2009). Mining developments and social impacts on communities: Bowen Basin case studies. *Rural Society*, *19*(3), 211–228. Retrieved from http://pubs.e-contentmanagement.com/doi/abs/10.5172/rsj.19.3.211

PewResearchCenter. (2015). U.S. Survey Research: Collecting survey data. Retrieved from http://www.pewresearch.org/methodology/u-s-survey-research/collecting-survey-data/

Pirie, W. (2006). Spearman rank correlation coefficient. In *Encyclopedia of Statistical Sciences*.

Porter, M. E., & Kramer, M. R. (2011). Creating shared value. *Harvard Business Review*, *89*(1), 62–77.

Prno, J., & Scott Slocombe, D. (2012). Exploring the origins of "social license to operate" in the mining sector: Perspectives from governance and sustainability theories. *Resources Policy*, *37*(3), 346–357. doi:10.1016/j.resourpol.2012.04.002

Que, S., & Awuah-Offei, K. (2014). Framework for mining community consultation based on discrete choice theory. *Int. J. Mining and Mineral Engineering*, *5*(1), 59–74.

Que, S., Awuah-Offei, K., & Samaranayake, V. A. (2015). Classifying critical factors that influence community acceptance of mining projects for discrete choice experiments in the United States. *Journal of Cleaner Production*, *87*(15), 489–500. doi:http://dx.doi.org/10.1016/j.jclepro.2014.09.084

Reed, M., Graves, A., Dandy, N., Posthumus, H., Hubacek, K., J, M., … Stringer, L. C. (2009). Who's in and why? A typology of stakeholder analysis methods for natural resource management. *Journal of Environmental Management*, *90*(5), 1933–1949.

Revelt, D., & Train, K. (1998). Mixed Logit with Repeated Choices: Households' Choices of Appliance Efficiency Level. *Review of Economics and Statistics*, *80*(4), 647–657. doi:10.1162/003465398557735

Rio Tinto. (2012). *Kennecott Utah Copper Sustainable Development Report*.

Robert E. Odeh, M. F. (1975). *Sample size choice: charts for experiments with linear models*. New York: Marcel Dekker.

Rockwell, B. W. (2000). *The Goldfield Mining District, Nevada:An Acid-Sufate Bonanza Gold Deposit* (pp. 1–22). Denver, Colorado.

Rodriguez, R. (2006). Correlation. In *Encyclopedia of Statistical Sciences*.

Rotheroe, N., Keenlyside, M., & Coates, L. (2003). Local agenda 21; articulating the meaning of sustainable development at the level of the individual enterprise. *Journal of Cleaner Production*, *11*(5), 537–548. doi:10.1016/S0959-6526(02)00075-6

Rowe, J., & Enticott, R. (1998). The role of local authorities in improving the environmental management of SMEs: some observations from partnership programmes in the West of England. *Journal of Eco-Management Auditing*, *5*(75-87).

Russo, M., & Fouts, P. (1997). A resource-based perspective on corporate environmental performance and profitability. *Academy of Management Journal*, *40*, 534–59.

Ryan, M., & Gerard, K. (2003). Using discrete choice experiments to value health care: current practice and future prospects. *Applied Health Economics and Policy Analysis*, *2*, 55–64.

SAS. (2007a). Analysis of Variance: Nonparametric One-Way Analysis of Variance.

SAS. (2007b). Exact Statistics. Retrieved from http://support.sas.com/onlinedoc/913/docMainpage.jsp

SAS. (2007c). Frequency Tables and Statistics.

SAS. (2007d). The % ChoicEff Macro.

SAS. (2007e). The % MktDups Macro.

SAS. (2007f). The % MktEx Macro.

SAS. (2007g). The % MktLab Macro.

SAS. (2007h). The % MktRuns Macro.

SAS. (2007i). The CORR Procedure: Getting Started.

SAS. (2007j). The CORR Procedure: Overview. Retrieved from http://support.sas.com/onlinedoc/913/docMainpage.jsp

SAS. (2007k). The NPAR1WAY Procedure Example 52.2: The Exact Wilcoxon Two-Sample Test.

SAS. (2007l). The UNIVARIATE Procedure Example 3.10: Computing Confidence Limits for Quantiles and Percentiles.

SAS. (2007m). The UNIVARIATE Procedure Example 3.12: Testing for Location.

SAS. (2014). The LOGISTIC Procedure:STRATA Statement.

Sassoon, M. (1998). Social Benefits and Costs: Workshop Report. In G. McMahon (Ed.), *Mining and the Comrmunity: Results of the Quito Conference*. The World Bank.

Schaefer, A. (2004). Corporate sustainability – integrating environmental and social concerns? *Corporate Social-Responsibility and Environmental Management*, *11*(4), 179–87.

Schlotzhauer, S. D. (2009). *Elementary Statistics Using SAS*.

Schooten, M., Vanclay, F., & Slootweg, R. (2003). Conceptualising socail change processes and social impacts. In E. Elgar (Ed.), *The international handbook of social impact assessment: conceptual and methodological advances* (pp. 74–91). cheltenham: Becher and Vanclay (eds).

Scott, A. (2002). Identifying and analysing dominant preferences in discrete choice experiments: an application in health care. *Journal of Economic Psychology*, *23*, 383–398.

Shapiro, S. S., Wilk, M. B., & Chen, H. J. (1968). A Comparative Study of Various Tests for Normality, *63*(324), 1343–1372.

Shrivastava, P. (1995). The role of corporations in achieving ecological sustainability. *Academy of Management Review*, *20*, 936–60.

Solley, W. B., Pierce, R. H., & Perlman, H. A. (1999). *Estimated Use of Water in the United States in 1995*. Denver, CO.

StatWizards LLC. (n.d.). *Discrete choice models*.

Stokes, M., Davis, C., & Koch, G. (2012). *Categorical Data Analysis Using SAS* (Third., pp. 1–620). SAS Institute Inc.

Stuart, A. (2006). Kendall's Tau I. In *Encyclopedia of Statistical Sciences*.

Taylor, C. (2013, June). Margin of Error Formula. Retrieved from http://statistics.about.com/od/Formulas/ss/Margin-Of-Error-Formula.htm

Tey, Y. S., Li, E., Bruwer, J., Abdullah, A. M., Cummins, J., Ismail, M. M., & Darham, S. (2012). Qualitative Methods for Effective Agrarian Surveys: A Research Note on Focus Groups. *Journal of Sustainable Agriculture*, *6*(1), 60–65.

Thomas, P., & Palfrey, C. (1996). Evaluation: stakeholder-focused criteria. *Social Policy and Administration*, *30*(2), 125–142.

Thompson, P. (1996). Editorial commentary: stakeholding as state strategy. *Renewal*, *4*(2), 3–11.

Thomson, I., & Boutilier, R. G. (2011). Social license to operate. In *SME Mining Engineering Handbook* (pp. 1779 – 1796). P.Darling Ed.

Thurstone, L. (1927). A law of comparative judgement. *Psychological Review*, *34*, 273–286.

Train, K. (2002). *Discrete Choice Methods with Simulation*. Cambridge University.

Walekhwa, P. N., Mugisha, J., & Drake, L. (2009). Biogas energy from family-sized digesters in Uganda: Critical factors and policy implications. *Energy Policy*, *37*(7), 2754–2762. doi:10.1016/j.enpol.2009.03.018

Walt, G. (1994). *Can interest groups influence government policy? Health policy: an introduction to process and power*. London: Zed Publications.

Willis, A. (2003). The Role of the Global Reporting Initiative ' s Sustainability Reporting Guidelines in the Social Screening of Investments. *Journal of Business Ethics*, *43*, 233–237.

Willis, K., Scarpa, R., Gilroy, R., & Hamza, N. (2011). Renewable energy adoption in an ageing population: Heterogeneity in preferences for micro-generation technology adoption. *Energy Policy*, *39*(10), 6021–6029. doi:10.1016/j.enpol.2011.06.066

Winslott Hiselius, L. (2005). Using choice experiments to assess people's preferences for railway transports of hazardous materials. *Risk Analysis : An Official Publication of the Society for Risk Analysis*, *25*(5), 1199–214. doi:10.1111/j.1539-6924.2005.00660.x

Witt, J., Scott, A., & Osborne, R. H. (2009). Designing choice experiments with many attributes. An application to setting priorities for orthopaedic waiting lists. *Health Economics*, *18*, 681–696. doi:10.1002/hec

Wood, D. J. (1991). Cororate social performance revisited. *Academy of Management Review*, *16*(4), 691–718.

Wood, D. J. (2010). Measuring Corporate Social Performance: A Review. *International Journal of Management Reviews*, *12*(1), 50–84. doi:10.1111/j.1468-2370.2009.00274.x

Zar, J. H. (2010). *Biostatistical Analysis* (Fifth edit.). Prentice Hall.

Zhou, X. H., Gao, S., & Hui, S. L. (1997). Methods for comparing the means of two independent log-normal samples. *Biometrics*, *53*(3), 1129–35.

VITA

Sisi Que, was born in Hubei, China in 1984. She went on to Wuhan University China, where she studied Water Engineering and received her Bachelor of Engineering in June 2008. In 2011, she started graduate studies at New Mexico State University at Las Cruces, New Mexico, where she studied Environmental Engineering and obtained an MS degree in August 2011. She came to Missouri University of Science & Technology in 2012 to pursue her PhD in Mining Engineering under the guidance of Dr. Kwame Awuah-Offei.