# Incidence and Costs of Pinhole Leak Corrosion and Corporate Cost of Capital Borrowing

Ewa Jadwiga Kleczyk

Dissertation submitted to the faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirement for the degree of

DOCTOR OF PHILOSOPHY
in
Agricultural and Applied Economics

Dr. Darrell Bosch, Co-Chairman
Dr. Kevin Boyle, Department Chair, Co-Chairman
Dr. Daniel Taylor
Dr. Andrea Dietrich
Dr. Christopher Parmeter

December 1, 2008
Blacksburg, Virginia

Keywords: Corrosion Detection and Control, Home Plumbing System, Drinking Water, Preference Elicitation Methodology (Conjoint Analysis), Debt Financing, Excess Bond Returns

Incidence and Costs of Pinhole Leak Corrosion and Corporate Cost of Capital Borrowing

Ewa Jadwiga Kleczyk

ABSTRACT

The first part of this doctorate dissertation examines the factors influencing the occurrence and costs of pinhole leak corrosion as well as the household decisions for corrosion prevention and plumbing material selection. Three mail surveys of households were used to elicit the experiences with leaks as well as the optimal corrosion prevention and material choices. Probability modeling (i.e. MNL) and linear regression analysis were used to analyze survey responses. Pinhole leak occurrences were found associated with pipe type installed, property age, pipe failure history, and dwelling distance from a water treatment plant. The number and location of pinhole leaks in the dwelling and the pipe type are associated with the financial costs of pinhole leaks. The corrosion prevention choices as well as the plumbing materials depended on the risk of corrosion and cost associated with each option. Previous experiences with pinhole leak impacted the decision for household choices. Faster responses to pinhole leak outbreaks by utility managers and policymakers in terms of advising homeowners on the best ways of responding to leaks would assist homeowners in reducing costs of pinhole leak repairs and associated damages.

The second part of this document deals with the debt financing issues. Debt financing decisions are made simultaneously by lenders and borrowers. Since lenders are unable to observe directly the firms' investment decisions, the banks offer contracts based up on firms' observable characteristics (i.e. wealth and size) and the prevailing market conditions. When deciding on the financing decisions, firms also take into account the changes in macroeconomic variables in order to lower the cost of borrowing. As a result, the goal for this article is to examine empirically the hypothesis of the effect of the debt determinant as well as the macroeconomic variables on the debt maturity structure. A reduced form of the simultaneous financing decisions model is estimated by employing several OLS estimation methods. The empirical findings offer strong support for firms with few growth options, large, and of low quality having more long-term debt in their capital structure. There was, however, no clear support for the impact of macroeconomic variables on debt maturity, as some variables were not statistically significant.

***This Doctorate Dissertation is dedicated to the memory of***
***Dr. G. V. Loganathan***

## ACKNOWLEDGMENTS

I am indebted to the assistance of my committee during this dissertation process: Dr. Darrell Bosch, Dr. Kevin Boyle, Dr. Daniel Taylor, Dr. Andrea Dietrich, and Dr. Christopher Parmeter. Dr. Bosch, I could not ask for a better advisor to lead me throughout this doctorate degree. You motivated me, supported me, pushed me, and believed in me even when everybody else doubted me. Without your support, I would not be able to finish this degree. Thank you from the bottom of my heart. Dr. Boyle, I would like to thank you for all you have done for me during my education at the University of Maine as well as Virginia Tech. Your leadership and guidance has been very important in shaping my education and career. Dr. Taylor, Dr. Dietrich, and Dr. Parmeter, thank you for your guidance during my doctorate studies. Finally, I would like to thank late Dr. G.V. Loganathan for the possibility to work with him. I miss greatly his mentorship. Thank you for teaching me so much during the couple of years that I knew you.

I would like to thank my parents and my brother. Without their constant support, I would not be able to obtain this doctorate degree. I love you very much and I hope I made you and Poland proud! (In Polish: Dziekuje za Wasze wsparcie podczas studiow doktoranckich na Virginia Tech. Mam nadzieje ze jestescie ze mnie dumni). I would like to thank Jackie and David Nutter for making me part of their family. Special thanks to Magdalena and Jaroslaw Kanscy as well as several other Polish Families from Blacksburg, Va. Thank you for your friendship and love (In Polish: Serdecznie dziekuje z calego serca).

I would like to thank Graduate Students in the Agricultural and Applied Economics and Economics Departments. Especially, I would like to thank Monica Licher, Sonia Aziz, Elton Mykarezi, Jaideep Chowdory, Jason Maupin, and Eftila Tanellari. Thank you for your support and belief in me ☺ I wish you best in your professorial paths. I would like also to thank Juneseok Lee of the Civil and Environmental Engineering Department. Finally, to all of the people I met at Virginia Tech…thank you for everything.

Big thanks to everyone at TargetRx, Inc. Especially, I would like to thank Monika Hodek, Patrick Howie, and Kathy Arber for their support and friendship during my doctorate education. I would like to thank the company for financially supporting my education. Final thanks go to Allison Miller, Ph.D. for her support and guidance during this thesis writing experience. Thank you for the daily ‘push’ and ‘encouragement’ words during these past few months.

## TABLE OF CONTENTS

## LIST OF TABLES

## LIST OF FIGURES

# CHAPTER 1

## 1.1 Introduction

### 1.1.1 Socioeconomic Impact of Pinhole Leak Corrosion on Households

Pinhole leaks are small holes in drinking water plumbing caused by pitting corrosion, a type of corrosion concentrated on a small area of the pipe. In most cases, very small pinhole leaks are hard to detect, especially if they appear in pipes running through walls and ceilings. Generally, it is recommended that pipes be replaced after three or four leaks (Gurner, 2003). Damage from pinhole leaks can include collapse of walls and ceilings and water can contribute to growth of mold on the surface of walls, floors, and ceilings. Mold exposure can cause allergic reactions, such as irritation of eyes, skin, and throat. In addition, costs of pinhole leaks include costs of fixing the plumbing system, repairing damage to property, and loss of revenue while plumbing systems are being repaired. Property damage may include mold damage, which led to an $8.9 billion dollar loss by the residential sector of insurance companies in 2001 (Mann, 2002).

While pinhole leaks are a nationwide problem, they are more common in certain regions of selected states including Maryland, California, Florida, and Ohio. In Maryland, the governor appointed a task force to examine the issue (Maryland Task Force Study, 2004). Some areas (for example, parts of Florida) have considered banning the use of copper piping in order to control the rising number of incidents (Gurner, 2003).

Due to the increased number of pinhole leaks and other types of pipe failures, it is important to learn the decision making process of households with regards to selecting a plumbing material when faced with plumbing material replacement. Consumers' decisions on plumbing material selection are dictated by various factors, including state and federal regulations, service providers, and preferences of individual households. In addition, insurance companies collect and utilize information on water damage and plumbing material failures in the decision-making process regarding new policy issuance and renewal of policies involving private properties. As a result, homeowners have an important stake in finding plumbing systems appropriate for their households, and therefore, they should acquire information on plumbing material attributes such as price and health impact.

Although the pinhole leak problem has been extensively studied during recent years, there have been no studies examining the impact of pinhole leak corrosion on individual households. When corrosion occurs, households have to deal with problems resulting from leak damage, including plumbing system repairs and the high financial, time, and emotional costs of dealing with the issue. There is also limited knowledge about the preferences of individual households for alternative plumbing materials and the driving forces of the decision-making process when the household is faced with the replacement of a home drinking water system. As a result, two articles are proposed to investigate the socioeconomic impact of pinhole leak corrosion on individual households and households' preferences for alternative plumbing material and corrosion prevention based on their past experiences. The first study investigates the rate of pinhole leak occurrences and the total costs related to the repairs needed after to pipe failure in the Maryland suburbs of Washington, D.C. The second study investigates the preferences for plumbing materials and preventive measures against corrosion utilizing stated preference Conjoint Analysis method.

### 1.1.2 Cost of Capital Borrowing in Corporate World

There has been significant discussion about the corporate management of financial policy to minimize the cost of capital borrowing. According to Stiglitz (1974), in efficient and integrated markets, the optimal financial management policy cannot decrease the costs of capital due to the interrelation between the different types of capital costs. Consequently, there is no gain from substituting debt for equity. Evidence, however, has been found against this claim, demonstrating that equity financing is related to the predictability of stock returns (Baker et al., 2003). Debt maturity has been found also to relate to market conditions, such as the interest and inflation rates, term-spread and dividend-yield (Bosworth, 1971; Barclay & Smith, 1995). Although plenty of articles provide support for the association between debt and returns, no studies have examined the cost of capital borrowing at different times of debt maturity.

The relationship between debt maturity and cost of capital borrowing is examined by Baker et al. (2003). In their paper, they analyze the variation in the maturity of debt due to debt market conditions (inflation and interest rates) and predict excess bond returns. They utilize aggregate annual time-series data and find a close relation between debt maturity and predictable variation in the excess bond returns. They are, however, unable to conclude whether firms actually reduce the overall cost of capital borrowing from timing the market because of difficulties in interpreting the regression results (Baker et al., 2003).

The third article extends the empirical analysis performed by Baker et al. (2003) to the pooled time- and cross-sectional data in order to examine the relationship between the excess bond returns and the corporate debt maturity structure. In addition, the analysis extends previous studies such as that by Barclay and Smith (1995) by incorporating the bond returns information with the corporate debt research. As a result, the study accounts for the relationship between debt maturity and the cost of borrowing at different maturities.

### 1.1.3 Common Theme

Although the three proposed articles' subjects are very different from each other, the common theme between them is the closing of the literature gap with regards to the subject matter and applied empirical methodology. For example, there is a large literature gap in the study of the socioeconomic impact of corrosion that includes the extent of pinhole leak occurrences, financial costs associated with repairs of pipe failures, and household preferences for home plumbing materials and corrosion prevention options. In the case of cost of capital borrowing, previously published articles did not attempt to investigate the impact of bonds and macroeconomic variables on the debt issuance at different maturity times. Additionally, each of the articles employs and builds on several quantitative techniques for empirical analysis of the underlying data. For example, as the finance area has not previously utilized individual firm panel data for studying the cost of capital borrowing, the individual level analysis is employed to investigate the determinants of long and short-term debt. On the other hand, CA study employs the Multinomial and Ordered Logit in analysis of household preferences for plumbing materials and corrosion preventive options. Finally, both of the corrosion articles required development and distribution of survey instruments, adding to the richness of the research.

### 1.2 Article #1: "Incidence and Costs of Home Plumbing Corrosion: An Investigation of Sample Selection Bias"

Since few studies have examined the rate of pinhole leak occurrence and the socioeconomic impacts of corrosion on individual households, this article addresses the literature gap by analyzing factors influencing the leak occurrence rate and costs of pinhole leak corrosion for individual households. Knowing the rate of pinhole leaks as well as the factors affecting their occurrence might help identify the residential areas most vulnerable for pinhole leak corrosion. For example, studying location of private residences relative to the treatment plants revealed that homes closer to the water treatment plant have a higher degree of leak incidents (Lee et al., forthcoming). Additionally, knowing the rate at which pinhole leaks occur helps determine the optimal time for plumbing system replacement. For example, the optimal replacement time for a plumbing material is predicted to be about 22 years after installation (Loganathan & Lee, 2005).

Knowing the total financial costs associated with damage repairs necessitated by pinhole leak corrosion as well as the factors driving them provide an additional insight into the corrosion impact on individual households. The financial costs for fixing leaking plumbing and the associated damage to ceilings, walls, and floors ranges between $500 and $25,000 (Kleczyk et al., 2005). The cost structure and the rate of pinhole leak occurrences may lead homeowners to evaluating an appropriate plumbing material for their residences. The decision for choosing best plumbing material is complicated and depends on evaluation of product attributes such as cost (material cost plus labor and installation cost) and health effects. Information on factors associated with pinhole leak incidence and costs can be useful to homeowners, policymakers, water utilities, housing authorities, and others in formulating policies and programs to reduce pinhole leak occurrences and the negative financial and emotional impacts of corrosion on individual households.

A mail survey of residential water users in the Maryland suburbs of Washington D.C. was conducted in 2004 to determine the prevalence of pinhole leaks in residences. This area was selected due to the large number of pinhole leaks reported by utility customers (Washington Suburban Sanitary Commission, no date). A total of 5,009 Maryland residents received the survey, and 1,128 responses were returned of which 1,120 responses were used in the analysis.

To analyze the causes of pinhole leak corrosion as well as the costs related to the pinhole leak corrosion, econometric models including a Binary Logit model and an Ordinary Least Square (OLS) regression corrected for a sample selection bias are employed. Results from the survey indicate that pinhole leaks in home plumbing systems are a frequently reported problem in the greater Washington, D.C. area. Based on the results of statistical analysis, pinhole leak occurrences depend on several factors, including the type of plumbing material and location of the house relative to the treatment plant. Pinhole leaks and the damage associated with their repairs can have high financial costs, even up to $18,000. The number of leaks and location of leaks in the dwelling impact repair costs. Accounting for the sample selection bias affects the level and statistical significance of variables related to cost of repairing pinhole leak damages.

### 1.3 Article #2: "Conjoint Analysis Study of Households' Preferences for Plumbing Materials"

Consumers' decisions on plumbing material selection are dictated by various factors, including state and federal regulations, service providers, and preferences of individual households. For example, general contractors are the primary decision-makers of plumbing material installation in new houses; while utility companies respond to corrosion threads by adding corrosion inhibitors into drinking water treatment. Consequently, all service providers influence consumer decisions regarding the best plumbing material for residential use.

Since homeowners have an important stake in finding plumbing systems appropriate for their households, they should not only rely on the expert advice but also acquire information on plumbing material attributes such as price, health impact, longevity, and corrosion resistance in order to make informed investment decisions about plumbing systems for their homes.  For example, health effects, longevity, taste and odor have been found to be the most important factors in consumers' evaluations of plumbing material for home use (Lee et al., 2005).  Additionally, households are willing to pay up to $4,000 when guaranteed a leak-free plumbing system for 50 years (Kleczyk et al., 2006).  Random utility models are usually adopted in preference elicitation and maximization (Champ et al., 2003). Information on consumer preferences for drinking water plumbing attributes can be useful not only to individual households but also to policymakers, program managers, water utilities, and firms with interests in drinking water infrastructure.

There are two types of commonly used preference elicitation methods that could be utilized in evaluation of households' preferences for drinking water infrastructure: revealed and stated preferences.  The main difference between these two methods is that stated preference approaches rely mainly on primary survey data while revealed preference approaches rely mostly on secondary data and draw their conclusions based on the actual decisions made by people (Champ et al., 2003). Many researchers tend to combine revealed preference and stated preference data as this approach allows for an extension of the behavioral analysis beyond the limited range of historical experience (Whitehead et al., 2005).

Economists utilize stated preference methods such as Conjoint Analysis (CA) and Contingent Valuation (CV) for non-market valuation analysis, including willingness-to-pay estimation. All of these methods are derived from the Random Utility Framework, which is based on the premises of preference ordering (Champ et al., 2003). On the other hand, Analytic Hierarchy Process (AHP) is widely employed by the Engineering field for determination of the most optimal attributes for given products.

For the purpose of this study, the stated preference framework method, CA, is employed to address the issues related to the drinking water infrastructure and to elicit information on home plumbing materials and the tradeoffs between risk of corrosion and financial spending for preventive options, two surveys of residents residing in a Southeastern Community were conducted. The first survey elicited information on the prevalence of pinhole leaks and other plumbing material failures, households' experiences with plumbing material failures, the cost of repairs and property damages due to the material failures, and household preferences for corrosion preventive measures. The follow-up survey of the Southeastern Community, sent only to those residents who agreed to participate in future studies, related to the plumbing material issues and elicited information on households' preferences for a set of hypothetical plumbing materials.

As it is important to learn household perceptions and preferences for drinking water infrastructure, the study objective is to investigate homeowners' preferences for plumbing materials and preventive techniques against corrosion based on households' experiences with plumbing material failures. The study employed CA questions in preference elicitation for corrosion prevention and plumbing materials. The data are analyzed by employing several regression methods including Multinomial Logit and Ordered Logit models. The model specifications include the attributes describing corrosion prevention alternatives and plumbing material as well as socioeconomic characteristics of the surveyed households. For example, the risk of corrosion was found to have no statistically significant impact on household decisions while a plumbing material described by 'no corrosion' and 'convenient installation' attribute level was found the most preferred in the set of alternative plumbing systems. Based on the regression results, the willingness to pay and compensating variation is computed to assess the level of financial impact the households are willing to allocate to for the best fitting home plumbing alternative.

### 1.4 Article # 3: "The Determinants of Corporate Debt Maturity Structure"

There has been a significant discussion about the corporate management of financial policy in order to minimize the cost of capital borrowing. According to Stiglitz (1974), in efficient and integrated markets, the optimal financial management policy cannot decrease the costs of capital due to the interrelation between the different types of capital costs. Consequently, there is no gain from substituting debt for equity. Evidence has been found, however, against this claim, demonstrating that equity financing is related to the predictability of stock returns as well as market variables such as inflation and interest rates (Bosworth, 1971; Barclay & Smith, 1995). Although the above papers provide support for the association between debt and returns, none offers information on the cost of capital borrowing at different times of debt maturity due to lack of analysis of returns data.

The relationship between debt maturity and cost of capital borrowing is examined by Baker et al. (2003). In their paper, they analyze the variation in the maturity of debt because of debt market conditions (inflation and interest rates) and predict excess bond returns. They utilize aggregate annual time-series data and find a close relation between debt maturity and predictable variation in the excess bond returns. According to their findings, firms issue long-term debt when bond returns are low (bond prices are high) and short-term debt when returns are high (bond prices are low). They are, however, unable to conclude whether firms actually reduce the overall cost of capital borrowing from timing the market due to difficulties in interpreting the predictions of the regression results (Baker et al., 2003).

This paper extends the empirical analysis performed by Baker et al. (2003) to the pooled time- and cross-sectional data in order to examine the relationship between the excess bond returns and the corporate debt maturity structure. In addition, the analysis extends previous studies such as that of Barclay and Smith (1995) by incorporating the bond returns information with the corporate debt maturity structure research. As a result, the study accounts for the relationship between debt maturity and the cost of borrowing at different maturities.

The goal of this paper is to examine the hypothesis of the effect of the excess bond returns on the debt maturity structure at the individual firm level data. It is predicted that firms issue long-term debt when bond returns are low. The results are expected to support the debt market timing theory as found by Baker et al. (2003). The notion of the corporate debt maturity structure is explained by the determinants of debt structure as well as debt market conditions. It is predicted that small and unregulated firms with more investment growth options have less long-term debt and managers use debt maturity to signal important information to the market. Furthermore, high and low risk quality firms issue more long-term debt compared to intermediate risk quality firms. Additionally, it is expected that the macroeconomic variables will be negatively related to the maturity of debt.

The research objective is achieved by investigating the relationship between the excess bond return as well as the determinants of the firm's debt and the corporate debt maturity structure. Several panel data estimation methods are employed to evaluate the relationship between debt maturity and its determinants including fixed and random effects regression and GLS regression (Wooldridge, 2002). The results obtained from this study add to the undeveloped literature of debt financing by extending the analysis of the impact of debt determinants on debt maturity by analyzing returns data and providing evidence for relationship between debt maturity and cost of borrowing at different maturities.

**1.5 Conclusions**

Although the three proposed articles' subjects are very different from each other, the common theme between the articles is the closing of the literature gap with regards to the subject matter: the socioeconomic impact of pinhole leak corrosion on individual households and excess bond returns impact on debt financing; and applied empirical methodology: utilization of pooled time- and cross-section data for cost of borrowing analysis and preference elicitation for plumbing material and prevention options. All three papers are innovative in the utilization of econometric techniques by utilizing cross-sectional and panel data estimation techniques.

**1.6 References**

Boyle, K. 2007. Personal Communication.

Baker, M., Greenwood, R., & Wurgler, J. (2003). The Maturity of Debt Issues and Predictable Variation in Bond Returns. *Journal of Financial Economics*: 261-291.

Baker, M., Ruback, R. S., & Wurgler, J. (2006). Behavioral Corporate Finance. In E. Ebcko, (Ed.), Handbook in Corporate Finance: Empirical Corporate Finance. North Holland: Elsevier.

Baker, M., Stein, S., & Wurgler, J. (2003). When Does the Market Matter? Stock Prices and Investment of Equity-Dependent Firms. *Quarterly Journal of Economics 118,* 969-1006.

Baker, M., & Wurgler, J. (2002). Market Timing and Capital Structure. *Journal of Finance 57*, 271-372.

Barclay, M. J., & Smith, Jr., C. W. (1995). The Maturity Structure of Corporate Debt. *Journal of Finance 50*, 609-631.

Bosworth, B. (1971). Patterns of Corporate External Financing. *Brookings Papers on Economic Activity 2*, 253–279.

Champ, P., Boyle, K., & Brown, T. (2003). *A Primer on Nonmarket Valuation*. Place: Kluwer Academic Publishers.

Edwards, M., J., Rushing, C., Kvech, S., & Reiber, S. (2004). Assessing Copper Pinhole Leaks in Residential Plumbing. *Water Science and Technology 49*(2), 83–90.

Dietrich, M., Edwards, M. A., Bosch, D. J., Duncan, S., Dwyer, S., & Loganathan, G. V. (2003). Toward *Sustainable Materials for Drinking Water Infrastructure*. Project Proposal funded by NSF. Blacksburg, Virginia: Virginia Tech.

Helm, R., Scholl, A., Manthey, L., & Steiner, M. (2004). Measuring Customer Preferences in New Product Development: Comparing Compositional and Decompositional Methods. *International Journal of Product Development 1*(*1*), 12-29.

Kleczyk, E. J. (2006). The Excess Bond Returns and Debt Maturity. *6th Global Conference on Business and Economics.* October 10-12, 2006. Harvard University, Boston, Massachusetts. Available online at www.trintrin.com/gebc/The%20Excess%20Bond%20Returns%20and%20Debt%20Maturity.doc. Accessed: October 2008.

Kleczyk, E. J., Bosch, D., Dwyer, S., Lee, J., & Loganathan, G. V. (2005). Maryland Home Drinking Water Assessment. *Virginia Water Resources Research Center, National Water Research Symposium: Balancing Water Law and Science*, October 10-12, 2005, The Inn at Virginia Tech and Skelton Conference Center, Virginia Tech, Blacksburg, Virginia.

Kleczyk, E. J., Tanellari, E., & Bosch, D. (2006). Corrosion in Home Drinking Water Infrastructure: Assessment of Causal Factors, Costs, and Willingness to Pay. *Water Quality Technology Conference and Exposition, American Water Works Association*, November 5-9, 2006, Denver, Colorado.

Lee, J., & Loganathan, G. V. (2005). Preference Analysis of Home Plumbing Material. *Virginia Water Resources Research Center, National Water Research Symposium: Balancing Water Law and Science*, October 10-12, 2005, The Inn at Virginia Tech and Skelton Conference Center, Virginia Tech, Blacksburg, Virginia.

Lee, J., Kleczyk, E., Bosch, D., Tanellari, E., Dwyer, S., & Dietrich, A. In Press. Case Study: Preference Trade-offs Towards Home Plumbing Attributes and Materials. *Water*

*Resource Planning Management Journal*. Special Edition in Memory of Dr. G. V. Loganathan.

Loganathan, G. V., & Lee, J., (2005). Decision tool for Optimal Replacement of Plumbing Systems. *Journal of Civil and Environmental Engineering Systems 22*(4), 189-204.

Stevens, T. H., Belkner, R., Dennis, D., Kitteredge, D., & Willis, C. (2000). Comparison of Contingent Valuation and Conjoint Analysis in Ecosystem Management. *Ecological Economics 32*, 63-74.

Stiglitz, J. E. (1974). On the Irrelevance of Corporate Financial Policy. *American Economic Review 64*(6), 851-866.

Whitehead, J. C., Pattanayak, S K., Van Houtven, G. L., & Gelso, B. R. (2005). Combining Revealed and Stated Preference Data to Estimate the Nonmarket Value of Ecological Services: An Assessment of the State of the Science. *Working Paper 05-19*. Available online at: http://econ.appstate.edu/RePEc/pdf/wp0519.pdf. Accessed October 2008.

# CHAPTER 2

# Incidence and Costs of Home Plumbing Corrosion: An Investigation of Sample Selection Bias

An earlier version of this article was published in the December 2008 edition of the *Journal of American Water Works Association 100(12).* E. Kleczyk and Dr. D. Bosch coauthored this article. The published article differs from Chapter 2 (extended version of the published article) by the extent to which the econometric model and economics theory are explained in the paper. The in-depth explanation of these concepts is not necessary to understanding of the article content.

An earlier version of this article was presented at the American Agricultural Economics Association Annual Meeting. Long Beach, CA. July 2006. The article is available online at http://agecon.lib.umn.edu/cgi-bin/view.pl.

## Abstract

This research study analyzes factors influencing the occurrence and costs of pinhole leak corrosion in households in suburban Washington, D.C. A mail survey of households is used to elicit whether leaks were experienced in the home, the number of pinhole leaks, and estimated repair costs for plumbing and associated damages. Probability modeling and linear regression analysis of survey responses with correction for sample selection bias are used to analyze survey responses regarding leak occurrences and costs. Pinhole leak occurrences are found to be associated with pipe type installed, property age, pipe failure history, and dwelling distance from a water treatment plant. The number and location of pinhole leaks in the dwelling and the pipe type are associated with the financial costs of pinhole leaks. For example, repair of leaks occurring in the basement generally are much lower compared to leaks occurring on the second floor. Faster responses to pinhole leak outbreaks by utility managers and policymakers in terms of advising homeowners on the best ways of responding to leaks would assist homeowners in reducing costs of pinhole leak repairs and associated damages. Expanding state-funded property insurance to cover property damage from leaks when private insurance coverage is inadequate would also benefit many homeowners.

**Keywords:** Corrosion Detection and Control, Home Plumbing System, Drinking Water, Heckman Sample Selection Bias

**JEL Classification Code:** Q25, C13

## 2.1 Introduction

Pinhole leaks are small holes in drinking water plumbing caused by pitting corrosion, a type of corrosion concentrated on a small area of the pipe. In most cases, very small pinhole leaks are hard to detect, especially if they appear in pipes running through walls and ceilings. Generally, it is recommended that pipes be replaced after three or four leaks (Gurner, 2003), although others have recommended two leaks (Edwards et al., 2004). Damage from pinhole leaks can include collapse of walls and ceilings and water can contribute to growth of mold on the surface of walls, floors, and ceilings. Mold exposure can cause allergic reactions, such as irritation of eyes, skin, and throat. Furthermore, copper itself can cause severe health problems, such as liver and kidney failure, if consumed in doses higher than 1.3 mg/l (EPA, 1992). In addition, costs of pinhole leaks include costs of fixing the plumbing system, repairing damage to property, and loss of revenue while plumbing systems are being repaired. On the other hand, water leaks increase water usage and therefore, utility revenue. Property damage may include mold damage, which led to an $8.9 billion dollar loss by the residential sector of insurance companies in 2001 (Mann, 2002).

While pinhole leaks are a nationwide problem, they are more common in certain regions of selected states including Maryland, California, Florida, and Ohio. In Maryland, the governor appointed a task force to examine the issue (Maryland Task Force Study, 2004). Some areas (for example, parts of Florida) have considered banning the use of copper piping in order to control the rising number of incidents (Gurner, 2003). However, due to lack of information on pinhole leak and corrosion problems, building regulators and policymakers have been slow to address the issue in some areas.

Although the pinhole leak problem has been extensively studied during recent years, there have been a small number of studies examining the impact of pinhole leak corrosion on individual households. When corrosion occurs, households have to deal with problems resulting from plumbing material leak damage. In many cases, repairs extend beyond replacing the part of failed pipe to repairing all related damage to the house. The process of repairs might take several weeks or months and cost households thousands of dollars. In many cases, the property insurance policy does not cover damages resulting from leaks and households are forced to cover the expenses by themselves. In addition to the financial and time costs, households experience emotional stress because of dealing with the corrosion solutions.

Few studies have examined the socioeconomic impacts of corrosion on individual households. This study addresses the gap by analyzing factors influencing the occurrence and costs of pinhole leak corrosion in individual households. Survey data and regression analysis are employed to assess pinhole leak occurrences in home plumbing systems and associated costs. The influence of sample selection bias on estimated coefficients and statistical significance of the estimated model is assessed. Information on factors associated with pinhole leak incidence and costs can be useful to policymakers, water utilities, housing authorities, and others in formulating policies and programs to reduce pinhole leak occurrences and the negative financial and emotional impacts of corrosion on individual households.

### 2.1.2 Study Objectives

Because few studies have examined the socio-economic impacts of corrosion on individual households, this study addresses the gap by analyzing factors related to the occurrence and costs of pinhole leak corrosion in individual households. The study objectives were met by (1) conducting a residential copper pitting survey in Washington D.C., an area known for a higher than average rate of copper pitting; (2) utilizing survey data in a weighted logistic model to predict probability of pinhole leak occurrence, and (3) analyzing the costs associated with pinhole leak damage using the ordinary least square regression (OLS) corrected for sample selection bias. Information on factors associated with pinhole leak incidence and costs can be useful to policymakers, water utilities, housing authorities, and others in formulating policies and programs to reduce pinhole leak occurrences and the negative financial and emotional impacts of corrosion on individual households.

## 2.2 Literature Review

### 2.2.1 Causes of Corrosion

Corrosion is “the chemical or electrochemical reaction between a material, usually a metal, and its environment that produces a deterioration of the material and its properties” (Corrosion Source, 2007). Pitting corrosion is a “form of localized corrosion as it does not spread laterally across an exposed surface rapidly but penetrates into the metal very quickly, usually at an angle of $90^{\circ}$ to the surface” (Brown, 2001). Theories as to the origins of pinhole leaks vary and there are few proven causes (Edwards et. al., 2004). Edwards suggests that removal of natural organic matter as mandated by tighter Environmental Protection Agency (EPA) drinking water standards can contribute to the problem in combination with other factors since natural organic matter can be an inhibitor to the corrosion-inducing chemical reactions. The natural organic matter can reverse copper corrosion by allowing the growth of organisms to reduce the oxygen and also by adsorption to the pipe wall (Edwards & Sprague, 2001). Pipe failures can also result from other factors, including faulty installation or wearing of copper due to friction from rubbing against a surface (Fleishman, 2001). Metallic plumbing materials such copper and stainless steel can be affected by uniform and pitting corrosion. Pitting corrosion might be affected by synergies among the following factors: source of water and the applied treatment, pipe type, plumbing material, water pressure and velocity, temperature, anthropogenic effects, and water chemistry including water hardness, alkalinity, dissolved oxygen, and corrosion inhibitors (Farooqi, 2006).

The corrosion occurrence rate varies significantly with the water chemistry. For example, fresh, surface, and saline water have different rates of corrosion occurrence. According to the Copper Development Association database (CDA) (1974-2004), forty-six percent of residential dwellings reporting pipe failures received drinking water from private wells while nineteen percent received water from surface water supply (Farooqi, 2006). The majority of residential dwellings receive water from large municipal distribution systems operated by private and/or public utility companies who draw water from either surface water or ground water aquifers. Ground and/or surface raw water and applied treatment could affect potential for corrosion (Farooqi, 2006). For example, surface water is more turbid and carries higher levels of organic contaminants and urban pollutants. On the other hand, ground water has higher concentrations of total dissolved solids (TDS) due to mineral pick up from soil and rocks (Tchobanoglous & Schroeder, 1985).

Pinhole leak corrosion is also dependent on the type of plumbing material of which there are four major types: copper, plastic (CPVC, PVC, PEX, Polybutylene), galvanized iron, and stainless steel. Copper pipes are the most widely used plumbing material in residential dwellings due to advantages such as durability, affordability, minimal health problems, and fire retardence. However, copper is susceptible to pinhole leak occurrences especially when the water has a pH lower than 6.5 ($pH < 6.5$) (Farooqi, 2006). Stainless steel has a high resistance to corrosion but has other disadvantages such as high cost (Farooqi, 2006). Galvanized iron pipes are used very rarely in residential plumbing due to problems with rust build up, leading to hydraulic malfunction of pipes (NSF, 2007). Plastic pipes are not subject to corrosion but have other problems such as a tendency to fail when exposed to extreme heat and cold (Farooqi, 2006). Pipe size also might have an impact on pinhole leak occurrences. Analysis of CDA database (1974-2004) revealed that the majority of pipe failures occurred in half-inch pipes (0.5 in) and three quarters pipe (0.75 in) (Farooqi, 2006).

Water pressure, velocity, temperature, and anthropogenic effects may also affect corrosion potential. In general, higher pressure and velocity of water are both deleterious (Nguyen, 2005), although higher velocities may increase the fusion rates of protective species within the pipe and protect against corrosion (Schock, 1999). Water temperature plays a significant role in copper corrosion through its impacts on solubility of solids, the formation of scales on the copper surface, brochanitite formation, and water velocity (Schock, 1999; Edwards et al., 2004). For example, the presence of sulfate, bicarbonate, and orthophosphate leads to short term reductions in copper solubility. However, over longer time periods, formation of basic brochantite can ultimately increase copper solubility (Powers, 2000). Overall, the impact of water temperature on pinhole leak corrosion is very complex and might depend on water chemistry.

Water chemistry, including water hardness, influences corrosion potential. Harder water tends to be less corrosive (Schock, 1999), although the form of hardness affects the level of corrosion resistance (Edwards et al., 1994). Water hardness may interact with hydraulic factors, such as formation of gaseous cavities, to influence corrosion (Cohen & Meyers, 2005).

Anthropogenic effects can also impact pinhole leak corrosion in copper pipes, including human errors during plumbing system installation and/or repair (CDA Reports, 1974-2004). Post-construction activities can cause pitting corrosion. For example, improperly flushed copper pipes might cause accumulation of debris at the bottom of the horizontal pipes, which can further induce pitting corrosion due to oxygen differential mechanism (Rossum, 1985).

### 2.2.2 Socioeconomic Impact of Corrosion

Only few studies have examined the socioeconomic impact of corrosion on individual households, however, accounting for economic as well as health and aesthetic problems resulting from corrosion alongside the chemical and biological causes of corrosion is critical in managing distribution systems to provide safe water to households. Cost of corrosion to public infrastructure is estimated to be billions of dollars and to include loss of about 15% of the water resource through leaks (Dietrich et al., 2004). Additionally, the total cost of pinhole leaks within the U.S. was estimated to be nearly $930 million per year. More than 50% of the costs are due to plumbing material failure in single-family homes while multi-family apartment dwellings and commercial buildings account for around 20% (Scardina et al., 2007). Repairs associated with plumbing failure can last up to several weeks and the repairs extend beyond replacing the part of failed pipe to repairing all related damage to the house. The extensive repairs may cost households thousands of dollars. Damages from pinhole leaks may be covered by a homeowners' insurance policy; however, in recent years, homeowners' insurance companies have suffered significant losses, even up to $9 billion in 2001. Water-pipe breakage and related problems are often not covered under new insurance regulations and homeowners are forced to fully cover all expenditures (Jones, 2002).

Insurance companies are assisted in screening undesirable risks by CLUE (Comprehensive Loss Underwriting Exchange) report, a comprehensive database of personal property information that includes information on claim date, its cause, and the amount paid out to the customer (Franscona, 2003). Insurance companies may limit their coverage of losses or decline to renew homeowner policies to customers who either had incidents of leaking water pipes or bought a house that was affected by the problem. The homeowner must either find another carrier, possibly at a much higher premium rate, or bear all costs associated with home plumbing failure. States may offer a last resort coverage plan for households who lost their coverage; however, premiums are very high, and loss coverage is limited with such plans (Dietrich et al., 2005).

Water-related problems may affect housing resale values. Difficulties in obtaining reasonably priced homeowners insurance coverage may discourage potential buyers and lower resale values. Resale values are also lowered through structural damage to the house and environmental contamination of drinking water and other aspects of the home. For example, in Washington, D.C., real estate agents have to disclose to potential buyers any environmental problems such as asbestos, radon gas, lead contamination of water, chronic pipe leaks, and flooding (Fleishmann, 2004).

It is difficult to assess the health costs associated with pinhole leaks due to the complex relationship of water quality and health. The health problems may be, however, examined through association to pinhole leaks and copper corrosion (Dietrich et al., 2005). Copper is an essential micronutrient and World Health Organization (WHO) recommended that adults consume 30-μg/kg bodyweight per day (WHO, 1996). Drinking water is a major source of dietary copper (Dietrich et al., 2004). Copper however may also cause severe health problems when it appears in water at levels higher than about 3 mg/L. The increased copper levels in water may cause acute symptoms such nausea, vomiting, and diarrhea in the gastrointestinal tract (Olivares et al. 1998, 2001; Pandit & Bhave, 1996; Pizarro et al. 1999, 2001). The stomach irritation is caused by copper ions producing the toxic effect. When copper is ingested in amount greater than 1g, an acute toxemia may result leading even to death (NRC, 2000).

Pinhole leaks may also lead to creation of mold on the surface of walls, floors and ceilings, which is very hazardous to human health. For example, diseases such as asthma can occur due to contact with mold. Mold exposure can also cause allergic reaction such as irritation of eyes, skin, nose, throat, and lungs. People with weakened immune systems may be more vulnerable to infections by molds and mold toxins. Finally, hypersensitivity pneumonitis may develop, following either short-term or long-term exposure to mold. The disease resembles bacterial pneumonia and is relatively uncommon (NIAQI, 2004).

Copper pipe corrosion presents problems for aesthetics of drinking water supplies, although research on water-related aesthetics of metallic materials is limited. Depending on individual taste acuity, different copper concentrations (1-5 mg/L) may produce unpleasant sensations described as metallic, acidic, astringent, salty, and bitter (Beguin-Bruhin et al., 1983; Zacarias et al., 2001). The major source of organoleptic problems is the water distribution system (Khiari et al., 1999) as it releases copper or iron when corroding (AWWA, 2003). “Metallic” taste is often the reported water problem (Suffet et al., 1995). For example, Virginia Tech research group found 70% of individuals participating in their study capable of detecting metallic taste in distilled water at copper concentrations below 1 mg/L while 75% of the panelists capable of detecting taste of copper at and below the health-based standards of WHO (Dietrich et al., 2004; Cuppett et al., 2006).

Beside taste, the metals can emit odor and therefore, negatively impact the water quality. Uncorroded metals are very unlikely to smell as they present extremely low vapor pressures (Hultgren, 1993). Likely corrosion products on the metal surface can easily evaporate (Weast, 1977) and therefore, cause unfavorable odor. Despite the potential for metal odors, no citations concerning “metallic” odors in water have been found so far (Dietrich et al., 2004).

The copper levels in drinking water above 1.3 mg/L may cause a blue-green discoloration of water and leaves a bluish-green stain on porcelain if water drips from a faucet. There is also a persistently strong smell of chlorine in the water supply when pinhole corrosion occurs in the copper plumbing system (NIAQI, 2004).

## 2.3 Methodology

### 2.3.1 Survey Design

A mail survey of residential water users in the Maryland suburbs of Washington D.C. was conducted in 2004 to determine the prevalence of pinhole leaks in residences. This area was selected because of the large number of pinhole leaks reported by utility customers (Washington Suburban Sanitary Commission, no date). A stratified random sample of 5,015 residents in the zip codes served by WSSC was chosen. Zip codes with high incidence of reported leaks to the WSSC and zip codes containing water treatment plant were sampled more heavily (Figure1 presents the surveyed area). The survey was distributed following the Dillman technique of mail surveying, which included mailing a questionnaire with postage-paid return envelope, sending a reminder card, and sending a second copy of the survey to nonresponders (Dillman, 1978). A total of 5,009 Maryland residents received the survey and 1,128 responses were returned of which 1,120 responses were used in the analysis. Eight responses were dropped from the study because they were incomplete. The survey analysis focused on the incidence of pinhole leaks; associated financial, time, and emotional costs; and potential causal factors.

The sample design was intended to fulfill two purposes: 1) provide an indication of the incidence of corrosion and pinhole leak problems across the study area; and 2) target areas with high reported rates of pinhole leaks in order to provide more information about pinhole leaks in residential dwellings. Due to financial constraints of the project, the sample consisted approximately of 5,000 households with 3,000 surveys allocated to the high frequency of incidents zip codes located nearby the Patuxent and Potomac water plants and the remaining 2,000 surveys distributed to the other zip codes reporting leaks in the WSSC database. These surveys were distributed based on the proportion of reported leaks in the zip code area relative to the total number of leaks in the remaining area. A minimum of 10 surveys was sent to every zip code reporting leaks.

Samples were weighted to correct for oversampling in zip codes with high numbers of reported pinhole leaks. Responses were weighted by the number of people over 18 years of age represented by each survey sample in each zip code. Weighting was done as follows: 1) *Calculate the number of residents over age 18 per sample survey in* *the study area*. There was an average of 216 residents per sample survey (1,083,323 residents divided by 5,009 surveys). 2) *Calculate the number of residents over age 18 per sample survey in an individual zip code area.* For example, a zip code area with a population of 35,000 that received 100 surveys would have 350 residents per sample. 3) *Calculate the weight for an individual zip code area as the ratio of the number of residents per sample in the individual study area to the number of residents per sample in the overall study area*. Responses from the zip code example would have a weight of 1.6 (= 350/216).

### 2.3.2 Analysis of Pinhole Leak Occurrences

Respondents were asked if they ever had a pinhole leak (small hole in pipe) in either hot or cold water pipes of their residence. Based on their responses (either 'yes' or 'no'), a probabilistic analysis of potential causal factors using a weighted binary logit model was performed to predict the probability of pinhole leaks occurrences. The logit regression predicts the probability of pinhole leak incidents associated with the independent factors. Weights used in the analysis are the sample weights for individual zip codes described earlier. To evaluate the logit regression, the explanatory variables are related to the probability of pinhole leak incidents.

If a respondent reported a pinhole leak, then a value of 1 was assigned to a response variable, $y_i$, if no leaks were reported than 0 was assigned. $y_i$ is a realization of a random variable ($Y_i$). $Y_i$ has a distribution described by a probability parameter, $\Pi_i$, for $n_i$ observations and takes the following form (Rodriguez, 2007):

$$Probability\ (Y_i = y_i) = ({}^{ni}{}_{yi})\Pi_i{}^{yi}(1 - \Pi_i)^{1-yi} \qquad (1).$$

Equation 1 presents that $Y_i$ takes values of either 1 or 0 with probabilities $\Pi_i$ and $(1 - \Pi_i)$ respectively.

Assuming that the logit of the probability is a linear function of predictors, the logit function is defined as follows:

$$logit(\Pi_i) = (z_i'\boldsymbol{\beta}), \qquad (2)$$

where $z_i$ is a vector of covariates (independent variables impacting the probability of pinhole leaks) and $\boldsymbol{\beta}$ is a vector of regression coefficients, which defines the systematic structure of the model, and $i = 1\ldots n$. Furthermore, $\beta_j$ represents the change in the logit of the probability associated with a unit change in the $j$-th predictor while holding all other predictors constant. When equation (2) is exponentiated, odds for $i$-th unit of a covariate are derived (Rodriguez, 2007):

$$\Pi_i/(1 - \Pi_i) = exp(z_i'\boldsymbol{\beta}). \qquad (3).$$

Solving equation 3 for the probability parameter, $\Pi_i$, results in the following model (Rodriguez, 2007):

$$\Pi_i = \Phi(z_i'\boldsymbol{\beta}) \quad (4)$$

where $\Phi(z'_i\boldsymbol{\beta})$ is the cumulative distribution of a logit random variable, $z_i$ explanatory variables, and $i = 1\ldots n$ (Griffiths et al., 1993). The cumulative distribution function is defined as:

$$\Pi_i = \Phi(z_i'\boldsymbol{\beta}) = 1/(1 + e^{(-zi'\boldsymbol{\beta})}) \quad (5).$$

The left hand side of the above expression is the probability scale while the right hand side is a non-linear function of the independent variables (predictors) (Rodriguez, 2007).

The binary logit regression does not depend on the restrictive assumptions of a least square regression such as linear relationship between dependent and independent variables, normally distributed endogenous variables and error terms, and homogeneity of variance (Garson, 2006). The logit probability density function is smooth, symmetric around zero, and bell-shaped. Maximum likelihood estimation is often employed in the logit function evaluation:

$$logL(\boldsymbol{\beta}) = \sum\{y_i \log(\Pi_i) + (n_i - y_i)\log(1 - \Pi_i)\} \quad (6)$$

where $\Pi_i$ depends on the covariates $z_i$ and a vector of $p$ parameters $\boldsymbol{\beta}$ through the logit transformation (Griffiths et al., 1993).

The dependent variable of the logit model is binary, taking a value of 1 when pinhole leaks occur and 0 otherwise. The following factors potentially associated with pinhole leaks were analyzed in the model: geographical location of dwelling relative to water treatment plant (approximated by estimated water travel time (water age) to zip code in which the dwelling is located) (provided by Robert Buglass, Principal Environmental Engineer, Washington Suburban Sanitary Commission, Laurel, Maryland), type of pipes installed (copper vs. others), history of pipe failure, plumbing age, and the source of water (approximated by a dummy variable with a value of 1 for respondents located in zip codes which predominantly receive their water from Patuxent Water Treatment Plant (PWTP) and 0 otherwise). History of pipe failure reflects whether the respondent had other types of failure in their drinking water pipes beside pinhole leaks. Age of plumbing material is a variable representing the actual age of the plumbing system installed in the household. Plumbing age was set equal to age of the dwelling unless the respondent indicated plumbing had been replaced in which case the number of years since replacement was used. In the questionnaire, the house age was determined using the following intervals: since 2000, 1990-1999, 1980-1989, 1970-1979, 1960-1969, 1950-1959, 1940-1949, before 1940. The plumbing age question was asked using the following intervals: after 2000, 1995-1999, 1990-1994, 1980-1989, 1970-1979, 1960-1969, and before 1960. The five-year intervals were combined to make ten-year intervals in order to match the ten-year interval format of house age question. Median value of the interval was computed and subtracted from the year 2006.

A variable not included in the pinhole leak occurrences regression analysis is the volume of water usage of each household annually. Information on water usage for each household was not asked in the survey. The influence of water use volume on pinhole leak occurrences was examined for a sample of respondents in a zip code with a high frequency of pinhole leaks. Data on a rolling average daily water use over the previous 12 months were obtained for 75 respondents with pinhole leaks and 72 respondents without pinhole leaks. A t-statistic test of mean water use for the groups with and without pinhole leaks indicated no statistically significant differences between the groups at a 5 percent significance level (t-statistic = 0.269). As a result, no further analysis was done of impact of the volume of water usage on pinhole leak occurrences.

### 2.3.3 Analysis of Costs Associated with Pinhole Leak Repair

The cost estimation method is applied to assess the factors driving costs of pinhole leak repairs and associated damage. The total costs of pinhole leak repairs consist of direct costs, which are represented by financial costs, as well as indirect costs associated with the time spent to deal with this issue. The indirect costs are translated into the forgone earnings, resulting from dealing with plumbing material failures. Total expenditures for plumbing material repairs are defined a function of the following variables: the number of leaks, type of pipe material, type of water conveyed in the pipe with the leak, place of leak in the dwelling, age of the dwelling, and source of water, as well as length of time in the current residency. A weighted ordinary least squares (OLS) regression is employed to estimate the total costs associated with pinhole leak repairs. The weights used in the regression are the sample weights for individual zip code areas described earlier.

The ordinary least square regression assumes i) error term ($e_i$) is independently and identically distributed in $i$ with a standard normal distribution; ii) $\{x_i: i\ldots n\}$ is independent of $e_i$; iii) error term ($e_i$ ) and unmeasured latent continuous random variable ($\delta_i$) are independent within respondent $i$. $\delta_i$ is distributed normally and represents the unmeasured characteristics of respondents (Briggs, 2002). The weighted ordinary least squares regression is specified as follows:

$$y = X\boldsymbol{\beta} + \boldsymbol{e} \quad (7)$$

$$\boldsymbol{e} \sim (0, \sigma^2 I_i) \quad (8)$$

where $y$ is the dependent variable weighted by observation, and $X$ is a vector of independent variables weighted by observation (Briggs, 2004).

The dependent cost variable (*y*) includes the financial and time costs of pinhole leaks. Financial costs included the total amount spent by the respondent, his/her landlord (if applicable), and/or the respondent's insurance company to repair the plumbing system and any other damage from pinhole leaks. Time spent by the respondent in dealing with pinhole leaks was translated into financial terms by multiplying the estimated time spent by the survey respondent by the estimated hourly wage rate for each zip code. The estimated hourly wage rate was calculated based on the inflation adjusted annual average income for each zip code as reported by Census Bureau divided by the assumed number of hours worked in a year, 2,080 (52 weeks * 40 hour/week; Census Bureau, 2000). The estimated total expenses (financial and time costs) ranged from $1,271 to $18,455. Fifty-six observations had missing values for either the time or money spent on pinhole leak repairs. As a result, the total number of unweighted observations for total costs is 392. The 392 observations represent the survey respondents who experienced pinhole leaks and accrued expenses from repairing damage associated with plumbing material failure.

Total expenses were regressed on number of pinhole leaks, type of pipe material (copper vs. others; iron vs. others), type of water conveyed in the pipe with the leak (hot vs. cold), place of leak in the dwelling, age of the dwelling, and source of water (Patuxent water treatment plant vs. other). Additionally, the impact of length of residency on the total estimated expenses was tested; however, this variable was found not to be a statistically significant factor, and therefore not included in the final model estimation. Regardless of the age of the plumbing material, the mean total expense for residents living in the house for longer than 10 years was $3,720 while for those residing less than 10 years mean expense was $3,496. T-statistic test mean indicated no statistically significant differences between the two groups at a 5 percent significance level (t-statistic = - 0.510). This finding is in agreement with the responses to the survey. Most of the respondents had experienced pinhole leaks during the 1990s and 2000s. As a result, there should not have been major differences in total expenses accrued by those residing for a longer time period compared to those residing a shorter period of time in their homes. No further analysis was performed for the impact of the length of residency on the total cost of the repairs associated with pinhole leak damage.

Unstandardized and standardized coefficients of each variable were reported for the weighted ordinary least squares regression. Unstandardized coefficients represent the actual effect of the independent variable on the dependent variable and represent the dollar amount of increase/decrease in total costs due to change in the independent variables by one unit. Standardized coefficients explain the impact of a one standard deviation increase in the independent variable on the dependent variable relative to other variables in the model. As a result, factor standardization allows for observing the relative importance of explanatory factors, while keeping the variable of interest in its original form (Kim & Feree, 1981).

### 2.3.4 Sample Selection Bias

Information on the total repair costs associated with pinhole leaks might not be available for all the survey respondents and the distribution of respondents over categories of the independent variables may not be random. As discussed below, some survey observations for pinhole leaks incidents were incorrectly classified (model predicted pinhole leaks for the respondent when none were observed or vice versa). The discrepancy between observed and predicted values could be related to the differences in unmeasured and measured characteristics between people reporting pinhole leaks and those not reporting.

There are several methods of correcting for the sample selection bias. These methods include censored Tobit and Heckman models (for additional information please see Sweeney (2006)). A Tobit model is used when the dependent variable is censored at point below which additional observations are not observable. The model, however, has some limitations such as determination of the probability of the censored observation as well as the value of the dependent variable for the same set of variables. The model also does not allow a full theoretical explanation of why the censored observations are censored. The Heckman procedure is an improvement on the Tobit model as it utilizes a different set of variables to determine the probability of censoring and the value of the dependent variable given that it is observed. It also allows for a greater theoretical development as it takes into account for selection and outcome being interrelated

(Sweeney, 2006).

A Heckman two-stage procedure was utilized to correct for the selection bias in order to estimate the unbiased effects of independent variables on total pinhole corrosion repair expenses (Briggs, 2004). In the first step of the Heckman procedure, the selection bias was computed. In the second step, the weighted OLS regression for total repair expenses was performed. The selection bias correction factor was added as an additional explanatory variable of pinhole leak repair costs. Since the addition of the sample selection bias variable impacts the estimated standard errors, the standard errors had to be corrected by utilizing the generalized least square regression (GLS).

The weighted binary logit model used to predict leak occurrences as described in equations (1) and (2) was employed to compare these dissimilarities between the two groups. Because the Heckman model assumes a normally distributed error term and the logit regression does not, the logit model estimates were converted into probit model estimates. Using the inverse cumulative distribution function of the normal distribution, logit individual probabilities (ranging from 0.17 to 0.81) were translated into the individual probit scores ($\xi i$) for each respondent. The probit scores represent the residuals between the observed and predicted probabilities and range between -0.94 and 0.86. The residuals of the selection equation were then utilized to construct a selection bias correction, which is equivalent to the inverse Mill's ratio. The following formulas were employed to calculate the sample selection bias correction factor:

$$\lambda_i = ((1/sqrt(2*3.142))*(exp(-\xi_i * \xi_i *0.5)))/ \Phi_i (\xi_i), \text{ for } y = 1 \quad (9)$$

$$\lambda_i = - ((1/sqrt(2*3.142))*(exp(-\xi_i * \xi_i *0.5)))/(1- \Phi_i (\xi_i)), \text{ for } y = 0 \quad (10)$$

where $\xi_i$ are the individual probit scores and $\Phi_i$ is the cumulative probability function. The selection bias correction measures the effects of all unmeasured characteristics of respondents related to pinhole leak corrosion. The bias correction variable is added as an explanatory variable. As a result, the remaining variation in the dependent variable after the removal of the effect of the known variables is explained by the unknown factors (Smits, 2003).

The inverse Mill's ratio for each respondent is as follows:

$$\lambda_i = \phi_i(Z) / \Phi_i(Z), \text{ for } y = 1 \quad (11)$$

$$\lambda_i = -\phi_i(Z) / (1 - \Phi_i(Z)), \text{ for } y = 0 \quad (12)$$

where $Z$ is the vector of explanatory variables for the respondent, respectively; $\phi_i$ is the density probability function; and $\Phi_i$ is the cumulative probability function. If the predicted probability of pinhole leak occurrences is close (but not equal) to unity for respondents with pinhole leaks, this implies that $\Phi_i(Z) \to 1$ and that $\phi_i(Z) \to 0$. As a result, the inverse Mills ratio term is close to zero for this respondent and the selection bias correction variable has no influence on the total cost of pinhole leak damage. In case of respondents without pinhole leaks occurrences, the high-predicted probability of corrosion implies high selection bias in absolute value. As a result, these respondents should have experienced pinhole leaks in their plumbing system. On the other hand, if the predicted probability is small for respondents with pinhole leaks occurrences, the inverse Mills ratio is large ($\Phi_i(Z) \to 0$ and $\phi_i(Z) \to 1$) and the selection bias correction variable strongly influences the total expenses of corrosion damage. For the respondents without pinhole leaks, the small-predicted probability implies the inverse Mill's ratio to be small in absolute terms. The low bias in absolute value means that respondents without pinhole leaks appropriately did not observe pinhole leaks in their plumbing (Briggs, 2004). For this analysis, the selection bias variable was calculated for 879 unweighted observations and ranged between -1.42 and 1.44.

In the second step of the Heckman procedure (substantial analysis), the weighted ordinary least square regression for total repair expenses was performed. The selection bias correction factor was added as an additional explanatory variable of pinhole leak repair costs. The ordinary least square model specification with the Heckman selection bias correction is provided by the following:

$$y = X\boldsymbol{\beta} + \lambda\boldsymbol{\theta} + \boldsymbol{e} \quad (13)$$

where $X$ is the vector of explanatory variables for the respondent with pinhole leaks, $\lambda$ is the sample selection bias correction for respondents with pinhole leaks, and $\theta$ is the coefficient of the selection bias defined as $\theta = \rho * \sigma$ where $\rho$ is the correlation coefficient and $\sigma$ is the standard deviation.

The F-test was employed to examine the importance of correction for selection bias by comparing models with and without sample selection bias correction variable. The F-test statistic is defined as follows:

$$F\text{-test statistic} = (SSE_r - SSE_u) / (SSE_u /(n-k)) \sim F(1, (n-k)) \quad (14)$$

where $SSE_r$ is the error sum of squares in the model without sample selection bias, $SSE_u$ is the error sum of squares in the model with sample selection bias, $n$ is the number of observations in the model with selection bias, and $k$ is the number of coefficients in the model with sample selection bias. The hypotheses are as follows:

$$H_0\text{: } y = X\boldsymbol{\beta} + \boldsymbol{e} \quad (15)$$

$$H_{1:}\ y = X\boldsymbol{\beta} + \lambda\boldsymbol{\theta} + \boldsymbol{e}. \quad (16)$$

If at a 5% significance level and (1, n-k) degrees of freedom, the critical F-value is less than the F-test statistic, then the $H_0$ is rejected and the sample selection bias correction should be included in the model (Griffiths et al., 1993).

Because the selection bias correction variable reflects all of the unmeasured characteristics related to pinhole leak occurrences, the remaining independent variables are free from this effect and unbiased. On the other hand, the standard errors of the ordinary least square estimates are heteroskedastic:

$$\boldsymbol{e} \sim (0, \sigma^2 V_i)) \quad (17).$$

With heteroskedastic errors, the least square estimates are still unbiased, but not efficient. As a result, the variance varies with each respondent (the true variance is either overestimated or underestimated). In order to correct for heteroskedasticity (obtain constant variance: $\boldsymbol{e} \sim (0, \sigma^2 I_i)$), a generalized least square regression (GLS) was employed (Smits, 2003).

In order to confirm the existence of selection bias in the total costs model specification, the correlation between the error terms of the selection and substantial equations should be estimated ($Corr(e, \delta) = \rho$). As presented in equation 9, the coefficient of selection bias estimated in the substantial analysis is defined in part by $\rho$. As a result, the correlation coefficient is calculated as follows:

$$\rho \quad = \quad sqrt(\theta^{\wedge 2} / \sigma^{\wedge 2}) \quad (18)$$

where $\theta^{\wedge}$ is the estimated coefficient of selection bias correction and $\sigma^{\wedge 2}$ is the estimated variance from the substantial analysis (Smits, 2003).

**2.4 Results**

Survey results indicate that pinhole leaks in copper home plumbing systems are a frequently reported problem in the Maryland suburbs of Washington, D.C. Four hundred and forty-eight respondents (40 percent of 1,120 respondents on an unweighted basis) reported incidents of pinhole leaks. On a weighted basis, 32 percent of respondents reported leaks. This number is higher compared to pinhole leak rates reported by other studies. For example, the average leak rate in non-hotspot areas nationwide is 8% (Scardina et al., 2007), while three hotspot areas reported rates between 11% and 21% (Scardina et al., 2007). Of those reporting leaks, most respondents (74% on a weighted basis) had one to four leaks; however, 8% reported seven or more leaks (Table 2.1). Forty-three percent of respondents (weighted basis) reported that the first leak occurred since 2000, 38 percent reported the first leak in the 1990s, 6 percent reported the first leak in the 1980s, and less than 1 percent reported the first leak before 1980 (Table 2.2).

The age distribution of dwellings with leaks is roughly proportional to the overall age distribution of dwellings. For example, on an unweighted basis (Table 2.3), dwellings built prior to 1970 account for 65% of the total observations and 63% of the leaks. The largest discrepancy is in the 1960-1969 period. Dwellings from that period account for 20% of the total observations but 31% of the leaks (unweighted basis).

Most leaks were located in the basement (61% on a weighted basis) followed by the first floor (37%) (Table 2.4). Only 19% of respondents (weighted basis) reported leaks occurring on the second floor or higher.

The total number of unweighted observations with complete information on all variables utilized in the estimation of pinhole leak occurrences was 879 (423 respondents with pinhole leaks and 456 respondents without pinhole leaks). The total unweighted number of observations with missing values is 241. Respondents without pinhole leaks were the control group in the logit regression analysis. The logistic model is significant at a 5 percent level with Log Likelihood equal to 1,103 (Table 2.5, Base model results). The logistic regression results are presented in form of probability values (Table 2.5).

The logit regression results are presented in form of odds ratio and probability values (Table 2.5). The estimated odds ratios indicate relationships between the dependent and independent variables. When the odds ratio is more than 1, there is a positive relationship between the dependent and independent variables and pinhole leaks are more likely to occur with a one unit increase in the value of the explanatory factor. An odds ratio less than 1 indicates a negative relationship between the causal factor and pinhole leaks. For example, an odds ratio of 0.58 for water travel time (water age) indicates that pinhole leaks are about half as likely to occur with a one-day increase in travel time. Finally, an odds ratio of 1 presents no relationships between leaks and its causes (UCLA Academic Technology Services, 2005).

The estimated probability coefficients (Table 2.5) can be interpreted as the impact of one unit increase (decrease) in the independent variable on the chance of pinhole leak occurrences while controlling for other variables in the model (UCLA Academic Technology Services, 2005). For example, one additional day of water age decreases the probability of pinhole leaks by 37%. This finding is in agreement with Rushing and Edwards (2004) who found houses located closer to the treatment plant to be exposed to higher levels of chlorine and therefore, to experience higher levels of pipe corrosion. Furthermore, copper pipes installed in the dwelling raise the chance of pinhole leak incidents by 68%; and history of pipe failures increases the probability of pinhole leaks by 68%. Having water from PWTP increases the probability of pinhole leaks by 85%. Age of plumbing material is not statistically significant.

A prediction accuracy analysis was performed to determine correct and incorrect estimates of the pinhole leak observations. Correct estimates are defined as those responses for which a probability greater than 0.5 is assigned to a respondent with pinhole leaks or a probability less than 0.5 is assigned to a respondent with no reported pinhole leaks. The analysis does not reveal, however, how close to 1.0 the correct predictions are nor how close to 0.0 are the errors (Garson, 2006). Overall prediction was 60% for the unweighted values with or without pinhole leaks, which implies an acceptable prediction level for the model (Table 2.6). Furthermore, 202 observations classified as no pinhole leaks should have been classified as observations with pinhole leaks, while 310 observations with pinhole leaks were predicted appropriately. As a result, the predicted number of pinhole leaks observations is 512 compared to the actual number of observations with pinhole leaks of 423. Thus, the model overpredicted the occurrence of pinhole leaks by about 17%. In a perfect model, all cases are on the diagonal and the percent of correctly predicted values is 100%.

A model for predicting pinhole leak occurrences was estimated based on the binary logit regression analysis. The proposed equation is adjusted by the 17% of overprediction resulting from the initial modeling. Engineers and water utility managers could employ the following model for pinhole leak predictions in their service areas:

$$\textit{pinhole leak probability} = \exp(PH) / (1+\exp(PH)) \quad (19)$$

where

$$PH = -0.170 + 0.677 * \textit{copper pipe} + 0.809 * \textit{history of pipe failure} - 0.532 * \textit{water travel age} \quad (20)$$

The -0.17 term corrects for 17% overprediction from the initial model. The pinhole leak probability variable represents the probability of pinhole leak occurrences and takes values between 0 and 1. The copper pipe variable is binary taking a value of 1 when copper pipes are used and 0 otherwise. History of pipe failure is also binary with a value of 1 when pipes failed previously. Finally, water travel age represents the number of days it takes to deliver drinking water to individual residential properties.

### 2.4.1 Costs associated with Pinhole Leaks

Households with pinhole leaks incur high financial, time, and emotional costs when dealing with pipe failures. Thirty percent of respondents with leaks (weighted data) reported expenditures of at least $500 for repairing leaks and associated damage (Table 2.7). Seven respondents cited $12,000 in repair expenses and one person reported more than $25,000 in damage. Many of the survey respondents pointed out that in addition to repairing or replacing water pipes, they also had to repair damage to ceilings, walls, and floors. In addition, a few respondents moved out of their homes during the renovation periods, which added to their repair expenses. Others mentioned the loss of priceless personal belongings such as family photos, clothes, and furniture damaged by pipe leaks.

Twenty-six percent of respondents with pinhole leaks (Table 2.8, weighted results) spent at least 20 hours dealing with pinhole leaks and related damage. Forty-three percent of respondents spent less than 10 hours. Fifty-six percent of respondents found the experience somewhat or very stressful. Survey respondents felt “aggravated,” “helpless,” and “worried” about future leaks. Lack of knowledge on the possible causes of pinhole leaks added to the overall anxiety felt by those with pinhole leaks. Furthermore, some respondents cited lack of third party involvement and responsibility for pinhole leak occurrences and related damage as additional stress. The respondents felt they had inadequate information to make decisions on water plumbing issues and were frustrated with insufficient assistance from local water utilities, insurance companies, and contractors.

### 2.4.2 Results with no Selection Bias Correction

The weighted ordinary least squares estimation was performed on the observations that reported pinhole leaks. The total unweighted number of observations utilized to estimate the effect of independent variables on the total costs is 392 after deleting 56 observations with missing data for total expenses of pinhole leak damage.

The number of leaks had a large positive impact on pinhole leak costs (Table 2.9). Copper plumbing and pipe failure on the second floor of a dwelling increased repair expenses. The average increase in total costs due to plumbing material failure on second floor is $1,012 (unstandardized coefficients in Table 2.9). Plumbing system malfunctions under the slab and in the basement were negatively related to the total repair expenses, resulting in an average $842 decrease in repair expenses, probably because these areas were easier to access for repair and/or because leaks in these areas represented less associated damage to the structure and furnishings of the dwelling. According to a report by NAHB (1992), pinhole leak occurrences under slab and underground (16 unweighted observations) might be affected by abnormally aggressive soil with a high capacity of obtaining moisture. Finally, the water treatment dummy is negatively related to the costs of repairs implying lower costs experienced by households obtaining water from PWTP. Iron pipes, cold water pipes, and plumbing material age were not related to pinhole leak damage costs.

### 2.4.3 Results with Selection Bias Correction

The effective unweighted sample size after employing the selection bias correction in the least squares regression is 367. The sample selection bias correction values are missing for 25 observations, which had missing observations for one or more explanatory variables. The selection bias correction coefficient is 0.52 and significant at the 0.05 significance level (Table 2.9). The sample selection bias correction is needed in the total expenses of corrosion damage model specification. The correlation coefficient between the error terms in the first and second step of the Heckman procedure is 0.30 and statistically significant at the 5% significance level. The F-test statistic, which tests the importance of selection bias correction in the model specification, is 7.86 (critical $F_{(1, 356)}$ value = 3.84). The above finding confirms the need for sample selection bias correction in the total expenses of corrosion damage model specification.

Including the selection bias caused coefficients and significance levels of other variables to change. Compared to the regression without selection bias correction, the magnitudes of the standardized coefficients increased in absolute value for copper pipes (from 0.09 to 0.17). The magnitudes of the standardized coefficients decreased in absolute value for number of leaks (from 0.33 to 0.16), leak under slab or in basement (from -0.69 to -0.48), and leak on the second floor (from 0.82 to 0.80).

Similarly to the OLS model without selection bias correction, number of leaks, copper plumbing and leaks on the second floor of a dwelling increased repair expenses while leaks under the slab and in the basement are negatively related to total repair expenses. Plumbing material age, cold water pipes, leaks on first floor, and iron pipes are still not significant in the corrected model.

Additionally, source of water (water from the PWTP) is negatively related to the repair costs of pinhole leaks. One possible explanation for finding the water from the Patuxent treatment plant being negatively related to the repair costs is differences in water treatment procedure between competing water treatment plants in the area (PWTP water might be less corrosive). However, the factors impacting water chemistry for water treatment plants were not specifically measured in this study. The treatment plant (and its associated impact on water chemistry) did not have a significant impact on the probability of pinhole leaks occurrences. Thus, the effect of treatment plant on repair costs is not strongly established in this analysis.

Water chemistry impacts water quality and therefore, the release of copper into water. Imran et al. (2005) found that alkalinity, chlorides, sulfates, sodium, and dissolved oxygen have a significant effect on the release of copper into water (corrosion) especially for blended finished waters from surface, ground and saline sources. Adverse effects on total copper release can be mitigated by controlling water quality of blended finished waters primarily by controlling alkalinity and pH (Xiao, 2004). Additionally, temperature and hydraulic retention time are statistically significant physical and operational parameters impacting water corrosiveness (Imran et al., 2005).

### 2.4.4 Sensitivity Analysis of Housing Age

Potentially, results of the regression analyses might be influenced by the large number of dwellings (498) built in the 1950-1969 period. An additional weighting was done to reduce the influence of the 1950-1969 period on regression results. The period from 1950-1969 was made the base period, and respondents living in homes built then received the weight based on their initial sampling weight. Respondents living in homes from other periods were weighted in inverse proportion to the ratio of homes in their age category to the base number of homes. For example, there were 26 respondents living in dwellings built since year 2000. Therefore, the weight for respondents living in dwellings built since 2000 is 498/26 = 19.15. This weight was multiplied by the initial sampling weight described above. For example, a respondent living in a zip code receiving a weight of 1.5 and in a dwelling built after 2000 would receive a final weight of 19.15 times 1.5 = 28.7. The unweighted sample size without selection bias correction is 392 (OLS regression) and with selection bias correction is 367 observations (GLS regression).

The results of the logistic model for observations weighted by dwelling age category (Table 2.5) are different from the base model. Probabilities associated with water age, history of pipe failure, source of water, and use of copper pipes do not differ much from the base model. The probability of pinhole leak occurrences changes in absolute magnitude for the copper pipes variable (68% vs. 73%) and water age variable (-37% vs. - 40%). The plumbing material age variable remains insignificant in the house-age weighted model. Weighting for the number of dwellings built between 1950 and 1969 has little impact on factors affecting leak occurrences.

In the OLS regression of total expenses with no correction for sample selection bias (Table 2.9 and 2.10), the results are very similar between the base model and the model weighted by the age category of dwelling. Similarly to the base model, number of leaks, copper plumbing and leaks on the second floor of a dwelling increased repair expenses while leaks under the slab and in the basement are negatively related to total repair expenses. For example, plumbing system malfunctions under the slab and in the basement resulted in the same average financial decrease in the repair costs of $842. The water source variable is negative and statistically significant at 10%. Plumbing material age, cold water pipes, and iron pipes are still not significant. Compared to the base model, the magnitude of the standardized coefficient increased slightly in absolute value for the number of leaks variable (from 0.33 to 0.34).

There are, however, differences between the OLS estimates without selection bias correction and the results accounting for the bias (Table 2.10). Similarly to population weighted OLS with sample selection correction model, the selection bias correction coefficient is 0.53 and statistically significant while the correlation coefficient between the error terms of the two steps of sample selection bias correction is 0.29 and statistically significant at the 0.05 significance level. The F-test statistic, which tests the importance of selection bias correction in the model specification, is 8.10. The above finding confirms the importance of including the selection bias correction in the model specification.

Similarly to the base model, water from the PWTP became statistically negatively related to the repair costs of pinhole leaks after accounting for the selection bias (Table 2.10). Compared to the regression without selection bias correction, the magnitudes of the standardized coefficients increased in absolute value for the source of water variable (-0.07 vs.-0.31). The magnitude of the standardized coefficient decreased in absolute value for number of leaks variable (0.34 vs. 0.16), leaks on second floor variable (0.82 vs. 0.81), and leaks under slab or in basement (from -0.69 to -0.48).

## 2.5 Conclusions

The objective of the study was to analyze factors influencing the occurrence and costs of pinhole leak corrosion in suburbs of Washington, D.C. This study addresses a gap in literature on socioeconomic impacts of pitting corrosion on individual households. Data were based on a survey of households in the Maryland suburbs of the Washington, D.C. conducted in 2004. Factors associated with frequency of pinhole leaks and repair costs of leaks were analyzed statistically.

Results from the survey indicate that pinhole leaks in home plumbing systems are a frequently reported problem in the greater Washington, D.C. area. This region was predicted to have unusually a high pinhole leak activity compared to the national average at the start of the project. Most pinhole leaks occurred in the mid-1990s and the 2000s. Most respondents had 1 to 4 leaks; however, 8% (weighted basis) reported 7 or more leaks.

Based on the results of the statistical analysis, pinhole leak occurrences depend on the following factors: plumbing material, location of the house relative to the treatment plant, and history of plumbing failure. Furthermore, the probability of pinhole leaks increases with use of copper pipes as the plumbing material. Dwellings located closer to the treatment plant have higher incidents of pinhole leaks possibly due to exposure to higher chlorine concentration. Leaks are also more likely to occur if a plumbing system has failed in the past. Plumbing material age was not found statistically significant in pinhole leak occurrences.

Pinhole leaks and the damage associated with their repairs can have high financial, emotional, and time costs. The total estimated costs for repairs associated with the leak occurrences can reach $25,000. Repair costs increase with the number of leaks, repairs of copper pipes, and location of leaks on the second floor of the dwelling. The costs are lower when repairing damage under the slab or in the basement of the property. Costs are positively correlated with source of water.

Policymakers and utility managers should emphasize faster responses to pinhole leak outbreaks in terms of determining likely causes of leaks and advising homeowners on the best ways of responding to leaks. States or regions with pinhole leak hotspots may wish to convene special task forces to look into the issue as was done in Maryland. Policymakers may also wish to consider expanding state-funded property insurance to cover property damage from pinhole leaks outbreaks in cases where private insurance coverage is inadequate.

In this study, the water chemistry was measured only indirectly based on differences in water treatment plant, which provided mixed results. More specific study of differences in water chemistry and its impact on corrosion and pinhole leak incidence would be beneficial. This extension refers also to houses with private wells. Since only 0.6% of the survey respondents indicated using water from private wells, the impact of well water on pinhole leak occurrences was not investigated. Additionally, the types of soil on which houses are built should be considered as another factor associated with pinhole leak corrosion. In this study, the basement/under slab locations of the house seem to be more prone for pinhole leak occurrences. Possibly these locations may be more prone to leaks due to the characteristics of soils on which the house is built.

### 2.6 Acknowledgements

Funding for this research was provided by the National Science Foundation, Award # DMI-0329474. The views expressed in this report are those of the authors and not of the National Science Foundation. Robert Schulman and Ying Zhang provided advice on sample design. Marc Edwards, Robert Buglass, and anonymous reviewers provided helpful comments on earlier versions of the paper.

### 2.7 References

American Water Works Association (AWWA). (2003). *Dawn the Replacement Era: Reinvesting in Drinking Water Infrastructure.* Washington, D.C.: AWWA Government Affairs Office.

Beguin-Bruhin, Y., Escher, F., Roth, H., & Solms, J. (1983). Threshold Concentration of Copper in Drinking Water. *Lebens Wiss Technologia 16*, 22-26.

Briggs, D. C. (2004). Causal Inferences and the Heckman Model. *Journal of Educational and Behavioral Statistics 29*(4), 397-420.

Brown, R. (2001). *Pitting Control.* Corrosion and Corrosion Control CHE534. Available online at http://www.egr.uri.edu/che/course/CHE534w/chapter4PittingCorrosion.htm. Accessed March 2007.

Census Bureau. (2000). *Quick Facts*. Available online at www.census.gov. Accessed March 2007.

*Copper Development Association (CDA) Reports*. (1974-2004). Available online at http://www.copper.org. Accessed March 2007.

Corrosion Source. (2007). *Handbook*. Available online at http://www.corrosionsource.com/handbook/glossary. Accessed June 2007.

Cohen, A., & Myers, J. (2005). *Copper-tube Corrosion in Domestic Water Systems: Identifying and Mitigating the Seven Forms of Corrosion*. Heating/Piping/Air Conditioning Engineering.

Cuppett, J. D., Duncan, S. E., & Dietrich, M. (2006). Evaluation of Copper Speciation and Water Quality Factors that Affect Aqueous Copper Tasting Response. *Chemical Senses 31*, 689 – 697.

Dillman, D. A. (1978). *Mail and Telephone Surveys*, New York: John Wiley and Sons.

Dietrich, A., Cuppett, J., Edwards, M., Powers, K., Duncan, S., Bosch, D., & Kleczyk, E. (2005). Corrosion of Copper Plumbing and Its Effects on Consumer Health and Economics. *Meeting of the Biocomplexity in the Environment Awardees*. Crystal City, Virginia.

Dietrich, A., Cuppett, J., Edwards, M., Powers, K., Duncan, S., Bosch, D. & Kleczyk, E. January 2005. Corrosion of Copper Plumbing and Its Effects on Consumer Health and Economics. *Proceedings of 2005 NSF DMII Grantees Conference*. Scottsdale, Arizona.

Dietrich, A., Glindemann, D., Pizarro, F., Gidi, V. Olivares, M., Araya, M. et al., (2004). Health and Aesthetic Impact of Copper Corrosion on Drinking Water. *Water Science and Technology 49(2),* 55-62.

Edwards M., Ferguson J. F., & Reiber S.H. (1994). The Pitting Corrosion of Copper. *Journal of AWWA 86(7),* 74-90.

Edwards, M., J. Rushing, C., Kvech, S., & Reiber, S. (2004). Assessing Copper Pinhole Leaks in Residential Plumbing. *Water Science and Technology 49(2),* 83–90.

Edwards, M., & Spraque, N. (2001). Matter and Copper Corrosion By-Product Release: A Mechanic Study. *Corrosion Science 43,* 1-18.

Edwards, M. (2005). *Frequently Asked Questions about Pitting Corrosion*. Washington Suburban Sanitary Commission (WSSC). Available online at http://www.wssc.dst.md.us/copperpipe/MarcEdwardsqa.cfm. Accessed: April 2006.

Farooqi, Owais M. (2006). *An Assessment and Modeling of Copper Plumbing Pipe Failures due to Pinhole Leaks*. M.S. Thesis. Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Fleishman, S. (2001, July). *And the Leaks Go On*. *Washington Post:* G01.

Fleishman, S. (2004). *D.C. Water Worries Sellers, Buyers*. *Washington Post:* A39.

*Franscona*. (2003). Available online at: www.franscona.com/resource/jag403clue.htm. Accessed June 2007.

Garson, D. G. (2006). *Logistic Regression*. Available online at http://www.2.chass.ncsu.edu/garson/PA765/logistic.htm. Accessed April 2006.

Griffiths, W. E., Hill, R. C., & Judge, G. G. (1993). *Learning and Practicing Econometrics*. New York: John Wiley and Sons.

Gurner, S. (2003). *Pinhole Leaks in Copper Pipes*. Available online at http://strathmore-belpre.org/Pinhole%20leaks.htm. Accessed April 2007.

Hultgren, R. (1993). Selected *Values of the Thermodynamic Properties of the Elements*. Metals Park, Ohio: American Society for Metals.

Imran, S. D, Diets, J.D., Taylor, J. S., Randall, A. A., & Cooper, C. D. (2005). Red Water Release in Drinking Water Distribution System. *Journal of the AWWA 97(7):* 93-100.

Jones, S. (2002). *Are There Claims on Your House? Washington Post.* Available online at: http://www.washingtonpost.com/ac2/wp-dyn?pagename=article&node=&contentId=A31. Accessed May 2006.

Khiari, D., Bruchet, A., Gitelman, T., Matia, L., Barrett, S., Suffett, I.H, et al. (1999). Distribution Generated Taste – and – Odor Phenomenon. *Water Science Technology 31(11),* 129-133.

Kim, J., & G. Feree. (1981). Standardization in Causal Analysis. *Sociological Methods Research 10(2),* 187-210.

Kleczyk, E. J., & Bosch, D. J. (2006). A Tiny Hole in Drinking Water Pipes Can Mean a Lot of Problems. *Virginia Water Central 38*. Virginia Water Resources Research Center. Virginia Tech. Blacksburg, Virginia. 8-10. Available online at www.vwrrc.vt.edu/central. Accessed May 2007.

Kleczyk, E. J. &. Bosch, D. J. (2006). *Corrosion in Home Plumbing Systems: Assessment of Costs and Causal Factor*. The 2006 AREUEA Mid-Year Meeting. Washington, D.C. Available online at www.areuea.org. Accessed May 2007.

Kleczyk, E. J. & Bosch, D. J. (2006). *Causal Factors and Costs of Home Plumbing Corrosion: An Investigation of Sample Selection Bias*. The American Agricultural Economics Association Annual Meeting. Long Beach, California. Available online at http://agecon.lib.umn.edu/cgi-bin/view.pl . Accessed May 2007.

Kleczyk, E. J., Bosch, D. J., Dwyer, S., Lee, J., &. Loganathan, G. V. (2005). Maryland Home Drinking Water Assessment. *The Proceedings of the 2005 Virginia Water Research Symposium.* Virginia Water Resources Research Center. Virginia Tech, Blacksburg, Virginia. 104-113. Available online at www.vwrrc.vt.edu. Accessed May 2007.

Kleczyk, E.J., Bosch, D. J., & Lee, J. (2006). Corrosion in Home Drinking Water Infrastructure: Assessment of Causal Factors. *The Proceedings of the 2006 Virginia Water Research Symposium*. Virginia Water Resources Research Center. Virginia Tech. Blacksburg, Virginia. Available online at www.vwrrc.vt.edu. Accessed May 2007.

Kleczyk, E. J., Tanellari, E., & Bosch, D. (2006). *Corrosion in Home Drinking Water Infrastructure: Assessment of Causal Factors, Costs, and Willingness to Pay*. 2006 Water Quality Technology Conference. American Water Works Association. Denver, Colorado.

Loganathan, G.V., & Lee., J. (2005). Decision Tool for Optimal Replacement of Plumbing Systems. *Journal of Civil and Environmental Engineering Systems 22 (4)*, 189-204.

Lee, J., Loganathan, G. V., Dwyer, S., Bosch, D. J., & Kleczyk, E. J. (2005). Preference Trade-Offs in Home Plumbing Material Selection. *The Proceedings of the 2005 Virginia Water Research Symposium*. Virginia Water Resources Research Center. Virginia Tech. Blacksburg, Virginia. 90-95. Available online at www.vwrrc.vt.edu/Proceedings.htm. Accessed March 2007.

Mann, A. (2002). *When Mold Takes Hold. USA WEEKEND.COM*. Available online http://www.usaweekend.com/02_issues/020721/020721moldapt.html. Accessed April 2006.

Maryland Task Force Study. (2004). *Final Report (2004)*. Annapolis, Maryland.

NAHB Research Center. (1992). *Investigative Report of Copper Plumbing Failures Under Concrete Slab*. Prepared for National Association of Home Builders, Washington, D.C.

NIAQI. (2004). *Health Risks Associated with Mold Exposure*. Available online at http://www.iaq-education.com/iaq-health.html. Accessed May 2005.

Nguyen, C. K. (2005). *Interactions between Copper and Chlorine Disinfectants: Chlorine Decay, Chloramine Decay, and Copper Pitting*. Masters' Thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

National Research Council (NRC). (2000). *Copper in Drinking Water.* National Academies Press. Washington, DC. Available online at http://books.nap.edu/books/0309069394/html. Accessed January 2008.

National Science Foundation (NSF). (2007). *Life Safer*. Available online at http://www.nsf.org. Accessed May 2006.

Olivares M., Araya M., Pizarro F., & Uauy R. (2001). Nausea Threshold in Apparently Healthy Individuals Who Drink Fluids Containing Graded Concentrations of Copper. *Regulatory and Toxicological Pharmacology 33*. 271-275.

Olivares M., Pizarro F., Speisky H., Lönnerdal B., & Uauy R. (1998). Copper in Infant Nutrition: Safety of WHO Provisional Guideline Value for Copper Content of Drinking Water. *Journal of Pediatric Gastroenterological Nutrition 26*, 251-257.

Pandit, A., & Bhave, S. (1996). Present Interpretation of the Role of Copper in Indian Childhood Cirrhosis. *American Journal of Clinical Nutrition 63,* 830S-835S.

Pizarro, F. Olivares, M, Araya, M., Gidi, V., & Uary, R. (2001). Gastrointestinal Effects Associated with Soluble and Insoluble Copper in Drinking Water. *Environmental Health Perspectives 190*, 949-952.

Pizarro F., Olivares M., Uauy R., Contreras P., Rebelo A., & Gidi V. (1999). Acute Gastrointestinal Effects of Graded Levels of Copper in Drinking Water. *Environmental Health Perspectives 107*, 117-121.

Powers, K. (2000). *Aging and Copper Corrosion By-Product Release: Role of Common Anions, Impact of Silica and Chlorine, and Mitigating Release in New Pipe.* M.S. Thesis. Virginia Polytechnic Institute and State University, Blacksburg, Virginia.

Rodriguez, G. (2007). *Logit Models for Binary Data.* Available online at: www.data.princeton.edu. Accessed November 2008.

Rossum J. R. (1985). Pitting in Copper Water Tubing, *Journal the AWWA, Research and Technology 77(2)*, 70-73.

Rushing, J. C., & Edwards, M. (2004). Effect of Aluminum Solids and Free $Cl_2$ on Copper Pitting. *Corrosion Science 46(12),* 3069-3088.

Scardina, P., Edwards, M., Bosch, D. J., Loganathan, G. V., & Dwyer, S. K. (2007). *Non-Uniform Corrosion in Copper Piping—Assessment*. Final Completion Report: 172 AwwaRF Research Grant # 3015. Blacksburg, Virginia, Virginia Tech.

Schock, M. R. (1999). Reasons for Corrosion Control other than the Lead and Copper Rule. *Journal of the New England Water Works Association 113(2),* 2-28.

Smits, J. (2003). *Estimating the Heckman Two-Step Procedure to Control for Selection Bias with SPSS*. Available online at http://home.planet.nl/~smits.jeroen. Accessed April 2006.

Suffet, I.H., Mallevialle, J., & Kawczynski, E. (1995). *Advances in Taste - and – Odor Treatment and Control*. AWWARF, Lyonnaise des Eaux-Dumez, Denver, Colorado.

Sweeney, K. (2006). *Implementing and Interpreting Sample Selection Models*. Available online at http://polisci.osu.edu/prl/Selection%20Models.pdf. Accessed September 2007.

Tchobanoglous, G., & Schroeder, E. D. (1985). *Water Quality.* Reading, Massachusetts: Addison-Wesley Publishing Company.

UCLA Academic Technology Services. (2005). *Annotated SPSS Output Logistic Regression*. Available online at www.ats.ucla.edu/STAT/SPSS/output/logistic.htm. Accessed April 2006.

U.S. Environmental Protection Agency (EPA). (1992). *Drinking Water Standards for Regulated Contaminants*. Available online at http://www.epa.gov/safewater/therul.html. Accessed April 2006.

Washington Suburban Sanitary Commission (WSSC). (n. d.). *Pinhole Leaks in Copper Pipes*. Available online at http://www.wssc.dst.md.us/copperpipe/pinhole_charts.cfm. Accessed April 2006.

Weast, R. C. (1977). *Handbook of Chemistry and Physics: Physical Constants of Inorganic Compounds.* 57th Edition. Cleveland, Ohio: CRC Press.

World Health Organization (WHO). (1996). *Trace Elements in Human Nutrition and Health.* Geneva, Switzerland.

Willis-Walton, S. (2006). *Homeowner Pinhole Leak and Drinking Water Survey*. Virginia Tech Center for Survey Research.

Willis-Walton, S. (2006). *Plumbers Survey: Examination of Pinhole Leaks in the United States and Targeted Geographic Areas*. Virginia Tech Center for Survey Research.

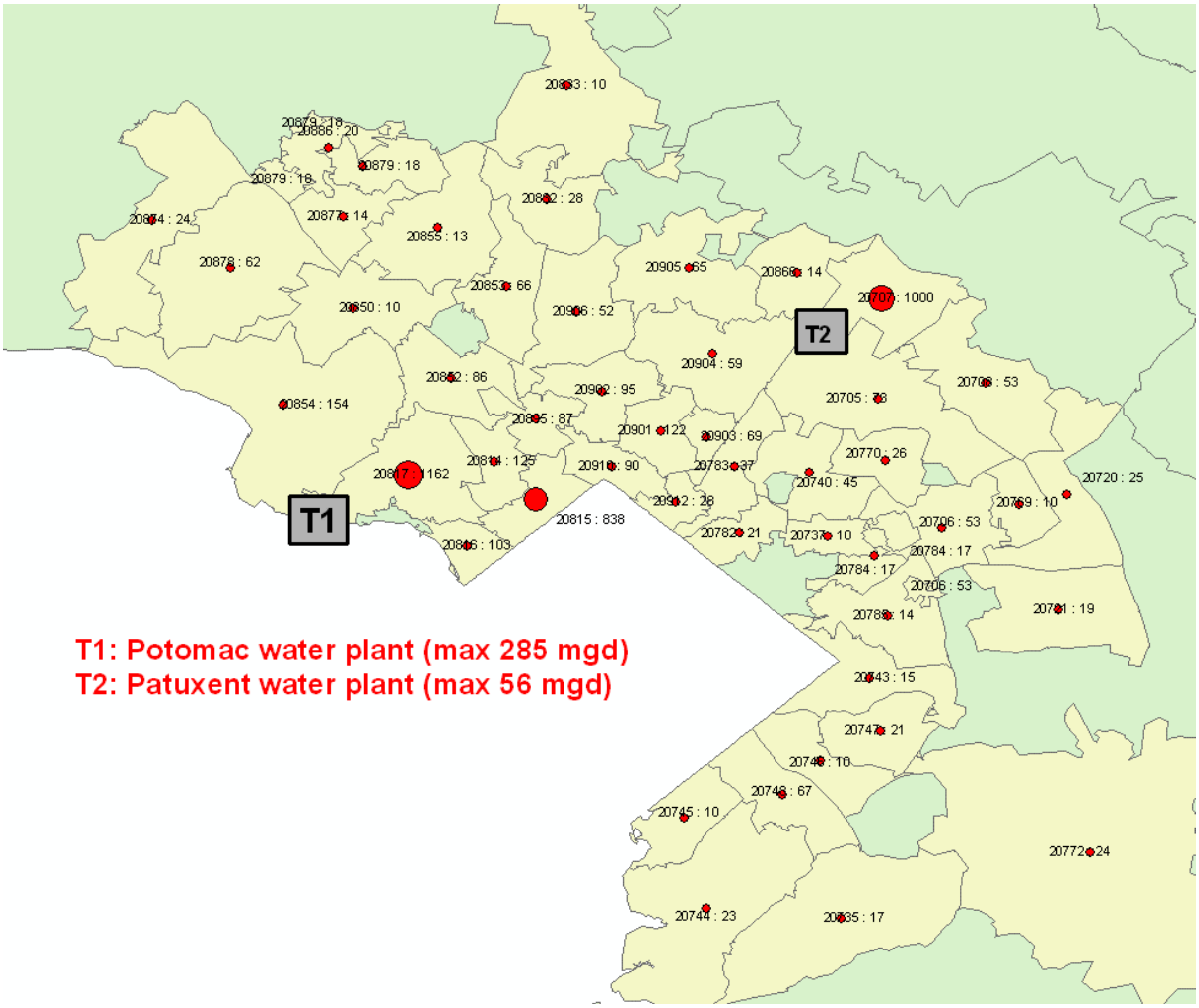
Xiao, W. (2004). *Effect of Source Water Blending on Copper Release in Pipe Distribution System: Thermodynamic and Empirical Models*. Doctoral Dissertation. University of Central Florida, Orlando, Florida.

Zacarias, I., Yanez, C. G., Arya, M., Oraka, C., Olivares, M., & Uauy, R. (2001). Determination of the Taste Threshold of Copper in Water. *Chemical Senses 26*, 85-89.

## 2.8 Appendix

### Appendix A

Figure 2.1: The Maryland Suburbs of Greater Washington, D.C. Area




*Note:* Figure 2.1 represents the Maryland suburbs of Washington D.C. There are two water treatment plants governed by WSSC: Potomac and Patuxent. The red dots on the map represent the number of surveys send out to each zip code in the study area. Zip codes 20817, 20707, and 20815 had the highest number of surveys sent due to the highest frequency of pinhole leak occurrences.

Table 2.1: Number of Pinhole Leaks Reported by Survey Respondents

| Number of leaks | Respondents by leak category (unweighted) | Percent (unweighted) | Respondents by leak category (weighted) | Percent (weighted) |
|---|---|---|---|---|
| 1-2 | 201 | 44.9 | 172.5 | 58.5 |
| 3-4 | 97 | 21.7 | 45.9 | 15.6 |
| 5-6 | 49 | 10.9 | 20.4 | 6.9 |
| 7-10 | 37 | 8.3 | 17.5 | 5.9 |
| > 10 | 25 | 5.6 | 6.9 | 2.4 |
| Don't know | 21 | 4.7 | 16.3 | 5.5 |
| Missing | 18 | 4.0 | 15.2 | 5.2 |
| Total | 448 | 100.0 | 294.7 | 100.0 |

*Note:* Number of pinhole leaks reported by mail survey respondents in the Maryland suburbs of Washington D.C. A total of 448 respondents reported experiencing pinhole leaks or 40% of 1,120 respondents on an unweighted basis. Unweighted responses represent the simple frequencies of respondents in each category of leak counts. Certain zip codes with higher leak reports were sampled more heavily; consequently, weighted responses represent a correction for oversampling. Samples are weighted by the number of people over 18 represented by each survey sample in a given zip code.

| Table 2.2: The Year Pinhole Leaks Were First Reported | | | | |
|---|---|---|---|---|
| Year pinhole leaks were first reported | Respondents by year leaks first reported (unweighted) | Percent (unweighted) | Respondents by year leaks first reported (weighted) | Percent (weighted) |
| Since 2000 | 204 | 45.5 | 126.4 | 42.9 |
| 1995 to 1999 | 124 | 27.7 | 75.1 | 25.5 |
| 1990 to 1994 | 50 | 11.2 | 37.3 | 12.6 |
| 1980 to 1989 | 29 | 6.5 | 18.1 | 6.1 |
| 1970 to 1979 | 7 | 1.6 | 1.5 | 0.5 |
| 1960 to 1969 | 1 | 0.2 | 0.1 | 0.0 |
| Before 1960 | 1 | 0.2 | 0.5 | 0.2 |
| Don’t know | 14 | 3.1 | 19.7 | 6.7 |
| Missing | 18 | 4.0 | 16.0 | 5.4 |
| Total | 448 | 100.0 | 294.7 | 100.0 |

*Note:* Year pinhole leaks first occurred in residents’ homes as reported by mail survey respondents in the Maryland suburbs of Washington D.C. Unweighted responses represent the simple frequencies of respondents in each period. Certain zip codes with higher leak reports were sampled more heavily; consequently, weighted responses represent a correction for oversampling. Samples are weighted by the number of people over 18 represented by each survey sample in a given zip code.

Table 2.3: Age of the Dwelling in which Pinhole Leaks Occurred

| Year house or apartment building was built | Total number of respondents by year residence built (unweighted) | Percent (unweighted) | Respondents with leaks by year residence built (unweighted) | Percent (unweighted) | Respondents with leaks by year residence built (weighted) | Percent (weighted) |
|---|---|---|---|---|---|---|
| Since 2000 | 26 | 2.3 | 1 | 0.2 | 0.1 | 0 |
| 1990-1999 | 52 | 4.6 | 8 | 1.8 | 8.7 | 2.9 |
| 1980-1989 | 136 | 12.1 | 52 | 11.6 | 52.1 | 17.7 |
| 1970-1979 | 124 | 11.1 | 47 | 10.5 | 45.5 | 15.4 |
| 1960-1969 | 225 | 20.1 | 82 | 31.3 | 59.9 | 20.3 |
| 1950-1959 | 273 | 24.4 | 140 | 15.4 | 75.8 | 25.7 |
| 1940-1949 | 135 | 12.1 | 69 | 8.3 | 34.3 | 11.6 |
| Before 1940 | 93 | 8.3 | 37 | 8.3 | 9.8 | 3.3 |
| Don't know | 51 | 4.6 | 10 | 2.2 | 8.5 | 2.9 |
| Missing | 5 | 0.4 | 2 | 0.4 | .2 | 0.1 |
| Total | 1120 | 100.0 | 448 | 100.0 | 294.7 | 100.0 |

*Note:* Year dwelling was constructed as reported by mail survey respondents in the Maryland suburbs of Washington D.C. Year of construction is reported separately for all respondents and for respondents with pinhole leaks. Unweighted responses represent the simple frequencies of respondents in each period. Certain zip codes with higher leak reports were sampled more heavily; consequently, weighted responses represent a correction for oversampling. Samples are weighted by the number of people over 18 represented by each survey sample in a given zip code.

| Table 2.4: Level of Dwelling on which Pinhole Leaks Occurred | | | | |
|---|---|---|---|---|
| Level of leak in dwelling | Respondents with leaks at indicated level (unweighted) | Percent (unweighted) | Respondents with leaks at indicated level (weighted) | Percent (weighted) |
| Under slab /underground | 16 | 3.6 | 18.3 | 6.2 |
| Basement | 304 | 67.9 | 179.8 | 61.0 |
| First floor | 173 | 38.6 | 109.3 | 37.1 |
| Second floor | 95 | 21.2 | 3.9 | 18.3 |
| Third floor | 8 | 1.8 | 0.8 | 0.3 |
| Fourth floor or higher | 7 | 1.6 | 1.4 | 0.5 |
| Don't know | 17 | 3.8 | 29.3 | 9.9 |
| Missing | 18 | 4.0 | 16.1 | 5.5 |

*Note:* Level of dwelling on which pinhole leaks occurred as reported by mail survey respondents in the Maryland suburbs of Washington D.C. Percent of respondents with leaks on each dwelling level is calculated based on the number of respondents with a leak on each level divided by the total number of respondents with leaks (448 unweighted, 294.7 weighted). Percents do not sum to 100 because some respondents indicated more than one level of dwelling had leaks. Unweighted responses represent the simple frequencies of respondents in each period. Certain zip codes with higher leak reports were sampled more heavily; consequently, weighted responses represent a correction for oversampling. Samples are weighted by the number of people over 18 represented by each survey sample in a given zip code.

Table 2.5: Effects of Selected Variables on Probability of Pinhole Leak Occurrences

| | Base model | | | Observations weighted by | | |
|---|---|---|---|---|---|---|
| Variables | Percent probability | Odds ratio | P-value | Percent probability | Odds ratio | P-value |
| Water age | -36.69 | 0.58 | 0.00 | -40.39 | 0.68 | 0.00 |
| Copper pipes | 68.17 | 2.14 | 0.00 | 72.79 | 2.68 | 0.00 |
| History of pipe failure | 68.03 | 2.13 | 0.00 | 68.64 | 2.19 | 0.00 |
| Plumbing material age | -49.90 | 1.00 | 0.23 | -49.98 | 0.99 | 0.82 |
| Water source (Patuxent plant) | 85.74 | 6.10 | 0.00 | 85.93 | 6.10 | 0.00 |

*Note:* In the base model, variables are weighted by sample weights, which correct for oversampling zip codes with higher leak reports. Model log Likelihood = 1103; Pseudo $R^2$ = 0.078; and n (unweighted) = 879. In the second model, observations are weighted by sample weights multiplied times a weight for age of dwelling. The 1950-1969 period was made the base period and respondents living in homes from other periods were weighted in inverse proportion to the ratio of homes in their age category to the base number of homes. Log Likelihood = 1099; Pseudo $R^2$ = 0.09; and n (unweighted) = 879.

| Table 2.6: Classification Table of Pinhole Leak Prediction Accuracy | | | |
|---|---|---|---|
| Observed | Predicted | | Percentage Correct |
| | No pinhole leaks | Pinhole leaks | |
| No pinhole leaks | 254 | 202 | 55 |
| Pinhole leaks | 113 | 310 | 73 |
| Overall Percentage | | | 64 |

*Note*: Comparison of predicted and observed pinhole leak occurrences. A leak is predicted (not predicted) if a probability of pinhole leaks greater than or equal to 0.5 (less than 0.5) is estimated for a respondent by the weighted logistic model. Observation frequencies reported in the table are unweighted.

Table 2.7: Costs of Repairing Pinhole Leaks and Associated Damage

| Amount of money | Number of observations (unweighted) | Percent (unweighted) | Number of observations (weighted) | Percent (weighted) |
|---|---|---|---|---|
| Less than $100 | 58 | 12.9 | 55.1 | 18.7 |
| $100 to $500 | 112 | 25.0 | 89.2 | 30.3 |
| $501 to $1,000 | 71 | 15.8 | 33.7 | 11.4 |
| $1,001 to $3,000 | 66 | 14.7 | 32.6 | 11.1 |
| $3,001 to $5,000 | 45 | 10.0 | 22.5 | 7.6 |
| More than $5,000 | 43 | 9.6 | 14.6 | 4.9 |
| Don't know | 34 | 7.6 | 29.7 | 10.1 |
| Missing | 19 | 4.2 | 17.4 | 5.9 |
| Total | 448 | 100.0 | 294.7 | 100.0 |

*Note:* Costs of repairing pinhole leaks and associated damage reported by mail survey respondents in the Maryland suburbs of Washington D.C. Unweighted responses represent the simple frequencies of respondents in each cost category. Certain zip codes with higher leak reports were sampled more heavily; consequently, weighted responses represent a correction for oversampling. Samples are weighted by the number of people over 18 represented by each survey sample in a given zip code.

| Table 2.8: Time Spent Dealing with Pinhole Leak Problems | | | | |
|---|---|---|---|---|
| Number of hours | Number of respondents (unweighted) | Percent (unweighted) | Number of respondents (weighted) | Percent (weighted) |
| Less than 10 | 173 | 38.6 | 125.4 | 42.6 |
| 11-20 | 88 | 19.6 | 46.2 | 15.7 |
| 21-40 | 60 | 13.4 | 31.0 | 10.5 |
| 41-80 | 45 | 10.0 | 33.5 | 11.4 |
| More than 80 | 26 | 5.8 | 11.5 | 3.9 |
| Don’t know | 36 | 8.0 | 29.8 | 10.1 |
| Missing | 20 | 4.5 | 17.2 | 5.8 |
| Total | 448 | 100 | 294.7 | 100 |

*Note:* Time spent dealing with pinhole leaks problems reported by mail survey respondents in the Maryland suburbs of Washington D.C. Unweighted responses represent the simple frequencies of respondents in each time category. Certain zip codes with higher leak reports were sampled more heavily; consequently, weighted responses represent a correction for oversampling. Samples are weighted by the number of people over 18 represented by each survey sample in a given zip code.

| Table 2.9. Total Costs of Repairing Pinhole Leaks and Associated Damage (Least Squares Regression) | | | | | | |
|---|---|---|---|---|---|---|
| | OLS regression without selection bias correction | | | GLS regression with selection bias correction | | |
| Variables | Unstandardized coefficients | Standardized coefficients | P-value | Unstandardized coefficients | Standardized coefficients | P-value |
| Constant | 293.83 | | 0.02 | 784.85 | | 0.88 |
| Number of leaks | 665.45 | 0.33 | 0.00 | 304.85 | 0.16 | 0.00 |
| Copper pipes | 340.46 | 0.09 | 0.02 | 625.38 | 0.17 | 0.00 |
| Iron pipes | -25.45 | -0.03 | 0.37 | -9.95 | -0.01 | 0.71 |
| Cold water pipes | -154.89 | -0.12 | 0.58 | -279.33 | -0.23 | 0.33 |
| Leak under slab and in basement | -842.16 | -0.69 | 0.00 | -568.27 | -0.48 | 0.01 |
| Leak on the first floor | -109.51 | -0.09 | 0.67 | -185.46 | -0.16 | 0.41 |
| Leak on the second floor | 1011.56 | 0.82 | 0.00 | 954.52 | 0.80 | 0.00 |
| Plumbing material age | -0.99 | -0.02 | 0.60 | -1.96 | -0.06 | 0.24 |
| Source of water | -286.12 | -0.06 | 0.07 | -1296.14 | -0.32 | 0.00 |
| Selection bias correction | N/A | N/A | N/A | 1390.32 | 0.52 | 0.0 |

*Note*: The correction for selection bias in the GLS model is based on an estimate of the effects of unobserved factors contributing to misclassification of observations with or without pinhole leaks. The unstandardized coefficients are explained as the dollar value increase/decrease in total costs due to change in the independent variables. Regression without selection bias correction: $R^2$ = 0.344; F-stat = 10.98; and n (unweighted) = 392. Regression with selection bias correction: $R^2$ = 0.294; F-stat = 27.21; and effective unweighted sample size after selection bias correction = 367.

Table 2.10. Total Costs of Repairing Pinhole Leaks and Associated Damage in the Age-Weighted Model (Least Squares Regression) with Observations Weighted by Age Category of Dwelling

| Variables | OLS regression without accounting for selection bias correction | | | GLS regression with accounting for selection bias correction | | |
|---|---|---|---|---|---|---|
| | Unstandardized coefficients | Standardized coefficients | P-value | Unstandardized coefficients | Standardized coefficients | P-value |
| Constant | 293.83 | | 0.02 | 677.47 | | 0.00 |
| Number of leaks | 665.45 | 0.34 | 0.00 | 312.20 | 0.16 | 0.00 |
| Copper pipes | 340.47 | 0.09 | 0.02 | 724.67 | 0.20 | 0.00 |
| Iron pipes | -25.46 | -0.03 | 0.37 | -10.12 | -0.01 | 0.69 |
| Cold water pipes | -154.90 | -0.13 | 0.58 | -268.98 | -0.23 | 0.34 |
| Leak under slab and in basement | -842.16 | -0.69 | 0.00 | -569.41 | -0.48 | 0.01 |
| Leak on the first floor | -109.52 | -0.09 | 0.67 | -195.94 | -0.17 | 0.38 |
| Leak on the second floor | 1011.57 | 0.82 | 0.00 | 963.93 | 0.81 | 0.00 |
| Plumbing material age | -1.00 | -0.03 | 0.60 | -1.52 | -0.05 | 0.36 |
| Source of water | -286.12 | -0.07 | 0.07 | -1290.79 | -0.31 | 0.00 |
| Selection bias correction | N/A | N/A | N/A | 1401.95 | 0.53 | 0.00 |

*Note:* Observations are weighted by sample weights multiplied times a weight for age of dwelling. The unstandardized coefficients are explained as the dollar value increase/decrease in total costs due to change in the independent variables. Age weighting is based on homes built between 1950-1969 period as the base period. Respondents living in homes built in other periods were weighted in inverse proportion to the ratio of homes in their age category to the base number of homes. Regression without selection bias correction: $R^2$ = 0.344; F-stat = 10.30; and n (unweighted) = 392. Regression with selection bias correction: $R^2$ = 0.295; F-stat = 27.35; and effective unweighted sample size after selection bias correction = 367.

## Appendix B

Maryland Home Drinking Water Assessment
Attn: Darrell Bosch
Department of Agricultural and Applied
Economics (0401)
Virginia Polytechnic Institute and State
University
Blacksburg, VA 24061

**Maryland Home Drinking Water Assessment**

July, 2004

Dear Water Consumer:

Safe, high quality drinking water is important to all of us. The accompanying survey asks about your attitudes and experiences regarding your drinking water and your home plumbing system, which includes hot and cold water pipes. With your help, we hope to learn more about consumer concerns related to their home drinking water and water supply system.

The survey is being conducted by researchers at Virginia Tech and is funded by a grant from the National Science Foundation. In addition, survey results will be shared with the Maryland Copper Pinhole Leak Task Force and will help them in planning strategies to deal with home plumbing corrosion. If you would like more information about the research project, please visit the project website at *http://www.vwrrc.vt.edu/Outreach/sps/project1.htm*.

Your responses are very important to the success of this survey. You are one of a select group of consumers chosen to represent water utility customers. Be assured that your responses will be treated confidentially. A code number has been inserted on the survey to remind nonrespondents to complete the survey. The responses from all consumers will be compiled and analyzed together.

The survey will take about ten minutes to complete. ***After completing the survey, please return it in the enclosed, postage-paid envelope or send it to the address above.*** If you would like to receive a summary of survey results, please include your name and contact information on the enclosed card and return it with the survey or mail it separately. If you have questions, please contact me at 540-231-5265 or by email at *bosch@vt.edu*.

Thank you very much for your help with this survey.

Sincerely,

Darrell Bosch

Darrell Bosch
Professor

**The first seven questions are about your home drinking water. (The water itself)**

1. If you were concerned about your home drinking water quality, who would you contact for information? (Please make up to 3 choices.)
   - ☐ Doctor's office
   - ☐ Extension service
   - ☐ Health department
   - ☐ Landlord
   - ☐ Neighbor or friend
   - ☐ Neighborhood association
   - ☐ Plumber or home supply store
   - ☐ Water utility company
   - ☐ Other (please specify) ______________________

2. How satisfied are you with the quality (taste, smell, and appearance) of your home drinking water?
   - ☐ Very satisfied
   - ☐ Somewhat satisfied
   - ☐ Somewhat dissatisfied
   - ☐ Very dissatisfied

3. How satisfied are you with the safety of your home drinking water?
   - ☐ Very satisfied
   - ☐ Somewhat satisfied
   - ☐ Somewhat dissatisfied
   - ☐ Very dissatisfied
   - ☐ Don't know

4. How important is **each item below** in describing drinking water from your home plumbing system. (Please check one box for each item.)

| | Very | Some-what | Not very | Not at all |
|---|---|---|---|---|
| Clearness | ☐ | ☐ | ☐ | ☐ |
| Cost | ☐ | ☐ | ☐ | ☐ |
| Health, Safety | ☐ | ☐ | ☐ | ☐ |
| Doesn't damage plumbing | ☐ | ☐ | ☐ | ☐ |
| Taste, Smell | ☐ | ☐ | ☐ | ☐ |

5. Do you use any of the following to improve your home drinking water? (Please check **all** that apply.)
- [ ] A filter for the entire home water system
- [ ] A filter as part of the refrigerator
- [ ] Commercially delivered drinking water
- [ ] Water softener
- [ ] A pitcher or bottle that filters drinking water
- [ ] A filter on faucet or under kitchen sink
- [ ] Store purchased drinking water
- [ ] Ultra violet (UV) system
- [ ] Don't use any of the above (skip to question 7)

6. If you use any item in question 5, why do you do so? Please check the **most important** reason.
- [ ] Improve safety of my drinking water
- [ ] Improve the clearness of my drinking water
- [ ] Improve taste or smell of my drinking water
- [ ] Sales information or presentation
- [ ] Other (please describe) ______________________

7. Where do you get your drinking water?
- [ ] Washington Suburban Sanitation Commission (WSSC)
- [ ] Other utility (please specify—e.g. Rockville City, Bowie City ______________________
- [ ] Private well
- [ ] Don't know

**The next questions are about your plumbing system for home drinking water. Drinking water pipes include hot and cold water pipes throughout your home.**

8. What kind of drinking water pipes do you have in your home? Drinking water pipes are typically 1/2 to 1 inch in diameter. (Please check **all** that apply.)
- [ ] Copper
- [ ] Iron (cast or galvanized)
- [ ] Plastic (CPVC, PEX, etc.)
- [ ] Other (please specify) ______________________
- [ ] Don't know

9. Since your dwelling was built, have some of your drinking water pipes been replaced or new pipe sections installed?
- ☐ Yes, all have been replaced
- ☐ Yes, some have been replaced or added
- ☐ No (skip to question 11)
- ☐ Don't know (skip to question 11)

10. If yes, when were the pipes replaced or new ones added? (Please check **all** that apply.)
- ☐ Since 2000
- ☐ 1995 to 1999
- ☐ 1990 to 1994
- ☐ 1980 to 1989
- ☐ 1970 to 1979
- ☐ 1960 to 1969
- ☐ Before 1960
- ☐ Don't know

11. Which of the following best describes your dwelling?
- ☐ Single family home that I own
- ☐ Apartment/condo that I own
- ☐ Single family home that I rent
- ☐ Apartment/condo that I rent
- ☐ Other (please specify) ____________________

12. About what year was your house or apartment building built?
- ☐ Since 2000
- ☐ 1990 to 1999
- ☐ 1980 to 1989
- ☐ 1970 to 1979
- ☐ 1960 to 1969
- ☐ 1950 to 1959
- ☐ 1940 to 1949
- ☐ Before 1940
- ☐ Don't know

13. Have you ever had a pinhole leak (small hole in pipe) in the hot or cold water pipes in your current home?
- ☐ Yes
- ☐ No
- ☐ Don't know

14. Have you had other kinds of failure in your drinking water pipes?

- ☐ Yes (Please specify)

- ☐ No

15. Have you ever had the quality of the drinking water in your current dwelling tested?

- ☐ Yes
- ☐ No

16. What is the best way to inform you about issues that may affect your home drinking water? (Please choose only **one**.)

- ☐ Automated phone system
- ☐ Cable TV community bulletin board
- ☐ E-mail
- ☐ Local government newsletter
- ☐ Local water utility newsletter
- ☐ Newspaper
- ☐ Radio news
- ☐ TV news
- ☐ Water bill insert
- ☐ Website

17. What year did you move into your current home? ____________

18. What is your zip code? ______________

19. What other issues or concerns you would like to share about your home drinking water? Please print.

**If you have had pinhole leaks in your current home, please answer the remaining questions.**
**If you have never had pinhole leaks, you have completed the survey. Thank you!**

20. About what year did you first have pinhole leaks that required repair in your current home?

- [ ] Since 2000
- [ ] 1995 to 1999
- [ ] 1990 to 1994
- [ ] 1980 to 1989
- [ ] 1970 to 1979
- [ ] 1960 to 1969
- [ ] Before 1960
- [ ] Don't know

21. In your current home, what is the total number of pinhole leaks requiring repair that you have had?

- [ ] 1-2
- [ ] 3-4
- [ ] 5-6
- [ ] 7-10
- [ ] More than 10 (please specify) ____________________
- [ ] Don't know

22. Did you report the leaks to any of the following? (Please check **all** that apply.)

- [ ] Health department
- [ ] Insurance company
- [ ] Landlord
- [ ] Neighborhood association
- [ ] Neighbor or friend
- [ ] Plumber or home supply store
- [ ] Water utility
- [ ] Did not report leaks to anyone
- [ ] Other (please specify) ________________________________

23. Were the leaks in a horizontal or vertical pipe?

- [ ] Horizontal pipes only
- [ ] Vertical pipes only
- [ ] Both
- [ ] Don't know

24. What type of pipes had pinhole leaks?

- ☐ Copper
- ☐ Plastic (CPVC, PEX, etc.)
- ☐ Iron (cast or galvanized)
- ☐ Don't know

25. In what area along the length of the pipe did leaks occur?

- ☐ Near a fitting
- ☐ Along a straight length of pipe
- ☐ Both
- ☐ Don't know

26. What level of your dwelling were leaks located? (Please check **all** that apply.)

- ☐ Under slab or underground
- ☐ Basement
- ☐ First floor
- ☐ Second floor
- ☐ Third floor
- ☐ Fourth floor or higher
- ☐ Don't know

27. Did the leaks occur in cold or hot water pipes?

- ☐ Cold water pipes
- ☐ Hot water pipes
- ☐ Both
- ☐ Don't know

28. How stressful has it been to deal with the pinhole leaks, damage and repairs in your home?

- ☐ Very stressful
- ☐ Somewhat stressful
- ☐ A little stressful
- ☐ Not stressful

29. About how much money was spent by you, your landlord, and/or your insurance company to make repairs to the plumbing system and any other damage to your current home from pinhole leaks?

- [ ] Less than $100
- [ ] $100 to $500
- [ ] $501 to $1,000
- [ ] $1,001 to $3,000
- [ ] $3,001 to $5,000
- [ ] More than $5,000 (please specify) ______________________
- [ ] Don't know

30. About how many hours have you spent dealing with pinhole leak problems in your current home?

- [ ] Less than 10
- [ ] 11-20
- [ ] 21-40
- [ ] 41-80
- [ ] More than 80 (please specify) ___________________
- [ ] Don't know

31. Have you any other comments about pinhole leak problems in home plumbing? Please print.

**Thank you for completing this survey!**

## Appendix C

Virginia Tech
VIRGINIA POLYTECHNIC INSTITUTE AND STATE UNIVERSITY

**Institutional Review Board**

Dr. David M. Moore
IRB (Human Subjects)Chair
Assistant Vice Provost for Research Compliance
CVM Phase II- Duckpond Dr., Blacksburg, VA 24061-0442
Office: 540/231-4991; FAX: 540/231-6033
email: moored@vt.edu

DATE: April 23, 2004

MEMORANDUM

TO: Darrell Bosch
~~Gordon E. Groover~~ Agricultural & Applied Economics 0401
Sharon K. Dwyer Center for Community Health 0113
G V. Loganathan Civil & Environmental Engineering 0105

FROM: David Moore

SUBJECT: **IRB Exempt Approval:** "Toward Sustainable Materials for Drinking Water Infrastructure" IRB # 04-243

I have reviewed your request to the IRB for exemption for the above referenced project. I concur that the research falls within the exempt status. Approval is granted effective as of April 22, 2004.

cc: File
Department Reviewer James Pease 0401
OSP 0170

# CHAPTER 3

# Conjoint Analysis Study of Households' Preferences for Plumbing Materials

Portions of this chapter were presented at the American Agricultural Economics Association Annual Meeting. Orlando, FL. July 2008.

**Abstract**

Consumers' decisions on plumbing material selection are dictated by various factors, including state and federal regulations, service providers, and individual household preferences. In addition, insurance companies collect information on water damage and plumbing material failures and use this information in decisions regarding the insuring of residential properties. As a result, homeowners have an important stake in finding plumbing systems appropriate for their households; therefore, homeowners should ensure that they are knowledgeable about plumbing material attributes such as price and health impact. In this article, the problem of household preferences for plumbing materials and the tradeoffs between risk of corrosion and the cost of preventive measures are studied. Two surveys of a Southeastern Community were conducted to obtain information on the prevalence of plumbing material failures, householders' experiences with plumbing material failures, the cost of repairs and property damages due to the material failures, and household preferences for corrosion preventive measures and to obtain opinions about a set of hypothetical plumbing materials.

The preferences for home plumbing materials and preventive measures were obtained by employing stated preference elicitation method of Conjoint Analysis (CA). The survey data were analyzed by employing Multinomial and Ordered Logit models. The several model specifications accounted for the socioeconomic characteristics of the households and differences in survey version presented to the respondents. The estimation results were employed in computation of hypothetical plumbing material preference ranking and probabilities of corrosion prevention option selection as well as the estimates of willingness to pay (*WTP)* and compensating variation.

The study results revealed that the cost of corrosion prevention had an impact on prevention decisions while the risk of corrosion had no impact on household decisions. Majority of households preferred to stay with the current plumbing materials rather than install an upgrade to reduce corrosion risk or purchase insurance against corrosion damages. The risk of corrosion had a statistically insignificant impact on individuals' decisions as the provided risk values were not large enough to induce corrosion prevention consideration. In case of preferences for hypothetical plumbing materials, plumbing material described by 'corrosion proof' and 'installation takes about 4 days' attributes was the most preferred among three hypothetical materials. This information on consumer preferences for drinking water infrastructure can be useful not only to individual households but also to policymakers and program and water utilities managers. The information gathered can be used to bridge the gap between public and water experts' perceptions of the problem of corrosion and can be used in the design of effective public policy for corrosion prevention.

**Keywords:** Conjoint Analysis, Plumbing Materials, Corrosion Prevention, Pinhole Leak Corrosion

**JEL Classification Codes:** Q51 and M31

### 3.1 Introduction

Consumers' decisions on plumbing material selection are dictated by various factors, including state and federal regulations, service providers, and individual household preferences. The regulations and standards of the federal, state, and local governments have major impacts on the plumbing material chosen for installation in a private house. For example, the use of plastic plumbing material, PEX, has been approved in all U.S. states except California and Massachusetts where the material installation requires jurisdiction acceptance. Similarly, in some parts of Florida, PEX is preferred due to the seriousness of pinhole leak problems (NSF, 2008). These regulations influence services provided by plumbers, material producers (e.g. pipe manufacturers, interior coating providers), and water utility companies. For example, general contractors are the primary decision-makers of plumbing material installation in new houses while utility companies respond to corrosion threats by adding corrosion inhibitors into drinking water treatment. Consequently, all service providers influence consumer decisions regarding the best plumbing material for the private property.

In addition, insurance companies collect and employ information on water damage and plumbing material failures in determining residential property insurance policy issuance and renewal. The information is stored in a Comprehensive Loss Underwriting Report (CLUE)) that represents a comprehensive database of personal property information relating to insurance claims on private property. It contains personal information on the insured, the claim number, the accident, as well as the amount paid out to the customer. Even an inquiry of home plumbing issues is recorded in the database. Consequently, homeowners may have difficulty in either purchasing or renewing home insurance in the year following any water damage related reports (Longcore, 2008).

Since homeowners have an important stake in finding plumbing systems appropriate for their households, they should rely not only on expert advice but also acquire information on plumbing material attributes such as price, health impact, longevity, and corrosion resistance in order to make informed investment decisions about plumbing systems for their homes. For example, health effects as well as taste, and odor have been found to be the most important factors in consumers' evaluations of plumbing material for home use (Lee et al., forthcoming). Additionally, households are willing to pay up to $4,000 when guaranteed a leak-free plumbing system for 50 years (Kleczyk et al., 2006). Information on consumer preferences for drinking water plumbing attributes can be useful not only to individual households but also to policymakers, program managers, water utilities, and firms with interests in drinking water infrastructure.

The public perceptions of corrosion risk and cost of prevention play a fundamental role in consumers' drinking water decisions. Homeowners' perceptions of risk and cost of prevention may affect households' decisions on plumbing material repairs and replacement as well as the type of material used. When informed about the attributes of each plumbing material alternative, consumers can decide on the most preferred plumbing system. Random utility models are usually adopted in preference elicitation and maximization. The decision of choosing an appropriate plumbing material is based on various plumbing material attributes, such as cost (material cost plus labor and installation cost), health effects, corrosion susceptibility, strength, property real estate values, and behavior in the case of a fire. On the other hand, the perception of risk for plumbing materials can be quantified through the assessment of willingness-to-pay (*WTP*) for a corrosion-free plumbing material or an improvement in the performance of existing plumbing material (Champ et al., 2003).

There are two types of commonly used preference elicitation methods that could be utilized in evaluation of households' preferences for drinking water infrastructure: revealed and stated preferences. The stated preference methods rely on study participants to compare and rank multiple objects. Different types of surveys are employed to elicit people's preferences for given objects. The survey questions are in the form of monetary amounts, choices, ratings, or other indications of preference. Based on respondents' preference valuations, the data is employed in empirical modeling to yield a measure of value for a given object (Champ et al., 2003).

Revealed preference methods draw statistical inferences on values from actual choices made within markets (Champ et al., 2003). The main difference between these two methods is that stated preference approaches mainly rely on primary survey data while revealed preference approaches rely mostly on secondary data and draw their conclusions based on the actual decisions made by people (Champ et al., 2003). Furthermore, as mentioned, stated preference data allows estimation of behavior beyond the range of historical experience while revealed preference data is limited to analyzing a range of behavior in response to a limited range of market changes. Many researchers tend to combine revealed preference and stated preference data as this approach allows for an extension of the behavioral analysis beyond the limited range of historical experience (Whitehead et al., 2005).

Economists utilize the stated preference methods, such as Conjoint Analysis (CA) and Contingent Valuation (CV), and the revealed preference methods, including Travel Cost (TC), Hedonic Prices (HP), and Defensive Behavior and Cost Damage models (DBCD), for non-market valuation analysis, including willingness-to-pay estimation. All of these methods are derived from the Random Utility Framework, which is based on the premises of preference ordering (desirability level of goods and services by individuals) (Champ et al., 2003). As a result, CA's objective is to estimate economic values for a technically divisible set of attributes of studied goods or services while CV elicits economic values placed on goods or services by survey respondents (Champ et al., 2003). TC is typically used to estimate the use values for different tourist sites, HP is employed for market housing valuation, and DBCD assesses either the preventive actions taken against hazardous outcomes or the resources allocated to treat the negative impact of hazardous activities (Champ et al., 2003). On the other hand, Analytic Hierarchy Process (AHP) is widely employed by the Engineering field for determination of the most optimal attributes for given products.

Often, these different preference elicitation methods are compared to each other with regards to the estimated preferences as well as the economic value. The differences in preferences elicited by different techniques are usually tested by validity measures such as construct and internal validity (Champ et al., 2003; Helm et al., 2004). These measures analyze and compare the predictive power of the employed models as well as the final results derived based on each studied method. Finally, the impact of the order of the different preference elicitation questions in a questionnaire on the respondents' preferences might be also investigated, as the survey responses and therefore preferences for corrosion prevention might be affected by the order of questions using different methods (Champ et al., 2003; and Boyle personal communication, 2007).

For the purpose of this drinking water infrastructure study, only the stated preferences framework, including CA, CV, and AHP methods, was adopted to examine issues related to the drinking water home delivery as well as to elicit information for home plumbing materials and the tradeoffs between risk of corrosion and financial spending for preventive options. Two surveys of residents residing in a Southeastern Community were conducted. The Southeastern Community was established in the 1980s and includes 3,300 homes with a total 6,600 residents. Most of the resident population is retired, so the community is a rather homogenous group. The community has experienced a high rate of pinhole leaks with the first incidents of plumbing failure being reported in 2001. The first survey elicited information on the prevalence of pinhole leaks and other plumbing material failures, households' experiences with plumbing material failures, the cost of repairs and property damages due to the material failures, and household preferences for corrosion preventive measures.

The first Southeastern Community survey employed Conjoint Analysis and Contingent Valuation to elicit the tradeoffs between risk of corrosion and the dollar amount spent on preventive measures against corrosion. In case of the CA questions, respondents were asked to choose between upgrading plumbing material, purchasing an insurance policy against corrosion while staying with the current plumbing material, and staying with current plumbing material only based on the provided risk of corrosion values as well as the dollar amount needed to pay for each option. In case of CV, the respondents were asked whether they would pay a certain amount of money to purchase an insurance policy against corrosion.

The follow-up survey of the Southeastern Community, sent only to those residents who agreed to participate in future studies, related to the plumbing material issues and elicited information on households' preferences for a set of hypothetical plumbing materials. CA and AHP methods were employed to elicit preferences (Helm et al., 2004). In case of the CA method, the survey respondents presented their preferences for plumbing materials based on the description of each material while in the AHP case, they pair-wise compared the attributes of each plumbing material while holding other attributes constant.

As it is important to learn household perceptions and preferences for drinking water infrastructure, the study objective is to investigate homeowners' preferences for plumbing materials as well as preventive techniques against corrosion (i.e. plumbing material upgrade, purchase of insurance against corrosion while staying with the current plumbing material, or staying with the current plumbing material with no action taken to reduce risk of corrosion damage) based on households' experiences with plumbing material failures. Conjoint analysis is employed in the estimation of household preferences as well as the factors impacting the decisions. Eftila Tanellari and Juneseok Lee have been investigating the preferences for prevention options and plumbing materials by analyzing the responses to Contingent Valuation and Analytic Hierarchy Process questions. The detailed AHP analysis and preference for plumbing materials' results are presented in Lee (2008).

The objective of eliciting household preferences for plumbing materials and preventive measures against corrosion is fulfilled by employing several regression methods including Multinomial and Ordered Logit models. Several model specifications are estimated, including equations with socioeconomic characteristics of the households as well as different survey versions presented to the households. Based on the regression results, willingness-to-pay and compensating variation measures are computed to assess the financial impact of corrosion on household preferences.

The study of water infrastructure issues in the Southeastern Community will present results for the preferences for preventive measures against corrosion as well as the preferences for hypothetical plumbing materials. In addition, the *WTP* estimates will inform of the households' willingness-to-pay for corrosion prevention when the risk of corrosion is known to households. Besides estimating the preferences for plumbing materials, the importance of pipe attributes will be also obtained to learn the most important factors in the decision-making process of plumbing material selection. The knowledge gained from the research presented in this study could be helpful in the design of public policy aimed at corrosion prevention. The research will provide information to federal and state officials, plumbers, plumbing material manufacturers, and utility company managers on the financial burden individual households are willing to take on to avoid corrosion. In addition, the study results should help in bridging the gap between the perceptions of the public and drinking water infrastructure experts regarding the problem of pinhole leaks and other corrosion related issues.

### 3.2 Literature Review

The household decision-making process with regards to choosing a plumbing material for a private residency is complicated and involves several factors, such as federal, state, and local standards and regulations, drinking water infrastructure service providers, insurance companies, and household's perception of corrosion risk and financial impact of corrosion prevention. The regulations and standards of the federal, state, and local governments have major impacts on the plumbing material chosen for installation in a private house. For example, the use of plastic plumbing material, PEX, has been approved in all U.S. states except for California and Massachusetts where the material installation requires jurisdiction acceptance. These regulations influence the services provided by plumbers, home builders, material producers, and water utility companies. General contractors are the primary decision-makers of plumbing material installation in new houses. On the other hand, utility companies respond to corrosion threats by adding corrosion inhibitors into drinking water treatment. Consequently, all service providers influence consumer decisions regarding the best plumbing material for residential use (Lee et al., forthcoming).

In addition, insurance companies employ information on water damage and plumbing material failures in determining residential property insurance policy issuance and renewal. The information stored in a Comprehensive Loss Underwriting Report (CLUE)) represents a comprehensive database of personal property information relating to insurance claims on private property. It contains personal information on the insured, the claim number, the accident, the amount paid out to the customer, and inquiries of home plumbing issues. Due to the information recorded in CLUE, homeowners may have difficulty in either purchasing or renewing home insurance. On average 3 to 4 reports/inquiries over a five-year span on water related issues may lead to non-renewal of the homeowner's policy (Longcore, 2008).

As discussed, all of the above drinking water infrastructure providers influence consumer decisions with regards to the plumbing system (Lee et al., forthcoming). To make an informed decision, homeowners need information on the various risks involved in choosing plumbing systems.  When informed about the plumbing material characteristics, the consumers are able to decide on an alternative most preferable to them based on the preference trade-offs among plumbing materials' attributes.  The Random Utility Theory is usually adopted to study the attribute trade-offs while maximizing the consumers' preference level (Champ et al., 2003).  In addition, the perception of risk for plumbing materials can be quantified through the assessment of the willingness-to-pay for a corrosion-free plumbing material or improvement in the performance of the existing plumbing material.  The estimate of willingness-to-pay reflects the socioeconomic characteristic and previous experiences of individual households (Champ et al., 2003; Lee et al., forthcoming).

Each household is forced to make a decision on a plumbing alternative when replacing an existing system or installing a plumbing system in a new house.  Each alternative has its own advantages and disadvantages that have an impact on health and the overall cost of installation and maintenance.  The problem becomes more complex as consumers think in terms of cost (material plus labor charges), taste and odor of the water, corrosion problem, longevity of the pipe system, fire retardance, convenience of installation or replacement, plumber or general contractor's opinions or expertise, and proven record in the market. Householders weight each of these attributes in order to choose the most preferred option for their houses (Lee et al., forthcoming).

For example, Lee et al. (2005), utilizing the AHP method, studied the preferences for plumbing materials of Virginia Tech potable water experts. Participants ranked the health effects, reliability, taste and odor, and longevity as the most important attributes when choosing a plumbing material.  Property value and fire resistance were listed at the bottom of the ranking.  These results showed that health as well as taste and odor, which are a surrogate for water purity, dominate preferences for plumbing materials.  Lack of reliability is related to mental stress resulting from repairing the damage associated with pipe corrosion and a constant worry about future leaks (Lee et al., 2005).

There are several plumbing material types for a householder to choose from when deciding on a plumbing material to be installed in a house: copper, plastic (CPVC and PEX), and stainless steel. According to Marshutz' survey (2000), copper is used in nearly

90% of homes in the U.S. followed by PEX (cross linked polyethylene) with a 7% installation rate, and CPVC (chlorinated polyvinyl chloride) with a 2% installation rate. Telephone surveys of plumbers show an increased use of plastic pipes due to easier handling in installation and lower material cost (Scardina et al., 2007).

Copper is the most widely used material in residential plumbing and has several advantages, including affordability, fire resistance, few health hazards, and durability. Woodson (1999) studied the performance of different plumbing material alternatives: copper, CPVC, and PEX. He found copper pipes generally perform well except for cases involving major leak problems (Woodson, 1999). Due to increased pinhole leak incidents reported in hotspot areas of the U.S. (eg. Washington, D.C. suburbs and Sarasota, Florida), many consumers replaced copper with other options. Concerns with copper pipes include a metallic taste, especially with long stagnation periods and increased absorption of residual disinfectant by the pipe walls. High levels of copper can cause nausea, vomiting, and diarrhea (ATSDR, 2004). Elevated copper levels in drinking water may increase lead levels when lead solder joints, lead service lines, or brass fixtures are present in plumbing material. It is advised to test for lead when testing for copper levels in drinking water as lead and copper enter drinking water under similar conditions (Lee, 2008).

PEX (polyethylene cross linked) is another type of plumbing material often used in residential plumbing. This material is used to make flexible plastic pipes. A different plumbing design characterized by individual pipe lengths is required for every fixture. The main advantage of PEX is the lack of joints requiring soldering, which decreases the probability of pipe failures. On the other hand, PEX plumbing has raised some concerns regarding possible leaching of MTBE (methyl tertiary butyl ether), ETBE (Ethyl tert-butyl ether), and benzene into drinking water.  Other concerns are the negative health impacts associated with PEX's reaction with chlorine, increased water odors (Durand & Dietrich, 2007), the material's ability to withstand fire, and its final disposal (PRNews Wire, 2004). In addition, PEX may become stiff in cold weather, which makes faulty pipe repairs more difficult.  PEX use has been approved in all U.S. states (Toolbase News, 2008) and has met all health standards set by NSF/ANSI-61 for potable water supply (NSF, 2008).

CPVC plumbing material is also employed in residential plumbing but presents many concerns. For example, it can become brittle when exposed to sunlight for an extended period of time and presents possible negative health effects from microbial growth in the inner pipe. Other possible concerns are cracking in the event of an earthquake, plastic water taste, and melting in the event of fire. The solvents used to join fittings and pipe lengths may contain volatile organic compounds (VOCs) requiring proper ventilation during installation and causing unpleasant odor problems. However, CPVC by itself has a low odor potential (Heim & Dietrich, 2007).

The last plumbing material type is stainless steel, which is often used in industrial applications. Stainless steel provides excellent resistance to corrosion due to the presence of 18% chromium and 8% nickel (Roberge, 2000). The stainless steel material is, however, expensive. Due to the cost, its use is limited to specialized industries for conveying chemicals or other similar applications (Lee, 2008). A concern with stainless steel pipes is the possibility of leaching chromium into drinking water; however, all U.S. states have approved stainless steel use (NSF, 2008; Roberge, 2000).

The economically sustainable optimal replacement time for home plumbing systems is about 22 years after installation (Loganathan & Lee, 2005). The estimate, however, is dependent on the source and type of the employed data (Loganathan and & Lee, 2005). When it is time to replace the plumbing system, the homeowners have to decide on a plumbing system to be installed in their homes. For example, several homeowners in a Southeastern Community replaced their copper pipes with PEX. According to them, PEX is less labor intensive in case of installation, resistant against corrosion, and less expensive compared to copper (Plumbing and Mechanical Magazine, 2007).

However, plumbing material replacement or repairs can be rather expensive. Farooqi and Lee (2005) conducted a survey of plumbers in the U.S. and found plumbers to charge their work on an hourly basis. The cost per hour varied form $45 to $75, and the total cost of plumbing material replacement ranged from $3,654 for PEX to $5,680 for copper pipes (Farooqi & Lee, 2005). Furthermore, fixing dry wall, floor tiles, or ceilings affected by plumbing material replacement is not part of the services provided by the plumber, and homeowners have to hire a general contractor to fix the water related damage. Kleczyk and Bosch (2008) have reported the additional costs associated with damage from pipe failures reaching as much as $25,000 and forcing household members to reside in temporary housing during the repair period.

On the other hand, Scardina et al. (2007) (also discussed in Kleczyk et al. 2006) investigated the willingness-to-pay for a leak-free plumbing material in households located in different parts of the U.S., such as Florida and California. They found 47% of all respondents willing to pay a positive amount to ensure that material would remain leak-free, 27% unwilling to pay any amount to ensure that material would remain leak-free, and 25% unsure about how much they would be willing to pay. About 6% of respondents were willing to pay at least $4,000 to ensure that material would remain leak free for 50 years. This amount is 10 times the suggested base material cost for re-plumbing a 2,000 square foot house. The mean willingness-to-pay estimate was higher for respondents with leaks compared to respondents who had no leaks, constituting $1,130 and $1,007 respectively. Finally, 45% of respondents with leaks and 41% of respondents without leaks were not willing to pay for leak-free plumbing materials (Dietrich et al., 2006; Kleczyk et al., 2006; Scardina et al. 2007).

### 3.3 Consumer Theory of Utility Maximization

To provide an understanding of the derivation of the preferences for plumbing materials as well as the preventive measures against corrosion, the consumer theory of utility maximization needs to be described. This framework guides two out of three preference elicitation methods (Conjoint Analysis and Contingent Valuation) and impacts the results obtained for each method. The basic problem of utility or preference maximization represents the set of all bundles (alternatives) satisfying the consumer's budget constraint. The consumer is assumed to have preferences on each consumption bundles within a consumption set $X$. Consumer preferences are an ordered set of bundles such that a bundle $\boldsymbol{x}$ is at least as good as a bundle $\boldsymbol{y}$ ($\boldsymbol{x} \geq \boldsymbol{y}$). Consequently, consumer preferences must satisfy the following properties (Varian, 1992: page 95):

1. Preferences are complete - for all $\boldsymbol{x}$ and $\boldsymbol{y}$ in $X$, either $\boldsymbol{x} \geq \boldsymbol{y}$ or $\boldsymbol{y} \geq \boldsymbol{x}$ or both. The preferences completeness assumption implies any two bundles can be compared to each other.

2. Preferences are reflexive - for all $\boldsymbol{x}$ in $X$, $\boldsymbol{x} \geq \boldsymbol{x}$.
3. Preferences are transitive - for all $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ in $X$, if $\boldsymbol{x} \geq \boldsymbol{y}$ and $\boldsymbol{y} \geq \boldsymbol{z}$, then $\boldsymbol{x} \geq \boldsymbol{z}$. The preferences transitivity assumption is necessary for preference maximization, as without it, sets of bundles might have no best alternatives.
4. Preferences are continuous - for all $\boldsymbol{y}$ in $X$, sets$\{x\text{: } \boldsymbol{x} \geq \boldsymbol{y}\}$ and $\{x\text{: } \boldsymbol{x} \leq \boldsymbol{y}\}$ are closed sets. It follows that $\{x\text{: } \boldsymbol{x} > \boldsymbol{y}\}$ and $\{x\text{: } \boldsymbol{x} < \boldsymbol{y}\}$ are open sets. For example, if $(x_i)$ is a sequence of consumption bundles at least as good as bundle $\boldsymbol{y}$ and converges to some bundle $\boldsymbol{x^*}$, then $\boldsymbol{x^*}$ is at least as good as $\boldsymbol{y}$.
5. Preferences are weakly monotonic - if $\boldsymbol{x} \geq \boldsymbol{y}$ then $\boldsymbol{x} \geq \boldsymbol{y}$, which implies that at least as much of everything is at least as good.
6. Preferences are strongly monotonic - if $\boldsymbol{x} \geq \boldsymbol{y}$ and $\boldsymbol{x} \neq \boldsymbol{y}$ then $\boldsymbol{x} > \boldsymbol{y}$, which implies that at least as much of every good, and strictly more of some good, is strictly better.
7. Preferences are locally non-satiated - given any $\boldsymbol{x}$ in $X$ and any $\varepsilon > 0$, then there is some bundle $\boldsymbol{y}$ in $X$ with $/\boldsymbol{x} - \boldsymbol{y}/ < \varepsilon$ such that $\boldsymbol{y} > \boldsymbol{x}$. The property of local non-satiation implies that improvement in selection of consumption bundle is attainable even if the change is restricted to only small changes in the consumption bundle.
8. Preferences are convex - given $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}$ in $X$ such that $\boldsymbol{x} \geq \boldsymbol{z}$ and $\boldsymbol{y} \geq \boldsymbol{z}$, then $t\boldsymbol{x} + (1 - t)\boldsymbol{y} \geq \boldsymbol{z}$ for $0 \leq t \leq 1$. The property of convex preferences implies that average bundles are preferred to extreme bundles.
9. Preferences are strictly convex - given $\boldsymbol{x} \neq \boldsymbol{y}$ and $\boldsymbol{z}$ in $X$ such that $\boldsymbol{x} \geq \boldsymbol{z}$ and $\boldsymbol{y} \geq \boldsymbol{z}$, then $t\boldsymbol{x} + (1 - t)\boldsymbol{y} > \boldsymbol{z}$ for $0 < t < 1$.

As mentioned above, the basic problem of preference maximization represents the set of all bundles (alternatives) satisfying the consumer's budget constraint. Each consumer is assumed to have preferences on each consumption bundle within a consumption set $X$. Consumer preferences are an ordered set of bundles such that a bundle $\boldsymbol{x}$ is at least as good as a bundle $\boldsymbol{y}$ ($\boldsymbol{x} \geq \boldsymbol{y}$). A set of all consumption bundles defined as indifferent to each other is called an indifference curve. The preference maximization problem is defined as maximization of a utility function based on a specified budget constraint requirement (Varian, 1992: page 98):

$$\mathit{Max}\ u(\mathbf{x}) \qquad (1)$$

$$\text{such that } px \leq m \qquad (2)$$

$$\boldsymbol{x} \text{ is in } X. \qquad (3)$$

where $u(\boldsymbol{x})$ represents the utility function and $px \leq m$ represents the budget constraint with $m$ being the fixed amount of money available to a consumer ($\boldsymbol{B} = \{\boldsymbol{x} \text{ in } X\text{: } \boldsymbol{px} \leq m\}$). The theoretical preference maximization can be stated by utilizing the Lagrangian function (Varian, 1992: page 100):

$$\ell = u(\mathbf{x}) - \lambda(\mathbf{px} - m) \qquad (4)$$

where $\lambda$ is the Lagrange multiplier. Differentiating the Lagrangian function with respect to $x_i$ gives the following first-order conditions (Varian, 1992: page 100):

$$\partial u(\mathbf{x})/\partial x_i - \lambda p_i = 0, \text{ for } i = 1,\ldots, k. \qquad (4)$$

When the $i^{th}$ first-order condition is divided by the $j^{th}$ first-order conditions the marginal rate of substitution between good $i$ and $j$ is obtained (Varian, 1992: page 100):

$$\partial u(\mathbf{x})/\partial x_i/\partial u(\mathbf{x})/\partial x_j = p_i/p_j, \text{ for } i,j = 1,\ldots, k. \qquad (5)$$

The fraction on the right represents the economic rate of substitution between $i$ and $j$. The utility maximization exercise implies that the marginal rate of substitution is equal to the economic rate of substitution (Varian, 1992).

### 3.4 Preference Elicitation Methods

In this section of the paper, three stated preference elicitation methods will be introduced: Conjoint Analysis (CA), Contingent Valuation (CV), and Analytic Hierarchy Process (AHP). Stated preference approaches are heavily dependent on the quality of questionnaire responses on the studied subject. The survey questions answers include information on monetary amounts, choices, rankings, and ratings, as well as other measures of people's preferences.  These responses are transformed and utilized in different types of preference models including Multinomial and Binary Logit regressions to provide a measure of a financial value for the studied topic (Champ et al., 2003).

### 3.4.1 Conjoint Analysis

The Conjoint Analysis (CA) is a stated-preference study that uses an experimental design to elicit people's preferences for goods and services.  Responses to survey questions regarding versions of good/service varying in attribute levels inform about the preferences for multiple states of the good/service.  The introduction of price as an attribute extends the application of the method into welfare analysis. Based on the preference function knowledge, simulation and optimization algorithms aid the process of determining the preference level for each product-attribute combination (Champ et al., 2003; Rao, 2007).

There are four major types of Conjoint Analysis: Rating-Based, Choice-Based, Self-Explicated, and Adaptive. In case of rating-based CA, preference ratings are collected from respondents using either full-attribute or partial-attribute profiles. In the choice-based methods, choice sets describing products/objects are used to collect preference data based on attributes and levels selected for each alternative. For example, a preferred timber harvesting practice is selected from among multiple alternatives described by set of forestry attributes and attribute levels (Champ et al., 2003). In the case of self-explicated methods respondents provide importance and desirability ratings for attributes. The preference levels are calculated using these importance/desirability values. For example, respondents indicate the importance of price, power steering, engine capacity etc. of an automobile on a five-point scale (Wedel & Kamakura, 2000). Different from decompositional CA studies (rating- and choice- based), this approach is compositional. Finally, the adaptive CA uses initial information on attribute level desirability and attribute's importance to design questions to elicit preference/choice data. Partial profiles are used to collect preference data (Rao, 2007). For example, a pair of profiles differing on a limited number of attributes (two to five) is selected for each subject on the basis of the self-stated part-worths (otherwise known as marginal utility functions). After the respondents provide preference intensity judgment of the pair, their part-worth estimates are updated, and the next pair is selected on the basis of the updated part-worths (Wedel & Kamakura, 2000).

CA has been applied successfully to several marketing decisions, including designing of new products, targeting market selections, pricing of new products, and studying competitive reactions. One of the advantages of CA is its ability to answer various "what if" questions when employed for analysis of hypothetical and/or real choice alternatives (Rao, 2008).

A typical CA study consists of four main steps (Rao, 2008):

1. development of stimuli based on a number of attributes
2. presentation of stimuli to an appropriate sample of respondents
3. estimation of marginal utility (part-worth) functions for the attributes as well as any heterogeneity among the respondents
4. employment of the estimates in tackling any managerial problems (e.g. forecasting, pricing, or product design)

The definitions of stimuli and part-worth functions are presented in the following paragraphs.

#### 3.4.1.1 Conjoint Analysis Terminology

Table 3.1 represents the terminology required to follow the CA description.

Table 3.1: Conjoint Analysis Terminology

| Conjoint Analysis Term | Definition |
|---|---|
| Item or Product | a product or service |
| Attributes | aspects on which the item is described |
| Levels | different values an attribute takes in the study |
| Profile | a description of an item on various attributes with corresponding levels |
| Choice Set | a small set of items (profiles) for a respondent to choose among |
| Stated Preference | judgment of liking one of the items by a respondent |
| Revealed Preference | preference as revealed by a respondent's behavioral choice |
| Stated Choice | stated judgment as to which item in the choice set is chosen by a respondent |

#### 3.4.1.2 Theoretical Model

The theoretical model guiding the CA preference elicitation methodology is the Random Utility Maximization (RUM). RUM is based on consumers' choices from a set of competing alternatives. Each respondent chooses the most preferred alternative from a set of alternatives while at the same time making tradeoffs between attributes of each alternative. Consequently, the utility model estimates represent utility differences across all alternatives contained in the choice set (Champ et al., 2003).

It is assumed that an individual chooses among a set of *J* product alternatives, a choice set, described by *r* attributes. $U_j$ is assumed to be the random utility for the *j*-th alternative decomposed into systematic (*v*) and random (*e*) components and is expressed as follows (Champ et al., 2003):

$$U_j = v(x_j, p_j; \beta)+e_j \qquad (6)$$

where $U_j$ is the indirect utility for profile *j* (unobservable in the real world), $x_j$ is a vector of attributes associated with profile *j*, $p_j$ is the cost of profile *j*, $\beta$ is a vector of preference parameters, and $e_j$ is a random error term. Although the individual choice is assumed deterministic, the researcher does not observe all information about each individual, and therefore, the error term represents the uncertainty about the final choice (Champ et al., 2003).

The probability of a consumer choosing an alternative *i* from a choice set containing competing alternatives is defined as follows:

$$P(i/C) = P(U_i > U_j) = P(v_i+e_i > v_j+e_j), \text{ for all } j \in C \qquad (7)$$

where *C* contains all alternatives in the choice set. In RUM, choices are made based on utility differences across alternatives:

$$P(i/C) = P(v_i - v_j > e_j - e_i), \text{ for all } j \in C \qquad (8).$$

As a result, any variables remaining the same across profiles, such as income and gender, are eliminated from the model (Champ et al., 2003).

Several probabilistic choice models can be derived from RUM, depending on the specific assumptions made about the distribution of the random error term. If error terms are distributed according to a bivariate normal distribution, a binary probit model is specified. If errors are independently and identically distributed, preferences are homogenous, and errors have the same scale parameter, the multinomial logit model applies and choice probability is defined as follows:

$$P(i/C) = exp(\mu\, v_i) \,/\, \Sigma_{(j \in C)}\, exp(\mu\, v_j) \qquad (9)$$

where $\mu$ is the scale parameter. The likelihood function for the multinomial logit model is defined as follows:

$$L = \Pi_{(n=1)}^{N}\, \Pi_{(i \in C)} P_n(i)^{yin} \qquad (10).$$

where $N$ represents the sample size. After logarithmically transforming the likelihood function, the multinomial logit model is estimated by solving for the $\beta$'s values that maximize the following log-likelihood function (Champ et al., 2003):

$$lnL = \sum_{(n=1)}^{N} \sum_{(i \in C)} y_{in} \{ \Sigma_{(k=1 \ldots K)} [\beta_k x_{ik}] + \beta_p p_i - ln \Sigma_{(j \in C)} (\Sigma_{(k=1,,,K)} [\beta_k x_{ik}] + \beta_p p_i) \}, \quad (11)$$

where $\sum_{(n=1)}^{N}$ represents the summation of responses over $n$ individuals , $\sum_{(i \in C)}$ represents the summation of responses over the alternatives in the choice set, and $y_{in}$ is the n$^{th}$ respondent's selection from the presented choices.

### 3.4.1.3 Conjoint Analysis Variables Description

CA reveals the underlying product preference function in terms of its attributes. A general product profile defined on $r$ attributes is presented as ($x_{j1}$, $x_{j2}$,…, $x_{jr}$) where $x_{jr}$ is the level for the $j$-th profile on the $k$-th attribute for a product profile. The CA model is usually based on an additive function. The preference score for the $j$-th profile, $y_j$, for a respondent is defined as follows (Rao, 2007):

$$y_j = U_1(x_{j1}) + U_2(x_{j2}) + \ldots + U_r(x_{jr}) \quad (12)$$

where $U_k(\cdot)$ is the component utility function specific to the $k$-th attribute often called part-worth utility function. The specifications of the $U$-function for any attribute depend upon its type: categorical or quantitative (Rao, 2007).

Categorical attributes are described by verbal descriptions such as high, medium, or low. The levels of a categorical attribute can be recoded into a set of dummy variables. In this case, the part-worth utility function is specified as a piecewise linear function of the dummy variables. The component-utility function for the $k^{th}$ categorical attribute is described as follows (Rao, 2007):

$$U_k(x_{jk}) = D_{k1}U_{k1} + D_{k2}U_{k2} + \ldots + D_{kr}U_{krk} \qquad (13)$$

where $r_k$ is the number of discrete levels for the $k$-th attribute; $D_{kr}$ is a dummy variable taking a value 1 if the value $x_{ik}$ is equivalent to the $k$-th discrete level of $x_r$ and 0 otherwise. $U_{kr}$ describes the component of the part-worth utility function for the $k$-th discrete level of $x_r$ (Rao, 2007).

Quantitative attributes are measured by either an interval or ratio scale. They are either coded into categories or used directly in the specification of the part-worth utility function for the attribute. In the latter case, the function is specified as a linear (vector) model (Rao, 2007):

$$U_k(x_{jk}) = \quad w_k x_{jk} \qquad (14)$$

where $w_k$ is the weight (positive / negative). A linear function is appropriate for an attribute deemed to be desirable or undesirable. It is often called a vector model for which the utility increases / decreases linearly with the numerical value of the attribute (Rao, 2007). When the quantitative attributes are coded as categories, the utility function is not assumed to be linear but rather nonlinear, which implies that the utility does not change (increase /decrease) linearly (constant change) with the numerical value of the attribute. As the attribute levels change, the utility impact might be larger or smaller depending on the coefficient on each category. When categorical variables are employed, the utility function is specified as a piecewise function with the dummy variables (Wierenga, 2008).

### 3.4.1.4 Conjoint Analysis Description

Based on RUM theory and depending on the attribute types employed, a utility model for a profile *j* can be built. The model is built from an additively separable and linear-in-parameters utility function[1]. The price attribute for each profile is represented separately to allow for welfare analysis. The utility model for profile *j* is specified as follows (Champ et al., 2003):

$$U_j = \Sigma_{(k=1)}[\beta_k x_{jk}] + \beta_p p_j + e_j \qquad (15)$$

where $\beta_k$ represents the preference parameter associated with attribute *k*, $x_{jk}$ is attribute *k* in profile *j*, and $\beta_p$ is the parameter on profile cost (Champ et. al, 2003). The parameter estimates ($\beta$'s) are defined as marginal utilities of attributes ($\beta_k = \delta U/\delta x_k$) and marginal utility of money ($\beta_p$) (Champ et al., 2003).

The marginal rate of substitution (*MRS*) between any two attributes *k* and *m* is defined as the ratio of two parameter estimates (Champ et al., 2003 and Jahle and Reny, 2000):

$$MRS = \beta_k / \beta_{m.} \qquad (16)$$

MRS is defined as the rate at which the consumer is willing to give up attribute *k* in exchange for attribute *m* such that they remain indifferent after the exchange (Jahle & Reny, 2000).

As mentioned earlier, incorporating price as a separate attribute in each profile allows for welfare analysis performance. Compensating variation (*CV*) examines the financial impact of change in an attribute and therefore, represents the amount of money given or taken away from a person to make him or her as well as off after a change as before a change. *CV* is computed from the estimated values of utilities for different alternatives. *CV* is defined as follows:

$$CV = (1/\lambda_y)[ln(\Sigma_{(j=1\ to\ J)}exp(V^1_j)) - ln(\Sigma_{(j=1\ to\ J)}exp(V^0_j)] \qquad (17)$$

[1] **Additive Separable Utility Function** – the utility function can be partitioned into mutually exclusive and exhaustive subsets (Jahle and Reny, 2000).

where $\lambda_y$ is the marginal utility of money, $V^0_j$ is the before change expected value of maximum utility, $V^1_j$ is the after change expected value of maximum utility. *CV* is defined as the difference in the expected values of maximum utility divided by the marginal utility of money. The marginal utility of money parameter is the coefficient on the price attribute variable multiplied by -1 to reflect the increasing utility of income (Champ et al., 2003).

### 3.4.1.5 Experimental Design

Any conjoint analysis depends upon the design of stimuli, which describe either profiles or choice sets. Employment of experimental design allows for manipulation of the independent variables (factors) associated with the attributes of goods or services over pre-specified factor levels. A factorial design combining every level of each attribute with every level of all other attributes is called a full-factorial experimental design. The set of profiles is in general very large and therefore, places burden on respondents. To limit the cognitive exhaustion of respondents, researchers utilize fractional (partial) factorial designs which only include a subset of all profiles (Champ et al., 2003; Rao, 2007).

In a study with $n$ attributes and $l_k$ levels for the $k$-th attribute, the total number of profiles amounts to $\Pi l_k$. This number can be very large as the number of attributes and their levels increases. The effective number of marginal utilities (part-worth) parameters to be estimated from conjoint data is $m = \sum(l_k - 1)$. To resolve the issue of a large number of estimated parameters, researchers generally construct factorial designs (Rao, 2007). Orthogonal arrays are usually employed for designing profiles for the ratings-based CA and for designing choice sets for the choice-based CA. The orthogonal arrays are derived out of the complete factorial of every attribute combination. Orthogonal arrays present the columns for the independent variables are "orthogonal" to one another. These vectors exhibit independence from each other, linearity, and separablility (Rao, 2008).

#### 3.4.1.5.1 Designs for Ratings-Based Methods

Orthogonal arrays are categorized by their resolution. Resolution describes the degree to which estimated "main effects" are confounded with estimated "higher-order level interactions" (2, 3, 4 or more) among the attributes (Rao, 2007). A "main" effect is defined as the difference between the average responses to each attribute level. In a multiple regression, main effects are represented by parameter estimates for the attribute levels. An interaction effect arises when the level of one attribute affects the response to the level of other attributes. Interaction effects are represented by the parameter estimates for the cross product of two or more variables and allow for analysis of the substitutability between attributes. If the interaction effect carries a positive sign, then the attribute levels are substitutes. On the other hand, a negative sign implies a complementary relationship (Champ et al., 2003). The resolution usually identifies which effects are confounded and which ones are estimable. For example, in a resolution-III design, some main effects are confounded with some 2-level interactions. In a resolution-V design, all main effects and two-factor interactions are estimable free of each other. Higher resolution designs require larger number of profiles and therefore a larger number of full-profiles to be administered to respondents (Champ et al., 2003; Rao, 2007).

Orthogonal arrays can be either balanced or unbalanced in terms of levels of attributes. The property of level balance implies that each level of an attribute occurs with the same frequency within each attribute in the design. An unbalanced design gives larger standard errors of the parameter estimates for those attributes that are less frequently administrated. An additional property of an orthogonal design is the proportional criterion which implies that the joint occurrence of any two levels of different attributes is proportional to the product of their marginal frequencies (Rao, 2007).

Various design efficiency measures can be described following the linear model definition as discussed earlier (Kuhfeld, Tobia, & Garratt, 1994):

$$Y = X\beta + e \qquad (18)$$

where $\beta$ is a px1 vector of parameters, $X$ is an nxp design matrix, and $e$ is the random error. The least squares estimate of $\beta$ is given by $(X'X)^{-1}X'Y$. The variance-covariance matrix of the parameter estimates (or partworths) of the attributes is proportional to $(X'X)^{-1}$. The efficiency of a design is based on the information matrix $X'X$. An efficient design has a smaller variance matrix and the Eigen-values of $(X'X)^{-1}$ provide a measure of the size of the matrix. Three efficiency measures can be defined based on Eigen-values:

1. A-efficiency: $1/(ntrace(((X'X)^{-1}/p))$
2. D-efficiency: $1/(n/( X'X)^{-1}/^{1/p})$
3. G-efficiency: $(p/n)^{0.5}/\sigma_M$,

where $p$ stands for the number of factor effects in the design, $n$ is the number of requested runs, and $\sigma_M$ is the minimum standard error possible. The minimum standard error is attained when a full-factorial design is used and a fractional design has efficiency less than 1. These three measures are useful for making efficiency comparisons between considered designs. Orthogonal designs for linear models are generally considered to be efficient because their efficiency measure is close to 1 (Kuhfeld, 2005).

D-efficiency measures the level of orthogonality and balance within the experimental design and is interpreted as the relative number of runs (in percent) required by an orthogonal design to achieve the same value of the determinant $|X'X|$. It is often employed to reduce the number of paired comparisons to the smallest number necessary for efficient estimation of preference weights (Huber & Zwerina, 1996; Kuhfeld, Tobias, & Garratt, 1994). A-efficiency minimizes the expected posterior squared errors (Toubia & Hauser, 2006) and is interpreted as the relative number of runs (in percent) that are required by an orthogonal design to achieve the same value of the trace *of* $(X'X)^{-1}$ (Huber & Zwerina, 1996; Kuhfeld, Tobias, & Garratt, 1994). Finally, G-efficiency is defined as a measure minimizing the maximum value of the standard error of the predicted response (StatSoft, Inc., 2008).

#### 3.4.1.5.2 Design for Choice-Based Conjoint Analysis

The probability of choosing an alternative in a choice-based study is generally modeled as a logit function in terms of the attribute differences of the item with respect to a base alternative in the choice set. Thus, the underlying model for a choice-based conjoint experiment is nonlinear and the considerations of choosing a design for a choice-based study are different than those for a ratings-based study. There are two additional properties considered when designing choice-based experiments: minimal level overlap and utility balance (Huber & Zwerina, 1996).

Minimal level overlap means that the probability of an attribute level repeating itself in each choice set should be as small as possible due to the contrasts between the attribute levels in the calibration of the logit model. In situations when the same level is repeated multiple times within the choice set, the choices made in that choice set do not contribute any information on the value of that attribute (Huber & Zwerina, 1996).

The property of utility balance implies that the utilities of the alternatives in a choice set are approximately equal. When a design is utility balanced, the variance of the probabilities of choice of alternatives within a choice set is reduced. Huber and Zwerina (1996) show that achieving such utility balance increases the efficiency of a design up to 50%. The process of swapping and re-labeling attribute levels of alternatives in an initial choice set accomplishes this objective. The initial choice sets are developed by employing orthogonal arrays, availability designs, and D-efficient designs. Kuhfeld (2005) developed a SAS OPTEX procedure for D-Efficient designs.

Burgess and Street (2003; 2005) modified the choice experiment by constructing choice set designs for choice experiments (with choice option amongst the alternatives) for binary attributes based on the multinomial logit (MNL) model for choice. Their designs are useful for a choice experiment that tests main effects and tests both main and two-attribute interactions. The methods lead to optimal and near-optimal designs with small numbers of choice sets for $2^K$ choice experiments (Burgess & Street, 2003; 2005).

### 3.4.1.6 Selected Data Collection Issues

When respondents are presented with partial profiles with missing information on some attributes in a ratings-based conjoint experiment, they tend to impute values for missing attributes. The process of such imputation has an effect on the estimated part-worth values. Bradlow et al. (2004) developed a mathematical model based on Bayesian learning and investigated the effects of such imputations. They found that learning in fact occurs and that the relative importance of attribute part-worths shifts when subjects evaluate partial profiles and the relative part-worth functions are sensitive to the order in which partial profiles are presented. They also found that the imputation process is sensitive to the available prior information on the product category (Bradlow et al., 2004).

Another issue in the CA data collection is whether respondents experience strong incentives to expand their cognitive resources or to devote adequate time and effort in providing responses (ratings or choices) to hypothetical stimuli presented as profiles or choice sets. The literature on experimental economics suggests that data collected without such incentive-compatibility may be inconsistent, erratic, and possibly untrustworthy. Ding, Agrewal and Liechty (2005) find that CA data collected when incentive of respondents are aligned outperforms those without such alignment in terms of out-of-sample predictive power. In fact, Wertenbroch and Skiera (2002) also show that willingness to buy estimates for products using conjoint analysis are lower when the incentive-compatibility constraint is not imposed. This finding has several implications for collecting conjoint data in practice.

### 3.4.1.7 Survey Question Design

In order to conduct the CA survey, the combinations of attribute levels are paired with ranking or rating questions in order to evaluate consumers' preferences for each alternative (Champ et al., 2003). Contingent ranking questions ask respondents to rank a set of profiles from most to least preferred. The framework assumes the most preferred profile to be chosen first from the choice set followed by the second ranked alternative chosen from the remaining choice set, and so forth. Theoretically, the additional information provided by the ranking exercise should result in lower standard errors of regression estimates; however, in practice ranking is a demanding exercise for respondents. The participants might get fatigued while proceeding through the sequence of choices. As a result, parameter estimates lack stability and reliability due to increased 'noise' in the data (Champ et al., 2003).

Rating scale questions require individuals to make rulings about the magnitude of utility related to profiles presented in an attribute-based experiment. Judgments directly transform utility to the rating scale, which then are regressed on the vector of responses to the attribute levels included in each profile. An ordered probit or logit[2] as well as an ordinary least square regression model can be used to estimate the utility model (Champ et al., 2003). The rating scale questions appeal to many researchers due to simplicity of the econometric analysis and the ease with which respondents answer rating questions. There are, however, issues related to this approach. For example, ratings must be adjusted to a common metric across all individuals. Additionally, a status quo (current choice) must be evaluated using the rating to judge whether an individual would rate a new alternative higher than the current choice. Finally, respondents might equally weight new alternatives, which would complicate econometric analysis (Champ et al., 2003).

### 3.4.2 Contingent Valuation

#### 3.4.2.1. Contingent Valuation Description

The contingent valuation method is a survey-based method for eliciting consumer values on goods and services. It provides a way to estimate values when markets do not exist and revealed preference methods cannot be used. In a CV study, the first step is to identify the changes in quantity or quality of the good to be evaluated. The change is assessed through a theoretical model, which is based on the difference of the level of utility in the current condition and the level of the utility with the new policy. The Hicksian surplus for a particular policy is defined as:

[2] Probit and Logit Models are both the probability-based models. They differ by the specification of the error term. Probit error terms are distributed according to a normal distribution, while logit errors are independently and identically distributed (Champ et al., 2003).

$$v(p^0, Q^0, y - c) = v(p^1, Q^1, y) \qquad (19)$$

where *v(.)* is indirect utility function, *p* is the price of the good or service, *Q* is the quality of the good or service, *0* denotes the initial level of the good, *1* denotes the quality of the good if the protection policy is not implemented, *y* is income and *c* is the Hicksian compensating variation or *WTP* (Champ et al., 2003). This represents the case when the purpose of the policy is to maintain the status quo, $Q^0$ where $Q^0 > Q^1$. Although it is difficult to establish the conditions $Q^0$ and $Q^1$, this information is generally included in the surveys via a description of the current conditions of the resource and how they change when the policy is implemented (Champ et al., 2003).

### 3.4.2.2 Population Sample Description

The next step is to specify how the proposed policy will affect people and who will be affected. This is important in establishing a representative sample when conducting the survey. Each respondent in the sample should have the same probability of being selected from the given population. Contingent valuation studies require large sample sizes due to the large variance of *WTP* responses (Mitchell & Carson, 1989). Boyle and Bouchard (2003) suggest that the larger the standard error of *WTP*, the larger should be the sample size to increase statistical efficiency.

### 3.4.2.3 Survey Question Design

Different formats are utilized in asking contingent valuation questions. The most common methods used are open-ended questions, the payment card method, and dichotomous choice questions. An open-ended question asks respondents to state the highest amount they are willing to pay for the policy that is being evaluated. The payment card method the respondents with a particular number of bids and asks them to pick the largest amount they would be willing to pay for the proposed policy. The dichotomous choice question asks respondents whether they are willing to pay $X for the change in resource. The dichotomous choice question has a take-it-or-leave-it approach in which the respondent can only answer “yes” or “no.” The initial bid amount asked varies among respondents (Champ et al., 2003).

When valuing a single program, survey respondents are asked about their *WTP* when change from $Q^0$ (the status quo) to some particular alternative $Q^1$ is introduced. There are single-bounded and multiple-bounded question formats to obtain the information when employing the closed-ended questions. As mentioned above, the single-bounded approach involves asking the respondents about their willingness to pay a specific dollar cost for a good/service. With a public program for example, the question is framed to ask whether the respondents would vote for or against the program if it involved a specific increase in their taxes. Their responses provide qualitative information on a bound on their *WTP* for the item: a lower bound if they answer “yes” and an upper bound if they answer “no” (Hanemann & Kanninen, 1998).

The double-bounded format, proposed by Hanemann (1985), follows up on the initial *WTP* question with a second question, involving a specific dollar cost to which the survey participants can respond with either a “yes” or “no” answer. The amount presented in the second bid depends on their response to their first bid. If the participants answered “no” to the first bid, then the second bid represents some lower amount. On the other hand, if they answered “yes”, the bid amount is increased. Consequently, if the participants answered “yes” to one of the questions and “no” to the other, this provides both upper and lower bounds on their *WTP*. Similarly, the triple-bounded format has additional follow-up question with a bid amount that depends on the responses to the first two questions. The additional bids can lead to sharper bounds on the estimate of *WTP*, although the additional improvement is rather small (Hanemann & Kanninen, 1998).

Other formulations of CV questions can also elicit information on *WTP*. For example, survey participants can be shown a payment card with various dollar amounts and asked to circle the one closest to their own value (Mitchell & Carson, 1981). Their response is interpreted not as an exact statement of *WTP* but as an indication of *WTP* lying between the highest number below the amount circled and the smallest number above it (Cameron & James, 1987). Equivalently, the respondents can be shown a list of dollar amounts and asked to mark off the amounts they would be willing to pay and those they would not (Welsh & Bishop, 1993). Similarly, respondents can be presented with alternative dollar ranges and asked to identify the one closest to their own valuation.

These approaches yield upper and lower bounds on *WTP* for the change in *Q* (Hanemann & Kanninen, 1998).

### 3.4.2.4 The Economic Foundations of a Statistical Model

Statistically, the CV responses are discrete dependent variables since they are measured on a nominal or ordinal scale. There is a variety of statistical models for analyzing such data. In the general case, the CV responses assume a finite number of values, which are indexed $j = 1,.., M$. For the *i*th observed response, the probability of taking a particular value is expressed as (Hanemann & Kanninen, 1998):

$$Probability\ \{response_i = j\} = H_j(A_i; Z_i\ ; \gamma) \qquad (20)$$

where $A_i$ is the commodity bid, $Z_i$ represents other covariates describing the survey respondent and the valued item. Finally, $\gamma$ is a vector of parameters to be estimated from the data. In order for the probabilities to be well-defined, the right hand side of the above question must return a value between zero and one, and sum over all possible outcomes $j = 1,\ldots, M$. Based on the specifications, the statistical response model is in the form of (Hanemann & Kanninen, 1998):

$$\text{Response} = \begin{cases} \text{'yes' if } T(A; Z; \gamma) - \eta \geq 0 \\ \text{'no' otherwise.} \end{cases} \qquad (21)$$

where $T(.)$ is some function of $A$ (the commodity bid) and $Z$ (other covariates), $\eta$ is some random variable with cumulative distribution function (cdf) of $F(.)$, and $\gamma$ represents both coefficients associated with $T(.)$ and parameters of the cdf.

It is further assumed that an individual consumer has a utility defined over both market commodities, denoted $X$, and some non-market item to be valued, denoted $Q$. There are also other arguments that shift the respondents' preferences for $X$ and $Q$, such as characteristics of the individual or the market goods denoted as $s$. The corresponding indirect utility function depends on the prices of the market goods, $p$, the individual's income, $y$, socioeconomic characteristics, $s$, and the non-market item $Q$ (Hanemann & Kanninen, 1998).

The other component of the indirect utility function is a stochastic component representing the notion of random utility maximization (RUM). In a RUM model, it is assumed that while the individual knows his or her preferences with certainty, some components are unobservable to the researcher and are treated as random (Hanemann 1984). These unobservables could be characteristics of the individual and/or attributes of the item. They stand for both variation in preferences among members of a population and the measurement error. The stochastic component of preferences is denoted by $\varepsilon$, so that the indirect utility function is $v(p, Q, y, s, \varepsilon)$ (Hanemann & Kanninen, 1998).

The individual is faced with the decision of securing a change from $Q^0$ to $Q^1 > Q^0$. The change is regarded as an improvement, so that $v(p, Q^1, y, s, \varepsilon) \geq v(p, Q^0, y, s, \varepsilon)$. The cost associated with the change is \$$A$. The respondent is then asked whether he or she would be in favor of purchasing $Q^1$ at the presented price. Per utility maximization problem, the "yes" answer is posted only if $v(p, Q^1, y\text{-}A, s, \varepsilon) \geq v(p, Q^0, y, s, \varepsilon)$, and the "no" answer otherwise. The above implies that (Hanemann & Kanninen, 1998):

$$Probability\{response\ is\ `yes'\} = Probability\{\ v(p, Q^1, y - A, s, \varepsilon) \geq v(p, Q^0, y, s, \varepsilon)\} \quad (22).$$

An equivalent way to express the outcome is to utilize the compensating variation measure, $C$, in the specification:

$$v(p, Q^1, y - C, s, \varepsilon) \geq v(p, Q^0, y, s, \varepsilon) \quad (23).$$

Thus, $C = C(p, Q^0, Q^1, y, s, \varepsilon)$ is the maximum *WTP* for the change from $Q^0$ to $Q^1$. It follows that respondents answer "yes" if the stated price is less than this WTP and 'no' otherwise (Hanemann & Kanninen, 1998).

In a RUM model, $C = C(p, Q^0, Q^1, y, s, \varepsilon)$ itself is a random variable as the respondent's *WTP* for the change in Q is not known by the researcher. When $G_C(.)$ is defined as the cumulative distribution function and $g_C(.)$ as the corresponding density function, the probability specification becomes (Hanemann & Kanninen, 1998):

$$Probability\{response\ is\ \text{'yes'}\} = 1 - G_C(A) \quad (24).$$

The cdf for the individual's random *WTP* value can be further specified as follows. Let $E\{C\} = \mu$ (in a regression model one would have $\mu = X\beta$), let $var\{C\} = \sigma^2$, and let $G(.)$ be the cdf of the standardized variant: $z = (C - \mu)/\sigma$; then (Hanemann & Kanninen, 1998):

$$Probability\{response\ is\ \text{"yes"}\} = 1 - G_C(A) = 1 - G[(A - \mu)/\sigma] \qquad (25).$$

For example, when $G(x) = \Phi(x)$, the standard normal cdf, a probit model is employed:

$$Probability\{response\ is\ \text{"yes"}\} = \Phi[(A - \mu)/\sigma] \qquad (26).$$

On the other hand, $G(x) = [1 + e^{-x}]^{-1}$, the standard logistic cdf, yields a logit model of:

$$Probability\{response\ is\ \text{"yes"}\} = 1/(1 + e^{(A - \mu)/\theta}) \qquad (27)$$

where $\theta = \sigma(3/\Pi)^{0.5}$ (Hanemann & Kanninen, 1998). Probit model is employed when errors are distributed according to a normal distribution. On the other hand, when errors are independently and identically distributed, the logit model should be employed for estimation (Champ et al., 2003)

#### 3.4.2.5 WTP Estimation Process

The linear and logarithmic models have been widely used in the CV analysis due to their simplicity. They impose, however, very stringent restrictions on the shape of *WTP* function. For example, the linear model implies that the income elasticity of *C* is zero while the logarithmic model implies that it is unity. When the linear model with $G_\eta(.)$ is a standard normal cdf, the specification becomes:

$$Probability\ \{response\ is\ \text{"yes"}\} = \Phi[\alpha - \beta A], \qquad (28)$$

which represents a probit model. Similarly when $G_\eta(.)$ is the standard logistic cdf, the specification becomes:

$$Probability\{response\ is\ \text{"yes"}\} = 1/(1 + e^{-\alpha + \beta A}), \qquad (29)$$

which implies that logit model becomes identical to the model in equation (10) (Hanemann & Kanninen, 1998).

### 3.4.2.6 Welfare Evaluation

As noted earlier per the RUM hypothesis, a person's *WTP* for the change in *Q* is a random variable with a cdf given by $G_C(.)$. The survey responses permit the estimation of this *WTP* distribution. This is, however, a basic difference between the conventional regression models for continuous dependent variables (e.g. responses to open-ended CV questions) and the quantal response models. The conventional regression models estimate the conditional mean of the dependent variable given the regressors. The quantal response models for discrete dependent variables estimate the conditional distribution of the dependent variable given the regressors (Hanemann & Kanninen, 1998).

The economic literature has generally focused on two summary statistics. One is the mean of the estimated *WTP* distribution:

$$C^+ = E\{C(q^0, q^1, y, \varepsilon)\} \qquad (30).$$

The other method is the median of the estimated *WTP* distribution defined as follows:

$$1 - G_C(C^*) = 0.5 \qquad (31).$$

It follows that *Probability {response is "yes"} = 0.5* when $A = C^*$, there is 50:50 chance that the individual would be willing to pay at least C* (Hanemann & Kanninen, 1998).

### 3.4.2.7 Experimental Design

In a discrete response CV experiment, the researcher has the opportunity to influence the data by setting the appropriate bids. To do this optimally, the researcher must first make an assumption about the form of the underlying *WTP* distribution and the value of its parameters. Moreover, the researcher must decide which particular welfare values should be estimated and therefore, select a particular criterion for the precision of the estimate. Those decisions reflect the researcher's interests and priorities. For example, there might be an interest in reducing the variance of the estimate of median *WTP*. This implies minimizing the asymptotic variance of the median (from the logit model) with respect to the bid values. In the experimental design literature, minimizing a function of parameter estimates is known as a C-optimal design criterion. Alternatively, the researcher could choose a D-optimal design, based on maximizing the determinant of the Fisher matrix[3]. The optimization minimizes the variance of the parameter estimates for the model in question. In the case of the logit model, the variance reduction corresponds to jointly minimizing the confidence intervals around the parameters $\alpha$ and $\beta$. This criterion would be of most use for a researcher who is interested in obtaining efficient parameter estimates for predictive purpose (Hanemann & Kanninen, 1998).

Since the distribution is unknown, the distribution estimates or guesses are employed in order to identify the bid points. This phenomenon represents the basic paradox of optimal experimental design in non-linear models. The optimal design depends on the true parameter values which should be known in order to obtain the best design for collecting the data. However, there would be no point in collecting the data if the true parameter values were known. This paradox is resolved by relying on prior information. In addition, the issue highlights a second important role of optimal design which is interpreting the results after the data collection and analysis by understanding, *ex post*, which bid points were influential or contributed the most statistical information (Hanemann & Kanninen, 1998).

---

[3] **Fisher Matrix** - The Fisher information measures the amount of information that an observable random variable $X$ carries about an unknown parameter $\theta$ upon which the likelihood function of $\theta$, $L(\theta) = f(X;\theta)$, depends. The likelihood function is the joint probability of the data ($X$) conditional on the value of $\theta$. Since the expectation of the score is zero, the variance is simply the second moment of the score, the derivative of the log of the likelihood function with respect to $\theta$ (Van Trees, 1968).

One way to resolve the paradox of optimal experimental design is to conduct the CV survey in stages using data from the earlier stages to optimize the design in the later stages. Most CV researchers do this at least to some degree. For example, a common practice is to employ a two-stage procedure; first perform a pre-test to obtain preliminary information about the *WTP* distribution, and then survey the full sample. The pre-test results are used to design optimal bids for the full survey based on the parameter values estimated from the pre-test data. Typically, the pre-test survey effort generates only a modest amount of initial information (Hanemann & Kanninen, 1998).

To take better advantage of the optimal design results available, a sequential estimation procedure is employed, preferably with more than two stages (Kannienen, 1993). After each iteration of the survey, the *WTP* distribution can be estimated using all available observations so that a new (optimal) bid design can be derived for use with the subsequent iteration of the survey. This approach asymptotically attains the efficiency of the theoretical optimal design results (Hanemann & Kanninen, 1998). With a sequential design, only two bid values are employed at each iteration, but the values change between iterations as information is accumulated. The number of bid values ultimately employed depends on the number of performed iterations. The difference in bid values of different iterations depends on the closeness of the initial parameter values to the true distribution (Hanemann & Kanninen, 1998).

### 3.4.3 Analytical Hierarchy Process

The AHP is a procedure for assessing preferences over a set of competing alternatives by using pair-wise preferences.  It is generally assumed that assessing pair-wise preferences is easier as it enables the researcher to concentrate judgment on taking a pair of elements and comparing them on a single property without thinking about other properties or elements (Saaty, 1990).  There are four axioms in the AHP that need to be satisfied: 1) reciprocity, which is defined as the ability to state the decision-maker’s preferences; 2) homogeneity, which is the bounded scale representing the stated preferences; 3) independence, defined as the criteria weight being independent of the considered alternatives; and 4) expectation, which implies a complete hierarchical structure (Saaty, 1990).

Based on Saaty (1980), the following steps are adopted in performing the analytical hierarchy process. First, the set of attributes for each alternative in consideration needs to be identified. The set of attributes as well as attributes levels are then pair-wise compared on a 1-9 scale. The AHP scale starts with a rating value of 1 - equally preferred and ends with 9 - extremely preferred. In the second step, a matrix equation is computed based on the pair-wise preference matrix for all alternatives:

$$\begin{bmatrix} w1/w1 & w1/w2 & \ldots & w1/wn \\ w2/w1 & w2/w2 & \ldots & w2/wn \\ \ldots & \ldots & \ldots & \ldots \\ wn/w1 & wn/w2 & \ldots & wn/wn \end{bmatrix} \begin{Bmatrix} w1 \\ w2 \\ \ldots \\ wn \end{Bmatrix} = n \begin{Bmatrix} w1 \\ w2 \\ \ldots \\ wn \end{Bmatrix} \quad (32)$$

where $\begin{Bmatrix} w1 \\ w2 \\ \ldots \\ wn \end{Bmatrix} = \{w\}$ is the required global preference vector (of weights). A reciprocal matrix is also computed and represents the off diagonal elements reciprocals of each other. As a result, the numerical score for each pair-wise comparison is defined as follows:

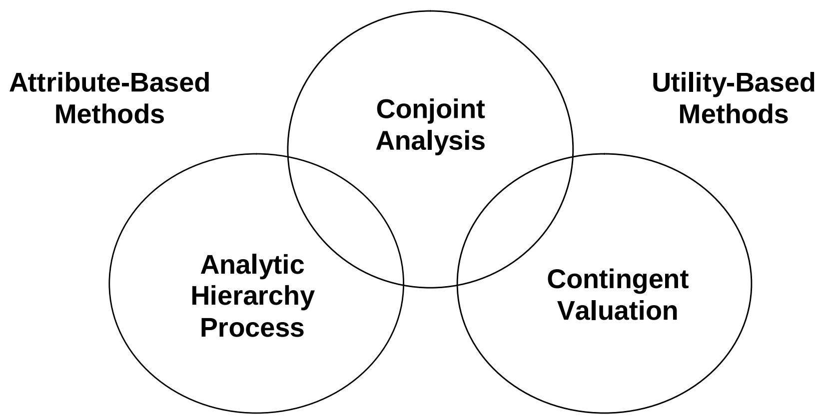
$$a(i,j) = \frac{weight\ \ for\, criterion, i, wi}{weight\ \ for\, criterion,\ j, wj} = 1/a(j,i) \quad (33).$$

The third step requires computing a rescaled preference matrix that is generated by dividing each attribute/attribute level pair-wise score in the matrix by that column’s sum. This computation allows for ranking of each attribute/level relative to all attributes/levels in the selection set. In the final step, the preferred alternative can be identified based on the ranking of all attributes and attribute levels (Saaty, 1980).

### 3.4.4 Stated Preference Methods Discussion and Comparison

In this section, the three preference elicitation methods are compared to each other. The Conjoint Analysis is the only method that can be used to compare results with Contingent Valuation and Analytic Hierarchy Process as it is an attribute-level methodology based on a utility maximization framework. Furthermore, CA and AHP are both attribute-based preference elicitation methods, which are often employed for multi-attribute optimization problems. On the other hand, CA and CV are both based on the utility maximization problem. Figure 3.1 presents the relationship between CA, CV, and AHP (Boyle, 2007).

Figure 3.1: CA relationship to CV and AHP




Although AHP is an entirely different method than CA, the AHP results are comparable with other methods of preference elicitation (Duke & Aull-Hyde, 2002). Zehedi (1987) proposed an underlying utility framework for AHP in which the process of selecting alternatives was consistent with maximizing a respondent's multi-attribute utility function. This finding allowed for synthesizing AHP and utility theory to solve the decision-making problem (Duke & Aull-Hyde, 2002). Both the CA and AHP methods are attribute-based preference elicitation techniques. Furthermore, CA is based on an additive part-worth (marginal) utility model while the AHP method employs a weighted additive utility model. The difference between the two methods is that CA is a decompositional approach while AHP is a compositional approach. This further implies that the CA part-worth utilities are obtained from responses provided to stimuli at different attribute levels while AHP starts with rating the levels of attributes and then aggregating the values into a final ranking of each attribute and then into a ranking of the goods in question. As a result, the total utility of an alternative is an additive value function summing up the weighted utilities of its attribute levels. The CA respondent's input is often presented in the form of ordinal or interval scales, which are then employed in the marginal utility estimates computation ($\beta$). In case of AHP, the survey respondents provide their degree of preference for one attribute/attribute level over another (a ratio scale ($w_i/w_j$)). In order to compare the CA and AHP results, the Marginal Rate of Substitution (MRS = $\beta_i/\beta_j$) computed for the Conjoint Analysis is compared to the ratio scale of preferences for the AHP methodology (Duke et al., 2002). The following Figure 3.2 presents the MRS and ratio scale relationship, which represent the point of comparison for both methods.

Figure 3.2: Comparison of CA and AHP.

Analytic Hierarchy Process

| | Attribute 1 | Attribute 2 | … | Attribute k |
|---|---|---|---|---|
| Attribute 1 | 1 | $w_1/w_2$ | … | $w_1/w_k$ |
| Attribute 2 | $w_2/w_1$ | 1 | … | $w_2/w_k$ |
| … | … | … | 1 | … |
| Attribute k | $w_k/w_1$ | $w_k/w_2$ | … | 1 |

Conjoint Analysis

$$Y_i = \alpha + \beta_1 X_{1i} + \ldots + \beta_k X_{ki} + u_i$$

$$\boldsymbol{MRS = \beta_1/\beta_k = w_1/w_k}$$

There are also several similarities between CA and CV, including the fact that both methods are based on the Random Utility Modeling (RUM) and describe the discrete choices in a utility maximizing theory. In addition, in both CA and CV the respondent is presented with the task of making one choice from a set of alternatives (CA might have multiple alternatives while CV might have a binary choice). CV, however, evaluates only the total value of the good or service, which implies difficulty in distinguishing the value of each attribute for the good. CA, which is an attribute-based method, is able to distinguish the value of each component of the good/service. As a result, the main attribute of CA is the repeated questions and profiles represented by bundles of attributes (Kuriyama et al., 1999). The experimental design aspect of CA is also an advantage as it creates bundles of attributes reflecting different states of the studied situation. As individuals are asked to choose their preferred alternative from a choice set, their choice reflects the trade-offs made between the good/service attributes. When a price or cost factor is included as an attribute, it is possible to estimate economic values associated with the other attributes. Including the price/cost attribute allows for willingness-to-pay and compensating variation computation, which present the financial impact of attribute levels changes. As noted earlier, CV also results in willingness-to-pay estimates as respondents directly present their willingness to pay for a described good/service based on the presented financial bids. In recent years many studies studied the differences and commonalities between the *WTP* estimates of both method types (Boxall et al., 1996).

Although attribute-based methods are very powerful tools, they currently present some unresolved problems. First, elicitation format of CA studies may not be incentive compatible, which results in respondents not reporting their true preferences. Second, the multinomial/conditional logit model, which is a statistical method used in modeling CA data, assumes the independence of irrelevant alternatives (IIA). If IIA holds, then the choice between alternatives is independent of alternatives that are not in the subset. When this assumption holds true, it simplifies the analysis of choice sets with a large number of alternatives. For example, when the choice set has more than 10 alternatives, only 10 alternatives can be selected for the estimation of sample decisions. When the IIA property is violated, more flexible and complex statistical methods are required. CV, on the other hand, does not require the assumption of IIA to hold, but it is limited to the estimation of the value of a good/service (Adamowicz et al., 1998).

### 3.5 Survey Design

To assess the preferences for plumbing materials and preventive measure choices for individual households, two surveys types were sent to a Southeastern Community. This community was established in 1980s and spans over 4,700 acres. There are about 3,300 homes, including condos and apartments, with 6,600 residents in total. Most of the resident population is retired, so the community is rather a homogenous group. There are two water utility suppliers (utility company 1 and utility company 2) in the area. The first incidents of pinhole leaks were reported in 2001. In 2006, most of the residents were switched to utility company 2, which resulted in an increased number of pinhole leaks. One of the reasons for the spike in the number of leaks was the difference in water chemistry between the two utility companies. The water quality differs in the chlorine level, hardness, and pH. For more details on the difference in water chemistry between the two utility companies please refer to Scardina and Edwards (2007). To resolve the problem, experts from Virginia Tech suggested adding corrosion inhibitor (phosphate) to the water treatment process, which resulted in drop of newly reported pinhole leak incidents (Lee, 2008).

The first survey acquires information on the incidents of pinhole leaks in the residential area, the adoption rate of preventive measures against corrosion and the homeowners' preferences for corrosion risk, and the costs associated with a leak free environment. To learn the tradeoffs between risk of corrosion and financial burden, CA and CV questions were included. The second survey elicits preferences for three hypothetical plumbing materials with different attribute levels. AHP and CA are conducted to gather plumbing preference information. To avoid an exposure bias[4], the plumbing materials are hypothetical. The sample of respondents is based on the first Southeastern Community survey respondents who were willing to participate in the follow-up questionnaire.

### 3.5.1 Southeastern Community Survey: Conjoint Analysis and Contingent Valuation

#### 3.5.1.1 Survey Distribution

In August 2007, a questionnaire was sent to 1600 households in the Southeastern Community. The community's Property Owners' Association provided a list of the residents' names and addresses and the sample was randomly selected from this list. Members of the board reviewed the survey questions. The Association encouraged

[4] **Exposure Bias** - prior information/experience with the studied issues, which might influence the study outcome.

participation of community residents in the Virginia Tech survey. The survey was distributed following the Dillman technique of mail surveying, which included mailing a questionnaire with postage-paid return envelope, sending a reminder card, and mailing a second copy of the survey to nonresponders (Dillman, 1978).

#### 3.5.1.2 Survey Pretest

The preference elicitation survey questions were pre-tested by surveying Virginia Tech graduate and undergraduate engineering students. Two preliminary surveys were conducted in July and November 2006 to test the survey clarity, length, and content. To motivate students to participate in the surveys, a 20-minute presentation regarding drinking water infrastructure problems from both engineering and economics viewpoints was given. The names of the materials were not disclosed in the surveys to prevent the results from being affected by previous experiences with a home plumbing system. The researchers were aware that students are not real decision makers in their households, so the results would not reflect the true decisions of homeowners even though participants were provided with information and motivation for completing the survey (Lee et al., forthcoming).

Preliminary data analysis for Conjoint Analysis, Contingent Valuation, and AHP questions was performed. In case of CA and CV, health effects as well as taste and odor were found to be the driving factors for plumbing materials. In case of AHP, students ranked health effects as the most important plumbing material attribute, followed by taste and odor, corrosion resistance, longevity, and price. Resale value and fire resistance were ranked the lowest on the list. The Contingent Valuation results show that the majority of the plumbing material attributes have a significant effect on the willingness-to-pay of respondents. Some of the graduate and undergraduate students complained about the length of the survey while others did not understand and comprehend the questions. The AHP survey design was found the most challenging by the students and took the longest to complete. Based on the respondents' inputs and results of the preliminary analysis, the Southeastern Community survey was designed (Lee et al., forthcoming).

### 3.5.1.3 Conjoint Analysis and Contingent Valuation Questions

In order to learn the preferences for preventive measures against corrosion, the CA and CV questions were designed to collect data on the tradeoffs between the cost of plumbing system replacement and the risk of corrosion. The CV question employs a dichotomous choice question, which asks respondents whether they are willing to pay $X for an insurance policy, which guarantees to pay for any future costs associated with the failure of the home plumbing system. The respondents are allowed only to answer “yes” or “no” to the question. The insurance bid amount varies among respondents. The CA question employs the choice-based format. The respondent chooses the most preferred option from the three following alternatives: 1) upgrading plumbing material, 2) keeping current plumbing material but buying insurance, and 3) keeping current plumbing material and not purchasing insurance. A question example for each of the two methods is presented in the Appendix B (questions 28 and 29).

The experimental design methodology was utilized to design the order of the CA and CV questions in the questionnaire as well as the cost and pinhole leak probability values. The partial factorial experimental design was generated using a SAS Program. SAS code is provided in Appendix D. The design is orthogonal (based on Kuhfled (2005) OPTEX program), which guarantees factor impact estimation to be independent from any other factor influences. The number of independent factors in the design is five. Three factors have 5 levels, one factor has 2 levels, and one factor has 3 levels. Table 3.2 summarizes the factors used in the experimental design.

The factors A, B, C, D, and E are generated independently. As a result, the total possible number of combinations is: $2^1*5^3*3^1 = 750$. The design was set to produce only 50 treatments. The experimental design randomly chose 50 treatments that had the highest efficiency of design and were representative of the design space. The D-efficiency for this design is 99.46%. The highest D-efficiency is the objective of the experimental design (Engineering Statistics Handbook, 2007). The efficiency of the standard fractional factorial is 100%. The efficiency values are a function of the number of points in the design, the number of independent variables in the model, and the maximum standard error for prediction over the design points. .

Table 3.2: Factor and Levels for Conjoint Analysis and Contingent Valuation Experiential Design

| Factor | Description | Level Values |
|---|---|---|
| A | Payment for insurance in CV question | 1000, 5000, 7500, 10000,14000 |
| B | Payment for insurance in CA question | 1000, 5000, 7500, 10000, 14000 |
| C | Payment for plumbing upgrade | 400, 2000, 3000, 4000, 6000 |
| D | Levels of corrosion risk with plumbing upgrade in CA question | 0%, 3%, 6% |
| E | Order of CA and CV questions in survey | 1 for CV and 2 for CA |

### 3.5.2 Follow-up Southeastern Community Survey: Conjoint Analysis and AHP

#### 3.5.2.1 Survey Distribution

A follow-up survey was administered in October 2007 that utilized the CA and AHP preference elicitation methodology to learn household preferences for home plumbing materials. The follow-up survey was sent to three hundred sixty three Southeastern Community householders who responded to the first survey and who agreed to participate in future surveys.  The respondents were exposed to attributes of three hypothetical plumbing system materials, which were left unnamed to avoid a survey exposure bias. Materials were left unnamed, because most homeowners are familiar with at least one material type (copper, plastic, or epoxy coating) and positive or negative experiences with these materials could influence responses. The last survey question elicited the respondents plumbing material rankings. These rankings provide the basis for CA and AHP comparisons. Examples of CA and AHP questions are presented in the Appendix C.

### 3.5.2.2 Conjoint Analysis and AHP Questions

The follow-up survey of Southeastern Community includes the CA, AHP, and baseline[5] methods' questions. The CA questions include two stimuli, which are compared simultaneously. Each respondent rates each of the two alternatives on a scale from 1 to 9. The scale value of 1 indicates the plumbing material is not preferred while 9 indicate an extremely preferred plumbing system. The AHP questions require a pair-wise comparison of plumbing material attributes or attributes levels on 1-9 rating scale. The scale value of 1 indicates that the plumbing material or their attributes are equally preferred while value of 9 implies one of the plumbing material and their attributes being extremely preferred. The CA and AHP questions appeared in different orders to allow for a question order testing for the two methods. Changes in the order of questions in the survey might affect the preferences for the hypothetical plumbing materials.

In order to perform a comparative analysis between the two methods, a baseline method is also included in this survey. The baseline method asks each respondent to choose the most preferred material for home installation based on each material attribute provided in the question. This question appears last in all surveys. The material attributes are listed in Table 3.3.

The experimental design methodology was used to identify the most D-efficient attribute combination exposure. There are 9 independent factors with 4 factors having 3 levels and 5 remaining factors having 2 levels. The following are the independent factors describing the characteristics of the three plumbing materials: A, B, C, D, E, F, and G. The plumbing material attributes are presented in Table 3.3 while Table 3.4 summarizes all factors employed in the experimental design. The number of attribute levels in each stimulus is restricted to 7 while the number of CA questions is restricted to 3. As a result, the total possible number of combinations in the experimental design is: $2^5*3^4 = 10,363$. The design was set to produce only 2 treatments represented by two survey versions for each treatment. For example, one survey version has CA questions in first order followed by AHP questions while the second survey version has AHP question appearing in first position followed by CA questions. Within each treatment, the question order is

[5] **Baseline Method** is employed to aid the comparative analysis of CA and AHP results. It represents the true preferences of households for home plumbing materials.

randomized to avoid an exposure bias. Furthermore, the last question refers to the baseline methodology that directly compares materials A, B, and C (the names of materials are not disclosed to avoid an exposure bias). The D-efficiency for this design is 98.9%, which implies that the number of paired comparisons was reduced to the smallest number necessary for efficient estimation of preference weights (Huber & Zwerina, 1996; Kuhfeld, Tobias, & Garratt, 1994).

Table 3.3: Descriptions of Plumbing Materials Used to Describe Plumbing Materials for CA and Baseline Questions and Plumbing Material Attributes for AHP Questions

| Attributes | Material A | Material B | Material C |
|---|---|---|---|
| Corrosion Resistance | Corrosion proof | Same as material A | Some risk of corrosion |
| Taste / Odor | Compounds released from this material in drinking water plumbing may give a chemical or solvent taste or odor to the water. | Same as material A | Compounds released from this material in drinking water plumbing may give a bitter or metallic taste or odor to the water. |
| Health Effects | Material meets EPA Standards. There is a very small chance that compounds from this plumbing material that are released into drinking water may lead to microbial growth in water. Microbial growth may cause severe illness. | Same as material A | Material meets EPA Standards. There is a very small chance that compounds from this plumbing material that are released into drinking water may cause vomiting, diarrhea, stomach cramps, and nausea. |
| Convenience of Installation | No need to tear into the wall and/or floor. Installation takes around 4 days. | Need to tear into some sections of wall for installation. Installation takes 5-6 days. | Need to tear into the wall and/or floor to replace the existing system. 7-9 days required for installation. |
| Proven performance in market | Less than 10 years in the market | Less than 20 years in the market | More than 50 years in the market |
| Cost (labor + material) | \$9,000 ~ 14,000 depending on the size of house | \$6,500 ~ 13,000 depending on the size of house | \$9,000 ~ 16,000 depending on the size of house |
| Warranty | Warranty is 15 years for the material. | Warranty is 10 years for the material. | A 50 year manufacturer's warranty applies. Some exceptions apply (e.g. warranty reduces to one year if compounds in water corrode pipes). |

Table 3.4: Factor and Levels for Conjoint Analysis and AHP Experiential Design

| Factor | Description | Level Values |
|---|---|---|
| A | Corrosion resistance | 1. Corrosion proof<br>2. Some risk of corrosion |
| B | Taste/odor | 1.  Compounds released from this material into drinking water plumbing may give a chemical or solvent taste or odor to the water`<br>2. Compounds released  from this material in drinking water plumbing may give a bitter or metallic taste or odor to the water |
| C | Health effects | 1. Material meets EPA standards. A very small chance exists that compounds from this plumbing material that are released into drinking water may lead to microbial growth in water. Microbial growth may cause severe illness<br>2. Material meets EPS Standards. A very small chance exists that compounds from this plumbing material that are released into drinking water may cause vomiting, diarrhea, stomach cramps, and nausea |
| D | Convenience of installation | 1. No need to tear into the wall and/or floor. Installation takes about 4 days<br>2. Need to tear into some sections of wall for installation. Installation takes 5-6 days.<br>3. Need to tear into the wall and/or floor to replace the existing system. Installation takes 7 to 9 days. |
| E | Proven performance in market | 1. Less than 10 years in the market<br>2. Less than 20 years in the market<br>3. More than 50 years in the market |
| F | Cost (labor + material) | 1. \$9,000 ~ \$14,000 depending on the size of house<br>2. \$6,500 ~\$13,000 depending on the size of house<br>3. \$9,000 ~ \$16,000 depending on the size of house |
| G | Warranty | 1. Warranty is 15 years for the material<br>2. Warranty is 10 years for the material<br>3. A 50 year manufacturers' warranty applies. Some exceptions apply (e.g. warranty reduces to one year if compounds in water corrode pipes). |
| H | Order of CA and AHP survey | 1 for CA and 2 for AHP |
| I | Number of stimuli in each scenario | 1 for Option A and 2 for Option B |
| J | Number of questions for question set | 1, 2, 3 |

## 3.6 Empirical Analysis

The empirical analysis includes several econometric techniques. The analysis concentrates on estimating the household preferences for corrosion prevention as well as plumbing materials based on the CA questionnaire questions presented in first and follow-up surveys of Southeastern Community. The data estimation process employs the Multinomial Logit and Ordered Logit regressions based on which the household preferences for corrosion preventive measures and plumbing materials are derived. As mentioned earlier the analysis of the CV as well as AHP data have been performed by Eftila Tanellari and Juneseok Lee and are not reported in this chapter.

### 3.6.1 Homeowners' Preferences for Preventive Measures Against Corrosion and Plumbing Materials

#### 3.6.1.1 Household Preferences for Preventive Measures Against Corrosion: Conjoint Analysis Methodology for the Southeastern Community Survey

In order to analyze the households' preferences for corrosion prevention, the CA methodology is employed to analyze the first Southeastern Community survey data. The example of the question is presented in the Appendix B (question 29). In the CA question, each person chooses the most preferred alternative *i* out of a set of three choices 1) upgrade plumbing material, 2) keep current plumbing material but buy insurance, 3) keep current plumbing material and not purchase insurance. A Multinomial Logit model (MNL) is employed in the estimation of the preferences for preventive measures against corrosion.

The Multinomial Logit model is based on the premises of utility-maximizing behavior[6]. A household faces a choice among three corrosion prevention alternatives listed above. The utility obtained from choosing a preventive measure, $i$, by the $n$th household is $U_{ni}$. The decision maker chooses the option yielding the highest level of utility, which implies the following behavioral model: $U_{ni} > U_{nj}$ where $i \neq j$. The level of utility is not observed by the researcher, but the attributes of the alternatives ($x_{ni}$) in the

[6] Utility Maximizing Problem is discussed in section 3.3.

choice set are observed as well as the socioeconomic characteristics of the decision maker ($z_n$). Based on the known variables, a representative utility function can be specified as: $V_{ni} = V(x_{ni}, z_n)$ for all alternatives (Train, 2003).

Since not all aspects of the decision making process are observable, the utility function has to be decomposed into observable and random components: $U_{ni} = V_{ni} + e_{ni}$, where $V_{ni}$ is the observable component and $e_{ni}$ represents the unobservable component. Due to the discussed utility function structure, the probability of choosing a preventive measure, $i$, is:

$$P_{ni} = Probability(U_{ni} > U_{nj}) = Probability\ (V_{ni} + e_{ni} > V_{nj} + e_{nj}) = Probability(e_{nj} - e_{ni} < V_{ni} - V_{nj}),$$
$$for\ all\ i \neq j. \qquad (34)$$

Equation 47 can be re-written to represent a cumulative probability distribution form:

$$P_{ni} = Probability(U_{ni} > U_{nj}) = Probability\ (V_{ni} + e_{ni} > V_{nj} + e_{nj}) = \int_e I(e_{nj} - e_{ni} < V_{ni} - V_{nj}) f(e_n) de_n\ ,$$
$$for\ all\ i \neq j \qquad (35)$$

where $I(.)$ is the indicator function equal to 1 when the expression is parenthesis is true and 0 otherwise and $f(e_n)$ is the joint density of the random vector $e_n = (e_{n1} \ldots e_{nI})$. In order to employ the MNL model in the estimation of the household preferences for corrosion preventive options, the joint density of the random vector is assumed to be independently and identically distributed (IID). The critical part of the IID assumption is the unobserved factors being uncorrelated over the alternatives, as well as having the same variance for all alternatives (Train, 2003).

The household representative utility function for estimating the preferences for preventive options is linear in parameters $V_{ni} = \beta' x_{ni}$ and follows the logit choice probability specification:

$$P_{ni} = (exp(\beta' x_{ni})/(\textstyle\sum_i exp(\beta' x_{ni}),. \qquad (36)$$

where $P_{ni}$ represents the probability of selecting the $i^{th}$ preventive option. The probability is bounded between 0 and 1 with the choice probabilities of each alternative summing to one as well. McFadden (1974) showed that the log-likelihood function is globally concave for these choice probabilities in estimation parameters, $\beta$'s. With the above assumptions holding, the Multinomial Logit model is estimated by finding the values of the $\beta$'s that maximize the log-likelihood function:

$$lnL = \sum_{(n=1)}^{N} \sum_{(i \in C)} y_{in} \{ ( \Sigma_{(r=1 \text{ to } 3)} \beta_{rk} r_{jkn} + \beta_p p_{jn} + \beta_z z_n ) - ln\, \Sigma_{(j \in C)} ( \Sigma_{(r=1 \text{ to } 3)} \beta_{rk} r_{ikn} + \beta_p p_{in} + \beta_z z_n ) \} \quad (37)$$

where $\beta_{rk}$ is the preference parameter associated with risk of corrosion ($r$), $r_{kn}$ is risk of corrosion attribute $k$ for individual $n$, $\beta_{pk}$ is the parameter on profile cost, $p_{kn}$ is the cost attribute $k$, $\beta_z$ is the socioeconomic variable parameter, and $z_n$ represents the socioeconomic variable of Southeastern Community residents (Champ et al., 2003). This variable is constant over the CA set of choices for each household. As each household chooses the preferred alternative from a set of alternatives, inclusion of the socioeconomic variables helps explain the differences in utility of different prevention alternatives and therefore, the households' decision-making process (Train, 2003). Finally, $y_{in}$ represents the sample size and takes a value of *1* if respondent $n$ chooses profile $i$ and *0* otherwise (Champ et al., 2003).

The IID assumption for error terms results in assumption of the independence of irrelevant alternatives (IIA) put on the MNL. This assumption implies that the error term of one alternative provides no information about the error term of another alternative (Train, 2003). This assumption further implies that, all else being equal, a household's choice between two preventive alternatives is unaffected by the available choices. A test developed by Hausman and McFadden (1984) is used to test the assumption of the independence of irrelevant alternatives. Hausman and McFadden test (better known as a Hausman test) compares the estimates of the full model with all corrosion prevention alternatives, $\beta^{\wedge}_f$, which are consistent and efficient to the consistent but inefficient estimates of the restricted model version, $\beta^{\wedge}_r$. The null hypothesis of this assumption states that both estimates are equal to each other. The Hausman test is defined as follows:

$$W = (\hat{\beta}_r - \hat{\beta}_f)'\hat{\Sigma}^{-1}(\hat{\beta}_r - \hat{\beta}_f),\ W \sim X^2(y) \qquad (38)$$

where $y$ is the number of estimated coefficients and $\hat{\Sigma}^{-1}$ is the difference between the estimated covariance matrices from the two estimators. If IIA holds, $W$ is asymptotically distributed Chi-square with degrees of freedom equal to the rows in $\hat{\beta}_f$. Significant values of $W$ indicate that the IIA assumption is violated. When $W$ is negative, then $(\hat{\beta}_r - \hat{\beta}_f)'\hat{\Sigma}^{-1}(\hat{\beta}_r - \hat{\beta}_f)$ is not positive semi-definite and there is evidence that IIA holds (Hausman & McFadden, 1984).

The violation of Hausman test implies that the estimated parameters from logit equation possess the properties of unbiasedness and consistency but not the property of efficiency. Although coefficient estimates are correct, standard errors will be inefficient if IIA does not hold (Fry & Harris, 1998). In the case of IIA violation, several adjustments can be made including specifying different models, allowing for correlated errors, re-specifying the utility function to capture explicitly the error term correlation and allowing for remaining error to be independent, and finally specifying the current utility function while considering the model only as approximation of the problem at hand. The violation of the IIA assumption does not result in a severe disablement of the CA analysis when only average preferences are estimated. The IIA assumption violation impacts the forecasting predictions derived based on the MNL results (Train, 2003).

If the IIA assumption holds true for the Multinomial Logit, the following empirical regression is estimated based on the MNL likelihood function:

$$C = \sum_{(n=1)}^{N} \sum_{(i \in C)} (\Sigma_{(r=1 \text{ to } 3)} \beta_k x_{jkn} + \beta_p p_{jkn} + \beta_z z_n) + e \qquad (39)$$

where $C$ represents the plumbing material choice $C \in \{i = 1,2,3\}$, where $\beta_k$ is the preference parameter associated with risk of corrosion ($r$), $x_{jkn}$ is risk of corrosion attribute ($r$) in profile $j$ for individual $n$, $\beta_p$ is the parameter on profile cost, $p_{jkn}$ is the cost attribute for profile $j$, $\beta_z$ is the parameter on socioeconomic characteristics, $z_n$ represents the socioeconomic variable of Southeastern Community residents, and $e$ is the error term.

Although attributes of the alternatives vary over alternatives; the characteristics of each household do not differ over the corrosion preventive alternatives. As a result, socioeconomic variables can enter the model estimation to leverage and explain the differences in utility between corrosion preventive options in two ways: either by dividing the cost of corrosion prevention by the household income or by interacting the attributes of each corrosion prevention alternative with each of the socioeconomic variables. These methods of accounting for individual characteristics allow for affecting the differences in utility through their interaction with the attributes of the corrosion prevention alternatives (Train, 2003).

When estimating the MNL model, a scale parameter, $\sigma$, is often employed to scale the MNL regression coefficients to reflect the variance of the unobserved portion of utility. For a $V_{ni}$ linear in parameters with coefficient, $\beta^*$, the choice probability is represented as follows:

$$P_{ni} = (exp((\beta^*/\sigma)' x_{ni})/(\sum_i exp((\beta^*/\sigma)' x_{ni})) \quad (40)$$

The MNL coefficients ($\beta = \beta^*/\sigma$) provided by most of statistical software are already scaled by the appropriate value of $\sigma$. Due to the scale parameter, the result interpretation is affected. The MNL coefficient estimates indicate the effect of each observed variable relative to the variance of the unobserved factors as a larger variance leads to smaller coefficients. The scale parameter, however, does not affect the ratio of any two coefficients as it drops out from the computation. Furthermore, the welfare measures computed based on these coefficients are not affected by the adjustment (Train, 2003).

Based on the parameters estimated in the MNL model, the willingness-to-pay values are computed by dividing the preference weight for the attribute risk of corrosion by the marginal utility of money (-1 times the preference weight for cost):

$$WTP = \beta_{rk}/ -\beta_p \quad (41)$$

The compensating variation (*CV*) is also computed by comparing the before and after change in the attribute measuring expected values of maximized utility. The following presents the *CV* computation:

$$CV = (1/\text{-}\beta_p)[(v^0) - (v^1)] \qquad (42)$$

where $v^0$ and $v^1$ are the expressions of utility for the base (staying with current plumbing material and purchasing no insurance) and altered cases (Champ et al., 2003).

As mentioned earlier, the CA question was viewed by half of respondents in the first position with the CV question in the second position and the order was reversed for the other half. To test the impact of order of question on CA preference changes, a dummy variable is inputted into the proposed Multinomial Logit regression model for CA. The dummy variable takes a value of *1* when the CA question is presented first and *0* if the question is presented in second position in the survey. If the parameter coefficient on the dummy variable is statistically significant, then the order of CA question in the survey impacts the CA results and therefore, impacts the preferences, *WTP* and compensating variation for plumbing materials.

#### 3.6.1.2 Household Preferences for Plumbing Materials: Conjoint Analysis Methodology for the Follow-up Survey

The second Southeastern Community survey data analysis also employs the CA methodology to analyze the preferences for plumbing materials. A different procedure than that used in the first survey was employed; each respondent pair-wise rated the preferred plumbing material option. The rating scale ranges from 1 to 9 with 1 indicating a not preferred plumbing material option and 9 indicating the most preferred option. Similarly to household decision-making process of choosing a preventive measure against corrosion, the plumbing material rating exercise is based on the utility-maximizing behaviour as higher plumbing material rating results in increased level of utility and therefore, a higher preference level for a given alternative. The rating scale questions require individuals to make judgements about the magnitude of utility associated with plumbing material profiles. These plumbing material evaluations directly transform utility levels into a rating scale. As a result, an employment of rating models in which the rating value for each profile is regressed on a vector of attribute levels is justified. Error structure in rating models is assumed as additive nuisance parameters rather than structural as in RUM models (Champ et al., 2003).

To analyze the CA data, an Ordered Logit regression is employed. The Ordered Logit is based upon the idea of the cumulative logit, which relies on the cumulative probability. The cumulative probability $CP_{nl}$ is the probability that the *n*th individual is in the *l*th or higher plumbing material valuation category:

$$CP_{nl} = probability\ (R_l \leq l) = \sum_{(l=1\ to\ L)} probability(R_l = L) \qquad (43).$$

The cumulative probability is transformed into the cumulative logit:

$$logit\ CP_{nl} = log(CP_{nl}(1 - CP_{nl})) \qquad (44).$$

The ordered logit simply models the cumulative logit as a linear function of independent variables:

$$logit\ CP_{nl} = \alpha_l - \boldsymbol{\beta x_n} \qquad (45).$$

There is a different intercept for each level of the cumulative logit, but $\boldsymbol{\beta}$ remains constant across rating categories. In addition, the product of $\boldsymbol{\beta}$ and the independent variable, $\boldsymbol{x_n}$, is subtracted rather than added in the model. As a result, each $\alpha_l$ indicates the logit of the odds of being equal to or less than category *l* for the baseline group (when all independent variables are zero). The $\boldsymbol{\beta}$ represents the increase in the log-odds of being higher than category *l* as the independent variable increases by one-unit (Edner, 2005).

The empirical Ordered Logit model is represented by the following regression:

$$R = \alpha_l - \sum_{(n=1\ldots N)} \sum_{(i \in R)} [\sum_{(k=1\ to\ K)} \beta_k x_{jkn} + \beta_p p_{jkn}] + e \qquad (46)$$

where $R$ represents the ordered rating scale (1-9), where $\beta_k$ is the preference parameter associated with the plumbing material attributes , $x_{jkn}$ are the plumbing material attributes in profile *j* for individual *n*, $\beta_p$ is the parameter on profile cost, $p_{jkn}$ is the cost attribute for profile *j* and *e* is the error term (Champ et al., 2003).[7]

Although attributes of the plumbing materials vary over alternatives; the characteristics of each household do not differ over the alternatives. As a result, the socioeconomic variables need to enter the model estimation to leverage and explain the differences in utility levels between corrosion preventive options. These characteristics can enter the model through interaction with the plumbing material attributes (Train, 2003).

As mentioned earlier, the CA question was viewed by half of the respondents in the first position and by the other half in the second position. To test the impact of order of CA question presentation on preference changes, a dummy variable is inputted into the proposed ordered logit for CA. The dummy variable takes a value of *1* when the CA question is presented first and *0* if the question is presented second in the survey. If the parameter coefficient on the dummy variable is statistically significant, then the order of CA question in the survey impacts the households' preferences for plumbing materials.

**3.7 Southeastern Community Survey Results: Descriptive Statistics**

**3.7.1 Pinhole Leak Awareness and Incidents**

A total of 1,047 responses were received, which represents a 65% response rate. Seventy-six percent of respondents reported being very aware of pinhole leaks in their community, 21% said they were somewhat aware of the problem, and 2% said they were unaware of the problem. Nineteen percent reported learning about pinhole leak problems through their own experience, 65% through a neighbor or friend, 48% through the media (Channel 3 or The Village Connection), and 42% through the Copper Pinhole Leak Committee.

Two hundred twelve respondents (20%) reported incidents of pinhole leaks in drinking water pipes in their current homes; 780 respondents (74%) reported no incidents of pinhole leaks; and 32 respondents (3%) were not sure of any incidents. One hundred

[7] If the cost parameter is found statistically significant, the willingness-to-pay values will be computed by dividing the preference weight for the attribute levels by the marginal utility of money (-1 times the preference weight for cost).

twenty eight respondents (60% percent of the respondents with leaks) had 1 or 2 leaks; 47 respondents (22%) had 3 or 4 leaks; 17 (8%) had 5 or 6 leaks; 15 (7%) had 7 or more incidents.  Most respondents with leaks reported the first leak occurred since 2004 (Table 3.5).

The rate of pinhole leaks in Southeastern Community is higher compared to the average-leak rate in non-hotspot areas nationwide. The nationwide rate for non-hotspot corrosion areas is 8% per year. On the other hand, the incidence rate is on par with nationwide hotspot pinhole leak rate ranging from 11% to 21% per year (Scardina et al., 2007). The pinhole leak rate for hotspots and non-hotspots was derived based on a telephone survey of homeowners across the country including regions known to be 'hotspots' of pinhole leaks (Scardina et al., 2007). The Southeastern Community pinhole leak rate is also lower compared to the rate observed in the suburban area of Washington, D.C.  Forty percent of households surveyed by mail in 2004 reported incidents of pinhole leaks (Kleczyk & Bosch, 2008).

Table 3.5: The Year Pinhole Leaks Were First Reported by Southeastern Community Survey Respondents.

| Year | Respondents | Percent |
|---|---|---|
| Since 2004 | 151 | 71 |
| 2000 to 2003 | 44 | 21 |
| 1995 to 1999 | 6 | 3 |
| 1990 to 1994 | 1 | 0 |
| Before 1990 | 0 | 0 |
| Don't know | 3 | 1 |
| Missing/not reported | 7 | 4 |
| Total | 212 | 100 |

Respondents with pinhole leaks had somewhat older homes compared to respondents without leaks (Table 3.6).  Fifty-three percent of respondents without leaks lived in homes built since 2000 compared to 4% of respondents with leaks.  Five percent of respondents without leaks lived in homes built before 1990 compared to 23% of respondents with leaks.

| Table 3.6: Year House Was Built | | | | |
|---|---|---|---|---|
| | Respondents without leaks | | Respondents with leaks | |
| Year house Was built | Number | Percent[a] | Number | Percent[b] |
| Since 2000 | 441 | 53 | 9 | 4 |
| 1995 to 1999 | 240 | 29 | 76 | 36 |
| 1990 to 1994 | 102 | 12 | 75 | 36 |
| Before 1990 | 39 | 5 | 49 | 23 |
| Do not know | 1 | 0 | 2 | 1 |
| Missing/not reported | 12 | 1 | 1 | 0 |
| Total | 835 | 100 | 212 | 100 |

[a]Percent = number divided by total number of respondents without leaks (835).
[b]Percent = number divided by total number of respondents with leaks (212).

Most respondents with leaks had leaks in horizontal pipes while fewer had leaks in vertical pipes or pipe bends (Table 3.7). Most leaks were in the finished or unfinished basement followed by the crawl space and first floor respectively (Table 3.8).

Table 3.7: Pinhole Leaks by Type of Pipe

| Type of pipe | Number of observations | Percent[a] |
|---|---|---|
| Horizontal pipe | 183 | 86 |
| Vertical pipes | 25 | 12 |
| Pipe bend | 6 | 3 |
| Do not know | 9 | 4 |
| Total | 223 | 105 |

[a]Multiple choices per respondent were accepted. Percent = number reported divided by the total number of respondents with leaks (212).

Table 3.8: Location of Leaks in House

| Location | Number of observations | Percent[a] |
|---|---|---|
| Under slab or underground | 9 | 4 |
| Crawl space | 56 | 26 |
| Unfinished basement | 43 | 20 |
| Finished or walkout basement | 97 | 46 |
| First floor | 29 | 14 |
| Second floor or higher | 9 | 4 |
| Do not know | 1 | 0 |
| Total | 244 | 114 |

[a]Multiple choices per respondent were accepted. Percent = number reported divided by the total number of respondents with leaks (212).

Pinhole leaks occurred in cold water pipes in 138 cases, in both cold and hot water pipes in 14 of the cases, and in hot water pipes only in 33 cases (Table 3.9). Twenty respondents were not aware of the type of water pipes where leaks occurred.

| Table 3.9: Pinhole Leaks Occurring in Cold or Hot Water Pipes | | |
|---|---|---|
| Type of pipe | Number of observations | Percent |
| Cold water pipes | 138 | 65 |
| Hot water pipes | 33 | 16 |
| Both | 14 | 7 |
| Do not know | 20 | 9 |
| Missing/not reported | 7 | 3 |
| Total | 212 | 100 |

### 3.7.2 Pinhole Leak Repairs and Repair Costs

Seventy-seven respondents repaired the leak using a clamp (Table 3.10). In some cases, a clamp was used initially and the leaking section was replaced later as the leaks continued. One hundred thirty-three respondents repaired leaks by replacing the leaking pipe section. Copper plumbing material was most often used for repairing leaking plumbing material sections. Fifty respondents repaired the leak by re-plumbing the entire house. PEX was most often used for house re-plumbing activities. Nine respondents applied epoxy coating to their existing plumbing systems.

| Table 3.10: Method of Leak Repair | | |
|---|---|---|
| Repair method | Number of observations | Percent[a] |
| Clamp over leak | 77 | 7 |
| Replaced leaking pipe section with copper | 75 | 35 |
| Replaced leaking pipe section with CPVC | 5 | 2 |
| Replaced leaking pipe section with PEX | 7 | 3 |
| Replaced leaking pipe section-material not specified | 46 | 22 |
| Applied epoxy coating to all plumbing | 9 | 4 |
| Re-plumbed with copper | 5 | 2 |
| Re-plumbed with CPV | 4 | 2 |
| Re-plumbed with PEX | 32 | 15 |
| Re-plumbed-material not specified | 9 | 4 |
| Other | 7 | 1 |
| Don’t know | 3 | 1 |
| Total | 279 | 129 |

[a]Multiple choices per respondent were accepted. Percent = number reported divided by the total number of respondents with leaks (212).

Forty-two percent of respondents with leaks spent less than 10 hours dealing with pinhole leaks; 21% spent between 11 and 20 hours; and 32% spent 21 or more hours (Table 3.11). Twenty percent spent more than 40 hours dealing with pinhole leaks.

Table 3.11: Time Spent Dealing with Pinhole Leak Problems by Survey Respondents

| Number of hours | Number of respondents | Percent[a] |
|---|---|---|
| Less than 10 | 89 | 42 |
| 11-20 | 44 | 21 |
| 21-40 | 25 | 12 |
| 41-80 | 26 | 12 |
| More than 80 | 18 | 8 |
| Do not know | 7 | 3 |
| Missing/not reported | 3 | 1 |
| Total | 212 | 99 |

[a]Numbers do not sum to 100 due to rounding. Percent = number reported divided by the total number of respondents with leaks (212).

Twenty-nine percent of respondents with leaks reported that the expense of repairing pinhole leaks was less than $100 while 30% reported expenses between $100 and $500, and 37% reported more than $500 in expenses for pinhole leak repairs (Table 3.12). Seven respondents reported more than $10,000 in costs of repairs.

Table 3.12: Costs of Repairing Pinhole Leaks

| Amount | Number of observations | Percent |
|---|---|---|
| Less than $100 | 61 | 29 |
| $100 to $500 | 64 | 30 |
| $501 to $1,000 | 14 | 7 |
| $1,001 to $3,000 | 11 | 5 |
| $3,001 to $5,000 | 20 | 9 |
| $5,001 to $10,000 | 28 | 13 |
| $10,001 to $20,000 | 6 | 3 |
| More than $20,000 | 1 | 0 |
| Do not know | 3 | 1 |
| Missing/not reported | 4 | 2 |
| Total | 212 | 99 |

[a]Numbers do not sum to 100 due to rounding. Percent = number reported divided by the total number of respondents with leaks (212).

In addition to the expense of repairing leaks, 92% of respondents with leaks reported having to repair property damage caused by leaks. Forty percent of respondents with damage reported less than $100 of damage (Table 3.13) while 49% had over $100 in damage. Twelve respondents had over $5,000 in property damage.

Table 3.13: Costs of Repairing Property Damage from Pinhole Leaks

| Amount | Number of observations | Percent[a] |
|---|---|---|
| Less than $100 | 85 | 40 |
| $100 to $500 | 47 | 22 |
| $501 to $1,000 | 16 | 8 |
| $1,001 to $3,000 | 18 | 8 |
| $3,001 to $5,000 | 10 | 5 |
| $5,001 to $10,000 | 12 | 6 |
| More than $10,000 | 0 | 0 |
| Do not know | 8 | 4 |
| No property damage | 16 | 8 |
| Total | 212 | 101 |

[a]Totals do not sum to 100 due to rounding. Percent = number reported divided by the total number of respondents with leaks (212).

Thirty-six percent of respondents reporting leaks found the experience of pinhole leaks very stressful while 46% found it somewhat stressful (Table 3.14). Thirteen percent experienced little or no stress.

Table 3.14: Stress from Dealing with Pinhole Leaks

| Stress level reported by respondents with leaks | Number of observations | Percent[a] |
|---|---|---|
| Very stressful | 77 | 36 |
| Somewhat stressful | 98 | 46 |
| A little stressful | 21 | 10 |
| Not stressful | 7 | 3 |
| Do not know | 4 | 2 |
| Missing/not reported | 5 | 2 |
| Total | 212 | 99 |

[a]Totals do not sum to 100 due to rounding. Percent = number reported divided by the total number of respondents with leaks (212).

### 3.7.3 Pinhole Prevention and Water Treatment Devices

Thirty-five percent of respondents with leaks and 20% of respondents without leaks use some type of pinhole leak prevention strategy (Table 3.15). The most common strategy among those with leaks is preventive re-plumbing, which was used by 13% of those with leaks. Water softener/conditioner was the most common strategy used by those without leaks and constituted 9% of those respondents.

Table 3.15: Use of Pinhole Leak Prevention Devices

| | Respondents with pinhole leaks[a] | | Respondents without pinhole leaks[b] | |
|---|---|---|---|---|
| | Number | Percent | Number | Percent |
| Preventive re-plumbing | 28 | 13 | 16 | 2 |
| Preventive epoxy injection | 8 | 4 | 4 | 0 |
| Phosphate injection | 12 | 6 | 26 | 3 |
| Water softener/water conditioner | 11 | 5 | 79 | 9 |
| Copper Knight | 5 | 2 | 12 | 1 |
| Other | 19 | 9 | 64 | 8 |
| None used | 134 | 63 | 644 | 77 |
| Missing/not reported | 4 | 2 | 29 | 3 |
| Total | 295 | 139 | 874 | 105 |

[a]Multiple choices per respondent were accepted. Percent = number reported divided by the total number of respondents with leaks (212).
[b] Percent = number reported divided by the total number of respondents without leaks (835).

The cost of purchasing and installing the preventive measure varies between less than a $100 to more than $10,000 for respondents with pinhole leaks (Table 3.16). The highest number of respondents chose a cost category of $100-$499 (14 respondents with pinhole leaks). Twenty-three percent of respondents with pinhole leaks spent between $500 and $7,499 while 6 respondents spent between $10,000 and $15,000.

| Table 3.16: Cost of Purchasing or Installing Preventive Corrosion Measure | | | | |
|---|---|---|---|---|
| | Respondents with pinhole leaks | | Respondents without pinhole leaks | |
| | Number | Percent[a] | Number | Percent[b] |
| $1 - $99 | 3 | 1 | 7 | 1 |
| $100 - $499 | 14 | 7 | 21 | 3 |
| $500 - $999 | 10 | 5 | 33 | 4 |
| $1,000 - $2,499 | 12 | 6 | 17 | 2 |
| $2,500 - $4,999 | 10 | 5 | 25 | 3 |
| $5,000 - $7,499 | 14 | 7 | 19 | 2 |
| $7,500 – $9,999 | 1 | 0 | 0 | 0 |
| $10,000 -14,999 | 6 | 3 | 2 | 0 |
| $15,000 - $19,999 | 0 | 0 | 0 | 0 |
| 20,000 or more | 1 | 0 | 0 | 0 |
| Do not know | 9 | 4 | 16 | 2 |
| Missing/not reported | 132 | 62 | 640 | 82 |
| Total | 212 | 100 | 780 | 100 |

[a]Percent = number reported divided by the total number of respondents with leaks (212).
[b]Percent = number reported divided by the total number of respondents without leaks (835).

Sixty-seven percent of respondents use some type of water treatment for purposes other than pinhole leak prevention (Table 3.17). The most common treatment was a refrigerator filter, used by 63%. Thirty-two percent of respondents reported that they purchase drinking water. The most common reasons given for using water treatment devices are to improve taste or smell of drinking water (mentioned by 45% of respondents) and to improve safety of drinking water (mentioned by 33% of respondents).

| Table 3.17: Use of Water for Purpose Other than Corrosion Prevention | | |
|---|---|---|
| | Number | Percent [a] |
| Filter for entire home | 133 | 16 |
| Refrigerator filter | 523 | 63 |
| Water softener/water conditioner | 66 | 8 |
| Pitcher or bottle to filter water | 136 | 16 |
| Purchased drinking water | 265 | 32 |
| Filter on faucet or under kitchen sink | 117 | 14 |
| Ultra violet (UV) system | 2 | 0 |
| Other | 25 | 3 |
| None used | 249 | 30 |
| Missing/not reported | 19 | 2 |
| Total | 1,535 | 184 |

[a]Multiple choices per respondent were accepted. Percent = number reported divided by the total number of respondents (1,047).

### 3.7.4 Concerns about Water Safety and Quality

Eighty-two percent of respondents were somewhat or very satisfied with home drinking water quality (Table 3.18). Only 5% of respondents were not at all satisfied with water quality. Problems with water quality most frequently mentioned were related to taste, in particular to a chlorine taste.

Table 3.18: Satisfaction with Home Drinking Water Quality

| | Number | Percent [a] |
|---|---|---|
| Very satisfied | 471 | 45 |
| Somewhat satisfied | 390 | 37 |
| Somewhat dissatisfied | 119 | 11 |
| Not at all satisfied | 54 | 5 |
| Missing/not reported | 13 | 1 |
| Total | 1,047 | 99 |

[a]Totals do not sum to 100 due to rounding.

Survey respondents varied in concern about future pinhole leaks. Forty percent were somewhat or very concerned while 55% were not very or not at all concerned (Table 3.19).

Table 3.19: Concern with Future Pinhole Leaks

| | Number | Percent |
|---|---|---|
| Very concerned | 65 | 6 |
| Somewhat concerned | 359 | 34 |
| Not very concerned | 400 | 38 |
| Not at all concerned | 173 | 17 |
| Do not know | 29 | 3 |
| Missing/not reported | 21 | 2 |
| Total | 1047 | 100 |

In summary, pinhole leaks in copper home plumbing systems were a frequently reported problem in Southeastern Community. Twenty percent of respondents had experienced at least one pinhole leak in their current homes. Almost all pinhole leaks occurred since 2000. Most respondents had 1 or 2 leaks; however, 15 respondents reported 7 or more leaks. Most leaks occurred in cold water plumbing and in horizontal pipes. Most leaks occurred in the basement (finished or unfinished).

Pinhole leaks can have high financial, emotional, and time costs. Over one-third of respondents with leaks reported expenditures of at least $500 for repairing leaks and over one-fourth reported additional expenses of at least $500 to repair property damage caused by leaks. Almost one-third of those reporting leaks spent at least 20 hours dealing with pinhole leaks and related damage, and over three-fourths found the experience somewhat or very stressful. Almost one-third of respondents with leaks used some form of pinhole leak protection, including preventive re-plumbing. Twenty percent of respondents without leaks used some form of pinhole leak protection. Two-thirds of respondents use some type of water treatment device or purchase drinking water. Respondents mentioned improving water safety and water taste and smell as reasons for using such devices or purchasing bottled water. Over 80% of respondents were somewhat or very satisfied with drinking water quality in their homes.

### 3.7.5 Socioeconomic Characteristics of Southeastern Community Respondents

Almost all of the Southeastern Community respondents live in a single family home that they own (98%) (Table 3.20). Fifty-six percent of all respondents, 95% of respondents with pinhole leaks, and 55% of respondents without pinhole leaks live in homes built before the year 2000. In comparison, only 4% of respondents with pinhole leak experience and 55% of respondents without the experience live in houses built after the year 2000 (Table 3.21).

Table 3.20: Type of Residence for Southeastern Community Mail Survey Respondents

| | All Respondents | | Respondents with pinhole leaks | | Respondents without pinhole leaks | |
|---|---|---|---|---|---|---|
| | Number | Percent | Number | Percent | Number | Percent |
| Single family home which I own | 1023 | 98 | 208 | 98 | 766 | 98 |
| Single family home which I rent | 8 | 1 | 4 | 2 | 3 | 0 |
| Missing/not reported | 16 | 2 | 0 | 0 | 11 | 1 |
| Total | 1047 | 100 | 212 | 100 | 780 | 100 |

Note: Respondents who did not know whether they had pinhole leaks were not included in the comparisons of the respondents with and without pinhole leaks analysis.

| Table 3.21: Year House Was Built for Southeastern Community Mail Survey Respondents | | | | | | |
|---|---|---|---|---|---|---|
| | All Respondent | | Respondents with pinhole leaks | | Respondents without pinhole leaks | |
| | Number | Percent | Number | Percent | Number | Percent |
| Since 2000 | 450 | 43 | 9 | 4 | 429 | 55 |
| 1995 to 1999 | 316 | 30 | 76 | 36 | 218 | 28 |
| 1990 to 1994 | 177 | 17 | 75 | 35 | 92 | 12 |
| Before 1990 | 88 | 8 | 49 | 23 | 32 | 4 |
| Don't know | 3 | 0 | 2 | 1 | 1 | 0 |
| Missing/not reported | 13 | 1 | 1 | 0 | 8 | 1 |
| Total | 1047 | 100 | 212 | 100 | 780 | 100 |

Note: Respondents who did not know whether they had pinhole leaks were not included in the comparisons of the respondents with and without pinhole leaks analysis.

In Southeastern Community, the property values vary between less than $200,000 and more than $750,000. For example, 3% of respondents live in houses valued below $200,000 while 29% live in houses valued above $400,000. There are no drastic differences between the respondents with pinhole leaks and without, but percentage wise there are more respondents with pinhole leaks living in houses below $300,000 (31% vs. 24%) and above $500,000 (33% vs. 28%) (Table 3.22).

| Table 3.22: The House Value for Southeastern Community Mail Survey Respondents | | | | | | |
|---|---|---|---|---|---|---|
| | All Respondents | | Respondents with pinhole leaks | | Respondents without pinhole leaks | |
| | Number | Percent | Number | Percent | Number | Percent |
| Less than $200,000 | 28 | 3 | 9 | 4 | 16 | 2 |
| $200,001 - $250,000 | 97 | 9 | 24 | 11 | 66 | 8 |
| $250,001 - $300,000 | 151 | 14 | 33 | 16 | 111 | 14 |
| $300,001 - $350,000 | 109 | 10 | 19 | 9 | 88 | 11 |
| $350,001 - $400,000 | 122 | 12 | 23 | 11 | 95 | 12 |
| $400,001 - $500,000 | 144 | 14 | 16 | 8 | 122 | 16 |
| $500,001 - $750,000 | 161 | 15 | 39 | 18 | 112 | 14 |
| Over $750,001 | 146 | 14 | 31 | 15 | 106 | 14 |
| Don't know | 27 | 3 | 6 | 3 | 17 | 2 |
| Missing/not reported | 62 | 6 | 12 | 6 | 47 | 6 |
| Total | 1047 | 100 | 212 | 100 | 780 | 100 |

As presented in Table 3.23, fifty four percent of all respondents moved into their houses since 2000 while only 3% moved in before 1990.

| Table 3.23: The Move-In Year for Southeastern Community Mail Survey Respondents | | |
|---|---|---|
| | Number | Percent |
| Before 1990 | 30 | 3 |
| 1990-1994 | 99 | 9 |
| 1995-1999 | 250 | 23 |
| 2000 – 2005 | 395 | 30 |
| 2005 - 2007 | 254 | 24 |
| Missing/not reported | 19 | 2 |
| Total | 1047 | 100 |

A vast majority of the households consists of two household members, although the highest number reported by survey respondents is nine. Seventy-nine percent of the survey respondents are reported to be males and 88% to be white. They tend to be older with 69% being born between 1930 and 1949 (Table 3.24). Seventy-one percent of the Southeastern Community residents have either a 4-year college degree or graduate/post-graduate degree (Table 3.25). Fifty-seven percent of survey respondents have income higher than $55,000 (Table 3.26).

| Table 3.24: Southeastern Community Respondents' Year of Birth | | |
|---|---|---|
| | Number | Percent |
| 1916-1919 | 7 | 1 |
| 1920-1929 | 71 | 7 |
| 1930-1939 | 335 | 32 |
| 1940-1949 | 383 | 37 |
| 1950-1959 | 68 | 6 |
| 1960-1969 | 9 | 1 |
| 1970-1981 | 11 | 1 |
| Missing/not reported | 163 | 16 |
| Total | 1047 | 100 |

| Table 3.25: Household Education Level for Southeastern Community Mail Survey Respondents | | |
|---|---|---|
| | Number | Percent |
| Grade School | 2 | 0 |
| Some High School | 8 | 1 |
| High School Graduate (or GED) | 45 | 4 |
| Trade/Voc. School after High School | 25 | 2 |
| Some College | 141 | 13 |
| Completed Community College | 40 | 4 |
| Four-year College/University Graduate | 406 | 39 |
| Graduate School/Professional School | 340 | 32 |
| Missing/not reported | 40 | 4 |
| Total | 1047 | 100 |

| Table 3.26: Household Income for Southeastern Community Mail Survey Respondents | | |
|---|---|---|
| | Number | Percent |
| Less than $35,000 | 36 | 3 |
| $35,001 - $55,000 | 125 | 12 |
| $55,001 - $75,000 | 180 | 17 |
| $75,001 - $100,000 | 188 | 18 |
| $100,001 - $ 150,000 | 150 | 14 |
| $151,000 - $200,000 | 46 | 4 |
| Over $200,000 | 45 | 4 |
| Do not know | 44 | 4 |
| Missing/not reported | 233 | 22 |
| Total | 1047 | 100 |

## 3.8 Homeowners Preferences for Prevention Measures Against Corrosion and Plumbing Materials

### 3.8.1 Preferences for Preventive Measures Against Corrosion: Conjoint Analysis Results for the Southeastern Community Survey

The first Southeastern Community survey employed CA and CV questions to elicit preferences for the trade-offs between the risk of corrosion and the cost of prevention. The CA questions consisted of three-profile presentation that varied on the possible solutions to protect their households from corrosion. Each solution traded-off the risk of corrosion with a prevention cost. The following options were presented to each respondent to elicit the corrosion-cost trade-offs (the question text is presented in Appendix B):

1. Upgrading plumbing material

2. Keeping current plumbing material but purchasing insurance
3. Keeping current plumbing material only

The probabilities of corrosion risk presented to respondents varied between 0 and 0.10 while cost associated with the chosen corrosion prevention (i.e house re-plumbing and purchasing an insurance policy) varied between $0 and $14,000.

The empirical analysis of the first Southeastern Community survey includes several econometric techniques. First, the Multinomial Logit Model is used to derive willingness-to-pay estimates. The model specification involves either including only the risk of corrosion levels as well as the dollar costs of re-plumbing or purchasing an insurance policy or in addition including the socioeconomic variables. The purpose of two model specifications is to learn the impact of socioeconomic characteristics of the respondents on the decision either to purchase an insurance policy or to upgrade the plumbing material.

#### 3.8.1.1 Conjoint Analysis Regression Estimation

Based on the three profiles of possible solutions to protect their households from corrosion, 69% of respondents who responded to the CA question chose to keep current plumbing material and not purchase any insurance to protect their household from pinhole leaks and failure of the plumbing material. With regards to the residents who had experienced pinhole leaks, 23% would upgrade the plumbing material while 9% would purchase an insurance policy. In comparison, 12% of those who had not experienced pinhole leaks would upgrade a plumbing material and 12% would purchase an insurance policy. The percent of respondents who would stay with the current plumbing material is higher for people without pinhole leak experience than for those with pinhole leak experience (Table 3.27).

| Table 3.27: Conjoint Analysis Frequency Alternatives | | | | | | |
|---|---|---|---|---|---|---|
| | All Respondents | | Respondent with Pinhole Leaks | | Respondents without Pinhole Leaks | |
| Option | Number | Percent | Number | Percent | Number | Percent |
| Upgrading Plumbing Material | 147 | 14 | 49 | 23 | 92 | 12 |
| Keeping Current Plumbing Material but Purchasing Insurance | 124 | 12 | 20 | 9 | 93 | 12 |
| Keeping Current Plumbing Material and Not Purchasing Insurance | 715 | 69 | 124 | 59 | 559 | 72 |
| Missing/not reported | 61 | 6 | 19 | 9 | 36 | 5 |
| Total Number of Observations | 986 | 100 | 212 | 100 | 780 | 100 |

In this study, the Conjoint Analysis (CA) evaluates the trade-offs between the risk of corrosion and the cost associated with the chosen prevention method of either re-plumbing the entire house or purchasing insurance against corrosion by estimation of a Multinomial Logit Model (MNL). The model specification includes/excludes the socioeconomic characteristics of the respondents. The first model specification omits the socioeconomic variables' impact on the household decision and only accounts for the corrosion risk and cost associated with several prevention options. In the model specification, the independent variables (corrosion risk and the dollar amount spent on corrosion prevention) were entered as continuous covariates. Estimating the impact of continuous variables implies linear impact of each unit of the variable on the household decision[8].

The final explanatory variable included in the model specification is a dummy variable representing the order of the CA question in the survey (value of 1 when in first position and value of 0 if in the second position). This factor was entered as a covariate variable. The dummy variable was tested for the appropriate form of entering the

[8] The independent variables could be also specified as categorical factors. When a categorical independent variable is employed, each level of the factor is assumed to have a nonlinear impact on the dependent variable. For the purpose of this study, the independent variables are assumed to have a constant impact per unit change on the chosen prevention option.

regression model.  The variable was treated either as the intercept shifter in the equation or as an intercept shifter combined with an interaction effect with the independent variables (risk of corrosion as well as the cost of plumbing material upgrade and insurance purchase). Wald test[9] performed to compare the two model specifications revealed that the CA survey version dummy variable should enter the MNL model only as an intercept shifter. Wald test is often employed to test whether the restrictions put on some coefficients in the unrestricted model are appropriate or not. The test compares the coefficients of the restricted and unrestricted models and tests for differences between the two specifications (Baum, 2006). The Chi-square statistics derived based on the two model comparisons was 1.789 (p-value of 0.770) and was less than the critical Chi-square statistic of 9.490 with 4 degrees of freedom.

In order to employ the MNL estimation, the CA question responses need to be appropriately transformed to fit the data estimation requirement.  As each respondent saw three prevention options with specific attributes, each respondent's data was stacked in a format resulting in each prevention option being a separate observation.  As 985 respondents provided answers to the CA question, there are 2,955 rows of data employed in the MNL estimation process.  The cost variable for the chosen preventive option was also stacked to form only one variable cost of corrosion prevention.  The cost values varied between \$0 and \$14,000. Value of zero refers to staying with the current plumbing material as no financial investment is needed for this option.

This MNL model specification employs the continuous format of the explanatory variables while assuming that the impact of a one unit change in the independent variable has a linear impact on the household decision for corrosion preventive option. The first

---

[9] **Wald Test** - employed to test whether an effect of an independent variables on the dependent variables exists or not. Under this test the maximum likelihood estimate of the parameter(s) of interest, $\beta$, is compared with the proposed value $\beta_0$. The square of the difference is compared to a Chi-square distribution. The Wald statistic is defines as:

$$(\beta^{\wedge} - \beta^{\wedge}_{0})' \sum^{-1}(\beta^{\wedge} - \beta^{\wedge}_{0}),\ W \sim X^2_k$$

where $\sum^{-1}$ variance-covariance matrix  associated with both coefficient estimates. The obtained statistic is compared against a Chi-square distribution for derive statistical significance of the tested coefficient estimates. (Baum, 2006)

MNL model specification includes continuous independent variables. In order to evaluate the household decisions, the model significance should be assessed first. The null hypothesis for the model is that no statistically significant differences exist between the model without independent variables and the model with independent variables:

$$H_0: C = e \qquad (80)$$

$$H_1: C = \sum_{(n=1)}^{N} \sum_{(i \in C)} \Sigma_{(r=1 \text{ to } 3)} [\beta_k x_{jkn}] + \beta_p p_{jkn} + e \qquad (81)$$

where $C$ represents the plumbing material choice $C \in \{i = 1,2,3\}$, $\beta_k$ is the preference parameter associated with risk of corrosion ($r$), $x_{jkn}$ is risk of corrosion attribute ($r$) in profile $j$ for individual $n$, $\beta_p$ is the parameter on profile cost, $p_{jkn}$ is the cost attribute for profile $j$, and $e$ is the error term.

The presence of a relationship between the dependent variable and the combination of independent variables is assessed based on the Chi-square and log-likelihood statistics of the final model. Based on the analysis, the Chi-square statistic is 45.344 with a p-value of 0.000 and a log-likelihood value of 590.606[10]. The existence of a relationship between the independent variables and the dependent variable is supported by the results (Garson, 1998).

The next step in the CA analysis is to assess fit of the MNL model (Garson, 1998). A Pseudo R-square statistic, such as Nagelkerke's R-square[11], is computed to represent the improvement of the model's fit from inclusion of the independent variables. The measure ranges from zero (when the estimated parameters are not better at predicting choices than zero parameters) to one (when the estimated parameters perfectly predict the choices of the sampled households) (Train, 2003). The goodness-of-fit test statistic is 0.019 for this model specification and represents the percentage increase in the log-

---

[10] A **Likelihood-Ratio test** is a test relying on a test statistic computed by taking the ratio of the maximum value of the likelihood function under the constraint of the null hypothesis to the maximum with that constraint relaxed. If the null hypothesis is true, then $-2 \log \Lambda$ is asymptotically $\chi^2$ distributed with degrees of freedom equal to the difference in dimensionality of $\Theta$ and $\Theta_0$. The likelihood ratio is $\Lambda = \sup\{L(\Theta|x): \Theta \in \Theta_0\} / \sup\{L(\Theta|x): \Theta \in \Theta\}$ (Griffiths et al., 1993).

[11] **Pseudo R-square** is defined as $\rho = 1 - (LL(\beta\hat{})/LL(0))$, where $LL(\beta\hat{})$ is the value of the log-likelihood at the estimated parameters and $LL(0)$ is the value when all parameters are zero (Train, 2003).

likelihood function above the value taken at zero parameters. As a result, there is a 2% increase in the log-likelihood function above the function at zero and the estimated parameters do improve the model's fit. Pseudo R-square does not have exactly the same meaning as R-square in the linear regression. As a result, the Nagelkerke's R-square of 0.019 should not be interpreted as 2% of variation in the dependent variable is explained by the variation in the independent variables (Train, 2003).

Another measure of assessing the performance of a MNL regression is the classification of accuracy, which compares the predicted group membership (based on the logit model predictions) to the actual known group membership (the value for the dependent variable). The useful benchmark employed to characterize a regression model is a 25% improvement over the rate of accuracy achievable by chance alone. Even if the independent variables were not related to the groups defined by the dependent variable, correct predictions of the group membership for some percentage of the observations would be still expected (Garson, 1998).

The estimate of by chance accuracy is proportional to the by chance accuracy rate computed by summing the squared percentage of cases in each group (please refer to Table 3.28) (Garson, 1998). The proportional by chance accuracy rate for this model is 56.4%. To characterize the model as useful, the overall percentage accuracy rate produced by SPSS software at the last step of estimation is compared to 25% more than the proportional by chance accuracy, that is, equal to 70.5%. The model's classification accuracy rate is 72.5% (Table 3.29), which is greater than the proportional by chance accuracy criteria of 70.5%. As a result, the criterion for classification accuracy is satisfied.

Table 3.28: Case Processing Summary for MNL Model without Socioeconomic Characteristics

| Group Membership | Number of Observations | Marginal Percentage |
|---|---|---|
| Upgrading Plumbing Material | 147 | 14.9 |
| Keeping Current Plumbing Material but Buying Insurance | 124 | 12.6 |
| Keeping Current Plumbing Material and Not Purchasing Insurance | 715 | 72.5 |

Note: Chance Accuracy = $0.149^2 + 0.126^2 + 0.725^2 = 0.564$
25% Improvement over the Chance Accuracy = 0.564*0.25 = 0.141
Chance Accuracy Criteria = 0.564 + 0.141 = 0.705

Table 3.29: Classification Table for MNL Model without Socioeconomic Characteristics

| Observed | Predicted: Upgrading Plumbing Material | Keeping Current Plumbing Material but Buying Insurance | Keeping Current Plumbing Material and Not Purchasing Insurance | Percent Correct |
|---|---|---|---|---|
| Upgrading Plumbing Material | 0 | 0 | 147 | 0.0 |
| Keeping Current Plumbing Material but Buying Insurance | 0 | 0 | 124 | 0 |
| Keeping Current Plumbing Material and Not Purchasing Insurance | 0 | 0 | 715 | 100 |
| Overall Percentage | 0 | 0% | 100% | 72.5% |

Note 1: Percent correct is defined as the proportion of predicted observations in a category compared to the observed number of observations in that group.
Note 2: Overall percentage is proportion of predicted observations in a group compared to the total number of observations.

In order to evaluate the importance of independent variables in households' decisions with regards to (1) upgrading plumbing material, (2) keeping current plumbing material but buying insurance, and (3) keeping current plumbing material and not purchasing insurance, two types of tests are performed (Garson, 1998):

1. The likelihood ratio test evaluating the overall relationship between an independent variable and the dependent variable;
2. The Wald tests evaluating whether or not the independent variable is statistically significant in differentiating between the two groups (base and comparison group). For example, the test measures whether an effect of an independent variables on the dependent variables exists or not (is statistically different from zero).

If an independent variable has an overall relationship with the dependent variable, it might or might not be statistically significant in differentiating between pairs of groups defined by the dependent variable. The interpretation for an independent variable centers on its ability to distinguish between pairs of groups and the independent variable's contribution to changing the odds of being in one dependent variable group rather than the other. To make significance statements about an independent variable, both the correlation and the parameter estimate should be significant. There are several reasons why correlations and logit coefficient may differ in significance: (1) logit coefficients are partial coefficients; (2) logit coefficients reflect linear and nonlinear relationships; and (3) a significant parameter estimate means there is a relation of the independent variable to the dependent variable for the selected control groups only. This finding further implies that while the logit coefficient is statistically significant, the correlation between the independent variable and the dependent is not (Garson, 1998).

In this analysis, two comparisons between the dependent variable groups are made:

1. The upgrade to new plumbing material group (coded as group 1) is compared to keeping current plumbing material and not purchasing insurance (coded as group 3)
2. Keeping current plumbing material but purchasing insurance (coded as group 2) is compared to keeping current plumbing material and not purchasing insurance (coded as group 3).

The reference category (group 3) is the category against which all other categories are compared.

For an MNL model specification that employs the independent variables as covariates and tests the overall significance of the relationship between the independent variables and the dependent variable, the results show that the order of the CA question in the survey and the cost of plumbing material upgrade or insurance policy purchase have an overall relationship with the households' decision. On the other hand, there is a statistically insignificant relationship between the dependent variable and the level of corrosion risk (Table 3.30).

Table 3.30: Likelihood Ratio Tests for MNL Model without Socioeconomic Characteristics

| | Model Fitting Criteria | Likelihood Ratio Test | |
|---|---|---|---|
| Effect | -2 Log Likelihood of Reduced Model | Chi-Square | Significance |
| CA Question Order | 618.827 | 28.258 | 0.000 |
| Corrosion Risk Level | 594.131 | 3.525 | 0.172 |
| Cost of Plumbing Upgrade or Insurance Purchase | 602.864 | 12.258 | 0.002 |

Before performing the MNL model estimation, the assumption of independence of irrelevant alternatives is tested to find if employment of the MNL model is appropriate for this analysis. In order to test for the assumption, a Hausman test is performed to compare coefficient estimates obtained before and after the removal of one of the corrosion prevention choices. For example, $\hat{\beta}_r$ indicates a parameter estimate representing the restricted data set of those respondents who chose 1) to re-plumb the house or stay with the current material; 2) to purchase an insurance policy against corrosion or stay with the current plumbing material; and 3) to re-plumb or purchase an insurance policy against corrosion. $\hat{\beta}_f$ represents the more efficient estimates derived based on the full data set while $\sum^{-1}$ represents the difference between the corresponding variance-covariance matrices. Some parameters estimated in the full model may not be identified in the restricted model, in which case, $\hat{\beta}_f$ refers to estimates under the restricted model that are identified in both setups (Train, 2003).

Based on the above specification, the Hausman test statistic ($W$) is defined as:

$$(\hat{\beta}_f - \hat{\beta}_r)' \textstyle\sum^{-1}(\hat{\beta}_f - \hat{\beta}_r),\ W \sim X^2_k \qquad (47)$$

where $k$ represents the number of elements in the restricted $\beta$ vector. The null hypothesis of this Hausman test is that both estimators are consistent while under the alternative hypothesis, one of the alternatives is inconsistent. The difference between the two estimators is Chi-square statistic, $X^2$. If the difference is large, then the second estimate is likely to be inconsistent and the alternative hypothesis is correct. If the difference is small, then the null hypothesis of IIA assumption holding is likely to be satisfied (Beck, 2004; Train, 2003).

The results of Hausman test presented in Table 3.31 demonstrate no differences between the estimates from the restricted and fully specified models. The critical Chi-square test statistic is equal to 9.487 for 4 degrees of freedom. The critical Chi-square test statistic is greater than the Chi-square statistics derived in the Hausman test. In the case of the option to purchase an insurance policy against corrosion and the option to stay with the current plumbing material, the Chi-square statistic is negative, which implies not a semi-positive Hessian matrix for both models. The not semi-positive value of the Hessian matrix results in the IIA assumption being satisfied. Based on the reviewed results, the null hypothesis of IIA assumption holding true cannot be rejected and the employment of the MNL model specification is justified for this model estimation.

Table 3.31: Hausman Specification Test for Independence of Irrelevant Alternatives Assumption

| To Re-plumb the House Option Excluded | | To Purchase an Insurance Policy Against Corrosion Excluded | | To Stay with the Current Plumbing Material Option Excluded | |
|---|---|---|---|---|---|
| Chi-Square Statistics | P-value | Chi-Square Statistics | P-value | Chi-Square Statistics | P-value |
| $X^2(4) = 0.000$ | 1.000 | $X^2(4) = -0.006$ | 1.000 | $X^2(4) = -3.171$ | 1.000 |

An evaluation of the impact of the independent variables on choosing to re-plumb the house, stay with current material and purchase an insurance policy, or stay only with the current plumbing material is performed. As displayed in Table 3.32, the risk corrosion variable is not statistically significant for either upgrading plumbing material or purchasing an insurance policy option. The explanatory variable has a standard error above 2.0. Beta coefficients with standard errors greater than 2.0 should not be interpreted when describing the model results. The coefficient value as well as the overall model prediction, however, are not affected by the large standard errors (Garson, 1998)[12].

As displayed in Table 3.32, the independent variable representing the cost of corrosion prevention (i.e. re-plumb the house) is not statistically significant at 95% in distinguishing group upgrading plumbing material from group staying with the current plumbing material. With regards to the decision of purchasing insurance against corrosion, the independent variable representing the cost of an insurance policy is the only statistically significant variable at 95% in distinguishing households choosing to purchase an insurance policy against corrosion from households choosing to stay with the current plumbing material. The additional dollar amount spent on an insurance policy is equal to 1.0 times the odds of purchasing the insurance policy vs. staying with the current plumbing material. This implies that the odds of choosing either option are not affected

[12] Correlation between each of the independent variables was investigated. Neither of the independent variables have correlation coefficients above 0.8 (value above or equal to 0.8 implies a high degree of collinearity). The correlation coefficient for pairs of independent variables varies is absolute value from 0.013 to 0.077. This finding confirms the orthogonality of the experimental design, which states that each independent variable is not correlated with any of the factors in the design.

by increasing cost of the insurance. This result is driven by the close to zero value of the coefficient parameter on the cost of plumbing upgrade or insurance purchase variable. In addition, at the mean value of the cost of re-plumbing or insurance purchase of $3,483, the probability of purchasing an insurance policy against corrosion is 11%. The mean cost values for corrosion prevention (for either plumbing material upgrade or insurance purchase) are derived based on the cost values presented to the survey respondents. The corrosion prevention cost values varied between $0 and $14,000. The probability estimate is computed as follows:

$$\textit{Probability of event} = exp(z)/(1 + exp(z)) \qquad (48)$$

where *z* represents the equation intercept plus the sum of cross-products of the beta coefficients times the values of their respective independent variables (Garson, 1998). The impact of the order to the CA question appearing in the questionnaire is also evaluated. The independent variable is statistically significant at 95% for both levels of MNL model. For example, based on the Odds Ratio (Table 3.32), the CA question appearing in the first position increases by 1.3 times the odds of home re-plumbing and by 1.8 times the odds of purchasing an insurance policy against corrosion.

Table 3.32: Multinomial Logit Model Estimates for MNL Model without Socioeconomic Characteristics (Dependent Variable Represents the Corrosion Prevention Choice and the Independent Variables Represent the Risk of Corrosion and the Cost of Plumbing Material Upgrade and Insurance Purchase )

| Variable Name[d] | Beta Coefficient[a] | Exp(B) Odds Ratio[b] | Probability Estimate[c] | Wald Stats[e] | Std. Error[e] | P-Value[d] |
|---|---|---|---|---|---|---|
| Upgrading Plumbing Material | | | | | | |
| Intercept | -1.715 | | | 146.284 | 0.105 | 0.000 |
| CA Question Order | 0.245 | 1.278 | 0.187 | 5.147 | 0.169 | 0.002 |
| Risk of Corrosion | 2.888 | 17.995 | 0.165 | 1.692 | 2.221 | 0.193 |
| Cost of Plumbing Upgrade or Insurance Purchase | -0.000001 | 1.000 | 0.153 | 0.001 | $2.42E^{-8}$ | 0.980 |
| Keeping Current Plumbing Material and Purchasing Insurance | | | | | | |
| Intercept | -1.914 | | | 145.136 | 0.159 | 0.000 |
| CA Question Order | 0.588 | 1.801 | 0.209 | 4.344 | 0.185 | 0.001 |
| Risk of Corrosion | 2.734 | 15.396 | 0.139 | 1.272 | 2.424 | 0.259 |
| Cost of Plumbing Upgrade or Insurance Purchase | -0.000051 | 1.000 | 0.109 | 10.156 | $2.77E^{-11}$ | 0.037 |

[a] Beta coefficients represent the parameter estimates from the MNL model estimation.
[b] Odd ratio (Exp(B)) represents the impact of increase in the independent variable on the likelihood of choosing a given option (insurance or plumbing upgrade). For example, an odds ratio of 1 means a one unit increase in independent variable implies an equal likelihood of choosing or not choosing the prevention option.
[c] Probability estimate for continuous variables is computed at the mean value of the independent variable (e.g. the cost of insurance or plumbing upgrade) presented to all respondents. Probability estimate for a dummy variable is the increase in probability of choosing the prevention option for respondents in the '1' category. For example, respondents who saw CA question first were 0.187 times more likely to choose plumbing upgrade compared to those who saw it second.
[d] Wald statistics and p-value represent the significance level.
[e] Standard error represents the variation of the coefficient estimate.
[f] Model Statistics: Log-likelihood value is 590.606; Nagelkerke's R-square is 0.019. Independent variables take form of continuous and dummy variables with value of one zero otherwise. To avoid a dummy variable trap, one of the attribute levels was excluded from the analysis. The number of respondents included in the model is 985.

The coefficient estimates for upgrading plumbing material and purchasing insurance policy against corrosion are tested for equality. If there are no statistically significant differences between the coefficients, then the response of households with regards to the prevention option is the same regardless which option is chosen. In order to test for the differences between coefficients of the independent variables, the Wald test is employed for each pair of coefficient estimates. The Wald test is described as follows:

$$(\hat{\beta}_p - \hat{\beta}_i)' \sum^{-1} (\hat{\beta}_p - \hat{\beta}_i), W \sim X^2_k \qquad (49)$$

where $\hat{\beta}_p$ is the coefficient estimate for the upgrade material option while $\hat{\beta}_I$ is the coefficient estimate for the insurance purchase option. The variance matrix associated with each coefficient estimate is denoted by $\sum^{-1}$. The value computed from equation 49 is $X^2$, which represents the Chi-square statistic. If the estimate difference is small, then the null hypothesis, implying no statistically significant differences between pair of coefficients estimates, is likely to be true (Brill, 2004).

The results of Wald test presented in Table 3.33 demonstrate no differences between the coefficient estimates for upgrading plumbing material and purchasing an insurance policy. The critical Chi-square test statistic is equal to 3.940 for 1 degree of freedom. The critical Chi-square test statistic is greater than the Chi-square statistics derived in the Wald test. Based on the reviewed results, the null hypothesis that coefficient estimates are equal cannot be rejected for this MNL model. This finding implies that the response functions represented by the independent variables coefficients are the same cross the two logit equations.

Table 3.33: Wald Test for Testing Equality of Coefficients Estimated From Multinomial Logit Model

| Variable Name | Chi-Square Statistics | P-value |
|---|---|---|
| Intercept | 1.091 | 0.300 |
| CA Question Order | 1.874 | 0.170 |
| Risk of Corrosion | 0.002 | 0.960 |
| Cost of Plumbing Upgrade of Insurance Purchase | 0.103 | 0.750 |

Based on the results presented in Table 3.32, the marginal utility of money is also computed. In the linear MNL model, the cost (price) coefficient reflects the change in utility for a change in price of the good or service, capturing the marginal utility of money. However, the negative cost coefficient implies a higher price, resulting in lower utility values. In order to change the coefficient into the marginal utility of income rather than the disutility of price, the coefficient is multiplied by -1 to obtain a positive value of the coefficient (Louvier et al., 2000). The estimates are only computed for the independent variables that are statistically significant at a 95% confidence interval (Champ et al, 2003). Based on the MNL regression results presented in Table 3.32, the marginal utility of money can be computed only for the insurance policy purchase group as the cost of insurance coefficient is statistically significant at a 95% confidence interval. As a result, the marginal utility of dollar amount spent on corrosion prevention is 0.000051 and carries the expected positive signs. The estimate cannot be computed for the home re-plumbing group as the coefficient on plumbing upgrade cost is not statistically significant.

The marginal willingness-to-pay (*WTP*) values for a marginal change in attribute were planned for estimation. The values are usually computed by dividing the beta coefficients of the risk level values by the marginal utility of money (-1 times the preference weights for the cost of plumbing upgrade or insurance purchase):

$$WTP = \beta_{rk} / -\beta_p \qquad (50).$$

However, due to the statistically insignificant impact of corrosion risk on household decisions, the marginal *WTP* estimates were not computed for this analysis.

The *WTP* probability at the mean cost value of prevention options (insurance or plumbing upgrade) and the compensating valuation were evaluated for two survey versions (CA question first and CA question second) (Table 3.34 and Table 3.35). Since the risk of corrosion variable was not statistically significant in the MNL regression, coefficient of zero was applied. Overall, only the independent variable coefficients significant at 95% confidence interval were employed in these computations of trade-offs. Based on the parameter coefficients, the decision making probabilities are computed. The compensating variation (*CV*) is also computed by comparing the before and after change in expected values of maximized utility. The following equation presents the *CV* computation:

$$CV = (1/-\beta_p)[(v^1) - (v^0)] \qquad (51)$$

where $v^0$ and $v^1$ are the expressions of utility for the base (staying with current plumbing material and purchasing no insurance) and altered cases. The utility values for each alternative were derived based on the statistically significant parameter coefficients at 5% and 10% significance level (the cost of insurance purchase, the CA order in the questionnaire, and the intercepts) and the mean values of their associated independent variables (Champ et al., 2003).

Based on the CA regression analysis and the marginal utility of money computation, the associated probability of household choices and the compensating variation are provided. Only independent variable coefficients significant at a 95% confidence interval were employed in these statistical computations. The mean cost values of corrosion prevention (either home re-plumbing or corrosion insurance purchase) presented to the respondents as well as the mid-point cost of home re-plumbing and purchasing an insurance policy against corrosion obtained from experts are employed in the computation of probability estimates. The expert estimates of cost values for home re-plumbing and corrosion insurance purchase were obtained from interviewing corrosion experts (i.e. plumber, builders, and insurance sales executives) by Virginia Tech graduate students: Juneseok Lee, Owais Farooqi, and Ewa Kleczyk. The cost for home re-plumbing ranged from \$3,000 to \$4,500 while the range for insurance against corrosion purchase varied from \$5,000 to \$10,000. The mid-point of these corrosion prevention cost ranges was employed for the probability estimation. The reason for employing the mid-point values for cost prevention provided by corrosion experts was to learn the true likelihood of alternatives being chosen to prevent corrosion based on market costs of prevention or insurance options[13].

The probability associated with the mean cost values for upgrading the plumbing material and purchasing an insurance policy is presented in Table 3.34. At the mean cost

[13] The market value costs of home re-plumbing are based on online search of several plumbing service websites and the service quotes displayed on their websites (theplumber.com and Mister Renovator available online at: www.misterrenovator.com) as well as information provided by plumbers contacted by two Virginia Tech researchers: Juneseok Lee and Owais Farooqi. The range of values for cost of insurance against corrosion is provided based on conversations with an executive working in the insurance and banking industry. Per his request, his identity is not disclosed in this article. Additional source of information were obtained from articles found on Ezine website (http://ezinearticles.com) and insurance quotes available on the following website www.compu-quote.com.

of insurance or plumbing upgrade presented to all respondents ($3,483), the probabilities of plumbing material upgrade and purchasing an insurance policy against corrosion are 16% and 15%, respectively ,when the CA question is asked in the first position. On the other hand, when the CA question follows the CV question, the probabilities of plumbing upgrade and insurance purchase are 14% and 9%, respectively. The probability for choosing an insurance policy by a homeowner is lower when the CA question follows the CV question in the questionnaire.

At the mid-point of the range of the cost values of prevention provided by the corrosion experts, the probability of plumbing material upgrade is 16% while the probability of purchasing an insurance policy is 13% in the case of the CA question being asked in the first position. On the other hand, when the CA question follows the CV question, the alternative occurrences' probabilities are 14% and 8%, respectively.

As a result, the probability of either re-plumbing the house or purchasing an insurance policy is sensitive to the changes in cost values employed for probability estimates. Higher cost of insurance policy against corrosion decreases the probability of insurance being chosen as the preferred prevention option. Furthermore, the CA order question in the questionnaire impacts the results by varying the probability estimates of each prevention option being chosen by a household. The CA order question regression coefficients are statistically significant in both logit models (home plumbing material upgrade and purchasing an insurance policy).

Table 3.34: Probability Estimates of Household Decisions for MNL Model without Socioeconomic Characteristics (Dependent Variable Represents the Corrosion Prevention Choice and the Independent Variables Represent the Risk of Corrosion and the Cost of Material Upgrade and Insurance Purchase)

| Survey Version | Cost of Plumbing Upgrade of Insurance Purchase | Probability Estimates for Upgrading Plumbing Material | Probability Estimates for Keeping Current Plumbing Material and Purchase Insurance | Probability Estimates Keeping Current Plumbing Material |
|---|---|---|---|---|
| CA Question First | The Mean Cost Value: $3,483[a] | 15.84% | 15.31% | 68.85% |
| | Corrosion Experts' Estimates: $3,750 for Home Re-plumbing; $7,500 for Insurance Cost[b] | 16.29% | 12.83% | 70.86% |
| CA Question Second | The Mean Cost Value: $3,483[a] | 13.81% | 9.47% | 76.72% |
| | Corrosion Experts' Estimates: $3,750 for Home Re-plumbing; $7,500 for Insurance Cost[b] | 14.05% | 7.86% | 78.09% |

[a]Mean cost of either home re-plumbing or insurance purchase represents the weighted average value of the prevention option (insurance or plumbing upgrade) presented to all respondents.
[b] Mid-point of cost of corrosion prevention represents the cost range for plumbing material upgrade and insurance policy purchase provided by corrosion experts.

The compensating variation was computed at the derived probability levels for each option presented in Table 3.35. At the probability derived based on the weighted average values of cost for upgrading the plumbing system and purchasing an insurance policy presented to the homeowners, the compensating variation for the survey version with the CA question in the first position is ($10,500) for the cost of insurance; for the second survey version, the compensating variation is ($13,185). At the probability values derived based on the mid-point cost costs of prevention provided by the corrosion experts, the compensating variation for the survey version with the CA in the first position is ($11,378) and for the second survey version is ($13,771). When the expert cost values are employed as the independent variable value, the absolute increase in the compensating variation ($11,378 vs. $13,771) is lower compared to the increase for the weighted average cost of prevention seen by respondents ($10,500 vs. $13,185). This finding implies that as the cost of insurance purchase increases, the CA question order seems to have less impact on the overall compensating variation estimate.

The compensating variation derived based on mid-point corrosion cost probability is larger in absolute value as the cost values provided by corrosion experts were larger compared to the weighted average cost values seen by the survey respondents. In both situations, the negative values of compensating variation imply that the households view the alternative attributes as a utility reducer compared to the base option attributes, and therefore, they should be refunded additional money to switch to the new options (Champ et al., 2003). The lower (more negative) values for the CA question second category were seen after the CV question. This finding might imply that survey respondents were 'burned out' when evaluating the CA question with regards to the corrosion prevention options and the associated costs after having seen the CV question first. As a result, they were more prone for staying with the current plumbing material even at the lower levels of cost. The compensating variation was not computed for the home re-plumbing option vs. staying with the current material, as the coefficient on the cost of corrosion for this option was not statistically significant at significance level of 5%.

Table 3.35: Compensating Variation Estimates for MNL Model without Socioeconomic Characteristics (Dependent Variable Represents the Corrosion Prevention Choice and the Independent Variables Represent the Risk of Corrosion and the Cost of Material Upgrade and Insurance Purchase)

| Survey Version | The Dollar Amount Spent on Cost of Plumbing Upgrade or Insurance | Probability Estimates for Upgrading Plumbing Material | Probability Estimates for Keeping Current Plumbing Material and Purchase Insurance | Probability Estimates for Keeping Current Plumbing Material |
|---|---|---|---|---|
| CA Question First | The Mean Cost Value: $3,483[a] | | | |
| | ($10,500) | 15.84% | 15.31% | 68.85% |
| | Corrosion Experts' Estimates: $3,750 for Home Re-plumbing; $7,500 for Insurance Cost[b] | | | |
| | ($11,378) | 16.29% | 12.83% | 70.86% |
| CA Question Second | The Mean Cost Value: $3,483[a] | | | |
| | ($13,185) | 13.81% | 9.47% | 76.72% |
| | Corrosion Experts' Estimates: $3,750 for Home Re-plumbing; $7,500 for Insurance Cost[b] | | | |
| | ($13,771) | 14.05% | 7.86% | 78.09% |

[a] Mean cost of prevention (either home re-plumbing or insurance purchase) represents the weighted average value of the prevention option (insurance or plumbing upgrade) presented to all respondents.
[b] Mid-point of cost of corrosion prevention represents the cost range for plumbing material upgrade and insurance policy purchase provided by corrosion experts.

The second Multinomial Logit model specification includes the risk of corrosion and cost of plumbing upgrade or insurance purchase, and the socioeconomic characteristics. Socioeconomic characteristics include the following: house value; age of the house; experience with pinhole leaks; respondent's previous cost of plumbing repairs and/or replacement due to leaks incurred by the household, landlord, or insurance company[14] and type of plumbing material installed at the house (copper vs. other)[15]. All independent variables were entered as covariates into the regression model. The socioeconomic variables are in the form of either dummy variables (experience with pinhole leaks) or continuous variables (house value and the respondent's previous cost of

[14] Respondents previous cost of repairing plumbing material was obtained with the following question: 'What was the approximate cost to you, your landlord, and/or your insurance company to repair the leaks (plumbing repair expenses only)?' at the median values of each category answer. The last three categories were joined, and a value of $7,500 was assigned to designate the group.
[15] Variables such race, education level, gender were not included in the model, as little variation in these characteristics was observed for the sample of respondents.

plumbing repairs and/or replacement, and house age). As the socioeconomic variables could not be entered into the MNL model directly (they do not differ between prevention alternatives for each respondent), they were interacted with the corrosion risk and cost of plumbing material upgrade or insurance purchase. The interaction allows for measuring the impact of demographics on households' decisions as these variables affect the differences in utility through their interaction with the attributes of each alternative (Champ et al., 2003; Train, 2003).[16]

Estimating the MNL model specification with socioeconomic variables interacted with the corrosion risk and cost of plumbing upgrade or insurance purchase results in a nearly singular Hessian matrix. This indicates that some predictors should be excluded from the model estimation as the number of observations employed in the analysis is only 942 and is not sufficient to include a large number of socioeconomic variables in one equation. Multiple model specifications were evaluated and showed that estimated coefficients for the risk of corrosion variable as well as some of the interactions of the socioeconomic variables with the corrosion risk and the cost of plumbing upgrade or insurance purchase variables presented standard errors greater than 2.0. The nearly singular Hessian matrix in the estimation can be avoided by eliminating some of these variables from the final model estimation. As a result, only selected variables, interacted with either the cost or risk of corrosion, are included in the estimation of the final MNL model.

In order to evaluate which independent variables should be entered into the final MNL model, the socioeconomic characteristics were interacted with either the risk of corrosion or the cost of upgrading plumbing material or purchasing an insurance policy and regressed one by one on the corrosion prevention options. The analysis described in Table 3.36 shows that a statistically significant relationship is observed at 5% and 10% significance level between the corrosion prevention options and the following independent variables: home value interacted with both the risk of corrosion as well as

[16] Another way of incorporating socioeconomic variables is by dividing the cost of corrosion prevention by the household income. Since the cost coefficient decreases in income, the model represents cost becoming less important in households' decision-making process when controlled for other factors (Train, 2003). The model specification did not yield statistically significant results when estimated, so it was not reported in detail in this section. This model specification is briefly discussed in Appendix A.

the cost of plumbing upgrade or insurance purchase, house age interacted with the cost of plumbing upgrade or insurance purchase, pinhole leak occurrences interacted with both the risk of corrosion and the cost of plumbing upgrade or insurance purchase, and the respondent's previous cost of plumbing repairs and/or replacement interacted with risk of corrosion[17]. The statistically significant variables are included in the MNL estimation of the choice of corrosion prevention.

For the purpose of this study, it is hypothesized that a higher home value should increase the probability of choosing to either re-plumb the house or purchase an insurance policy against corrosion as a residential property is viewed as a form of investment. Age of the house should also increase the likelihood for corrosion prevention, as older houses tend to experience more plumbing material problems due to the age of the drinking water infrastructure (Kleczyk & Bosch, 2008). Furthermore, having an experience with pinhole leaks in the past should also increase the likelihood for corrosion prevention to avoid future problems as such experience may increase the respondent's subjective probability of future leaks (Kleczyk & Bosch, 2008). On the other hand, respondent's previous cost of plumbing repairs and/or replacement should negatively impact the probability of choosing plumbing material upgrade or insurance purchase as previous costs on plumbing material repairs or replacement can be viewed as reducing the probability of future leaks.

[17] Respondent's previous cost of plumbing repairs and/or replacement was not interacted with the cost of plumbing upgrade or insurance purchase as preliminary results of the pairing resulted in a negative marginal utility of income. As a result, this variable was only interacted with the corrosion risk variable.

Table 3.36: Likelihood Ratio Tests for MNL Model with Socioeconomic Characteristics
(Dependent Variable Represents the Corrosion Prevention Choice (Insurance or Plumbing Upgrade) and the Independent Variables Represent the Risk of Corrosion, the Cost of Plumbing Upgrade or Insurance, and Socioeconomic Characteristics Inserted One at the Time into Regression)

| | Corrosion Prevention Variable | | | | Interaction Variable[d] | | |
|---|---|---|---|---|---|---|---|
| Socioeconomic Variable Name | Corrosion Prevention Variable Name[a] | -2 Log Likelihood of Reduced Model[b] | Chi-Square[b] | p-Value[b] | -2 Log Likelihood of Reduced Model[c] | Chi-Square[c] | p-Value[c] |
| Copper Plumbing | Risk of Corrosion | 596.251 | 3.980 | 0.137 | 592.725 | 0.453 | 0.797 |
| | Cost of Plumbing Upgrade or Insurance Purchase | 604.172 | 11.901 | 0.003 | 593.079 | 0.808 | 0.668 |
| Home Value | Risk of Corrosion | $1.370e^3$ | 2.982 | 0.225 | $1.377e^3$ | 9.925 | 0.007 |
| | Cost of Plumbing Upgrade or Insurance Purchase | $1.376e^3$ | 8.763 | 0.013 | $1.374e^3$ | 7.001 | 0.030 |
| House Age | Risk of Corrosion | 918.017 | 4.598 | 0.100 | 916.639 | 3.220 | 0.200 |
| | Cost of Plumbing Upgrade or Insurance Purchase | 917.417 | 3.998 | 0.135 | 918.714 | 5.295 | 0.071 |
| Pinhole Leak Occurrences | Risk of Corrosion | 628.775 | 7.237 | 0.027 | 626.557 | 5.019 | 0.081 |
| | Cost of Plumbing Upgrade or Insurance Purchase | 635.436 | 13.899 | 0.001 | 633.725 | 12.187 | 0.002 |
| Respondent's Previous Cost of Plumbing Repairs and/or Replacement | Risk of Corrosion | 818.376 | 5.088 | 0.079 | 823.230 | 9.942 | 0.007 |
| | Cost of Plumbing Upgrade or Insurance Purchase | 826.201 | 12.913 | 0.002 | N/A[17] | N/A[17] | N/A[17] |

[a] Each MNL model employing a socioeconomic variable one at a time includes the corrosion prevention variable (either risk of corrosion or cost of plumbing upgrade or insurance purchase) as well as the interaction of the socioeconomic variable with the corrosion prevention variables.
[b] Overall corrosion prevention variable significance test (Likelihood Ratio, Chi-Square, and p-value test).
[c] Overall interaction variable significance test (Likelihood Ratio, Chi-Square, and p-value test).
[d] Interaction variable between socioeconomic variable and corrosion prevention variable

Based on the estimated final MNL model specification, the presence of a relationship between the dependent variable, corrosion prevention choice (insurance or plumbing upgrade), and the combination of independent variables, risk of corrosion, the cost of plumbing material upgrade or insurance purchase, and socioeconomic characteristics, is assessed based on the Chi-square and log-likelihood statistics of the final model. The Chi-square statistic is 76.804 with the significance level of p-value of 0.000 and a log-likelihood value of $2.232e^3$. The existence of a relationship between the independent variables and the dependent variable is supported by the results.

The next step in CA analysis is to assess fit of the model. A Pseudo R-square statistic, such as Nagelkerke's R-square, is computed to represent the improvement of the model's fit from inclusion of the independent variables (Train, 2003). The goodness-of-fit value is 0.036, which implies that there is a 4% increase in the log-likelihood function above the function at zero parameter values and the estimated parameters do improve the model's fit (Train, 2003).

The estimate of by chance accuracy is proportional to the by chance accuracy rate computed by summing the squared percentage of cases in each group (Table 3.37) (Garson, 1998). The proportional by chance accuracy rate for this model is 56.4%. To characterize the model as useful, the overall percentage accuracy rate produced by SPSS software at the last step of estimation is compared to 25% more than the proportional by chance accuracy that is equal to 70.5%. The model's classification accuracy rate is 72.5% (Table 3.38), which satisfies the criterion for classification accuracy.

Table 3.37: Case Processing Summary for MNL Model Socioeconomic Characteristics (Dependent Variable Represents the Corrosion Prevention Choice and the Independent Variables Represent the Risk of Corrosion, the Cost of Plumbing Upgrade or Insurance, and Socioeconomic Characteristics)

| Group Membership | Number of Observations | Marginal Percentage |
|---|---|---|
| Upgrading Plumbing Material | 147 | 14.9 |
| Keeping Current Plumbing Material but Buy Insurance | 124 | 12.6 |
| Keeping Current Plumbing Material and Not Purchasing Insurance | 715 | 72.5 |

Note: Chance Accuracy = $0.149^2 + 0.126^2 + 0.725^2 = 0.564$
25% Improvement over the Chance Accuracy = 0.564*0.25 = 0.141
Chance Accuracy Criteria = 0.564 + 0.141 = 0.705

Table 3.38: Classification Table for MNL Model with Socioeconomic Characteristics (Dependent Variable Represents the Corrosion Prevention Choice and the Independent Variables Represent the Risk of Corrosion, the Cost of Plumbing Upgrade or Insurance, and Socioeconomic Characteristics)

| Observed | Predicted | | | |
|---|---|---|---|---|
| | Upgrading Plumbing Material | Keeping Current Plumbing Material but Buying Insurance | Keeping Current Plumbing Material and Not Purchasing Insurance | Percent Correct |
| Upgrading Plumbing Material | 0 | 0 | 147 | 0.0 |
| Keeping Current Plumbing Material but Buying Insurance | 0 | 0 | 124 | 0 |
| Keeping Current Plumbing Material and Not Purchasing Insurance | 0 | 0 | 715 | 100 |
| Overall Percentage | 0 | 0% | 100% | 72.5% |

Note 1: Percent correct is defined as the proportion of predicted observations in a category compared to the observed number of observations in that group.
Note 2: Overall percentage is proportion of predicted observations in a group compared to the total number of observations.

The overall relationship between the dependent variable and the independent variables is also evaluated for the final MNL model specification. As shown in Table 3.39, for model specification including only those socioeconomic variables statistically significant when regressed alone on corrosion prevention choices, the overall relationship between the independent variables and the dependent variable is observed at the 95% and 90% confidence interval for the following independent variables: the cost of plumbing material upgrade, home value interacted with both the corrosion risk as well as the cost of plumbing upgrade and insurance purchase, pinhole leak occurrences interacted with the cost of plumbing material upgrade and insurance purchase, and respondent’s previous dollar amount spent on plumbing material repairs and replacement interacted with risk of corrosion. The remaining explanatory variables have overall statistically insignificant relationship with choice of corrosion prevention options.

Table 3.39: Likelihood Ratio Tests for MNL Model with Selected Socioeconomic Characteristics (Dependent Variable Represents the Corrosion Prevention Choice (Insurance or Plumbing Upgrade) and the Independent Variables Represent the Risk of Corrosion, the Cost of Material Upgrade and Insurance Purchase, and Selected Socioeconomic Characteristics )

| Variable Name | Model Fitting Criteria | Likelihood Ratio Test | |
|---|---|---|---|
| | -2 Log Likelihood of Reduced Model | Chi-Square | p-Value |
| Risk of Corrosion | $2.232e^3$ | 0.492 | 0.782 |
| Cost of Plumbing Upgrade or Insurance Purchase | $2.239e^3$ | 6.876 | 0.032 |
| Home Value* Cost of Plumbing Upgrade or Insurance Purchase | $2.239e^3$ | 7.840 | 0.020 |
| House Age*Cost of Plumbing Upgrade or Insurance Purchase | $2.235e^3$ | 3.284 | 0.194 |
| Pinhole Leak Occurrences* Cost of Plumbing Upgrade or Insurance Purchase | $2.238e^3$ | 6.538 | 0.038 |
| Respondent's Previous Cost of Plumbing Repairs and/or Replacement *Corrosion Risk | $2.237e^3$ | 5.538 | 0.063 |
| Home Value *Corrosion Risk | $2.239e^3$ | 7.648 | 0.022 |
| Pinhole Leak Occurrences* Corrosion Risk | $2.236e^3$ | 3.944 | 0.139 |

In this part of the MNL model discussions, the effects of selected variables on choices of plumbing upgrade or insurance purchase are assessed. For example, the risk level of corrosion was found to have no statistically significant impact in differentiating between households that are willing to either upgrade their plumbing material (coded as group 1) or keep current plumbing material and purchase insurance policy (coded as group 2) and those that stay with the current system (coded as group 3). The variable also presented standard errors greater than 2.0. In such a situation, it is recommended not to interpret the estimation results for that variable. Please note that the coefficient values as well as the overall model prediction does not change in response to dropping the corrosion variable; only the standard errors and associated statistics are impacted.

As presented by Table 3.40, there is, however, a statistically significant impact at 5% and 10% of the cost of plumbing upgrade or insurance purchase and its interaction effects (home value, home age, and past experience with pinhole leaks) in distinguishing those households choosing to upgrade plumbing material from those staying with the current plumbing material. The odds ratios (Exp(B)) represents the factor by which odds(event) is multiplied for a unit increase in the independent variable; however, the effect of the independent variable is not properly gauged if the covariate is involved in interaction effects. Before exponentiating, the coefficient must be adjusted to include the interaction coefficient terms. Exp(B*) is the estimate of the odds ratio including the main and interaction effect of the independent variable, which varies according to the interaction variable (Garson, 1998).

Based on the updated Exp(B*) definition, the dollar cost of insurance or plumbing upgrade by itself has the same odds impact on choosing to upgrade the plumbing material vs. staying with the current material. The associated probability derived based on the mean value of cost of upgrading plumbing material and purchasing an insurance policy presented to the survey respondents ($3,483) is 12%. When the odds ratio computation includes the home value, house age, or pinhole leak, then the odds are still equal to 1, implying the same odds of choosing to upgrade the plumbing material vs. staying with the current material. The associated probabilities ('probability estimate', Table 3.40) derived based on the mean of cost of prevention presented to households is 15% when accounting for the interaction with the home value and house age[18], and 13% when accounting for pinhole leak occurrences.

Although these interaction effects are statistically significant in the regression at 5% and 10% significance level, none of these joint coefficients[19] is statistically different from 0.[20] This finding should not be surprising, as these socioeconomic interaction

---

[18] The interaction variable between the cost of material upgrade and insurance purchase with the house age is not statistically significant when evaluating the overall relationship between the independent variable and the prevention options. This result should be taken lightly in result interpretation.

[19] A **Joint Coefficient Parameter** is defined as the coefficient of cost of upgrading plumbing material and purchasing insurance and its interaction variable multiplied by the value of the socioeconomic variable.

[20] House Value interacted with Cost of Upgrading Plumbing Material and Purchasing Insurance: joint coefficient is $-2.327e^{-5}$, Walt statistic is 0.159 with p-value of 0.690.

variables have positive coefficients and in combination with the cost of plumbing material upgrade bring the joint coefficient values closer to zero. The closer the joint coefficient to zero, the lower is the impact of the cost of plumbing material upgrade on household prevention decisions.

In addition, the respondent's previous cost of plumbing repairs and/or replacement interacted with the risk of corrosion is also statistically significant at 10% confidence interval. The odds ratio of choosing to upgrade plumbing material vs. to stay with current plumbing material is 0.999, which implies a slightly lower likelihood of the prevention option being chosen compared to staying with the current plumbing material when the expenditures increase by one unit. The probability of choosing to upgrade plumbing system is 18% when accounting for the interaction variable. Although the interaction effect by itself has a statistically significant impact on the corrosion prevention decision, the joint coefficient derived based on the coefficients of both the risk of corrosion and respondent's previous cost of plumbing repairs and/or replacement is statistically insignificant at 95% confidence interval, as the risk of corrosion variable is not statistically significant at 5% significance level[21].

The findings discussed above imply that households residing in more valuable homes are more likely to upgrade drinking water infrastructure in order to protect their investment from losing its financial value (assuming that owning a house is considered as an investment). Households residing in older houses are also more likely to choose the option of upgrading their plumbing materials, as the age of the house approximates the age of the plumbing material. The older the house is, the older the plumbing material is and, therefore, older houses are more prone for future corrosion issues (Kleczyk & Bosch, 2008). Pinhole leaks occurrences also increase the probability of home plumbing, as past pipe failure implies faulted pipes and therefore, future problems with home drinking water infrastructure (Kleczyk & Bosch, 2008). Finally, households that have

---

House age interacted with Cost of Upgrading Plumbing Material and Purchasing Insurance: joint coefficient is $-2.642e^{-5}$, Walt statistic is 0.337 with p-value of 0.560.
Pinhole Leaks Occurrences with Cost of Upgrading Plumbing Material and Purchasing Insurance: joint coefficient -0.00003, Walt statistic is 0.262 with p-value of 0.610.
[21] Respondent's Previous Cost of Plumbing Repairs and/or Replacement interacted with Risk of Corrosion is 2.234; Wald statistic for is 0.196 with p-value of 0.660.

repaired their plumbing materials in the past and accrued financial costs are less likely to choose to re-plumb their house, as they have already made financial investments into corrosion prevention of the currently installed plumbing material in their homes.

With regards to the decision to stay with the current plumbing material and purchase an insurance policy against corrosion (Table 3.41), two independent variables are statistically significant at either 5% or 10% significance level: home value and the respondent's previous cost of plumbing repairs and/or replacement interacted with risk of corrosion. When the odds ratio computation includes the home value, then the odds ratio is equal to 1, implying the same odds of choosing to purchase an insurance against corrosion vs. staying with the current material with an increase in the house value. On the other hand, when odds ratio includes the respondent's previous cost of plumbing repairs and/or replacement, the odds ratio is slightly less than 1 (0.999) implying a slightly lower likelihood of the prevention option being chosen. The probability associated with the house value interacted with the risk of corrosion as well as for the respondent's previous cost of plumbing repairs and/or replacement is 15% for both interaction variables. Although the interaction effects by themselves have a statistically significant impact on the corrosion prevention decision, the joint coefficient derived based on the coefficients of both the risk of corrosion and either the home value or respondent's previous cost of plumbing repairs and/or replacement is statistically insignificant at 95% confidence interval as the risk of corrosion variable is not statistically significant at 5% significance level[22].

The above findings imply that having a more valuable house increases the probability of purchasing an insurance policy as households are protecting their investment from declining in value. On the other hand, accruing financial costs associated with plumbing material repairs decreases the likelihood of purchasing an insurance policy as by fixing the faulty pipes, the households are protecting themselves from future pipe failures and the associated financial burden in the future.

[22]Home Value interacted with Risk of Corrosion: joint coefficient is 2.205; Wald statistic for is 1.115 with p-value of 0.290.
Respondent's Previous Cost of Plumbing Repairs and/or Replacement interacted with Risk of Corrosion is 1.883; Wald statistic for is 0.813 with p-value of 0.370.

Table 3.40: Multinomial Logit Model Estimates for MNL Model with Socioeconomic Characteristics (Dependent Variable Represents the Corrosion Prevention Choice (Insurance or Plumbing Upgrade) and the Independent Variables Represent the Risk of Corrosion, the Cost of Material Upgrade and Insurance Purchase, and Socioeconomic Characteristics) – Upgrading Plumbing Material

| Variable Name[d] | Beta Coefficient[a] | Exp(B) Odds Ratio[b] | Probability Estimate[c] | Wald Stats[e] | Std. Error[e] | P-Value[d] |
|---|---|---|---|---|---|---|
| Upgrading Plumbing Material | | | | | | |
| Intercept | -1.621 | | | 178.007 | .121 | .000 |
| Risk of Corrosion | 2.498 | 12.159 | 0.177 | .245 | 5.050 | .621 |
| Cost of Plumbing Upgrade or Insurance Purchase | -0.0001 | 1.000 | 0.122 | 6.527 | 0.00005 | .011 |
| Home Value* Cost of Plumbing Upgrade or Insurance Purchase | 1.76E-10 | 1.000 | 0.154 | 7.502 | 6.42E-11 | .006 |
| House Age* Cost of Plumbing Upgrade or Insurance Purchase | 5.74E-06 | 1.000 | 0.148 | 3.060 | 0.00000 | .080 |
| Pinhole Leak Occurrences*Cost of Plumbing Upgrade or Insurance Purchase | 0.00007 | 1.000 | 0.128 | 6.588 | 0.00003 | .010 |
| Home Value *Corrosion Risk | 6.93E-06 | 1.000 | 0.177 | .571 | 0.00001 | .450 |
| Pinhole Leak Occurrences* Corrosion Risk | 4.146 | 63.188 | 0.181 | 1.212 | 3.766 | .271 |
| Previous Cost of Plumbing Repair Material Fixing or Replacement* Corrosion Risk | -0.001 | .999 | 0.175 | 3.126 | .001 | .077 |

[a] Beta coefficients represent the parameter estimates from the MNL model estimation.
[b] Odd ratio (Exp(B)) represents the impact of increase in the independent variable on the likelihood of choosing a given option (insurance or plumbing upgrade). For example, an odds ratio of 1 means a one unit increase in independent variable implies an equal likelihood of choosing or not choosing the prevention option.
[c] Probability estimate for continuous variables is computed at the mean value of the independent variable (e.g. the cost insurance or plumbing upgrade) presented to all respondents. Probability estimate for a dummy variable is the increase in probability of choosing the prevention option for respondents in the '1' category.
[d] Wald statistics and p-value represent the significance level.
[e] Standard error represents the variation of the coefficient estimate.
[f] Model Statistics: Log-likelihood value is $2.232e^{3}$; Nagelkerke's R-square is 0.036. Independent variables take form of continuous and dummy variables with value of one zero otherwise. To avoid a dummy variable trap, one of the attribute levels was excluded from the analysis. The number of respondents included in the model is 942.

Table 3.41: Multinomial Logit Model Estimates for MNL Model with Socioeconomic Characteristics (Dependent Variable Represents the Corrosion Prevention Choice and the Independent Variables Represent the Risk of Corrosion, the Cost of Material Upgrade and Insurance Purchase, and Socioeconomic Characteristics) - Keeping Current Plumbing Material and Purchasing Insurance

| Variable Name[d] | Beta Coefficient [a] | Exp(B) Odds Ratio[b] | Probability Estimate[c] | Wald Stats[e] | Std. Error[e] | P-Value[d] |
|---|---|---|---|---|---|---|
| Keeping Current Plumbing Material and Purchasing Insurance | | | | | | |
| Intercept | -1.827 | | | 188.773 | .133 | .000 |
| Risk of Corrosion | 2.205 | 9.070 | 0.147 | .169 | 5.358 | .681 |
| Cost of Plumbing Upgrade or Insurance Purchase | -0.00005 | 1.000 | 0.121 | .452 | 0.0001 | .501 |
| Home Value* Cost of Plumbing Upgrade or Insurance Purchase | 7.80E-11 | 1.000 | 0.134 | .942 | 8.03E-11 | .332 |
| House Age* Cost of Plumbing Upgrade or Insurance Purchase | -0.000001 | 1.000 | 0.116 | .085 | 0.000005 | .771 |
| Pinhole Leak Occurrences*Cost of Plumbing Upgrade or Insurance Purchase | 0.00002 | 1.000 | 0.122 | .216 | 0.00004 | .642 |
| Home Value *Corrosion Risk | 0.00002 | 1.000 | 0.147 | 6.262 | 0.00001 | .012 |
| Pinhole Leak Occurrences* Corrosion Risk | 6.349 | 28.47 | 0.153 | 2.016 | 4.471 | .156 |
| Previous Cost of Plumbing Repair Material Fixing or Replacement** Corrosion Risk | -0.001 | .999 | 0.146 | 3.176 | .001 | .075 |

[a] Beta coefficients represent the parameter estimates from the MNL model estimation.
[b] Odd ratio (Exp(B)) represents the impact of increase in the independent variable on the likelihood of choosing a given option (insurance or plumbing upgrade). For example, an odds ratio of 1 means a one unit increase in independent variable implies an equal likelihood of choosing or not choosing the prevention option.
[c] Probability estimate for continuous variables is computed at the mean value of the independent variable (e.g. the cost insurance or plumbing upgrade) presented to all respondents. Probability estimate for a dummy variable is the increase in probability of choosing the prevention option for respondents in the '1' category.
[d] Wald statistics and p-value represent the significance level.
[e] Standard error represents the variation of the coefficient estimate.
[f] Model Statistics: Log-likelihood value is $2.232e^3$; Nagelkerke's R-square is 0.036. Independent variables take form of continuous and dummy variables with value of one zero otherwise. To avoid a dummy variable trap, one of the attribute levels was excluded from the analysis. The number of respondents included in the model is 942.

The coefficient estimates for the independent variables for equations representing either upgrading the plumbing material vs. staying with the current system or purchasing an insurance policy vs. staying with the current plumbing material, are tested for equality. If there are no statistically significant differences between the coefficients, then the effect of each variable on the response of households with regards to the prevention options is the same regardless which option is chosen. The Wald test is employed to test each pair of coefficient estimates pair. The results of Wald test presented in Table 3.42 demonstrate no differences between the coefficient estimates for upgrading plumbing material and purchasing an insurance policy for all of the independent variables. The critical Chi-square test statistic at 5% significance level is equal to 3.940 for 1 degree of freedom and is greater than the Chi-square statistics derived in the Wald test for most of the independent variables. Based on the reviewed results, the null hypothesis that coefficient estimates are equal cannot be rejected for this MNL model. This finding implies that the response functions represented by the independent variables coefficients are the same across the two Logit equations.

Based on the results presented in Tables 3.40 and 3.41 the marginal utility of money is computed. In the linear MNL model, the cost coefficient reflects the change in utility for a change in price of the good or service. However, as the negative cost coefficient implies that a higher price results in lower utility values, the coefficient has to be multiplied by -1 to reflect the marginal utility of money instead of the marginal disutility of price (Louvier et al., 2000). The marginal utility of money to upgrade the plumbing material is positive in the case of upgrading plumbing material (0.0001), which is in accordance with economic theory. In the case of keeping the current material and purchasing an insurance policy, the marginal utility of money for purchasing the insurance policy was not computed, as the coefficient parameter was not statistically significant at the 95% confidence interval.

Table 3.42: Wald Test for Testing Equality of Coefficients Estimated From Multinomial Logit Model in Predicting Responses to Purchasing Insurance or Plumbing Upgrade to Prevent Loss from Corrosion Damage

| Variable Name | Chi-Square Statistics | P-value |
|---|---|---|
| Intercept | 1.560 | 0.210 |
| Risk of Corrosion | 0.002 | 0.960 |
| Dollar Amount Spent on Corrosion Prevention | 0.261 | 0.610 |
| Home Value*Dollar Amount Spent on Corrosion Prevention | 1.059 | 0.300 |
| House Age*Dollar Amount Spent on Corrosion Prevention | 2.379 | 0.120 |
| Pinhole Leak Occurrences* Dollar Amount Spent on Corrosion Prevention | 1.195 | 0.270 |
| Home Value *Corrosion Risk | 1.383 | 0.240 |
| Pinhole Leak Occurrences* Corrosion Risk | 0.170 | 0.683 |
| Cost of Plumbing Material Fixing or Replacement* Corrosion Risk | 0.026 | 0.870 |

The marginal willingness-to-pay (*WTP*) values for a marginal change in corrosion risk cannot be computed due to the statistically insignificant impact of corrosion risk on household decisions. However, the *WTP* probability at the mean value of prevention options (either insurance or plumbing upgrade) for cost values seen by respondents as well as the compensating variation was evaluated (Table 3.43). Overall, only the independent variable coefficients significant at 95% confidence interval were employed in these computations of trade-offs. Based on the resulting estimates, the decision making probabilities are computed. The compensating variation (*CV*) is also computed by comparing the before and after change in expected values of maximized utility (Champ et al., 2003).

Based on the CA regression analysis and the marginal utility of money computation, the associated probability of household choices and the compensating variation are provided. Only independent variable coefficients significant at 95% confidence interval were employed in these statistical computations. The weighted average values of plumbing upgrade or insurance purchase cost presented to the respondents as well as the mid-point cost of estimated by corrosion experts are employed in the probability estimates computation.

The probability associated with the mean cost values for upgrading the plumbing material and purchasing an insurance policy is presented in Table 3.43. The mean cost values for corrosion prevention represent the weighted average of cost values for home plumbing upgrade and insurance purchase seen by survey participants. The mid-point values of cost provided by the corrosion experts are derived based on the cost for home re-plumbing ranging from $3,000 to $4,500 and the insurance against corrosion purchase varying from $5,000 to $10,000.

At their mean cost values, the probabilities of plumbing material upgrade and purchasing an insurance policy against corrosion are 17% and 11%. At the midpoint cost values of prevention provided by corrosion experts, the probability of plumbing material upgrade is also 17% while the probability for purchasing an insurance policy is 11%. At the mean cost value of corrosion prevention presented to all respondents as well as the midpoint of the cost range of either insurance or plumbing upgrade costs provided by the experts, the probability of household choosing either plumbing upgrade or insurance are nearly the same (17.01% vs. 16.63% for plumbing upgrade and 11.39 vs. 11.45% for insurance). The small differences in probabilities result from two factors: 1) the small value of the estimated coefficient for cost of plumbing material upgrade; and 2) the small differences between the mean cost values of plumbing material upgrade presented to survey respondents and the mid-point values of plumbing material upgrade provided by the corrosion experts.

Table 3.43: Probability Estimates of Household Decisions for MNL Model with Socioeconomic Characteristics (Dependent Variable Represents the Corrosion Prevention Choice and the Independent Variables Represent the Risk of Corrosion, the Cost of Material Upgrade and Insurance Purchase, and Socioeconomic Characteristics)

| Cost of Plumbing Upgrade or Insurance Purchase | Probability Estimates for Upgrading Plumbing Material | Probability Estimates for Keeping Current Plumbing Material and Purchasing Insurance | Probability Estimates for Keeping Current Plumbing Material |
|---|---|---|---|
| The Mean Cost Value: \$3,483[a] | 17.01% | 11.39% | 71.60% |
| Corrosion Experts' Estimates: \$3,750 for Home Re-plumbing; \$7,500 for Insurance Cost[b] | 16.63% | 11.45% | 71.92% |

[a] Mean cost of prevention (either home re-plumbing or insurance purchase) represents the weighted average value of the prevention option (insurance or plumbing upgrade) presented to all respondents.
[b] Mid-point of cost of corrosion prevention represents the cost range for plumbing material upgrade and insurance policy purchase provided by corrosion experts.

The compensating variation was also computed at the derived probability levels for each corrosion prevention option presented in Table 3.43. At the probability estimates derived based on the weighted average values of cost for upgrading the plumbing system and purchasing an insurance policy presented to the homeowners, the compensating variation is (\$5,460) for the cost of corrosion prevention when choosing to upgrade plumbing material; at the probability values derived based on the mid-point cost values of prevention provided by the corrosion experts, the compensating variation is (\$5,529).

In both situations, the negative values of compensating variation imply that the households view the alternative attributes as reducing the utility level when compared to the base option attributes. The households should be compensated for the decrease in utility due to the change in alternatives in order to be switched to the new prevention option (Champ et al., 2003). The difference in the compensating variation is also not very large as the probabilities of choosing to re-plumb the plumbing system were similar between the different cost values. The compensating variation was not computed for purchasing an insurance policy against corrosion vs. staying with the current material as the coefficient on the cost of corrosion prevention for this option was not statistically significant at a significance level of 5%.

When comparing the MNL model specifications with/without socioeconomic independent variables, the models provide statistically significant results in discriminating between options to upgrade the plumbing material, stay with the current plumbing material but purchase an insurance policy, or stay with the current material and choose no preventive measures. For both MNL models specifications, the corrosion risk variable was not a statistically significant factor in the decision-making process; however the corrosion risk impact was accounted for indirectly when the socioeconomic variables were interacted with the risk of corrosion.

Several of the socioeconomic variables were found statistically significant: the house value, house age, past pinhole leak occurrences, and the respondent's previous cost incurred for plumbing repairs and/or replacement. These socioeconomic variables interacted with either the risk of corrosion or the cost of plumbing material upgrade or insurance purchase had a statistically significant impact in differentiating between the households who chose to either re-plumb their house or purchase insurance vs. those who chose to stay with the current plumbing material. Adding the socioeconomic variables made the impact of the cost of plumbing upgrade on household decisions statistically significant when deciding to upgrade plumbing material. When cost is the only independent variable, the cost factor differentiates between the households that prefer to purchase an insurance policy against corrosion vs. those that choose to stay with the current plumbing system, but it is not significant for plumbing material upgrade. As a result, the socioeconomic characteristics impact the model results when included in the specification.

The probabilities derived based on either the mean cost value of plumbing upgrade or insurance provided to respondents or the cost values provided by the corrosion experts are similar in value across the different model specifications. For the model without the socioeconomic variables, the probability of choosing to upgrade the plumbing material is between 14% and 16% while for the model with the socioeconomic variable the probability increases to 17%. In the case of purchasing an insurance policy, the probability varied between 9% and 15% for the MNL model specification without the socioeconomic variables and 11% with these factors taken into account. The compensating variation estimates are negative for both model specifications and varied between ($5,460) and ($5,529) in the case of switching household to home re-plumbing and between ($10,500) and ($13,771) in the case of purchasing an insurance policy. The negative values of the compensating variation imply that in order for the households to switch between the current plumbing material and a prevention option, they would have to be reimbursed the specified amounts to keep them at the same level of utility.

### 3.8.2 Household Preferences for Plumbing Material: Conjoint Analysis for the Follow-up Southeastern Community Survey

#### 3.8.2.1 Summary of Descriptive Results

Every respondent to the first Southeastern Community survey was asked to participate in the follow-up survey. Three hundred sixty three respondents agreed to participate and 245 responded to the follow-up survey. Each respondent evaluated three Conjoint Analysis scenarios describing a set of two plumbing materials and answered questions comparing material attributes (AHP method[23]). Each plumbing material was described by the following attributes: corrosion resistance, taste and odor, health effects, convenience of installation, proven performance on the market, plumbing material cost, and warranty length. To avoid a respondent exposure bias, the plumbing materials were not named but instead referred to by their vector of characteristics as shown in Table 3.44.

[23] Juneseok Lee from the Civil and Environmental Engineering Department at Virginia Tech performed the AHP analysis. For more information see Lee (2008).

Table 3.44: Descriptions of Plumbing Materials

| Attributes | Material A | Material B | Material C |
|---|---|---|---|
| Corrosion Resistance | Corrosion proof | Same as material A | Some risk of corrosion |
| Taste / Odor | Compounds released from this material in drinking water plumbing may give a chemical or solvent taste or odor to the water. | Same as material A | Compounds released from this material in drinking water plumbing may give a bitter or metallic taste or odor to the water. |
| Health Effects | The material meets EPA Standards. There is a very small chance that compounds from this plumbing material that are released into drinking water may lead to microbial growth in water. Microbial growth may cause severe illness. | Same as material A | The material meets EPA Standards. There is a very small chance that compounds from this plumbing material that are released into drinking water may cause vomiting, diarrhea, stomach cramps, and nausea. |
| Convenience of Installation | No need to tear into the wall and/or floor. Installation takes around 4 days. | Need to tear into some sections of wall for installation. Installation takes 5-6 days. | Need to tear into the wall and/or floor to replace the existing system. 7-9 days required for installation. |
| Proven performance in market | Less than 10 years in the market. | Less than 20 years in the market. | More than 50 years in the market. |
| Cost (labor + material) | \$9,000 ~ 14,000 depending on the size of house | \$6,500 ~ 13,000 depending on the size of house | \$9,000 ~ 16,000 depending on the size of house |
| Warranty | Warranty is 15 years for the material. | Warranty is 10 years for the material. | A 50 years manufacturer's warranty applies. Some exceptions apply (e.g. warranty reduces to one year if compounds in water corrode pipes). |

Each respondent was asked to compare a pair of plumbing materials and evaluate each plumbing material based on a 1-9 preference scale. For example, Material A might be rated as 6 while Material B might be rated as 1. The 1-9 preference scale had a verbal preference assigned to each categorical value. Preference values of 1, 3, 5, 7, and 9 were assigned to 'Not Preferred', 'Moderately Preferred', 'Strongly Preferred', 'Very Strongly Preferred', and 'Extremely Preferred', respectively. Two hundred thirty respondents fully answered all CA questions, and each viewed three pairs of two plumbing materials resulting in 1,380 preference responses.

As each presented plumbing material had all attributes listed and there was no randomization of attributes level across the plumbing materials, the preference score was easily identified with the preferred plumbing material by comparing the attributes levels with the plumbing material descriptions. All preference responses to each plumbing material were then summed, and the plumbing material with the highest number of 'Extremely Preferred' responses and with lowest number of 'Not Preferred' responses was selected as the most preferred plumbing material. Table 3.45 presents the descriptive statistical summary of preference valuation break down of the 1,380 responses for plumbing materials. Material C is the least preferred type of plumbing material (211 not preferred responses) while Material A is the most preferred material among homeowners, (39 extremely preferred responses).

Table 3.45: Preference Valuation of Plumbing Materials

| | Preference Response Value | | | | |
|---|---|---|---|---|---|
| Plumbing Material | Not Preferred | Moderately Preferred | Strongly Preferred | Very Strongly Preferred | Extremely Preferred |
| Material A | 103 | 134 | 99 | 85 | 39 |
| Material B | 148 | 151 | 85 | 56 | 20 |
| Material C | 211 | 156 | 56 | 31 | 8 |
| Total | 460 | 441 | 240 | 172 | 67 |

In addition to evaluating three sets of two plumbing material scenarios, each respondent selected the most preferred plumbing material across all three materials displayed at the same time. Table 3.46 presents that Material A is chosen as the preferred plumbing material by more than 50% of respondents. Material C is the least often chosen as the preferred plumbing material (17.8%). These two separate measures yield the same result of Material A being the most preferred plumbing material across the three alternatives.

Table 3.46: Baseline Plumbing Material Chosen as Most Preferred

| Plumbing Material | Frequency | Percent |
|---|---|---|
| Material A | 116 | 50.4 |
| Material B | 48 | 20.8 |
| Material C | 41 | 17.8 |
| Neither | 4 | 1.7 |
| Missing | 22 | 9.7 |
| Total | 230 | 100 |

#### 3.8.2.2 Conjoint Analysis Results

As mentioned earlier, there are two estimation methods that could be employed in plumbing material preference valuation exercise: Ordinary Least Squares as well as Ordered Logit regressions. For this analysis, only the Ordered Logit regression is utilized in the plumbing material estimation of preferences as the differences between the preference valuation categories[24] are assumed not constant. The Ordered Logit model is estimated at the aggregate response level. The aggregate level analysis implies that average value coefficients are estimated for the participating sample of respondents. The individual level analysis, which implies estimating separate regressions for each of the respondents, was also attempted; however, the individual-level part-worth estimates are not reliable due to the partial factorial design, leaving few degrees of freedom for estimation at the individual level (Wedel & Kamakura, 2000). Due to the low degree of freedom count and lack of stable and statistically significant results for each respondent's regression, the individual-level analysis was not reported. Each respondent evaluated three pairs of plumbing materials, which provided 6 observations for an individual regression analysis. As the number of independent variables was greater than the number of observations for each respondent, inadequacy in the number of degrees of freedom and parameter instability was observed.

This CA analysis provides information on the preferences of homeowners for plumbing materials and the attributes that drive their decision when making purchasing decision with regards to the type of home plumbing system. Each respondent evaluated a set of two plumbing material portfolios at one time for a total of six portfolios using the valuation metrics 1-9 described earlier. Each of the plumbing materials has a set of attributes described in Table 3.44. Each of the material attribute level is employed as the independent variable in the material preference analysis. They are coded as dummy variables taking a value of 1 when the plumbing material characteristics are a part of the product portfolio and zero otherwise. There are six different survey versions, which differ by the placement of the CA questions within the questionnaire as well as the order of the plumbing material pairs' occurrences. To test for the CA order in the survey as

[24] **Preference Valuation Category –** plumbing material preference measure with scale from 1 to 9 where 1 implies 'Not Preferred' category and 9 implies 'Extremely Preferred' category.

well as the order of the CA questions, dummy variables representing each survey version are created. Finally, the socioeconomic characteristics (reported in the first survey) are also included in the Ordered Logit model. These characteristics represent household home value (continuous variables), age of the house (continuous variable), plumbing material type (dummy variable), pinhole leak occurrences in the past (dummy variable), and respondent's previous cost of plumbing material repairs and replacement (continuous variable).[25]

In this section, the estimation results of the aggregate level Ordered Logit model are presented. As stated in the earlier section, each respondent evaluated three scenarios describing a set of two plumbing materials. Each plumbing material was described by the following attributes: corrosion resistance, taste and odor, health effects, convenience of installation, proven performance on the market, plumbing material cost, and warranty length. Each respondent was asked to compare a pair of plumbing materials and evaluate each plumbing material based on a 1-9 preference scale. The independent variable that includes the attribute-levels of each plumbing material and the survey versions are coded as dummy variables. In order to avoid the dummy variable trap[26], one attribute level of plumbing material and sixth survey version are omitted from the regression. In order to test whether the survey version has an impact on the placement of each valuation category, the survey version dummy variable is inputted into the equation. On the other hand, to test the impact of survey versions on the actual responses to the plumbing material attributes, the dummy variables are interacted with levels of each attribute. The attribute levels related to Material C represent the omitted attribute levels in the estimated regression. As a result, all Material A and Material B attribute levels are compared to the Material C's attribute levels in the regression. The findings' discussion includes case processing summary, model fitting information, goodness-of-fit measure, and parameter estimates results summary.

[25] Variables such race, education level, gender were not included in the model, as little variation in these characteristics was observed for the sample of respondents.
[26] **Dummy Variable Trap** – the regression intercept as well as a number of dummies equal to the number of characteristics cannot be included in the equation because it results in exact collinearity and unique estimates cannot be estimated. To avoid it employ one less dummy variable than the number of categories that exist for each characteristic (Griffith et al., 1993).

The estimated Ordered Logit model involves the attribute levels of plumbing materials, survey version dummies, as well as survey version interaction variables with the attribute levels. As the total number of respondents is rather small (only 230), there are not enough degrees of freedom to include all interaction variables between the attribute levels and the survey versions. As a result, the Ordered Logit models with survey dummy variables entering one at a time are estimated to measure the impact of survey version characteristics on the plumbing material preferences. The statistically significant attribute levels, dummy variables, and the interaction variables are identified and entered into the final model estimation.

Based on the model specifications with survey versions entering one at a time, two survey versions were found statistically significant as intercept shifters as well as interaction variables. Survey version 3 and 5 as well as interaction variable between 'corrosion proof' and survey version 3 and 5 were statistically significant and were inputted into the final Ordered Logit model estimation with all plumbing material attributes levels. As both survey versions 3 and 5 are inputted into the Ordered Logit model, the remaining survey versions, 1, 2, 4, and 6, are treated as the survey base employed to compare the impact of survey version 3 and 5.

The first part of the model evaluation involves the case processing description (Table 3.47). As presented, 33% of responses fall into the 'Not Preferred' preference valuation while only 5% fall within the 'Extremely Preferred' category.

Table 3.47: Frequency Distribution of Responses for Valuations of Plumbing Material Scenarios Based on the Ordered Logit Model

| Category | Observations | Marginal Percentage |
|---|---|---|
| Not Preferred | 461 | 33.4% |
| Moderately Preferred | 442 | 32.0% |
| Strongly Preferred | 239 | 17.3% |
| Very Strongly Preferred | 173 | 12.5% |
| Extremely Preferred | 67 | 4.8% |

Model fitting information for the Ordered Logit model indicates the parameters for which the model-fit is calculated. Comparison of two model versions is conducted: model with only intercept (null hypothesis model) and model with independent variables (final model). The intercept only version of the model describes a model without any predictors in the regression. The final model includes the specified independent variables arrived at through an iterative process maximizing the log-likelihood of the outcomes (Aaron, 2005). The final model should improve upon the intercept only model. This test is conducted by analyzing the differences in the -2(log likelihood) values associated with the models. The -2 log likelihood is the product of -2 and the log likelihoods of the null model (fitted final model). The likelihood of the model is used to test whether all predictors' regression coefficients in the model are simultaneously zero and in tests of nested models (Aaron, 2005).

The model fitting information presents that the Chi-square statistic is 114.136 with a p-value of 0.000 and a log-likelihood value of 182.641 (Table 3.48). As a result, the existence of a relationship between the independent variables and the dependent variable is supported and the final model can be estimated. The goodness-of-fit measure is employed as well in the overall model evaluation. The Nagelkerke's R-square is 0.084, which implies that 8% of variation in the dependent variable is explained by the variation in the independent variables. Finally, the correlations between the predicted response values and the actual response values are compared. A Pearson correlation statistic is employed to evaluate how closely related the predicted and actual responses are to each other (Griffiths et al., 1993). The correlation results suggest that the predicted and actual responses are not very highly correlated with each other. The Pearson correlation statistics is only 0.245 with significance level of 0.000.

Table 3.48: Model Fitting Information for Ordered Logit Regression with Categorical Answers (Dependent Variable Represents the Plumbing Material Valuation and the Independent Variables Represent the Plumbing Material Attributes)

| Model | -2 Log Likelihood | Chi-Square | Significance |
|---|---|---|---|
| Intercept Only | 269.778 | | |
| Final | 182.641 | 114.136 | 0.000 |

In evaluating the Ordered Logit model, threshold represents the response variable in the regression. A different intercept is provided for the different levels of the cumulative logit model. The beta coefficient of the independent variables does not change and the value of each is subtracted from the intercept. Each threshold level indicates the logit of the odds of being equal to or less than the baseline category when all independent variables are zero (Aaron, 2005). The baseline group is set to ‘Extremely Preferred’. The beta estimate represents that a one unit increase in the independent variable increases/decreases the log-odds of being higher than a specific preferred valuation category. Because the beta coefficient is not indexed by each category, a one unit increase affects the log-odds the same regardless of which threshold value is considered (Aaron, 2005).

As represented in Table 3.49, the regression estimates reveal that when compared to the baseline category (‘Extremely Preferred’), the categories ‘Moderately Preferred’, ‘Strongly Preferred’, and ‘Very Strongly Preferred’ have higher threshold estimates. A category ‘Not Preferred’ has a statistically insignificant negative coefficient estimate. Since the estimate is not statistically significant at the 95% confidence interval, it is not included in comparison analysis between the categories.

The threshold values are evaluated. These values inform the expected cumulative distribution of categorical preference values for individuals with the independent variables set to zero (Aaron, 2005). This threshold represents a natural tendency for all the responses to all the scenarios presented to respondents when the independent variables are suppressed. Scenarios are shown in the survey presented in Appendix C. When these coefficients are exponentiated, the cumulative odds for each category are obtained (Table 3.49). By employing the following equation, (odds/(1+odds)), the cumulative probabilities are computed (Aaron, 2005). Table 3.49 represents the odds ratios and cumulative probabilities (columns 3 and 4 in Table 3.49). For example, the ‘Moderately Preferred” category is 3.7 times more likely to be selected by the respondent compared to the ‘Extremely Preferred’ category when all independent variables are set to zero. On the other hand, the ‘Not Preferred’ category is 0.91 times as likely to occur compared to the baseline category when no independent variables are considered.

Table 3.49: Threshold Estimates from Ordered Logit Regression with Categorical Answers (Dependent Variable Represents the Plumbing Material Valuation and the Independent Variables Represent the Plumbing Material Attributes)

| Variable Name | Coefficient Estimates | Significance | Odds Ratio Estimates | Cumulative Probability Estimates | Distribution Estimates |
|---|---|---|---|---|---|
| Not Preferred | -0.096 | 0.374 | 0.908 | 0.476 | 0.476 |
| Moderately Preferred | 1.315 | 0.000 | 3.725 | 0.788 | 0.312 |
| Strongly Preferred | 2.289 | 0.000 | 9.865 | 0.908 | 0.120 |
| Very Strongly Preferred | 3.742 | 0.000 | 42.182 | 0.977 | 0.069 |
| Extremely Preferred | | | | 1.000 | 0.023 |

The independent variable coefficient estimates are statistically significant only for two attribute levels: risk of corrosion variable represented by ‘corrosion proof’ attribute level and convenience of installation represented by ‘no need to tear into the wall and/or floor. Installation takes around 4 days’ (Table 3.50). Other independent variables were considered redundant in the model estimation. The independent variable coefficients represent how the log-odds of these thresholds increase/decrease with one unit of the independent variable. The positive value indicates that one unit of independent variable increases the odds of being in a higher category (Aaron, 2005). For example, the ‘corrosion proof’ attribute level increases the odds of choosing a higher preference category by 0.654 compared to the independent variable represented by the some risk of corrosion attribute level. Installation of plumbing material taking about 4 days increases the odds of choosing a higher preference category by 0.559 compared to the installation taking between 7 and 9 days.

The impact of the survey version of respondents’ preference valuation was also tested. Survey versions 3 and 5 have a statistically significant impact at either 5% or 10% significant level on the categorical preference answers and therefore the odds of being in the higher category. Survey version 3 is described by CA questions being presented at the beginning of the questionnaire with Materials B and C being compared in the first round followed by comparison of Materials A and B, and finally followed by comparison of Materials A and C. Survey version 5 is described by AHP questions appearing first in the questionnaire. The CA plumbing material comparisons follow the following order: 1) Materials A and C, 2) Material B and C, and 3) Material A and B.

Results in Table 3.50 present that survey version 5 decreases the odds of choosing a higher preference category by 0.470 while survey version 3 increases the odds ratios by 0.521. The interaction variables between the survey versions are also statistically significant at 5% significance level. The joint coefficient for corrosion proof attribute level and survey version 3 is 0.017, which implies that the odds ratio of choosing a higher preference category is increased by 0.02 for respondents who answered CA questions to survey version 3[27]. This joint coefficient is, however, not statistically significant different from zero and confirms the small magnitude observed on the joint coefficient estimate. On the other hand, the joint coefficient between the attribute level of ‘corrosion proof’ and survey version 5 is 1.452[28], which implies seeing survey version 5 increases the odds of choosing a higher category of preference by 1.452. As a result, the survey versions seen by the respondents do impact the preference valuation for plumbing materials.

Besides evaluating the directional impact of the independent variables on the preference level of the households, the impact of the statistically significant independent variables on the preference category is evaluated for all three plumbing materials. The statistically significant attribute levels were computed together with the thresholds levels by plumbing material into odds ratios and then into the probability values for survey version 3 and 5 and all other remaining survey versions. As the attribute levels describing each of the three hypothetical materials are known, the regression results can be organized by plumbing materials. For example, for base survey version (including survey versions 1, 2, 4, and 6), Material A is described by attribute level called ‘corrosion proof’ as well as ‘installation takes around 4 days’. The coefficient estimates for the statistically significant attribute levels are employed to compute preference valuation categories for each material type. In case of the Material A computation of the preference valuation category called ‘Moderately Preferred’, the following represents the estimate computation: 1.315 - 0.654 - 0.559= 0.102, where 1.315 is the moderately preferred coefficient, 0.654 is the corrosion proof coefficient, and 0.559 is the

[27] Survey Version 3 * Corrosion Proof: Wald statistic = 0.002 and p-value = 0.900.
[28] Survey Version 5 * Corrosion Proof: Wald statistic = 20.537 and p-value = 0.000.

convenience of installation coefficient; and the odds ratio computation: exp(0.102) = 1.107.

Table 3.50: Ordered Logit Regression Estimates with Categorical Answers (Dependent Variable Represents the Plumbing Material Valuation and the Independent Variables Represent the Plumbing Material Attributes (Without Socioeconomic Variables))[a]

| Variable Name | Coefficient Estimate[b] | Standard Error[c] | Wald-Stats[d] | P-Value[d] |
|---|---|---|---|---|
| Threshold Values (For All Independent Variables Set to Zero) | | | | |
| Not Preferred | -0.096 | 0.108 | 0.790 | 0.374 |
| Moderately Preferred | 1.315 | 0.115 | 131.554 | 0.000 |
| Strongly Preferred | 2.289 | 0.125 | 333.413 | 0.000 |
| Very Strongly Preferred | 3.742 | 0.164 | 521.510 | 0.000 |
| Independent Variables (Variables that Improve Overall Model Significance)[e] | | | | |
| Corrosion proof | 0.654 | 0.143 | 20.943 | 0.000 |
| Need to tear into some sections of wall for installation. Installation takes around 4 days. | 0.559 | 0.119 | 22.016 | 0.000 |
| Survey Version 3 | 0.521 | 0.220 | 5.592 | 0.018 |
| Survey Version 5 | -0.470 | 0.263 | 3.187 | 0.074 |
| Corrosion proof *Survey Version 3 | -0.637 | 0.269 | 5.620 | 0.018 |
| Corrosion proof *Survey Version 5 | 0.798 | 0.314 | 6.447 | 0.011 |

[a] The number of observations included in the model is 1086. Independent variables take form of dummy variables with value of one when the characteristic was present in the plumbing material profile and zero otherwise. To avoid a dummy variable trap, one of the attribute levels was excluded from the analysis.
[b] Coefficient estimates show how much increase in the likelihood of being in a higher category results from a one unit increase in the independent variable.
[c] Standard error represents the variation of the estimate.
[d] Wald statistics and p-value represent the significance level.
[e] Model Statistics: Log-likelihood value is 182.641 with chi-square of 114.136 and p-value of 0.000; Nagelkerke's R-square is 0.084. The Pearson correlation statistic comparing the actual and estimated values is only 0.245 with a p-value of 0.000.

As presented in Tables 3.51, 3.52, and 3.53, for all survey versions, Material A has the lowest values of the coefficient estimates for all preference categories compared to Materials B and C. Material C has the highest values of preference valuation. For example, for survey version 3, the coefficient for "Not Preferred" for Material A is -1.193 compared to 0.020 for Material C indicating that Material A is less likely to be evaluated as "Not Preferred" compared to Material C. When compared across the survey versions, Materials A, B, and C have the smallest coefficient values for survey version 5, while these materials have the highest coefficient values for survey version 3.  This finding implies that respondents who saw survey version 5 had a higher level of preferences for all three materials compared to the remaining survey versions.

On the other hand, Material C has the highest values of odds ratios for each preference category while Material A has the lowest for all survey versions.  The odds ratios that present the likelihood of a preference category being selected are compared to the base category. For example, for survey version 3, the category 'Very Strongly Preferred' is 47.4 times as likely to be selected compared to the base category for Material C while for Material A, it is only 14.1 times as likely compared to the 'Extremely Preferred Category'. A lower likelihood for each preference category is more preferred as it implies that the 'Extremely Preferred' category has a higher chance of being chosen. When compared across all survey versions, survey version 5 presents the lowest odds ratios for Materials A, B and C while survey version 3 presents the highest values for these three plumbing materials. This finding implies that there is a lower chance of being in the lowest category for survey version 5 compared to survey version 3.

Based on the marginal distribution of the probability estimates, across all survey versions, Material A has a larger probability estimate for 'Strongly Preferred' to 'Extremely Preferred' category preference.  On the other hand, Material C has a higher probability estimates for categories 'Not Preferred' and 'Moderately Preferred'.  When compared across survey versions, Materials A, B, and C have the highest frequency of estimates falling into 'Strongly Preferred' to 'Extremely Preferred' categories for survey version 5. All three materials have the highest frequency of estimates falling into 'Not Preferred' and 'Moderately Preferred' categories across all survey versions.  Based on the overall results, Material A is the most preferred material followed by Material B. Material C is the least preferred plumbing material.

As presented above, the plumbing material valuation results vary by a survey version, which implies that seeing certain question order might influence households' plumbing material assessment. For example, survey version 5 presented the lowest odds ratios and highest frequency distribution of the 'Extremely Preferred' answer category. This outcome might have been a result of the survey version 5 being described by CA questions appearing after the AHP pair-wise attribute comparisons. The pair-wise comparisons of plumbing material attributes might have provided a learning opportunity for establishing the relative importance of each attribute level. In his CA study, Bradlow et al. (2004) found that in fact, learning of attributes occurs and the relative weights of each attribute level change during the portfolio evaluation process. As a result, the established importance of each attribute level based on the AHP exercise could have aided the plumbing material assessment in the CA section of the questionnaire, resulting in in-depth differentiation of the alternatives.

Bradlow et al. (2004) also found that the relative importance of each attribute level and therefore the portfolio assessment depends on the order in which the product profiles appear in the survey. Consequently, as Materials A and B have some of the same attribute levels ( e.g. ‘corrosion proof’), comparing each of these materials first against Material C (which does not share the same attribute levels) might have allowed for clear differentiation between the attribute levels of each material and therefore, a higher degree of differentiation within the household preference valuation.

On the other hand, survey version 3 presented the lowest preference valuation for plumbing materials compared to the other survey versions. Having presented the CA questions at the beginning of the questionnaire could have complicated the comparison of plumbing types by removing the learning opportunity that was provided by the AHP comparison and, therefore, resulting in lower preference valuation of each material. According to Saaty (1980), it is more difficult to assess the preference level of a portfolio of attributes compared to a pair of attributes evaluated at a time. Furthermore, Materials A and B (which share some of the same attribute levels) were presented as second in the CA section and the similarity of the attribute levels across the products might have negatively impacted respondents’ ability to clearly discriminate between the plumbing materials. The impaired valuation of each portfolio of attributes might have resulted in overall lower preference valuation of plumbing materials.

Table 3.51: Ordered Logit Regression Results' Analysis by Plumbing Material Type (Dependent Variable Represents the Plumbing Material Valuation and the Independent Variables Represent the Plumbing Material Attributes (No Socioeconomic Variables)) – Survey Versions 1, 2 4, and 6.

| | Material A | Material B | Material C |
|---|---|---|---|
| | Coefficient Estimates[a] | | |
| Not Preferred | -1.309 | -0.750 | -0.096 |
| Moderately Preferred | 0.102 | 0.661 | 1.315 |
| Strongly Preferred | 1.076 | 1.635 | 2.289 |
| Very Strongly Preferred | 2.529 | 3.088 | 3.742 |
| Extremely Preferred | | | |
| | Odds Ratio Estimates | | |
| Not Preferred | 0.270 | 0.472 | 0.908 |
| Moderately Preferred | 1.107 | 1.937 | 3.725 |
| Strongly Preferred | 2.933 | 5.129 | 9.865 |
| Very Strongly Preferred | 12.541 | 21.933 | 42.182 |
| Extremely Preferred | | | |
| | Cumulative Probability Estimates | | |
| Not Preferred | 0.213 | 0.321 | 0.476 |
| Moderately Preferred | 0.525 | 0.659 | 0.788 |
| Strongly Preferred | 0.746 | 0.837 | 0.908 |
| Very Strongly Preferred | 0.926 | 0.956 | 0.977 |
| Extremely Preferred | 1.000 | 1.000 | 1.000 |
| | Marginal Probability Estimates Distribution | | |
| Not Preferred | 0.213 | 0.321 | 0.476 |
| Moderately Preferred | 0.313 | 0.339 | 0.312 |
| Strongly Preferred | 0.220 | 0.177 | 0.120 |
| Very Strongly Preferred | 0.180 | 0.120 | 0.069 |
| Extremely Preferred | 0.074 | 0.044 | 0.023 |

[a]Coefficient estimates are built up from the statistically significant estimates for the attribute levels and threshold values.  Coefficients are compared to the base "Extremely Preferred" level.

Table 3.52: Ordered Logit Regression Results' Analysis by Plumbing Material Type (Dependent Variable Represents the Plumbing Material Valuation and the Independent Variables Represent the Plumbing Material Attributes (No Socioeconomic Variables)) – Survey Version 3.

| | Material A | Material B | Material C |
|---|---|---|---|
| | Coefficient Estimates[a] | | |
| Not Preferred | -1.193 | -0.634 | 0.020 |
| Moderately Preferred | 0.218 | 0.777 | 1.431 |
| Strongly Preferred | 1.192 | 1.751 | 2.405 |
| Very Strongly Preferred | 2.645 | 3.204 | 3.858 |
| Extremely Preferred | | | |
| | Odds Ratio Estimates | | |
| Not Preferred | 0.303 | 0.530 | 1.020 |
| Moderately Preferred | 1.244 | 2.175 | 4.183 |
| Strongly Preferred | 3.294 | 5.760 | 11.078 |
| Very Strongly Preferred | 14.083 | 24.631 | 47.371 |
| Extremely Preferred | | | |
| | Cumulative Probability Estimates | | |
| Not Preferred | 0.233 | 0.347 | 0.505 |
| Moderately Preferred | 0.554 | 0.685 | 0.807 |
| Strongly Preferred | 0.767 | 0.852 | 0.917 |
| Very Strongly Preferred | 0.934 | 0.961 | 0.979 |
| Extremely Preferred | 1.000 | 1.000 | 1.000 |
| | Marginal Probability Estimates Distribution | | |
| Not Preferred | 0.233 | 0.347 | 0.505 |
| Moderately Preferred | 0.322 | 0.338 | 0.302 |
| Strongly Preferred | 0.213 | 0.167 | 0.110 |
| Very Strongly Preferred | 0.167 | 0.109 | 0.062 |
| Extremely Preferred | 0.066 | 0.039 | 0.021 |

[a]Coefficient estimates are built up from the statistically significant estimates for the attribute levels and threshold values. Coefficients are compared to the base "Extremely Preferred" level.

Table 3.53: Ordered Logit Regression Results' Analysis by Plumbing Material Type (Dependent Variable Represents the Plumbing Material Valuation and the Independent Variables Represent the Plumbing Material Attributes (No Socioeconomic Variables)) – Survey Version 5.

| | Material A | Material B | Material C |
|---|---|---|---|
| | Coefficient Estimates[a] | | |
| Not Preferred | -1.637 | -1.078 | -0.424 |
| Moderately Preferred | -0.226 | 0.333 | 0.987 |
| Strongly Preferred | 0.748 | 1.307 | 1.961 |
| Very Strongly Preferred | 2.201 | 2.760 | 3.414 |
| Extremely Preferred | | | |
| | Odds Ratio Estimates | | |
| Not Preferred | 0.195 | 0.340 | 0.654 |
| Moderately Preferred | 0.798 | 1.395 | 2.683 |
| Strongly Preferred | 2.113 | 3.695 | 7.106 |
| Very Strongly Preferred | 9.034 | 15.800 | 30.387 |
| Extremely Preferred | | | |
| | Cumulative Probability Estimates | | |
| Not Preferred | 0.163 | 0.254 | 0.396 |
| Moderately Preferred | 0.444 | 0.582 | 0.728 |
| Strongly Preferred | 0.679 | 0.787 | 0.877 |
| Very Strongly Preferred | 0.900 | 0.940 | 0.968 |
| Extremely Preferred | 1.000 | 1.000 | 1.000 |
| | Marginal Probability Estimates Distribution | | |
| Not Preferred | 0.163 | 0.254 | 0.396 |
| Moderately Preferred | 0.281 | 0.329 | 0.333 |
| Strongly Preferred | 0.235 | 0.205 | 0.148 |
| Very Strongly Preferred | 0.222 | 0.153 | 0.091 |
| Extremely Preferred | 0.100 | 0.060 | 0.032 |

[a]Coefficient estimates are built up from the statistically significant estimates for the attribute levels and threshold values. Coefficients are compared to the base "Extremely Preferred" level.

The second specification of the Ordered Logit model includes the socioeconomic variables alongside of the attributes for plumbing material. As the socioeconomic characteristics do not vary by a respondent, they are interacted with the attributes levels of each attribute. As the total number of respondents is rather small (230), there are not enough degrees of freedom to include all interaction variables between the attribute levels and the household characteristics. As a result, the Ordered Logit model was first estimated with socioeconomic variables entering one at a time to measure the impact of household characteristics on the plumbing material preferences. The statistically significant interaction variables were then included in the final model estimation.

When the socioeconomic variables were entered in the Ordered Logit model one at a time, ‘corrosion proof’ as well as ‘installation takes about 4 days’ were the two attribute levels appearing statistically significant in many of the model specifications. The coefficient value for corrosion attribute varied from 0.651 to 1.450 and the convenience of installation coefficient varied from 0.554 to 0.754. The only statistically significant interaction effect was observed between attribute level of ‘corrosion proof’ and respondent’s previous cost of plumbing materials repairs or replacement (coefficient estimate = 0.00001; standard error = 0.00000; Wald-statistic = 15.773; p-value = 0.000). This interaction effect was entered into the final model estimation alongside of other plumbing material attributes.

The overall model evaluation is presented for the model including all observations and differences for regression with a socioeconomic factor included one at a time. As before, the first part of the model evaluation involves the case processing summary based on the total number of observations included in the model estimation. The results are presented in Table 3.54[29]. As presented 33% of responses fall into the ‘Not Preferred’ preference valuation while only 5% fall within the ‘Extremely Preferred’ category.

Table 3.54: Case Processing Summary for Ordered Logit Regression with Categorical Answers (Socioeconomic Variables Included)

| Category | Observations | Marginal Percentage |
|---|---|---|
| Not Preferred | 358 | 33.4% |
| Moderately Preferred | 335 | 31.2% |
| Strongly Preferred | 195 | 18.2% |
| Very Strongly Preferred | 133 | 12.4 % |
| Extremely Preferred | 51 | 4.8% |

In evaluating the Ordered Logit model, thresholds represent the response variable in the regression. A different intercept value presents a different level of the cumulative logit model. Each threshold level indicates the logit of the odds of being equal to or less than the baseline category when all independent variables are zero (Aaron, 2005). The baseline group is set to ‘Extremely Preferred’. The beta coefficient, however, does not change and it is subtracted from the intercept. As represented in Table 3.55, the estimates

[29] The total number of observations is 1072 as not all respondents provided information on their socioeconomic characteristics.

reveal that when compared to the baseline category ('Extremely Preferred'), categories: 'Moderately Preferred', 'Strongly Preferred', and 'Very Strongly Preferred' have higher frequencies. A category 'Not Preferred' has a statistically significant coefficient estimate at significance level of 10%. The threshold coefficients are the same for the overall model and the model with each socioeconomic characteristic entered one at a time.

The threshold values, which inform the expected cumulative distribution of categorical preference values for individuals with the independent variables set to zero, are evaluated (Aaron, 2005). Table 3.55 represents the odds and probabilities. For example, the 'Moderately Preferred" category is 3.67 times more likely to be selected by the respondent than the 'Extremely Preferred' category when all independent variables are set to zero. On the other hand, the 'Not Preferred' category is only 0.86 times as likely to occur compared to the baseline category when no independent variables are considered.

Table 3.55: Threshold Estimates from Ordered Logit Regression with Categorical Answers (Dependent Variable Represents the Plumbing Material Valuation and the Independent Variables Represent the Plumbing Material Attributes and the Socioeconomic Characteristics)

| Variable Name | Coefficient Estimates | Significance | Odds Ratio Estimates | Cumulative Probability Estimates | Distribution Estimates |
|---|---|---|---|---|---|
| Not Preferred | -0.147 | 0.100 | 0.863 | 0.463 | 0.463 |
| Moderately Preferred | 1.300 | 0.000 | 3.669 | 0.786 | 0.323 |
| Strongly Preferred | 2.317 | 0.000 | 10.145 | 0.910 | 0.588 |
| Very Strongly Preferred | 3.790 | 0.000 | 44.256 | 0.978 | 0.390 |
| Extremely Preferred | | | | 1.000 | 0.025 |

Based on Table 3.56, the independent variable coefficient estimates are statistically significant only for two attribute levels: corrosion proof and installation takes about 4 days. For example, the 'corrosion proof' variable increases the odds by 1.145 of choosing a higher preference category compared to the variable set at 'some risk of corrosion'. Installation of plumbing material taking about 4 days increases the odds of choosing a higher preference category by 0.575 compared to the installation taking between 7 and 9 days. The only socioeconomic variable entered into the regression is the respondent's previous cost of plumbing repairs and/or material fixing or replacement and is statistically significant when interacted with corrosion proof attribute level. The joint coefficient is 1.197 (1.145+0.0001*$522[30]) and is statistically significant at 5% significance level[31]. This coefficient value further implies that the interaction variable increases the odds by 1.197 of choosing a higher preference category compared to the variable set at 'some risk of corrosion'. This finding can be explained as households, who have accrued cost of plumbing material repairs in the past, value the 'corrosion proof' attribute level more compared to the 'some risk of corrosion' attribute level. Plumbing material with low corrosion risk would imply decrease in the future costs of plumbing material repairs.

As in the previous version of the Ordered Logit model, effects of statistically significant independent variables on the preference category for all three plumbing materials are evaluated. The statistically significant attribute levels were computed together with the thresholds levels by plumbing material into odds ratios and probability values. As attribute levels describing each of the three hypothetical materials are known, the regression results can be organized by plumbing materials. For example, Material A is described by attribute level called 'corrosion proof' and 'installation takes about 4 days'. The coefficient estimates for the statistically significant attribute levels are employed in the material based preference category computation. In case of Material A, the computation for preference valuation category of 'Moderately Preferred', the following represents the estimate computation: 1.300- 1.145- 0.575 -0.0001*$522 = -0.472; and the odds ratio computation: exp(-0.472) = 0.624 (Table 3.57).

[30] $522 is the mean cost value of the previous cost spent on plumbing material repairs and replacement.
[31] Cost of Plumbing Material Fixing or Replacement * Corrosion Proof: Wald statistic = 5.684 and p-value = 0.020.

Table 3.56: Ordered Logit Regression Estimates with Categorical Answers (Dependent Variable Represents the Plumbing Material Valuation and the Independent Variables Represent the Plumbing Material Attributes and Socioeconomic Variables Interacted with Attribute Levels) [d]

| Variable Name | Coefficient Estimate[b] | Standard Error[c] | Wald-Stats[d] | P-Value[d] |
|---|---|---|---|---|
| Threshold Values (For All Independent Variables Set to Zero) | | | | |
| Not Preferred | -0.147 | 0.089 | 2.705 | 0.100 |
| Moderately Preferred | 1.300 | 0.098 | 176.801 | 0.000 |
| Strongly Preferred | 2.317 | 0.114 | 415.544 | 0.000 |
| Very Strongly Preferred | 3.790 | 0.164 | 532.389 | 0.000 |
| Independent Variables for Model Specification with Socioeconomic Variable Interactions[e and f] | | | | |
| Corrosion Proof | 1.145 | 0.502 | 5.190 | 0.023 |
| Need to tear into some sections of wall for installation. Installation takes around 4 days. | 0.575 | 0.134 | 18.331 | 0.000 |
| Respondent's previous cost of plumbing repairs and/or replacement * Corrosion Proof | 0.0001 | 0.00006 | 4.644 | 0.031 |

[a] The number of observations included in the model is 1086. Independent variables take form of dummy variables with value of one when the characteristic was present in the plumbing material profile and zero otherwise. To avoid a dummy variable trap, one of the attribute levels was excluded from the analysis.
[b] Coefficient estimates show how much increase in the likelihood of being in a higher category results from a one unit increase in the independent variable.
[c] Standard error represents the variation of the estimate.
[d] Wald statistics and p-value represent the significance level.
[e] Only those independent variables are included which are statistically significant at either 5% or 10%. None of the model specifications with socioeconomic variables entering one at a time had statically significant coefficients when interacted with the attribute levels of plumbing materials.
[f] The number of observations included in the model is 1072. Independent variables take form of dummy variables with value of one when the characteristic was present in the plumbing material profile and zero otherwise. To avoid a dummy variable trap, one of the attribute levels was excluded from the analysis. Model Statistics: Log-likelihood value is 1565.522 with chi-square of 119.384 and p-value of 0.000; Nagelkerke's R-square is 0.101. Pearson correlation statistics comparing the actual and estimated values is only 0.330 with a p-value of 0.000.

As presented in Table 3.57, Material A has the lowest values of estimates for all preference categories, compared to Materials B and C. Material C has the highest values of preference valuation. Threshold values with smaller absolute values imply smaller differences between preference valuation categories and the base category in the likelihood of that preference category being selected. For example, Material B has a smaller absolute threshold value compared to Material A for the "Not Preferred" category, implying a smaller difference between 'Not Preferred' and 'Extremely Preferred' for Material B (-1.344) compared to Material A (-1.919).

Material C has the highest values of odds ratios for each preference category while Material A has the lowest. The odds ratios that present the likelihood of a preference category being selected are compared to the base category. For example, the category 'Strongly Preferred' is 10.145 times as likely to be selected as the base category for Material C while for Material A it is only 1.724 times as likely. A lower odds ratio for each preference category is more preferred as it implies that the 'Extremely Preferred' category has a higher chance of being chosen relative to other categories.

Based on the marginal distribution probability estimates, Material A has a larger probability estimate for 'Strongly Preferred' to 'Extremely Preferred' category preference. On the other hand, Material C has higher probability estimates for category 'Not Preferred'. Based on these results, Material A is again the most preferred material followed by Material B. Material C as previously found is the least preferred plumbing material.

| Table 3.57: Ordered Logit Regression Results' Analysis by Plumbing Material Type (Dependent Variable Represents the Plumbing Material Valuation and the Independent Variables Represent the Plumbing Material Attributes and the Socioeconomic Characteristics) | | | |
|---|---|---|---|
| | Material A | Material B | Material C |
| | Coefficient Estimates[a] | | |
| Not Preferred | -1.919 | -1.344 | -0.147 |
| Moderately Preferred | -0.472 | 0.103 | 1.300 |
| Strongly Preferred | 0.545 | 1.120 | 2.317 |
| Very Strongly Preferred | 2.018 | 2.593 | 3.790 |
| Extremely Preferred | | | |
| | Odds Ratio Estimates | | |
| Not Preferred | 0.147 | 0.261 | 0.863 |
| Moderately Preferred | 0.624 | 1.108 | 3.669 |
| Strongly Preferred | 1.724 | 3.064 | 10.145 |
| Very Strongly Preferred | 7.522 | 13.367 | 44.256 |
| Extremely Preferred | | | |
| | Cumulative Probability Estimates | | |
| Not Preferred | 0.128 | 0.207 | 0.463 |
| Moderately Preferred | 0.384 | 0.526 | 0.786 |
| Strongly Preferred | 0.633 | 0.754 | 0.910 |
| Very Strongly Preferred | 0.883 | 0.930 | 0.978 |
| Extremely Preferred | 1.000 | 1.000 | 1.000 |
| | Distribution Estimates | | |
| Not Preferred | 0.128 | 0.207 | 0.463 |
| Moderately Preferred | 0.256 | 0.319 | 0.323 |
| Strongly Preferred | 0.249 | 0.228 | 0.124 |
| Very Strongly Preferred | 0.250 | 0.176 | 0.068 |
| Extremely Preferred | 0.117 | 0.070 | 0.022 |

[a]Coefficient estimates are built up from the statistically significant estimates for the attribute levels and threshold values. Coefficients are compared to the base "Extremely Preferred" level.

As in the previous model specifications, Material A is the most preferred plumbing material when the CA data is estimated, employing an Ordered Logit Model with and without socioeconomic characteristics. Material C is the least preferred plumbing material. Two plumbing material attributes are important in making the decision on type of pipes to be installed in a house: plumbing material installation time and corrosion risk. The regression coefficients as well as the computed odds ratios and probability estimates differ between the model specification with and without the socioeconomic variables.

For example, for Material A, the odds ratios are lower for all preference categories in the case of model specification with socioeconomic variables (e.g. category" 'Very Strongly Preferred' has odds ratios ranging from 9.034 to 14.083 for model without socioeconomic variables and 7.522 for model including socioeconomic variables). This finding implies that the socioeconomic variables impact the discrimination level between the plumbing material preference valuations. For example, if a household has experienced previous cost of plumbing repairs and/or replacement, their preference valuation level is lower for a more corrosion prone plumbing material compared to material with an attribute level of 'corrosion proof'.

The marginal distribution of probability estimates has higher values for lower preference categories for Material C in the case of model specification without socioeconomic variables. For example, for Material C, 'Not Preferred' has probability distribution estimate ranging between 0.396 and 0.505 depending on survey version (without) compared to 0.463 (with socioeconomic variables). The marginal distribution estimates for higher preference valuation categories are lower for Material A and B for model without socioeconomic variables. For example, for Material A, 'Extremely Preferred' has a probability distribution estimate ranging from 0.068 to 0.100 (without socioeconomic) compared to 0.117 (with socioeconomic variables). As a result, the inclusion of socioeconomic variables raises the level of preference for Materials A and B, while it decreases the level of preference for Material C.

In conclusion, although the inclusion of socioeconomic variables does not change the final preference ranking of the plumbing materials, it increases the estimated level of preference for Material A and B by increasing the marginal probability distribution of estimates for the higher preference categories (i.e. 'Strongly Preferred'). The increase is the most pronounced in the case of Material A (model with socioeconomic variables) for which the 'Extremely Preferred' category has a probability distribution estimate almost twice as large compared to the model specification without socioeconomic variables in case of survey version 3 (0.117 vs. 0.068). The respondent's previous cost of plumbing material repairs and replacement impacts positively the preference level for plumbing materials described by ‘corrosion proof’ attribute level. This finding implies that Materials A and B are more highly preferred when socioeconomic factors are taken into consideration. Households experiencing high costs of fixing corrosion related damage in the past are more likely to prefer and choose materials with lower corrosion levels. The decreased corrosion level implies lower future plumbing material failures and therefore, lower costs associated with repairs of water-related damage.

### 3.11 Conclusions and Discussion

Due to the fact that homeowners have an important stake in finding plumbing systems appropriate for their households, they should not only rely on expert advice but also acquire information on plumbing material attributes such as price, health impact, longevity, and corrosion resistance in order to make informed investment decisions about plumbing systems for their homes. Information on consumer preferences for drinking water plumbing attributes can be useful not only to individual households but also to policymakers, program managers, water utilities, and firms with interests in drinking water infrastructure.

This study addressed the issues of household plumbing material decisions as well as preferences for choosing corrosion prevention options manifested in the form of either house re-plumbing or purchasing of insurance against corrosion. The information was elicited by two surveys of residents residing in a Southeastern Community by employing stated preference methods (CA, CV, and AHP). Economists often utilize the stated preference methods (CA and CV) for non-market valuation analysis including willingness-to-pay estimation (Champ et al., 2003). On the other hand, AHP is widely used by the Engineering field for determination of the most optimal attributes (Helm et al., 2004). Results of the CA analysis are reported in this study.

In order to address the issues related to the household preferences for plumbing materials, the study objective was to investigate homeowners’ preferences for plumbing materials as well as preventive techniques against corrosion (plumbing material upgrade, purchase of an insurance policy against corrosion while staying with the current plumbing material, or staying with the current plumbing material) based on households’ experiences with plumbing material failures.

Two surveys of residents residing in a Southeastern Community were conducted. The first survey elicited information on the prevalence of pinhole leaks and other plumbing material failures, households' experiences with plumbing material failures, the cost of repairs and property damages due to the material failures, and household preferences for corrosion preventive measures. The follow-up survey, sent only to those residents who agreed to participate in future studies related to the plumbing material issues, elicited information on households' preferences for a set of hypothetical plumbing materials.

The objective of eliciting household preferences for plumbing materials and preventive measures against corrosion employed CA questions, which were analyzed using Multinomial Logit and Ordered Logit regression models. The data analysis employing MNL model specification was tested for violation of independence of irrelevant alternatives assumption. The several estimated model specifications included equations accounting for the socioeconomic characteristics of the households as well as differences in survey versions presented to the households (differences in the order of CA question occurrence in the survey layout).

Overall, the Southeastern Community survey revealed high level of awareness of pinhole leak problem among residents of the community. Twenty percent of the households reported actual pinhole leak incidents. The percent of pinhole leak reports was on par with other hot-spot corrosion places in the U.S., but above the rate of pinhole leak occurrences in non-hot spots (Scardina et al., 2007). The pinhole leak problem was more prevalent in houses built before the 1990s with copper pipes installed as the plumbing system. This finding is in an agreement with the Maryland Pinhole Leak Survey conducted by Kleczyk and Bosch (2008).

The total repair expenses due to the pinhole leaks varied between $100 and $5,000 with several reports of more than $5,000 in repairs. Similar results were found by Kleczyk et al. (2006) of selected communities in the East, Southeast, Midwest, and West regions. Over 50% of surveyed respondents spent more than $100 on repairs with estimates as high as $12,000. In comparison, in their Maryland Pinhole Leak Survey, Kleczyk and Bosch (2008) found costs from the plumbing material failure repairs to reach even $25,000. Unlike the present survey, however, the study by Kleczyk and

Bosch (2008) did not separate the costs associated with pipe failure and property damage. This Southeastern Community survey accounted for this factor, which might have resulted in the differences between the two studies. Furthermore, many households in the Southeastern Community cited using a preventive measure against corrosion including whole house re-plumbing and installation of water softeners. Over 80% of residents of the Southeastern Community were satisfied with the water quality in their homes.

The Conjoint Analysis of the first Southeastern Community survey data revealed that the corrosion risk level did not have a statistically significant impact on the decision to re-plumb the whole house, purchase an insurance policy while staying with the current plumbing system, or stay with the current plumbing material. The decision was impacted by either the cost of house re-plumbing or the cost of insurance policy against corrosion. When the socioeconomic variables were included in the analysis, a couple of factors including the house value, house age, and the respondent’s previous cost associated with fixing or replacing plumbing material impacted the household decisions. Overall, based on the corrosion risk values and the cost of the preventive measures, most households (69%) preferred to stay with the current plumbing material.

Four socioeconomic variables impacted the decisions for plumbing upgrade or insurance purchase: the house value, house age, pinhole leak occurrences in the past, and the respondent’s previous cost of plumbing repairs and/or replacement. These socioeconomic characteristics had an impact on the decision with regards to choosing to either re-plumb the house or purchase an insurance policy. For example, higher home value was found to increase the probability that respondents would choose a prevention option (plumbing upgrade or insurance) against costs of corrosion damage in order to avoid loss on their property investment. On the other hand, if households have accrued plumbing material repairs or replacement costs in the past, they might be less prone to either re-plumb their houses or purchase an insurance policy against corrosion, as they have already invested in corrosion prevention by either repairing or replacing their drinking water infrastructure. House age positively impacted the decision to upgrade plumbing material as the house age approximates the age of the plumbing material. For example, Lee et al. (forthcoming) found that plumbing material should be replaced 22 years after pipe installation. Furthermore, Kleczyk and Bosch (2008) found that the house age as well as the past experiences with pinhole leaks are the drivers of either pinhole leak corrosion and/or costs related to pinhole leak damage. This finding implies that at some point households should preventively re-plumb their houses to prevent future leaks and future financial burden (Kleczyk & Bosch, 2008).

The probability of choosing a plumbing upgrade or insurance purchase was rather consistent between different model specifications. The compensated variation analysis revealed that households would have to be financially compensated in order to switch from the current plumbing material to either upgrade the plumbing system or purchase an insurance policy. For example, to induce a household to switch from the current plumbing material to home re-plumbing, the homeowners would have to be compensated between $5,460 and $5,529. The compensation is even higher when switching from the current plumbing material to purchasing an insurance policy against corrosion. The compensation varies from $10,500 to $13,771. The negative compensating variation values imply that the attributes of the alternatives were viewed as utility reducers, requiring compensation to keep the household on the same level of utility.

The test for the independence of irrelevant alternatives assumption revealed justification for employment of the MNL model specification. Finding the confirmation for this assumption implies that inclusion of new alternatives into the choice set of prevention has no effect on the relative probabilities of the choices. As a result, households' preferences for prevention against corrosion are transitive and allow for finding the most preferred (optimal) approach of corrosion prevention (utility-maximizing behavior). When taking into account that the population residing in the Southeastern Community is a very homogenous group in terms of their geographic, demographic and socioeconomic characteristics and that these residents share similar experiences with corrosion and pinhole leak problems, this result is not surprising.

The equality of coefficients estimates between the equation for upgrading the plumbing material vs. staying with the current material and the equation of purchasing an insurance policy against corrosion vs. staying with the current plumbing material was also tested. The test results revealed statistically equal coefficients for risk of corrosion as well as cost of corrosion prevention option (either upgrading plumbing material or purchasing an insurance policy) between the two equations. The equality implies the same response function for independent variables between the models regardless of the chosen prevention option. The homogeneity of the Southeastern Community residents is manifested in the form of preferences for corrosion prevention as well as the responsiveness to the external factors impacting the corrosion prevention decisions.

The CA of the follow-up survey data of residents in the Southeastern Community revealed that among three hypothetical plumbing materials (A, B, and C), the households preferred Material A followed by Material B. Material C was the least preferred material in the set. This result was derived based on each of the respondents' preference evaluation of the different plumbing material groupings. The preference ranking of the materials was the same across both Ordered Logit model specifications. Furthermore, the CA results were in agreement with the survey baseline method, which ranked Material A as the most preferred and Material C as the least preferred. The baseline ranking of plumbing materials was obtained from households' comparisons of all three plumbing materials at the same time.

The plumbing material attributes that were important in the decision-making process included: corrosion risk and time length of plumbing material installation. In both cases the attribute level rankings were in agreement with the transitivity assumption of preferences and the lower corrosion risk attribute level as well as shorter amount of time required for plumbing material installation was more preferred to the more corrosion risk prone and longer installation period attribute levels.

Only one socioeconomic variable had a statistically significant impact on the chosen plumbing material: cost of plumbing material repairs and replacement incurred by the respondent. This variable was statistically influential when interacted with corrosion attribute levels. Although it did not change the preferences for plumbing materials, the variable skewed the preference valuations favorably towards plumbing materials described by 'corrosion proof' attribute level. This finding implies that the more each household had previously spent on repairs associated with plumbing material failures, the more they preferred a plumbing material with lower corrosion level to avoid future expenditures on drinking water system repairs.

There are several implications for further research that would improve the analysis and impact the precision of the final results for the CA of households' preference for corrosion prevention as well as plumbing materials. Future CA studies eliciting preferences for preventive measures should include higher risk values of corrosion to test the responsiveness of households to higher corrosion levels in considering corrosion preventive options as protection for their homes. For example, the risks of corrosion values presented in the first Southeastern Community survey were too small to induce a choice of corrosion prevention option by the survey respondents as they varied from no corrosion risk to only 6% chance of corrosion. With such low levels of corrosion, 69% of respondents did not choose a prevention option, although the pinhole leak occurrence rate for the Southeastern Community was 25%.

In the case of the CA for preferred hypothetical plumbing materials, there are several improvements proposed to the survey design. The information set of plumbing material attributes might have not been the most complete and objective description of the pipe characteristics. Households with copper plumbing materials installed in their houses were more likely to identify Material C as copper and therefore might have evaluated it based on their experiences and not based on the comparison with other plumbing materials. This finding, however, is not unexpected, as part of the research question was to examine the impact of previous experiences with plumbing material failures on household decisions for corrosion prevention and plumbing material choices. Furthermore, in his AHP study, Lee (2008) noticed that some of the householders provided a high degree of preference for a specific plumbing material in the survey, but in reality installed other types in their homes (Lee, 2008). As a result, there is a mismatch between the stated preferences derived based on the homeowners' survey and the actual behavior exhibited by the households.

As some of the survey respondents (15) had difficulty following and understanding the survey questions, the format of CA questions should also be rethought and simplified. Working with either an educator or health official would allow gearing the reading level of the survey to match the capabilities of the respondents. Finally, as many of Southeastern Community residents were older than the national age average (72 years old vs. 36 years old (U.S. Census Bureau, 2008)) and had problems with following and understanding the survey, surveying a more nationwide representative sample of households could eliminate the issues resulted from misinterpretation and misunderstanding of the questionnaire questions. These improvements would result in increased precision of the outcomes and therefore more appropriate policy implications concerning the drinking water infrastructure.

The above survey results inform policy makers, utility managers, and home plumbing systems producers on the homeowners' preferences for plumbing materials and the trade-offs between the risk of corrosion and cost of a leak-free environment based on their experiences with pipe failures in the past. The cost of alternative preventive measures, corrosion risk, and convenience of plumbing material installation drive the decisions of homeowners regarding their plumbing system. As a result, policy makers should take into consideration the implications of new federal and state regulations on the interactions between drinking water and drinking water plumbing. Furthermore, their regulations and standards should accurately test the different types of plumbing materials used in the drinking water infrastructure as well as their chemical and physical interactions with chemicals used to treat drinking water.

For example, Edwards et al. (2004) suggested that removal of natural organic matter mandated by tighter EPA drinking water standards contributed to the pinhole leak problem in combination with other factors, including faulty installation, since natural organic matter is an inhibitor to the corrosion-inducing chemical reactions. To deal with this problem, Bosch et al. (2006) found that almost 60% of water utilities added corrosion inhibitors such as phosphate to water treatment. The inhibitors were added to protect water service lines, to comply with the lead and copper rule proposed by EPA, and to give protection to residential customers. Similarly, after adding phosphate to the water treatment process by utility companies who distribute water to the Southeastern Community as suggested by Virginia Tech researchers, Dr. Edwards and Dr. Scardina, the Southeastern Community reported a decrease in the number of pinhole leak reports (Scardina & Edwards, 2007).

Furthermore, the cost associated with employment of different prevention options as well as the convenience of installation has an impact on households' decisions concerning choosing a plumbing material for their houses. As a result, service providers (i.e. plumbers and material manufactures) should be sensitive to households' financial constraints and convenience of plumbing installation for homeowners. For example, 33%

of Southern Community respondents with pinhole leaks spent at least $500 repairing damaged plumbing material while more than 75% of survey participants with pinhole leaks experience at least moderate level of stress. In their Maryland study of pinhole leak corrosion, Kleczyk and Bosch (2008) estimated the total cost[32] of fixing damage related to pinhole leaks to range from roughly $1,300 to more than $18,000. As a result, when plumbing services are expensive, the service providers should concentrate on installing plumbing materials that are convenient to install and present a low failure rate to minimize future financial spent on plumbing material repairs.

Finally, water professionals and policy makers should work on public policy that would address public preferences for drinking water infrastructure. Results of this Southeastern Community analysis results can provide information to policy experts and water utility managers who are dealing with extensive corrosion problems in their areas. Information will fill the gaps of knowledge about corrosion occurrences, the financial impact of plumbing material repairs on households, and households' preferences for drinking water infrastructure as well as the ability of householders to pay for different corrosion prevention options.

### 3.10 Acknowledgements

This article is written in the memory of Dr. Loganathan who died tragically on April 16, 2007 during the Virginia Tech campus shooting. The author also would like to acknowledge the financial support provided by the National Science Foundation under the grant DMI-0329474 and the American Water Works Association Research Foundation under the project #3015. Finally, the author would like to thank Alison Miller and Patricia Murphy who have helped in the article editing process. The views expressed in this report are those of the author and not of the National Science Foundation nor of the American Water Works Association Research Foundation.

[32] Total cost of repairing pinhole leak damage includes the financial and time costs.

### 3.11 References


Aaron. G. (2005). *Ordered Logit Model*. Available online at: http://www.uoregon.edu/~arrong/teaching/G4075_Outline/node27.html. Accessed: June 2008.

Adamowicz, W., Boxall P., Williams, M., & Louvier, J. (1998). Stated Preference Approaches for Measuring Passive Use Values: Choice Experiments and Contingent Valuation. *American Journal of Agricultural Economics 80:* 64-75.

Agency for Toxic Substances and Disease Registry (ATSDR). (2004). *Toxicological Profile for Copper*. Available Online at: http://www.atsdr.cdc.gov/toxprofiles/tp132.html. Accessed: June 2008.

Baum, C. F. (2006). *An Introduction to Modern Econometrics Using Stata.* Stata Press.

Beck, N. 2004. Limited Dependent Variables. Department of Politics. New York University. Available online at: http://weber.ucsd.edu/~tkousser/Beck%20Notes/ldvs.pdf. Accessed: June 2008.

Boyle, K. (2007). Personal communication.

Boyle, K. J., & Bouchard, R. (2003). Water Quality Effects on Property Values in Northern New England. *Lakelines 23*(3), 24-27.

Bosch, D., Kleczyk, E., Lee, J., & Tanellari, E. (2008). *Southeastern Community SurveyReport*. Department of Agricultural and Applied Economics. Virginia Tech. Blacksburg, VA.

Bosch, D., Kleczyk, E., Lee, J., & Tanellari, E. (in press). Home Drinking Water Assessment. *Drinking Water and Backflow Prevention Journal*.

Boxall, P. C., Adamowicz, W. L., Swait J., Williams, M., & Louvier, J. (1996). A Comparison of Stated Preference Methods for Environmental Valuation. *Ecological Economics 18*, 243-253.

Breedlove, J. (1999). RL30242: Natural Resources: Assessing Nonmarket Values Through Contingent Valuation. *CRS Report for Congress*. Available online at: http://www.ncseonline.org/nle/crsreports/natural/nrgen-24.cfm#Validity%20and%20Bias. Accessed: June 2008.

Bradlow, E., Hu, Y., & Ho, T. H. (2004). A Learning-Based Model for Imputing Missing Levels in Partial Conjoint Profiles. *Journal of Marketing Research 41*, 369-381.

Briggs, D. C. (2004). Causal Inferences and the Heckman Model. *Journal of Educational and Behavioral Statistics 29*(4), 397 - 420.

Brill, E.J. (2004). Comparing Social Groups: Wald Statistics for Testing Equality Among Multiple Logit Models. *International Journal of Comparative Sociology*. Published in Goliath Business Knowledge in Demand. Available online at: http://goliath.ecnext.com. Accessed: October 2008.

Burgess, L., & Street, D. (2005). Optimal Designs for Choice Experiments with Asymmetric Attributes. *Journal of Statistical Planning and Inference 134,* 288-301.

Burgess, L., & Street, D. (2003). Optimal Designs for $2^K$ Choice Experiments. *Communications in Statistics: Theory and Methods 32*(11), 2185-2206.

Cameron, T., A., & James, M., D. (1987). Efficient Estimation Methods for 'Close-Ended' Contingent Valuation Survey. *Review of Economics and Statistics 69*(2), 269-276.

Champ, P., Boyle, K., & Brown, T. (2003). *A Primer on Nonmarket Valuation*. Boston: Kluwer Academic Publishers.

Compu – Quate (2008). Available online at: www.compu-quote.com. Accessed: October 2008.

Dietrich, A., T., Heim, H., Johnson, Y., Zahng, M., Edwards, G. V., Loganathan, et al., (July 2006). *Plumbing Materials: Costs, Impacts on Drinking Water Quality, and Consumer Willingness to Pay*. Proceedings of 2006 NSF Design, Service, and Manufacturing Grantees Conference. St. Louis, Missouri. Available online at http://www.dmigranteeconference.org/paper.htm. Accessed: October 2008.

Dillman, D. A. (1978). *Mail and Telephone Surveys*, New York: John Wiley & Sons.

Ding, M., Agarwal, R., & Liechty, J. (2005). Incentive-Aligned Conjoint Analysis. *Journal of Marketing Research 4,* 67-82.

Duke, J., & Aull-Hyde, R. (2002). Identifying Public Preferences for Land Preservation Using the Analytic Hierarchy Process. *Ecological Economics 42,* 131-145.

Duke, J. M., Ilvento, T. W., & Hyde, R. A. (2002). *Public Support for Land Preservation: Measuring Relative Preferences in Delaware*. Department of Food and Resource Economics. University of Delaware. FREC Research Report No. 02-01.

Durand, D., & Dietrich, A. (2007). Contributions of Silane Cross-Linked PEX Pipe to Chemical/Solvent Odors in Drinking Water. *Water Science &and Technology 55*(5), 153–160.

Edner, P. (2005). *Applied Categorical and Nonnormal Data Analysis: Ordered Logit and Probit Models. Eduction 231C.* Available online at: http://www.gseis.ucla.edu/courses/ed231c/notes2/ologit.html. Accessed: October 2008.

*Engineering Statistics Handbook*. (2007). Available online at: http://museum.nist.gov/exhibits/timeline/item.cfm?itemId=77. Accessed: October 2008.

Griffiths, W. E. Hill, R.C., & Judge, G. G. (1993). *Learning and Practicing Econometrics*. New York: John Wiley and Sons.

Environmental Protection Agency (EPA) (2006). *Groundwater and Drinking Water Consumer Fact Sheet on Copper*. Available Online at http://www.epa.gov/safewater/contaminants/dw_contamfs/copper.html. Accessed: October 2008.

Farooqi, O., & Lee, J. (2005). *Plumber Telephone Surveys.* Virginia Tech. Blacksburg, VA.

Fry, T. R., & Harris, M. N. (1998). Testing for Independence of Irrelevant Alternatives: Some Empirical Results. *Sociological Methods and Research 26*(3), 401-421.

Goodman, J. A. & Jaworski, E. R. *Clue Reports*. Frascona, Joiner, Goodman and Greenstein, P.C. Available Online at: http://www.frascona.com/resource/jag403clue.htm. Accessed: September 2008.

Gullickson, A. (2005). Lecture Outline for Sociology G4075: *Introduction to Social Data Analysis II.* Available online at: http://www.columbia.edu/~ag2319/teaching/G4075_Outline/G4075_Outline.html. Accessed: May 2008.

Hausman, J., & McFadden, D. (1984). Specification Tests for the Multinomial Logit Model. *Econometrica 52***,** 1219–1240.

Helm, R., Scholl, A., Manthey, L., & Steiner, M. (2004). Measuring Customer Preferences in New Product Development: Comparing Compositional and Decompositional Methods. *International Journal of Product Development 1*(1),12-29.

Hanemann, W. M. (1984). Welfare Evaluation in Contingent Valuation Experiments with Discrete Responses. *American Journal of Agricultural Economics 66:* 332 – 341.

Hanemann, W. M., & Kanninen, B. (1998). *The Statistical Analysis of Discrete Response CV Data*. California Agricultural Experiment Station. Giannini Foundation of Agricultural Economics. Department of Agricultural and Resource Economics and Policy. Division of Agricultural and Natural Resources. University of California at Berkley.

Heim, T., & Dietrich, A. (2007). Sensory Aspects and Water Quality Impacts of Chlorinated and Chloraminated Drinking Water in Contact with HDPE and CPVC Pipe. *Water Research 55*(5), 757 –764.

Heiman, A., Just, D. R., and Zilberman, D. 2000. The Role of Socioeconomic Factors and Lifestyle Variables in Attribute and Demand Genetically Modified Foods. *Journal of Agribusiness 18(3):*149 – 260.

Huber, J., & Zwerina, K. (1996). On the Importance of Utility Balance in Efficient Designs. *Journal of Marketing Research 33*, 307-317.

Jahle, G. A., & Reny, P. J. (2000). *Advanced Microeconomic Theory*. Boston: Addison-Wesley.

Kanninen, B. J. (1993). Design of Sequential Experiments for Contingent Valuation Studies. *Journal of Environmental Economics and Management 25,* S1 – S11.

Kleczyk, E., & D. Bosch. (in press). Incidence and Costs of Home Plumbing Corrosion. *Journal of American Water Works Association*.

Kleczyk, E. J., Tanellari, E., & Bosch, D. J. (November 2006). *Corrosion in Home Drinking Water Infrastructure: Assessment of Causal Factors, Costs, and Willingness to Pay*. 2006 Water Quality Technology Conference. American Water Works Association. Denver, Colorado. Available Online at: http://www.techstreet.com/cgi-bin/detail?product_id=1320028. Accessed: October 2008.

Kuhfeld, W. F. (2005). Marketing Research Methods in SAS: Experimental Design, Choice, Conjoint, and Graphical Techniques. *SAS 9.1 Edition, TS-722*. Cary, NC: SAS Institute Inc.

Kuhfeld, W. F., Tobias, R. D., & Garratt, M. (1994). Efficient Experimental Designs with Marketing Research Applications. *Journal of Marketing Research 31*: 545-557.

Kuriyama, K., Takeuchi, K., & Washida, T. (1999). *Repeated Dichotomous Choice Contingent Valuation for Multi-Attribute Environmental Resources: A Comparison Between Contingent Valuation and Choice Experiment Format.* Working Papers. School of Political Science and Economics. Waseda University. Tokyo, Japan.

Lee, J. (2008). *Two Issues in Premier Plumbing Contaminants Intrusion at Service Line and Choosing Alternative Plumbing Material*. Doctoral Dissertation. Virginia Polytechnic Institute and State University.

Lee, J., Loganathan, G. V., Bosch, D., Dwyer, S., & Kleczyk, E. (October 2005). *Preference Analysis of Home Plumbing Material*. Virginia Water Resources Research Center, National Water Research Symposium: Balancing water law and science, The Inn at Virginia Tech and Skelton Conference Center, Virginia Tech, Blacksburg, Virginia.

Lee, J., Loganathan, G.V., Bosch, D., Dwyer S., & Kleczyk, E. (August 2006). *Consumer Concerns and Preference Tradeoffs in Home Plumbing Systems.* 8th Annual International Symposium on Water Distribution Systems Analysis. Cincinnati, Ohio.

Lee, J., Kleczyk, E., Bosch, D., Tanellari, E., Dwyer, S., & Dietrich, A. (Accepted for Publication on October 16, 2008). Case Study: Preference Trade-offs Towards Home Plumbing Attributes and Materials *Water Resource Planning Management Journal.* Special Edition in Memory of Dr. G.V. Loganathan.

Longcore, R. (2008). What is CLUE Report and Why Should You Care? *Ezine Articles.* Available Online at: http://ezinearticles.com. Accessed: October 2008.

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

Marshutz, S. (2001). Hooked on Copper. *Reeves Journal*. Available Online at: http://www.reevesjournal.com/CDA/ArticleInformation/features/Features_Index/1,3816,27-820,00.html. Accessed: October 2008.

Myers, J. L., & A. D. Well. (2003). *Research Design and Statistical Analysis*. Second Edition. Lawrence Erlbaum

McFadden, D., Train, K., & Tye, W. B. (1981). An Application of Diagnostic Tests for the Independence from Irrelevant Alternatives Property of the Multinomial Logit Model. *Transportation Research Board Record 637*: 39-46.

Mister Renovator. (2008). Available online at: www.misterrenovator.com. Accessed: October 2008.

Mitchell, R. C., & Carson, R. T. (1989). *Using Surveys to Value Public Goods: The Contingent Valuation Method*. Washington D.C.: Resource for the Future.

National Science Foundation (NSF). (2008). *NSF Standard Accepts New Stainless Steel Materials in Drinking Water Applications*. Available Online at: http://www.nsf.org/business/newsroom/press_release.asp?p_id=12241. Accessed: October 2008.

Parson, G. R., & Massey, D. M. (2003). ARUM Model of Beach Recreation. In N. Hanley, W. D. Shaw, R. E. Wright, & J. C. Coleman (Eds.), The New Economics of Outdoor Recreation. Cheltenham, UK: Edward Elgar.

*Frustrated by Pinhole Leaks in Their Copper Plumbing, Homeowners Find Relieve with PEX.* (July 2007). *Plumbing and Mechanical Magazine 25*(5), 19.

*PRNews Wire*. (2004). Available Online at: http://www.prnewswire.com/cgibin/stories.pl?ACCT=109&andSTORY=/www/story/11-18-2004/0002464315ENDDATE. Accessed: October 2008.

Rao, R. S, Kumar, G.C., Prakasham, R.S. & Hobbs, P.J. (2008). The Taguchi Methodology as a Statistical Tool for Biotechnological Applications: A Critical Appraisal. *Biotechnology Journal 3,* 510 – 523.

Rao, V. R. (March 2008). *Conjoint Analysis*. Web-Seminar Sponsored by the Agribusiness Management and Economics Section of the AAEA. Cornell University. Johnson Graduate School of Management.

Rao, V. R. (2007). *Development in Conjoint Analysis*. Working Paper. Cornell University. Johnson Graduate School of Management. Forthcoming In B. Wierenga (Ed.), Marketing Decision Making.

Roberge**,** P. R. (2000). Searching the Web for Corrosion Intelligence. *Corrosion Reviews 18*(1), 23-40.

Robinson, J. (2008). The Average Cost of Texas Homeowners Insurance and How You Can Find Cheap Texas Home Insurance Rates. *Ezine Articles.* Available online at: http://ezinearticles.com. Accessed: October 2008.

Saaty, T. L. (1980). *The Analytic Hierarchy Process*. New York: McGraw-Hill.

Saaty, T. L. (1990). How to Make a Decision: The Analytic Hierarchy Process. *European Journal of Operational Research 48*, 9-26.

Scholl, A., Manthey, L.. Helm, R., & Steiner. M. (2005). Solving Multiattribute Design problems with Analytic Hierarchy Process and Conjoint Analysis: An Empirical Comparison. *European Journal of Operational Research 164*, 760 – 777.

Scardina, P., Edwards, M., Bosch, D. J., Loganathan, G. V., & Dwyer, S. K. (2007). *Non-Uniform Corrosion in Copper Piping—Assessment.* Final Project Completion Report to American Water Works Association Research Foundation. Blacksburg, Virginia: Virginia Tech.

Scardina, P., & Edwards, M. (2007). *Preliminary Investigation of Copper Pipe Failures*. Report submitted to the Southeastern Community.

*SPSS Annotated Output: Multinomial Logistic Regression* (2005) Department of Statistics at University of California – Los Angeles. Available online at: http://www.ats.ucla.edu/stat/spps/output/mlogit.htm. Accessed: October 2008.

Srinivasan, V., & Park, C. S. (1997). Designing Conjoint Choice Experiments Using Managers' Believes. *Journal of Marketing Research 38, 340*-444.

StatSoft, Inc. (2008). *Experimental Design: Industrial DOE*. Available online at: http://www.statsoft.com/textbook/stexdes.html. Accessed: October 2008.

Stevens, T. H., Belkner, R., Dennis, D., Kittredge, D., & Willis, C. (2000). Comparison of Contingent Valuation and Conjoint Analysis in Ecosystem Management. *Ecological Economics 32*, 63-74.

The Plumber. (2008). Available online at: www.theplumber.com. Accessed: October 2008.

Toolbase News (2008). Available Online at: http://www.toolbase.org/pdf/techinv/homerunplumbingsystems_techspec.pdf. Accessed: October 2008.

Toubia, O & Hauser, J. R. (2006). *On Managerial Efficient Experimental Designs*. Columbia Business School. Columbia University. New York, N.Y.

Train, K. (2003). *Discrete Choice Methods with Simulation.* Cambridge University Press. Available online at: http://elsa.berkeley.edu/books. Accessed: October 2008.

U.S. Census Bureau. (2000). *Population Finder.* Available Online at: http://www.census.gov/. Accessed: October 2008.

Van Trees, H. L. (1968). *Detection, Estimation, and Modulation Theory, Part I*. New York: Wiley.

Varian, H. R. (1992). *Microeconomics Analysis*. (3rd ed.). New York: W.W Norton and Company.

Welsh, M. P., & Bishop, R.C. (1991). Multiple Bounded Discrete Choice Models. In W-133 Benefits and Cost Transfer in Natural Resource Planning. Western Regional Research Publication complied by J. Bergstorm, Department of Agricultural and Applied Economics, University of Georgia.

Wedel, M., & Kamakura, W. A. (2000). *Market Segmentation: Conceptual and Methodological Foundation* (2nd ed.). Boston: Kulwer Academic Publisher.

Wertenbroch, K. & Skiera, B. (2002). Measuring Consumers' Willingness to Pay at the Point of Purchase. *Journal of Marketing Research 39*, 228-241.

Wierenga, B. (2008). *Handbook of Marketing Decision Models.* Springer Publisher.

Whitehead, J.C., Pattanayak, S. K., Van Houtven, G. L., & Gelso, B. R. (2005). Combining Revealed and Stated Preference Data to Estimate the Nonmarket Value of Ecological Services: An Assessment of the State of the Science. *Working Paper 05-19*.

Available online at: http://econ.appstate.edu/RePEc/pdf/wp0519.pdf. Accessed: October 2008.

Willis-Walton, S. (2006). *Plumbers Survey: Examination of Pinhole Leaks in the United States and Targeted Geographic Areas*. Virginia Tech Center for Survey Research.

Woodson, R.D. (1999). *Plumber's Standard Handbook.* New York: McGraw-Hill.

Zahedi, F. (1987). A Utility Approach to the Analytic Hierarchy Process. *Mathematical Modeling 9,* 387–395.

## 3.12 Appendix

### Appendix A: Additional Analysis

**Conjoint Analysis with Cost of Corrosion Prevention Divided by Household Income**

This MNL model specification includes socioeconomic variables represented by the cost of corrosion prevention divided by the household income. As mentioned earlier, socioeconomic variables are constant over the alternatives for each household, so the difference between utilities from choosing different prevention options is equal to zero. There are two ways in which the socioeconomic variables can enter the MNL model and influence choice of prevention options: (1) through interaction with alternative attributes and (2) through dividing the cost variable by income (Train, 2003). The interaction of socioeconomic variables with other explanatory variables was discussed in section 3.8.1.1.

In order to evaluate the results of MNL model accounting for socioeconomic variables, the model significance should be assessed first. The presence of a relationship between the dependent variable and the combination of independent variables is assessed based on the chi-square and log-likelihood statistics of the final model. Based on the analysis, the chi-square statistic is 3.867 with the significance level of p-value of 0.144 and a log-likelihood value of 913.788. The existence of a relationship between the independent variables and the dependent variable is not supported by the results. Pseudo R-square is 0.002, which implies that there is a 0.2% increase in the log-likelihood function above the function at zero and the estimated parameters improve the model's fit only slightly. In addition, there is no overall significant relationship between the independent variable: risk of corrosion and the cost associated with plumbing material upgrade and insurance purchase/household income and the dependent variable of corrosion prevention options (Table 3. A1).

Table 3.A1: Likelihood Ratio Tests for MNL Model without Socioeconomic Characteristics (Cost/Income in a Continuous Variable Format)

| | Model Fitting Criteria | Likelihood Ratio Test | |
|---|---|---|---|
| Effect | -2 Log Likelihood of Reduced Model | Chi-Square | Significance |
| Risk of Corrosion | 916.282 | 2.494 | 0.287 |
| Cost of Home Re-plumbing and Insurance Purchase /Household Income | 917.752 | 3.964 | 0.138 |

As presented in Table 3.A2, there are no statistically significant results for any of the preventive options. As a result, the analysis of marginal utility of income and compensating variation of the MNL model specification is not performed. Based purely on the statistically significant intercepts, the probability of upgrading plumbing material is 14%, the probability of purchasing an insurance policy is 14% and the probability of staying with the current plumbing material system without purchasing insurance is 72% (Table 3.A2).

Table 3.A2: Regression Estimates for MNL Model without Socioeconomic Characteristics (Cost/Income in a Continuous Variable Format)

| Variable Name | Beta Coefficient [a] | Exp(B) [b] | Probability Estimate [c] | Wald Stat [d] | Std. Error [e] | P-Value [d] |
|---|---|---|---|---|---|---|
| Upgrading Plumbing Material [f] | | | | | | |
| Intercept | -1.644 | | 0.140 | 551.015 | 0.070 | 0.000 |
| Risk of Corrosion | 2.372 | 10.718 | 0.172 | 0.942 | 2.443 | 0.332 |
| Cost of Home Re-plumbing and Insurance Purchase /Household Income | $9.843E^{-11}$ | 1.000 | 0.137 | 0.596 | 0.000 | 0.441 |
| Keeping Current Plumbing Material and Purchasing Insurance [f] | | | | | | |
| Intercept | -1.668 | | 0.140 | 526.388 | | |
| Risk of Corrosion | 2.796 | 16.384 | 0.174 | 1.158 | 2.598 | 0.282 |
| Cost of Home Re-plumbing and Insurance Purchase /Household Income | $-2.493E^{-10}$ | 1.000 | 0.136 | 2.640 | 0.000 | 0.114 |

[a] Beta coefficients represent the parameter estimates from the MNL model estimation.

[b] Odd ratio (Exp(B)) represents the impact of increase in the independent variable on the likelihood of choosing a given option (insurance or plumbing upgrade). For example, an odds ratio of 1 means a one unit increase in independent variable implies an equal likelihood of choosing or not choosing the prevention option.

[c] Probability estimate for continuous variables is computed at the mean value of the independent variable (e.g. the cost of insurance or plumbing upgrade) presented to all respondents. Probability estimate for a dummy variable is the increase in probability of choosing the prevention option for respondents in the '1' category.

[d] Wald statistics and p-value represent the significance level.

[e] Standard error represents the variation of the coefficient estimate.

[f] Independent variables take the form of continuous variables. Risk of corrosion represents the risk of corrosion occurrence when using this corrosion prevention method and the ratio of cost of plumbing material upgrade and insurance purchase over the household income represents the cost of corrosion prevention adjusted for the annual income to measure the relative impact of cost on corrosion prevention. The number of observations included in the model is 986. Model Statistics: Log-likelihood value is 913.788; Nagelkerke's R-square is 0.002

## Appendix B: Southeastern Community First Survey

National Home Drinking Water Assessment
Attn: Darrell Bosch
Department of Agricultural and Applied Economics (0401)
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

August 2007

Dear Tellico Village Resident:

Funded by the National Science Foundation, Virginia Tech researchers are investigating pinhole leak damages in home drinking water plumbing in the U.S. Cooperating with Tellico Village Public Works staff, Loudon and Monroe County water utilities, and Tellico Village Copper Pipe Pinhole Leak Committee (Tellico Village Committee), we are gaining a better understanding of pinhole leak problems. Tellico Village Committee encouraged us to survey residents about their concerns and experiences regarding their drinking water and home plumbing.

With this survey, we seek to understand the costs of pinhole leaks. Also, we wish to understand the decision-making processes residents use to cope with their particular circumstances. Please complete the survey even if you have never had problems with your plumbing. Your responses will assist in providing a more accurate picture of drinking water plumbing in Tellico Village. Survey results will help water utilities and policy experts plan strategies to deal with home plumbing corrosion.

Be assured that your responses will be treated confidentially. Individual responses will not be reported. All responses from residents will be compiled, analyzed, and reported together. Questions that refer to specific repair choices or insurance decisions are provided as examples only and are not connected to any group or business now or in the future.

The survey will take about fifteen minutes to complete. ***After completing the survey, please return it in the enclosed, postage-paid envelope or send it to the address above.*** If you would like to receive a summary of results, please include your contact information on the enclosed card and return it with the survey or mail it separately. If you have questions, contact me at 540-231-5265 or by email at *bosch@vt.edu*.

Thank you very much for your help with this important survey.

Sincerely,

Darrell Bosch

Darrell Bosch, Professor

1. Which Tellico Village neighborhood do you live in? (CIRCLE ONE NUMBER)
    1. Chatuga
    2. Chota
    3. Coyatee
    4. Kahite
    5. Mialaquo
    6. Tanasi
    7. Tommotley
    8. Toqua
    9. Don't know

2. How satisfied are you with the quality of tap water in your house (CIRCLE ONE NUMBER)?
    1. Very satisfied
    2. Somewhat satisfied
    3. Somewhat dissatisfied (Specify) ________________________________
    ________________________________________________
    4. Not at all satisfied (Specify) ________________________________
    ________________________________________________

3. Have you had any problems with the following attributes of your tap water? (CIRCLE ALL THAT APPLY)
    1. Taste
    2. Odor
    3. Color
    4. Other (Specify) ________________________________________
    5. No problems

4. How aware are you of pinhole leak problems (small holes in home drinking water pipes) in the area where you live? (CIRCLE ONE NUMBER)
   1. Very aware and have had pinhole leaks
   2. Very aware but have not had pinhole leaks
   3. Have heard of problem and know some details
   4. Have heard of problem but do not know any details
   5. Not at all aware of this problem

5. How did you hear about the pinhole leak problem? (CIRCLE ALL THAT APPLY)
   1. Own experience
   2. Neighbor or friend
   3. Media (Channel 3, The Village Connection)
   4. Copper Pipe Pinhole Leak Committee of Property Owner's Association
   5. Other (Specify) ________________________________________________

6. What type of drinking water pipes do you have in your home? (CIRCLE ALL THAT APPLY)
   1. Copper
   2. CPVC (Plastic pipe)
   3. PEX (Polyethylene plastic tubing)
   4. Other (Specify) ________________________________________________
   5. Don't know

7. Have you ever had to replace or repair any of the drinking water pipes in your house? (CIRCLE ONE NUMBER)
   1. Yes (please specify why repair/replacement was made) ____________________
   2. No
   3. Don't know

8. If yes, when were the pipes replaced or new ones added? (CIRCLE ALL THAT APPLY)
   1. Since 2004
   2. 2000 to 2003
   3. 1995 to 1999
   4. 1990 to 1994
   5. Before 1990
   6. Don't know

9. Have you ever had a pinhole leak in the hot or cold drinking water pipes in your current home? (CIRCLE ONE NUMBER)
   1. Yes
   2. No
   3. Don't know

10. Have you had other kinds of failure in your drinking water pipes?
   1. Yes (Specify) ______________________________________________
      ______________________________________________
   2. No

*If you have never had pinhole leaks, please go to question 25. If you have had pinhole leaks in your current home, please continue to the next question.*

11. What material was the pipe with the leaks made of? (CIRCLE ALL THAT APPLY)
   1. Copper
   2. CPVC
   3. PEX
   4. Brass (Specify: leak around valve stem, brass cracked, or pinhole leak on the brass) ______________________________________________
      ______________________________________________
   5. Other (Specify) ______________________________________________
      ______________________________________________
   6. Don't know

12. In approximately what year did you **first** have a pinhole leak in your drinking water pipes? (CIRCLE ONE NUMBER)
    1. Since 2004
    2. 2000-2003
    3. 1995 to 1999
    4. 1990 to 1994
    5. Before 1990
    6. Don't know

13. How many pinhole leak incidents have you had in your home? (CIRCLE ONE NUMBER)
    1. 1 – 2
    2. 3 – 4
    3. 5 – 6
    4. 7 – 10
    5. More than 10 (Specify) ___________
    6. Don't know

14. Over what period of time did these leaks occur? (CIRCLE ONE NUMBER)
    1. Less than one year
    2. 1 to 3 years
    3. 3 to 6 years
    4. More than 6 years
    5. Don't know

15. Were the leaks in a horizontal, vertical pipe, or pipe bend? (CIRCLE ALL THAT APPLY)
    1. Horizontal pipes
    2. Vertical pipes
    3. Pipe bend
    4. Don't know

16. Did the pinhole leaks occur underground, in your basement, on the first floor or higher? (CIRCLE ALL THAT APPLY)
    1. Under slab or underground
    2. Crawl space
    3. Unfinished basement
    4. Finished or walkout basement
    5. First floor
    6. Second floor or higher
    7. Don't know

17. Did the leaks occur in cold or hot water pipes? (CIRCLE ONE NUMBER)
    1. Cold water pipes
    2. Hot water pipes
    3. Both cold and hot water pipes
    4. Don't know

18. Who fixed your leak problems? (CIRCLE ALL THAT APPLY)
    1. Fixed them myself
    2. Household member (other than myself)
    3. Relative, friend, or neighbor
    4. Professional plumber or contractor
    5. Other (Specify) ________________________________________
    6. Don't know

19. How did you fix the leaks? (CIRCLE ALL THAT APPLY)
    1. Clamp over leak
    2. Replaced only leaking section of pipe (please specify material used) _________
       ________________________________________
    3. Applied coating or epoxy to all plumbing
    4. Entirely replaced or replumbed (please specify material used for replumbing)
       ________________________________________
    5. Other (Specify) ________________________________________
       ________________________________________
    6. Don't know

20. What was the approximate cost to you, your landlord, and/or your insurance company to repair the leaks (plumbing repair expenses only)? (CIRCLE ONE NUMBER)
    1. Less than $100
    2. $100 – $500
    3. $501 – $1,000
    4. $1,001 – $3,000
    5. $3,001 – $5,000
    6. $5,001 – $10,000
    7. $10,001 – $20,000
    8. More than $20,000 (Please specify) ____________________
    9. Don't know

21. About how much money was spent by you, your landlord, and/or your insurance company to make repairs *for property damage to your home* resulting from leaks? (CIRCLE ONE NUMBER)
    1. Less than $100
    2. $100 – $500
    3. $501 – $1,000
    4. $1,001 – $3,000
    5. $3,001 – $5,000
    6. $5,001 – $10,000
    7. More than $10,000 (Please specify) ____________________
    8. Don't know

22. Did the leaks cause the need for you to replace or repair anything else in your home or to lose valuable family heirlooms? (CIRCLE ONE NUMBER)
    1. Yes (please describe repairs/replacements/losses) ________________________
       ________________________________________________________________
    2. Damage or losses have not been replaced or repaired yet
    3. No damage or losses
    4. Don't know

23. About how many hours in total would you estimate you have spent dealing with pinhole leak problems in your home? (CIRCLE ONE NUMBER)
    1. Less than 10
    2. 11 – 20
    3. 21 – 40
    4. 41 – 80
    5. More than 80 (Please specify) ____________
    6. Don't know

24. Would you say that your overall experience with pinhole leaks was (CIRCLE ONE NUMBER)
    1. Very stressful
    2. Somewhat stressful
    3. Not very stressful
    4. Not at all stressful
    5. Don't know

**ALL RESPONDENTS (WITH OR WITHOUT PINHOLE LEAKS) PLEASE CONTINUE ANSWERING HERE.**

25. Do you use any of the following applications to **prevent** plumbing corrosion or pinhole leaks in your home? (CIRCLE ALL THAT APPLY)
    1. Preventive replumbing (replumbed without experiencing any leaks) (Specify material used) ______________________________________________
    ______________________________________________
    2. Applied coating/epoxy to all plumbing without experiencing any leaks
    3. In-line injection of phosphate (corrosion inhibitors) into home water system
    4. Water softener/water conditioning system
    5. Copper Knight
    6. Other (Specify) ______________________________________________
    ______________________________________________
    7. I have not used any preventive device (Go to question 27)

26. If you used preventive replumbing or other corrosion prevention devices, what is the approximate cost of purchasing and installing the devices since you have moved in your home? (CIRCLE ONE NUMBER)
    1. $1 - $99
    2. $100 - $499
    3. $500 - $999
    4. $1,000 - $2,499
    5. $2,500 - $4,999
    6. $5,000 - $7,499
    7. $7,500 – $9,999
    8. $10,000 -14,999
    9. $15,000 - $19,999
    10. 20,000 or more (Specify) ____________________
    11. Do not know

27. What is your approximate total monthly maintenance expenditure on corrosion prevention (including service contracts for corrosion treatment devices, water softeners, other chemicals, etc.)? (CIRCLE ONE NUMBER)
    1. $0
    2. $1 -$24
    3. $25 - $49
    4. $50 - $74
    5. $75 - $99
    6. $100 or more
    7. Do not know

28. Consider the current condition of drinking water plumbing systems in Tellico Village. The risk is about 10 percent that a home will develop a pinhole leak problem. If a home has pinhole leaks, costs of plumbing repair/replacement could range from $6,500 to $14,000 and property damage costs could range from $0 to $20,000. You currently have the option of staying with your current plumbing system and fixing any problems which may arise. Or suppose an insurance program covered all future costs of leaks and repairs, including plumbing repairs and damage to home and personal property. This insurance program will cost you a one-time payment of $14,000? Would you buy the insurance? (CIRCLE ONE NUMBER)
    1. Yes, I would buy the insurance
    2. No, I would not buy the insurance

29. Again consider the current condition of drinking water plumbing systems in Tellico Village homes. You have the opportunity to keep your current plumbing with no insurance, buy the insurance described in question 28, or upgrade your current plumbing to reduce risk of corrosion and pinhole leaks. Please consider each of these three alternatives.

| | 1. Upgrade Current Plumbing System | 2. Current Plumbing and Insurance | 3. Current Plumbing and No Insurance |
|---|---|---|---|
| Risk of financial loss due to leaks | 5% | No risk | 10% |
| Cost | $6,000 | $14,000 | 0 |

Which alternative would you choose? (CIRCLE ONE NUMBER)
    1. Upgrade current plumbing
    2. Keep current plumbing and buy insurance
    3. Keep current plumbing with no insurance

30. How concerned are you that your home plumbing will develop pinhole leaks in the future? (CIRCLE ONE NUMBER)
    1. Very concerned
    2. Somewhat concerned
    3. Not very concerned
    4. Not at all concerned
    5. Don't know

31. Do you use any of the following for purposes other than corrosion prevention? (CIRCLE ALL THAT APPLY)
    1. Filter for the entire home water system
    2. Filter as part of the refrigerator
    3. Water softener/water conditioning system
    4. Pitcher or bottle that filters drinking water
    5. Store purchased drinking water or bottled water
    6. Filter on faucet or under kitchen sink
    7. Ultra violet (UV) system
    8. Other (Specify) ______________________________
    9. I do not use anything to improve home drinking water

32. If you use any of the items in question 31, please check the reason why you do so. (CIRCLE ALL THAT APPLY)
    1. Improve safety of my drinking water
    2. Improve the clearness of my drinking water
    3. Improve taste or smell of my drinking water
    4. Sales information or presentation
    5. Other (Specify) ______________________________

The following questions will help us better characterize the Tellico Village Community for statistical reference in our survey analysis. All responses will remain confidential. Overall results will be reported in summary form only.

33. Which of the following describes your home? (CIRCLE ONE NUMBER)
    1. Single family home which I own
    2. Single family home which I rent

34. Approximately what year was your house built? (CIRCLE ONE NUMBER)
    1. Since 2000
    2. 1995 to 1999
    3. 1990 to 1994
    4. Before 1990
    5. Don't know

35. What month and year did you move into your current home? (FILL IN THE BLANK)

    __________ of ________
    (month) (year)

36. What is the highest level of formal education you have completed? (CIRCLE ONE NUMBER)
    1. Grade School
    2. Some High School
    3. High School Graduate (or GED)
    4. Trade/Voc. School after High School
    5. Some College
    6. Completed Community College
    7. Four-year College/University Graduate
    8. Graduate School/Professional School

37. Counting yourself, how many people live in your home now? (FILL IN THE BLANK)

    ________________

38. What is your gender? (CIRCLE ONE NUMBER)
    1. Male
    2. Female

39. What is the approximate current market value of your home? (CIRCLE ONE NUMBER)
    1. Less than $200,000
    2. $200,001 - $250,000
    3. $250,001 - $300,000
    4. $300,001 - $350,000
    5. $350,001 - $400,000
    6. $400,001 - $500,000
    7. $500,001 - $750,000
    8. Over $750,001
    9. Don't know

40. Please choose the bracket that includes your best estimate of your household income before taxes last year: (CIRCLE ONE NUMBER)
    1. Less than $35,000
    2. $35,001 - $55,000
    3. $55,001 - $75,000
    4. $75,001 - $100,000
    5. $100,001 - $ 150,000
    6. $151,000 - $200,000
    7. Over $200,000
    8. Do not know

41. In what year were you born? OPTIONAL (FILL IN THE BLANK)
    19______

42. Do you consider yourself to be: OPTIONAL (CIRCLE ONE)
    1. Caucasian
    2. African American
    3. Asian
    4. Hispanic
    5. Other (Please specify) _____________

Do you have additional comments regarding pinhole leak problems in home plumbing?

Are there other problems you would like to share about your home drinking water?

**Thank you for completing this survey!**

## Appendix C: Southeastern Community Follow-up Survey

National Home Drinking Water Assessment
Attn: Darrell Bosch
Department of Agricultural and Applied Economics (0401)
Virginia Polytechnic Institute and State University
Blacksburg, VA 24061

October 2007

Dear Tellico Village Resident:

Thank you for agreeing to be contacted again for a follow-up survey on home drinking water plumbing. Virginia Tech researchers are conducting this follow-up survey with funding from the National Science Foundation.

In this survey, we seek to understand your decision-making process in the choices of home drinking water plumbing systems and materials. Plumbing materials have important implications for costs, reliability, aesthetics, taste, odor, and health impacts of drinking water. The main purpose of the survey is to elicit your preferences for home drinking water plumbing attributes and materials. The information you provide will be very useful to policymakers, water utility managers, educators, researchers, and other professionals concerned with the quality of drinking water.

Please be assured that your responses will be treated confidentially. Individual responses will not be reported. All responses from residents will be compiled, analyzed, and reported together. ***After completing the survey, please return it in the enclosed, postage-paid envelope or send it to the address above.*** If you would like to receive a summary of results, please include your contact information on the enclosed card and return it with the survey or mail it separately. If you have questions, contact me at 540-231-5265 or by email at *bosch@vt.edu*. We greatly appreciate your help with this important survey.

Sincerely,

Darrell Bosch

Darrell Bosch
Professor

**Introduction**

This survey asks you to compare plumbing materials, which may differ according to the following attributes:

*Corrosion resistance* — ability of the material to resist corrosion.
*Taste/odor* — effect of the material on the taste or odor of drinking water.
*Health* — possible health effects of the material on those who consume drinking water.
*Convenience of installation* — ease of installing the plumbing material in the home.
*Proven performance in market* — time the plumbing material has been successfully used in the market.
*Cost (material and labor)* — material and labor cost to install the plumbing material in the home.
*Warranty* — length of time the plumbing material is guaranteed against failure.

**Section 1. Comparing plumbing material options**

In this section of the survey, you will see three questions consisting of two groups of messages describing plumbing material options. Imagine you were choosing plumbing materials for your house. After reading the messages in each question, please indicate your preference for each plumbing material option.

**Question 1: First group of plumbing material options.**
How would you rate each of the two plumbing material options described on the next page?

| Preference Level | Score |
|---|---|
| Not preferred | 1 |
| Moderately preferred | 3 |
| Strongly preferred | 5 |
| Very strongly preferred | 7 |
| Extremely preferred | 9 |

Circle one number showing your preference level for each option.

**Option A**

1 3 5 7 9

**Option B**

1 3 5 7 9

| | **Option A** | **Option B** |
|---|---|---|
| **Corrosion Resistance** | Corrosion proof | Same as material A |
| **Taste / Odor** | Compounds released from this material in drinking water plumbing may give a chemical or solvent taste or odor to the water. | Same as material A |
| **Health Effects** | Material meets EPA Standards. A very small chance exists that compounds from plumbing made of this material that are released into drinking water may lead to microbial growth in water. Microbial growth may cause severe illness. | Same as material A |
| **Convenience of Installation** | No need to tear into the wall and/or floor. Installation takes around 4 days. | Need to tear into some sections of wall for installation. Installation takes 5 to 6 days. |
| **Proven Performance in Market** | Less than 10 years in the market. | Less than 20 years in the market. |
| **Cost (labor + material)** | \$9,000 ~ \$14,000 depending on the size of house. | \$6,500 ~ \$13,000 depending on the size of house. |
| **Warranty** | Warranty is 15 years for the material. | Warranty is 10 years for the material. |

**Question 2: Second group of plumbing material options.**
How would you rate each of the two plumbing material options described on the next page?

| Preference Level | Score |
|---|---|
| Not preferred | 1 |
| Moderately preferred | 3 |
| Strongly preferred | 5 |
| Very strongly preferred | 7 |
| Extremely preferred | 9 |

Circle one number showing your preference level for each option.

**Option A**

1 3 5 7 9

**Option B**

1 3 5 7 9

| | Option A | Option B |
|---|---|---|
| **Corrosion Resistance** | Corrosion proof | Some risk of corrosion |
| **Taste / Odor** | Compounds released from this material in drinking water plumbing may give a chemical or solvent taste or odor to the water. | Compounds released from this material in drinking water plumbing may give a bitter or metallic taste or odor to the water. |
| **Health Effects** | Material meets EPA Standards. A very small chance exists that compounds from this plumbing material that are released into drinking water may lead to microbial growth in water. Microbial growth may cause severe illness. | Material meets EPA Standards. A very small chance exists that compounds from this plumbing material that are released into drinking water may cause vomiting, diarrhea, stomach cramps, and nausea. |
| **Convenience of Installation** | No need to tear into the wall and/or floor. Installation takes around 4 days. | Need to tear into the wall and/or floor to replace the existing system. Installation takes 7 to 9 days. |
| **Proven Performance in Market** | Less than 10 years in the market. | More than 50 years in the market. |
| **Cost (labor + material)** | \$9,000 – \$14,000 depending on the size of house. | \$9,000 – \$16,000 depending on the size of house. |
| **Warranty** | Warranty is 15 years for the material. | 50 years manufacturer's warranty. Some exceptions apply, e.g. warranty reduces to one year if compounds in water corrode pipes. |

**Question 3: Third group of plumbing material options.**
How would you rate each of the two plumbing material options described on the next page?

| Preference Level | Score |
|---|---|
| Not preferred | 1 |
| Moderately preferred | 3 |
| Strongly preferred | 5 |
| Very strongly preferred | 7 |
| Extremely preferred | 9 |

Circle one number showing your preference level for each option.

**Option A**

1 3 5 7 9

**Option B**

1 3 5 7 9

| | Option A | Option B |
|---|---|---|
| **Corrosion Resistance** | Corrosion proof | Some risk of corrosion |
| **Taste / Odor** | Compounds released from this material in drinking water plumbing may give a chemical or solvent taste or odor to the water. | Compounds released from this material in drinking water plumbing may give a bitter or metallic taste or odor to the water |
| **Health Effects** | Material meets EPA Standards. A very small chance exists that compounds from this plumbing material that are released into drinking water may lead to microbial growth in water. Microbial growth may cause severe illness. | Material meets EPA Standards. A very small chance exists that compounds from this plumbing material that are released into drinking water may cause vomiting, diarrhea, stomach cramps, and nausea. |
| **Convenience of Installation** | Need to tear into some sections of wall for installation. Installation takes 5 to 6 days. | Need to tear into the wall and/or floor to replace the existing system. Installation takes 7 to 9 days |
| **Proven Performance in Market** | Less than 20 years in the market. | More than 50 years in the market. |
| **Cost (labor + material)** | \$6,500 ~ \$13,000 depending on the size of house. | \$9,000 ~ \$16,000 depending on the size of house. |
| **Warranty** | Warranty is 10 years for the material. | 50 years manufacturer's warranty. Some exceptions apply, e.g. warranty reduces to one year if compounds in water corrode pipes. |

### Section 2. Comparing Plumbing Attributes

In this section, you will compare the attributes of the plumbing materials. Using the scale in the table below, please indicate which attribute is more important to you by circling the number. For example, if ***cost (material and labor)*** *is equally preferred to* ***health***; circle 1. Similarly, if ***convenience of installation*** is *moderately preferred* to ***taste and odor***, circle 3 in *convenience of installation side.*

**Table 1. Standard numerical scores**

| Preference Level | Score |
|---|---|
| Equally preferred | 1 |
| Moderately preferred | 3 |
| Strongly preferred | 5 |
| Very strongly preferred | 7 |
| Extremely preferred | 9 |

Example

| Cost (material + labor) | | | | | | | | Health |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

| Convenience of installation | | | | | | | | Taste/Odor |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Comparison (a) | Corrosion resistance | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 | Cost (material + labor) |
| Comparison (b) | Health | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 | Proven performance in market |
| Comparison (c) | Convenience of installation | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 | Warranty |
| Comparison (d) | Warranty | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 | Proven performance in market |
| Comparison (e) | Cost (material + labor) | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 | Taste/Odor |
| Comparison (f) | Health | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 | Cost (material + labor) |
| Comparison (g) | Taste/Odor | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 | Warranty |
| Comparison (h) | Warranty | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 | Health |

Circle the number showing your preference (1 = equally preferred; 9 = extremely preferred

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Comparison (i) | Proven performance in market | | | | | Corrosion resistance | | | |
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |
| Comparison (j) | Health | | | | | Taste/Odor | | | |
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |
| Comparison (k) | Taste/Odor | | | | | Corrosion resistance | | | |
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |
| Comparison (l) | Proven performance in market | | | | | Cost (material + labor) | | | |
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |
| Comparison (m) | Cost (material + labor) | | | | | Convenience of installation | | | |
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |
| Comparison (n) | Convenience of installation | | | | | Taste/Odor | | | |
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |
| Comparison (o) | Health | | | | | Corrosion resistance | | | |
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

Circle the number showing your preference (1 = equally preferred; 9 = extremely preferred

| Comparison (p) | Warranty | | | | | Corrosion resistance | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

| Comparison (q) | Convenience of installation | | | | | Health | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

| Comparison (r) | Corrosion resistance | | | | | Convenience of installation | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

| Comparison (s) | Proven performance in market | | | | | Taste/Odor | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

| Comparison (t) | Convenience of installation | | | | | Proven performance in market | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

| Comparison (u) | Cost (material + labor) | | | | | Warranty | | | |
|---|---|---|---|---|---|---|---|---|---|
| | 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

### Section 3. Comparing Plumbing Materials

In this section you are asked to compare plumbing materials with respect to specific attributes: Cost, Taste and odor, Health, Corrosion resistance, Convenience of installation, Warranty, and Proven performance in the market. Using the scale in the table below, please indicate which option is preferred to you by circling the number. For example, if ***Option A*** *is equally preferred to* ***Option B,*** circle 1.

Example: **Option A** and **Option B** are *equally preferred* when we consider taste and odor effects only. Therefore, number 1 is circled in this example.

**Table 2. Standard numerical scores**

| Preference Level | Score |
|---|---|
| Equally preferred | 1 |
| Moderately preferred | 3 |
| Strongly preferred | 5 |
| Very strongly preferred | 7 |
| Extremely preferred | 9 |

**Taste and odor**

| | |
|---|---|
| Option A | Compounds released from this material in drinking water plumbing may give a chemical or solvent taste or odor to the water. |
| Option B | Compounds released from this material in drinking water plumbing may give a bitter or metallic taste or odor to the water. |

| Option A | | | | | | | | Option B |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

### I. Corrosion resistance

The following question asks you to compare two options for corrosion resistance.

| | |
|---|---|
| Option A | Some risk of corrosion |
| Option B | Corrosion proof |

| Option A | | | | | Option B | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

**Taste and odor**

| | |
|---|---|
| Option A | Compounds released from this material in drinking water plumbing may give a chemical or solvent taste or odor to the water. |
| Option B | Compounds released from this material in drinking water plumbing may give a bitter or metallic taste or odor to the water. |

| Option A | | | | | Option B | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

**III. Health**

The following question asks you to compare two options for health effects.

| | |
|---|---|
| Option A | Material meets EPA Standards. A very small chance exists that compounds from this plumbing material that are released into drinking water may lead to microbial growth in water. Microbial growth may cause severe illness. |
| Option B | Material meets EPA Standards. A very small chance exists that compounds from this plumbing material that are released into drinking water may cause vomiting, diarrhea, stomach cramps, and nausea. |

| Option A | | | | | Option B | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

**IV. Cost (material + labor)**

The following question asks you to compare three options for cost of installation

| | |
|---|---|
| Option A | \$9,000 ~ \$14,000 depending on the size of house. |
| Option B | \$6,500 ~ \$13,000 depending on the size of house. |
| Option C | \$9,000 ~ \$16,000 depending on the size of house. |

Comparison (a) Please circle the number to show your preference for option A or B

| Option A | | | | | Option B | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

Comparison (b) Please circle the number to show your preference for option B or C

| Option B | | | | | Option C | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

Comparison (c) Please circle the number to show your preference for option C or A

| Option C | | | | | Option A | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

**V. Convenience of installation**

The following question asks you to compare three options for convenience of installation.

| | |
|---|---|
| Option A | No need to tear into the wall and/or floor. Installation takes around 4 days. |
| Option B | Need to tear into some sections of wall for installation. Installation takes 5 - 6 days. |
| Option C | Need to tear into the wall and/or floor to replace the existing system. 7 - 9 days required for installation |

Comparison (a) Please circle the number to show your preference for option A or B

| Option A | | | | | Option B | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

Comparison (b) Please circle the number to show your preference for option B or C

| Option B | | | | | Option C | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

Comparison (c) Please circle the number to show your preference for option C or A

| Option C | | | | | Option A | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

**VI. Proven performance in the market**

The following question asks you to compare three options for proven performance in the market.

| | |
|---|---|
| Option A | Less than 10 years in the market. |
| Option B | Less than 20 years in the market. |
| Option C | More than 50 years in the market. |

Comparison (a) Please circle the number to show your preference for option A or B

| Option A | | | | | Option B | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

Comparison (b) Please circle the number to show your preference for option B or C

| Option B | | | | | Option C | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

Comparison (c) Please circle the number to show your preference for option C or A

| Option C | | | | | Option A | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

**VII. Warranty**

The following question asks you to compare three options for warranty of material.

| | |
|---|---|
| Option A | Warranty is 15 years for the material |
| Option B | Warranty is 10 years for the material. |
| Option C | 50 years manufacturer's warranty. Some exceptions apply, e.g. warranty reduces to one year if compounds in water corrode pipes. |

Comparison (a) Please circle the number to show your preference for option A or B

| Option A | | | | | Option B | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

Comparison (b) Please circle the number to show your preference for option B or C

| Option B | | | | | Option C | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

Comparison (c) Please circle the number to show your preference for option C or A

| Option C | | | | | Option A | | | |
|---|---|---|---|---|---|---|---|---|
| 9 | 7 | 5 | 3 | 1 | 3 | 5 | 7 | 9 |

**Section 4. Choosing a plumbing material**

Using the information in the table below and on the next page that describes each material, please select the material you would choose if you have to replace your current home plumbing system. Circle the number by the material you would choose.

1) Material A 2) Material B 3) Material C

| | Material A | Material B | Material C |
|---|---|---|---|
| **Corrosion Resistance** | Corrosion proof | Same as material A | Some risk of corrosion |
| **Taste / Odor** | Compounds released from this material in drinking water plumbing may give a chemical or solvent taste or odor to the water. | Same as material A | Compounds released from this material in drinking water plumbing may give a bitter or metallic taste or odor to the water. |
| **Health Effects** | Material meets EPA Standards. A very small chance exists that compounds from this plumbing material that are released into drinking water may lead to microbial growth in water. Microbial growth may cause severe illness. | Same as material A | Material meets EPA Standards. A very small chance exists that compounds from this plumbing material that are released into drinking water may cause vomiting, diarrhea, stomach cramps, and nausea. |
| **Convenience of Installation** | No need to tear into the wall and/or floor. Installation takes around 4 days. | Need to tear into some sections of wall for installation. Installation takes 5-6 days. | Need to tear into the wall and/or floor to replace the existing system. Installation takes 7 to 9 days. |

**Continued on next page.**

| | Material A | Material B | Material C |
|---|---|---|---|
| **Proven performance in market** | Less than 10 years in the market. | Less than 20 years in the market. | More than 50 years in the market. |
| **Cost (labor + material)** | \$9,000 ~ \$14,000 depending on the size of house. | \$6,500 ~ \$13,000 depending on the size of house. | \$9,000 ~ \$16,000 depending on the size of house. |
| **Warranty** | Warranty is 15 years for the material. | Warranty is 10 years for the material. | 50 years manufacturer's warranty. Some exceptions apply, e.g. warranty reduces to one year if compounds in water corrode pipes. |

***Please make any comments about plumbing material alternatives here.***

***Thank you for completing this survey!***

**Appendix D: Statistical Software Programs**

Experimental Design for Southeastern Community Survey

```
options errors =3 nocenter pagesize = 60;

libname out "C:\Documents and Settings\ekleczyk";

title1 "Test Conjoint Design";

proc factex;
factors x1-x5  /nlev=5;/
model res = 5;
size design = min;
output out=out.Design1;
 x1 nvals=(1 2 3 3 3)
 x3 nvals=(1000 5000 7500 10000 14000)
 x4 nvals=(400 2000 3000 4000 6000)
 x5 nvals=(0 0.03 0.06 0.06 0.06);
run;

proc optex data = out.design1 seed=26547;
class x1-x5;
model x1-x5;
generate n=1600 method=m_fedorov iter=40;
output  out=out.design2;
run;

proc print data = out.design2;
run;

proc corr data=out.design2;
var x1-x5;
run;
```

## Appendix E: Southeastern Community Survey IRB Exempt Approval

VirginiaTech

**Office of Research Compliance**
1880 Pratt Drive (0497)
Blacksburg, Virginia 24061
540/231-4358 Fax: 540/231-0959
E-mail: ctgreen@vt.edu
*www.irb.vt.edu*
FWA00000572( expires 7/20/07)
IRB # is IRB00000667.

DATE: August 4, 2006

MEMORANDUM

TO: Darrell J. Bosch
Ewa Kleczyk
Eftila Tanellari

FROM: Carmen Green

SUBJECT: **IRB Exempt Approval:** "Evaluating Preferences for Plumbing Material Attributes: Comparison of CV, CA and AHP" , IRB # 06-425

I have reviewed your request to the IRB for exemption for the above referenced project. I concur that the research falls within the exempt status. Approval is granted effective as of August 4, 2006.

As an investigator of human subjects, your responsibilities include the following:

1. Report promptly proposed changes in previously approved human subject research activities to the IRB, including changes to your study forms, procedures and investigators, regardless of how minor. The proposed changes must not be initiated without IRB review and approval, except where necessary to eliminate apparent immediate hazards to the subjects.

2. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

VirginiaTech

**Office of Research Compliance**
1880 Pratt Drive (0497)
Blacksburg, Virginia 24061
540/231-4358 Fax: 540/231-0959
E-mail: ctgreen@vt.edu
*www.irb.vt.edu*

FWA00000572( expires 7/20/07)
IRB # is IRB00000667.

DATE: October 5, 2006

MEMORANDUM

TO: Darrell J. Bosch
Ewa Kleczyk
Eftila Tanellari

FROM: Carmen Green

SUBJECT: **IRB Amendment 1 Approval:** "Evaluating Preferences for Plumbing Material Attributes: Comparison of CV, CA and AHP" , IRB # 06-425

This memo is regarding the above referenced protocol which was previously granted approval by the IRB on August 4, 2006. You subsequently requested permission to amend your IRB application. Approval has been granted for requested protocol amendment, effective as of October 5, 2006.

As an investigator of human subjects, your responsibilities include the following:

1. Report promptly proposed changes in previously approved human subject research activities to the IRB, including changes to your study forms, procedures and investigators, regardless of how minor. The proposed changes must not be initiated without IRB review and approval, except where necessary to eliminate apparent immediate hazards to the subjects.
2. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

VirginiaTech

**Office of Research Compliance**
1880 Pratt Drive (0497)
Blacksburg, Virginia 24061
540/231-4358 Fax: 540/231-0959
E-mail: ctgreen@vt.edu
*www.irb.vt.edu*

FWA00000572( expires 7/20/07)
IRB # is IRB00000667.

DATE: November 20, 2006

MEMORANDUM

TO: Darrell J. Bosch
Ewa Kleczyk
Eftila Tanellari

FROM: Carmen Green

SUBJECT: **IRB Amendment 2 Approval:** "Evaluating Preferences for Plumbing Material Attributes: Comparison of CV, CA and AHP" , IRB # 06-425

This memo is regarding the above referenced protocol which was previously granted approval by the IRB on August 4, 2006. You subsequently requested permission to amend your IRB application. Approval has been granted for requested protocol amendment, effective as of November 20, 2006.

As an investigator of human subjects, your responsibilities include the following:

1. Report promptly proposed changes in previously approved human subject research activities to the IRB, including changes to your study forms, procedures and investigators, regardless of how minor. The proposed changes must not be initiated without IRB review and approval, except where necessary to eliminate apparent immediate hazards to the subjects.
2. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

VirginiaTech

**Office of Research Compliance**
Institutional Review Board
2000 Kraft Drive, Suite 2000 (0497)
Blacksburg, Virginia 24061
540/231-4991 Fax 540/231-0959
e-mail moored@vt.edu
*www.irb.vt.edu*
FWA00000572( expires 1/20/2010)
IRB # is IRB00000667

DATE: June 1, 2007

MEMORANDUM

TO: Darrell J. Bosch
Sharon K. Dwyer
Eftila Tanellari

FROM: David M. Moore

SUBJECT: **IRB Exempt Approval:** "Toward Sustainable Materials for Drinking Water Infrastructure-Tellico Survey", OSP #477056, IRB # 07-301

I have reviewed your request to the IRB for exemption for the above referenced project. I concur that the research falls within the exempt status. Approval is granted effective as of June 1, 2007.

As an investigator of human subjects, your responsibilities include the following:

1. Report promptly proposed changes in previously approved human subject research activities to the IRB, including changes to your study forms, procedures and investigators, regardless of how minor. The proposed changes must not be initiated without IRB review and approval, except where necessary to eliminate apparent immediate hazards to the subjects.

2. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

VirginiaTech

**Office of Research Compliance**
Institutional Review Board
2000 Kraft Drive, Suite 2000 (0497)
Blacksburg, Virginia 24061
540/231-4991 Fax 540/231-0959
e-mail moored@vt.edu
*www.irb.vt.edu*
FWA00000572( expires 1/20/2010)
IRB # is IRB00000667

DATE: July 2, 2007

MEMORANDUM

TO: Darrell J. Bosch
Sharon K. Dwyer
Eftila Tanellari

FROM: David M. Moore

SUBJECT: **IRB Amendment 1 Approval:** "Toward Sustainable Materials for Drinking Water Infrastructure-Tellico Survey", OSP #477056, IRB # 07-301

This memo is regarding the above referenced protocol which was previously granted approval by the IRB on June 1, 2007. You subsequently requested permission to amend your IRB application. Approval has been granted for requested protocol amendment, effective as of June 29, 2007.

As an investigator of human subjects, your responsibilities include the following:

1. Report promptly proposed changes in previously approved human subject research activities to the IRB, including changes to your study forms, procedures and investigators, regardless of how minor. The proposed changes must not be initiated without IRB review and approval, except where necessary to eliminate apparent immediate hazards to the subjects.
2. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

VirginiaTech

**Office of Research Compliance**
Institutional Review Board
2000 Kraft Drive, Suite 2000 (0497)
Blacksburg, Virginia 24061
540/231-4991 Fax 540/231-0959
e-mail moored@vt.edu
*www.irb.vt.edu*
FWA00000572( expires 1/20/2010)
IRB # is IRB00000667

DATE: July 24, 2007

MEMORANDUM

TO: Darrell J. Bosch
Sharon K. Dwyer
Eftila Tanellari

FROM: Carmen Green

SUBJECT: **IRB Amendment 2 Approval:** "Toward Sustainable Materials for Drinking Water Infrastructure-Tellico Survey", OSP #477056, IRB # 07-301

This memo is regarding the above referenced protocol which was previously granted approval by the IRB on June 1, 2007. You subsequently requested permission to amend your IRB application. Approval has been granted for requested protocol amendment, effective as of July 24, 2007.

As an investigator of human subjects, your responsibilities include the following:

1. Report promptly proposed changes in previously approved human subject research activities to the IRB, including changes to your study forms, procedures and investigators, regardless of how minor. The proposed changes must not be initiated without IRB review and approval, except where necessary to eliminate apparent immediate hazards to the subjects.
2. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

VirginiaTech

**Office of Research Compliance**
Institutional Review Board
2000 Kraft Drive, Suite 2000 (0497)
Blacksburg, Virginia 24061
540/231-4991 Fax 540/231-0959
e-mail moored@vt.edu
*www.irb.vt.edu*
FWA00000572( expires 1/20/2010)
IRB # is IRB00000667

DATE: September 21, 2007

MEMORANDUM

TO: Darrell J. Bosch
Sharon K. Dwyer
Eftila Tanellari

FROM: Carmen Green

SUBJECT: **IRB Amendment 3 Approval:** "Toward Sustainable Materials for Drinking Water Infrastructure-Tellico Survey", OSP #477056, IRB # 07-301

This memo is regarding the above referenced protocol which was previously granted approval by the IRB on June 1, 2007. You subsequently requested permission to amend your IRB application. Approval has been granted for requested protocol amendment, effective as of September 21, 2007.

As an investigator of human subjects, your responsibilities include the following:

1. Report promptly proposed changes in previously approved human subject research activities to the IRB, including changes to your study forms, procedures and investigators, regardless of how minor. The proposed changes must not be initiated without IRB review and approval, except where necessary to eliminate apparent immediate hazards to the subjects.
2. Report promptly to the IRB any injuries or other unanticipated or adverse events involving risks or harms to human research subjects or others.

# CHAPTER 4

# The Determinants of Corporate Debt Maturity Structure

An earlier version of this article was published in the Fall 2007 edition of the *International Journal of Business & Economics*, Volume 6, Issue 1, pages 80-88. The published article does not discuss the economic theory guiding the econometric models, including the efficient market and principal-agent theories. The theoretical framework is not necessary to understand the empirical results and conclusions of the analysis.

An earlier version of this article was presented at the 6th Global Conference on Business and Economics at Harvard University, Boston, MA. October 2006. The text is available online at
www.trintrin.com/gebc/The%20Excess%20Bond%20Returns%20and%20Debt%20Maturity.doc.

## Abstract

Debt financing decisions are made simultaneously by lenders and borrowers. Since lenders are unable to observe directly the firms' investment decisions, the banks offer contracts based up on firms' observable characteristics such as wealth and size. The contracts offered to each firm differ on the cost of capital borrowing as well as the debt maturity. When deciding on the financing decisions, firms also take into account the changes in macroeconomic variables in order to lower the cost of borrowing. As a result, the goal for this article is to examine empirically the hypothesis of the effect of the debt determinant as well as the macroeconomic variables on the debt maturity structure. A reduced form of the simultaneous financing decisions model is estimated by employing several OLS estimation methods. The empirical findings offer strong support for firms with few growth options, large, and of low quality having more long-term debt in their capital structure. Macroeconomic variables such as excess bond returns, term spread, and inflation also affect the debt maturity decisions, which imply that managers time the debt market. On the other hand, there is no evidence in support of the impact of short-term interest rates on the debt maturity, confirming the hypothesis of the firms' inability to profit from market timing. Furthermore, there is little evidence in support of regulatory impact on the debt structure.

**Keywords:** Debt Maturity, Capital Structure, Panel Data Analysis

**JEL Classification Codes:** G32, G39

### 4.1. Introduction

According to Stiglitz (1974) and Modigliani and Miller (1958), in efficient and integrated markets, the financial management policy cannot decrease the costs of capital due to the interrelation between the different types of capital costs. Consequently, there is no gain from substituting debt for equity. However, evidence has been found against this claim, demonstrating that equity financing is related to the predictability of stock returns. For example, firms issue equity when the equity premium is low in order to time an inefficient market and reduce the cost of capital borrowing and/or to optimize capital structure together with expected returns (Baker et al., 2003). On the other hand, there is substantially less literature on debt financing. For example, according to Bosworth (1971), debt maturity is related to market conditions, such as the interest and inflation rates. Furthermore, Barclay and Smith (1995) find firms with higher information asymmetries to issue short-term debt. They determine a positive relationship between debt-maturity and dividend yield as well as a negative relationship between the maturity of debt and the term-spread (Barclay & Smith, 1995). Although the above papers provide support for the association between debt and returns, none offers information on the cost of capital borrowing at different times of debt maturity due to lack of analysis of returns data.

The relationship between debt maturity and cost of capital borrowing is examined by Baker et al. (2003). In their paper, they analyze the variation in the maturity of debt due to debt market conditions (inflation and interest rates) as well as the excess bond returns. Excess bond returns are an index of investment-grade corporate bonds over commercial papers (Baker et al., 2003). In their article, the authors utilize aggregate annual time-series data and find a close relation between debt maturity and predictable variation in the excess bond returns. According to their findings, firms issue long-term debt when bond returns are low and short-term debt when returns are high. They explain the above evidence as managers timing an inefficient capital market using public information to guide their debt maturity decisions. They are, however, unable to conclude whether firms actually reduce the overall cost of capital borrowing from timing the market as a result of difficulties in interpreting the predictions of the regression results (Baker et al., 2003).

This paper extends the empirical analysis performed by Baker et al. (2003) to the pooled time- and cross-sectional data in order to examine the relationship between the excess bond returns and the corporate debt maturity structure. In addition, the analysis extends previous studies such as that conducted by Barclay and Smith (1995) by incorporating the bond returns information in the corporate debt maturity structure research. As a result, the study accounts for the relationship between debt maturity and the cost of borrowing at different maturities.

The goal of this paper is to examine the hypothesis of the effect of the macroeconomics variables, excess bond returns, as well as the corporate debt determinants on the debt maturity structure at the individual firm level data. The Principal – Agent Theory and the Efficient Market Hypothesis provide the theoretical framework for the analysis. The Principal - Agent Theorem describes the relationship between the lenders (banks) and borrowers (firms) in the credit market as well as the process of loan issuance. As firms' investment decisions are not directly observable by the banks, the terms of the loan (interest rate, loan size, and length) are usually based on the firms' observable characteristics such as wealth and size. The Efficient Market Hypothesis states that it is impossible in an efficient market to profit from market timing in the long-run. Market timing is defined as the opportunity to obtain higher return on investment by simply observing and responding to the changes in macroeconomic variables (Baker et al., 2003).The Efficient Market Hypothesis widely recognized by economists has been questioned by financial analysts. Several empirical studies including studies by Baker et al. (2003), Bosworth (1971) and Malkiel (2004), however, have found support for profiting from market timing in the short-run.

Based on these theorems, it is predicted that small and unregulated firms with more investment growth options have less long-term debt. When the bond market cannot distinguish between high - and low – quality firms, defined by the firm's abnormal earnings, the high-quality firms will signal their wealth through issuance of less underpriced short-term debt. Both investors and low - quality firms realize the high – quality firm's behavior, forcing the high - quality firms to issue more short-term debt (Kale & Noe, 1990). Similarly, firms with large potential information asymmetries are likely to issue short-term debt due to their larger information costs associated with long-term debt (Flannery, 1986). On the other hand, when the high – quality firms are faced with low risk of debt refinancing, they are more likely to issue short-term debt while less creditworthy firms will issue longer-term debt. The very poor firms cannot issue long-term debt as they face a higher likelihood of reporting low profits and facing higher borrowing costs (Berlin, 2006).

Additionally, according to Baker et al. (2003) debt maturity is also related to debt market conditions, such as interest and inflation rates as well as a term spread. It is expected that the macroeconomic variables are negatively related to the maturity of debt. Following Baker et al. (2003), it is predicted that firms issue long-term debt when bond returns are low. Finding support for the impact of macroeconomics variables including excess bond returns on debt maturity decisions implies that in the short-run period firms lower the cost of capital borrowing from timing the market. In the long-term, however, this hypothesis should not be supported by the model's results as suggested by the Efficient Market Hypothesis.

The research objective is achieved by investigating the relationship between the excess bond returns, macroeconomics variables, as well as the determinants of the firm's debt and the corporate debt maturity structure. In order to perform the analysis, the pooled time- and cross-sectional (panel) data for year 1991- 2000 from COMPUSTAT as well as the Federal Reserve are utilized. The variables used in the econometric analysis are: the dependent variable: share of long-term debt, and the explanatory variables: excess bond returns, market-to-book ratio, regulatory dummy variable, log of market value of firm, abnormal earnings, term spread, inflation, and short-term interest rates.

There are several Ordinary Least Squares (OLS) estimation methods employed to estimate the relationship between debt maturity and its determinants. First, the pooled OLS estimation is utilized; however, there is a problem of unobserved error term dependence, which results in biased and inconsistent findings (Wooldridge, 2002). In order to correct for this issue, a cross-sectional OLS regression method is utilized in combination with the time-series mean of each variable for an individual firm. However, this method only accounts for the variation across firms without analyzing the time-variation (Barclay & Smith, 1995).

In order to investigate the time-series dispersion, fixed and random effects regression models are employed. In the fixed effects regression, the cross-sectional variation is arbitrarily correlated with the explanatory variables. In the random effects regression model, it is assumed, on the other hand, that the cross-sectional variation is random, and it is a part of the unobserved variation included in the error term. The Hausman specification test is used to test for the sufficiency of random effects estimation. Finally, in order to correct for the existing heteroskedasticity and serial correlation across panels, the generalized least squares estimation is used to obtain efficient estimators (Wooldridge, 2002).

The results obtained from the OLS estimation provide evidence in support of a negative relationship between the debt maturity and the excess bond returns. Furthermore, the market-to-book ratio, abnormal earnings, as well as term spread and inflation are negatively related to the share of long-term debt. However, no statistically significant relationships were found for regulatory dummy and short-term interest rate although the coefficients of these variables had the predicted signs.

The results confirm the hypothesis of debt determinants and their impacts on the structure of corporate debt. The findings imply that firms with few investment growth opportunities, large in size, and of high quality have a large share of long-term debt in their financing structure. There is, however, an ambiguous result with regard to hypothesis of market timing lowering the cost of capital borrowing. On one side, managers time the debt market and incorporate the debt market conditions as well as the excess bonds returns in their debt financing decisions. On the other hand, the short-term interest rate does not have a statistically significant impact on a firm's financing decision as suggested by the Efficient Market Hypothesis. The research results add to the undeveloped literature of debt financing by extending the analysis of the impact of debt determinants on debt maturity and by analyzing returns data and providing evidence for relationship between debt maturity and cost of borrowing at different maturities.

## 4.2. Theoretical Framework

The theoretical framework guiding this article is the Principal – Agent Problem Model as well as the Efficient Market Hypothesis. These hypotheses include the Agency Cost Theorem, Signalling Theorem, and Market Timing Hypothesis.

### 4.2.1 Principal-Agent Theory

In the Principal – Agent problem, the principal wants to induce a person, the agent, to take some actions. In most cases, the principal is unable to observe directly the agent's action, but instead observes some output determined by the agents' actions. The principal's problem is to design an incentive payment to induce the agent to take the best action for the principal. The simplest example of the Principal-Agent Problem is a manager (principal) and a worker (agent). In the area of finance, the Board of Investors (principal) and the chief financial officer (CFO) of a company (agent) as well as credit institutions (principals) and corporate institutions (agents) are the examples. In the manager and worker case, the manager wants the worker to exert as much effort as possible in order to produce as much output as possible, while the worker rationally wants to make a choice that maximises his/her own utility given the effort and incentive payment scheme (Varian, 1992).

The Principal - Agent problem version follows Stiglitz and Weiss (1992) who provides the basic insights into the role of informational asymmetries[33] in credit markets. In their model, they take into account adverse selection, which includes signalling[34] and screening[35], as well as incentive problems that induce the agent to behave according to the principal's plan. In the model, the lenders vary simultaneously loan collateral requirements and interest rates in order to affect both the mix of applicants and encourage successful loan applicants to apply for loans. The loan collateral represents the approximation of the loan quality, as not all firms are able to afford high collateral

[33] **Information Asymmetry** - a study of decisions in transactions where one party has more or better information than the other (Stiglitz and Weiss, 1992).
[34] **Signalling** - in a situation with information asymmetry, some agents signal their types, thus believably transferring information to the principal and resolving the information asymmetry (Spence, 1973).
[35] **Screening** - the underinformed principals can induce the agents to reveal their information. They can provide a menu of choices in such a way that the choice depends on the private information of the agent (Stiglitz and Weiss, 1992)

requirements. The authors establish several equilibriums: pooling equilibriums in which high risk and low risk firms borrow at the same loan terms as well as separating equilibriums with low- risk firms obtaining loans on terms different from the high-risk firms. Additionally, Stiglitz and Weiss (1992) introduce rationing concept in which the lender might reject investors based on their perceived risk level due to the possible losses from loan contracts (Stiglitz & Weiss, 1992).

#### 4.2.1.1 Borrowers

In the Finance area, the basic Principal - Agent Model describes the behaviour of lenders and firms with regard to loan offers in the credit market. A borrower represented by a CFO invests funds lent by a bank in either a successful or an unsuccessful project with different probabilities of occurrences (please refer to Table 4.1 for term/notation definitions). The investment is made by utilizing either safe or risky investment techniques, which yield some level of return. An unsuccessful project is assumed to yield always a zero return. Lenders represented in the article as banks are unable to observe directly the chosen investment technique, which represents the asymmetric information faced by the banks with regard to the firm's characteristics (Stiglitz & Weiss, 1992).

Table 4.1: Principal – Agent Theory Terminology

| Variables | Definitions |
|---|---|
| Project Return | 1. Safe investment technique: $R^s$<br>2. Risky investment technique: $R^r$<br>3. Assumption: $R^s < R^r$ |
| Project Success Probability | 1. Safe investment technique: $p^s$<br>2. Risky investment technique: $p^r$<br>3. Assumption: $p^s > p^r$ |
| Wealth | 1. Initial wealth of $W_0$<br>*a.* collateralizable wealth: $C_0$<br>*b.* non-collateralizable wealth: $H_0$<br>2. Not invested wealth return: $W = W_0(1 + i^*)$<br>3. Firms differ on the amount of wealth:<br>*a.* Rich firms: $W_r$<br>*b.* Poor firms: $W_p$<br>c. Assumption: $W_r > W_p$ |

In this Principal - Agent Problem, all project costs are normalized at unity to simplify the analysis and each borrower undertakes at most one project. The borrower has an initial wealth of $W_0$ with collateralizable wealth, $C_0$. It is assumed that all borrowers have the same utility functions[36]. As a result, differences in wealth induce dissimilarity in the indifference curve slopes between the required interest payments on a loan and the required collateral. In order to receive a loan contract to finance an investment project, the bank requires the borrower to provide collateral of *C-hat* and pay a loan interest rate of *r*. Based on the above assumptions, the borrower's end-of-period wealth (it is the end of the project) is defined as follows:

$$Y = max\ \{W + R - (1+r), W - C\}, \qquad (1)$$

where *R* is the project return, *W* represents firms wealth, $W = W_0(1 + i^*)$, *C* represents the bank's collateral, $C = C\text{-}hat(1 + i^*)$, $Y_1 = W + R - (1+r)$ is the wealth of a successful project, and $Y_0 = W - C$ is the wealth following an unsuccessful project (Stiglitz and Weiss, 1992)[37].

The borrower's objective is to maximize his/her expected utility subject to the end-of-period wealth ($Y_1$ or $Y_0$) and the project success probabilities:

$$Max\ E\{U\} = U_1(Y_1)p + U_0(Y_0)(1-p) \quad (2)$$

$$Subject\ to\ Y = max\ \{W + R - (1 + r), W - C\} \quad (3)$$

Following Stiglitz and Weiss (1992), the borrower is assumed risk averse[38] with decreasing absolute risk aversion[39]. The assumption of decreasing absolute risk aversion

[36] The utility theory is explained in Appendix A.
[37] The analysis for firms profits based on the investment decisions are discussed in section on the Efficient Market Hypothesis.
[38] **Risk Aversion** - a concept related to the behavior of investors under uncertainty. It is the reluctance of a person to accept a bargain with an uncertain payoff rather than another bargain with a more certain, but possibly lower, expected payoff (Rubinstein, 1976).
[39] **Absolute Risk Aversion** - the Arrow-Pratt measure of the curvature of the utility function. Denote a utility function by *u(c)*, the Arrow-Pratt measure of absolute risk aversion is defined by:

$$R_A = -u''(c)/u'(c).$$

This measure is invariant to affine transformation of the utility function, which implies that the transformation does not affect the preferences expressed by *u()*. If $R_A()$ is decreasing in *c*, then *u()* displays implies wealthier ($W_0$) borrowers choose riskier techniques. The slope of the indifference curve is obtained from differentiating the expected utility, $E\{U\}$, with respect to $r$ and $C$:

$$E\{U\} = U_1(W + R - (1+r))p + U_0(W - C))(1-p)/dr,\ \ dC \qquad (4)$$

$$dE\{U\}= -U_1{}' pdr - U_0{}'(1 - p)dC = 0 \quad (5)$$

$$-U_1{}' pdr = U_0{}'(1 - p)dC \qquad (6)$$

$$-dr/dC = U_1{}'(1 - p)/U_0{}'p \qquad (7)$$

where

$$U_b{}' = U'(Y_b) = dU(Y_b)/dY_b,\ b = 0, 1. \quad (8)$$

$U'(Y_b)$ represents the marginal utility of wealth based on the stated outcome of project success denoted as $b$ (Stiglitz and Weiss, 1992).

Per the utility maximization problem, the borrower always chooses the technique yielding the highest expected utility. The borrower is, however, indifferent between the two techniques along the locus of $\{C, r\}$. The expected utility of risky investment is equal to the expected utility of the safe investment:

$$EU^r = EU^s \qquad (9)$$

Substituting the definition of the expected utility for risky and safe projects results in the following equation:

$$U(Y^r{}_1)p^r + U(Y_0)(1 - p^r) = U(Y^s{}_1)p^s + U(Y_0)(1 - p^s) \quad (10)$$

where $Y^r{}_1$ denotes the end-of-period wealth of a borrower who uses the risky technique that is successful and $Y^s{}_1$ denotes the end-of-period wealth of a borrower who uses the safe technique that is successful. The locus of $\{C, r\}$ is the switch line represented by combinations of interest rate and collateral at which the borrower is indifferent between

decreasing absolute risk aversion. If $R_A()$ is increasing in c, then $u()$ displays increasing absolute risk aversion. If $R_A()$ is constant with respect to changes in c, then $u()$ displays constant absolute risk aversion (Rubinstein, 1976).

undertaking safe and risky projects (see Figure 4.1). The switch line is positively sloped which can be shown by differentiating equation (10) with respect to *r* and *C*:

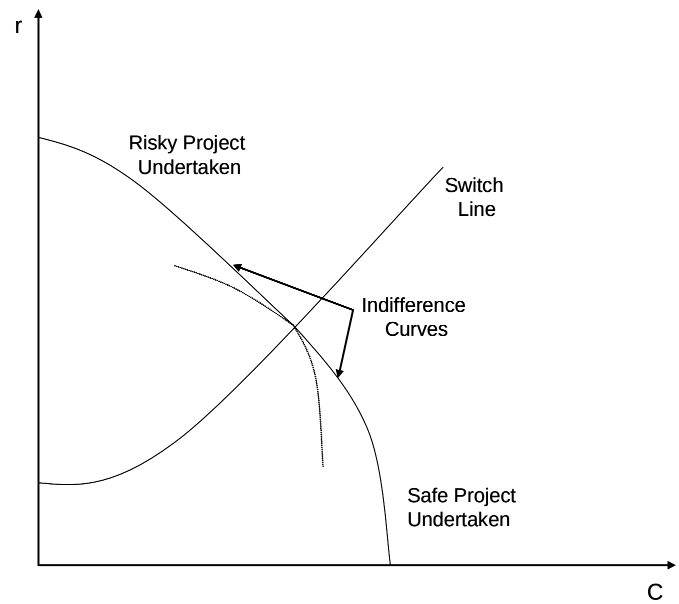
$$U'(Y^r{}_1)\, p^r dr + U'(Y_0)(1 - p^r)dC = U'(Y^s{}_1)\, p^s dr + U'(Y_0)(1 - p^s)dC \quad (11)$$

$$(-U'(Y^r{}_1)\, p^r + U'(Y^s{}_1)\, p^s)dr = (-U'(Y_0)(1 - p^s) + U'(Y_0)(1 - p^r))dC \quad (12)$$

$$dr/dC = U'(Y_0)(p^s - p^r)/(U'(Y^s{}_1)p^s - U'(Y^r{}_1)p^r) > 0. \quad (13)$$

The borrower uses a risky technique when the investment yields return above the switch line. The safe technique is utilized below the switch line where interests rates are low and collateral requirements are high. By specifying the levels of *r* and *C*, the bank indirectly controls the technique chosen by the investing firm (Stiglitz & Weiss, 1992).

Figure 4.1: Decision to Undertake the Safe vs. Risky Project



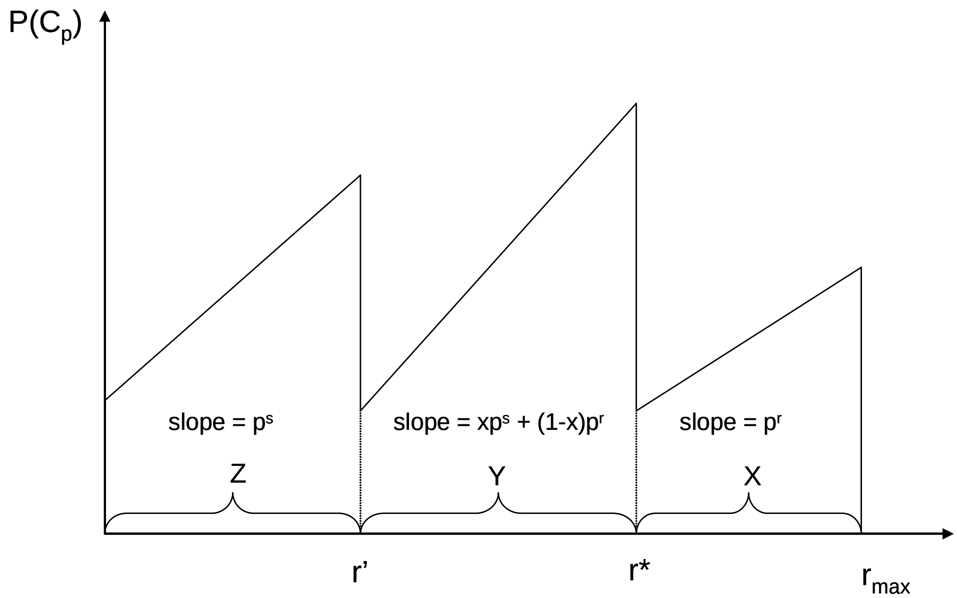

Note1: X-axis represents the amount of collateral required by the banks; Y-axis represents the level of interest rate.
Note 2: Figure 4.1 is based on Stiglitz and Weiss (1992) graphical representation.

Amongst other characteristics such as firm's size, the bank is able to observe the firm's wealth. There are two types of firms identified based on their amount of wealth: rich firms ($W_r$) and poor firms ($W_p$). The wealth differential amongst the firms impacts the amount of collateral that is offered for a loan. The collateral differential implies that the more wealthy firms are able to provide more collateral compared to the poor firms:

$$W_p < W_r \Rightarrow C_p < C_r. \quad (14)$$

The wealth differential and the assumption of decreasing absolute risk aversion imply the switch line for the rich firm to lie below the poor firm switch line. Based on the placement of the rich and poor switch lines, there are three regions identified that examine the possible loan contracts issued by the bank (Figure 4.2). In region X, the rich and poor firms use only risky techniques; in region Z, both firm types use only the safe technique; and in region Y, the poor firms use the safe technique and the rich firms employ the risky technique. The indifference curve of the rich firms through any point in regions X and Z is always flatter than that of the poor. The flatter indifference curve implies that the rich firms need lower interest rate reduction to compensate them for any increase in the amount of required collateral. In region Y, however, where the rich use the riskier technique and the poor the safer one, the indifference curve of the rich is steeper than that of the poor (Stiglitz & Weiss, 1992).

Figure 4.2: Loan Contract Regions: X, Y, and Z




Note1: Collateral is fixed at $C_p$ (the minimum amount of collateral for poor firms); $x$ represents a fraction of those borrowing at $C_p$; at $r'$ the rich firms switch to risky projects; at $r^*$ the poor firms switch to risky projects; and above $r_{max}$ no firms borrow; $P(C_p)$ represents the probability of reaching the minimum collateral level required by banks..

Note2: Figure 4.2 is based on Stiglitz and Weiss (1992) graphical representation.

#### 4.2.1.2 Lenders

In the Principal - Agent Problem, lenders often represented by banks are assumed to be risk neutral. If a firm employs an investment technique *m* to execute on its project, the expected return to the bank (often referred to as profits) represented by *v* is:

$$v = p^m(1 + r) + (1 - p^m)C \qquad (15)$$

where $p^m$ represents the success probability of chosen investment technique. For a given project, the bank's iso-return curve (otherwise known as the iso-profit curve) is a straight line. Regardless of whether the individual firm uses the safe or the risky technique, the borrower's indifference curve is always steeper than the bank's iso-return locus. The reason for the disparity in slopes is the borrower's risk aversion:

$$dr/dC = -(1 - p^m)/p^m \qquad (16)$$

which is always less than the slope of the indifference curve given by

$$-dr/dC = U_0'(1 - p^m)/U_1' p^m \quad (17)$$

since $Y_1 > Y_0$ and $U'' < 0$ (Stiglitz and Weiss, 1992).

#### 4.2.1.3 The Principal-Agent Problem Equilibrium

Banks know that the credit market consists of poor and rich borrowers, but due to asymmetric information on their side, they cannot ascertain each firm type. On the other hand, they know that the firm type can be revealed by choice of credit contracts the borrowers apply for. This concept represents the market's signaling and screening solution to the problem of incomplete information. For example, an applicant trying to obtain a contract with collateral requirements in excess of $C_p$ (amount of collateral greater than the minimum requirement of collateral for poor firms) must be a rich firm (Stiglitz and Weiss, 1992).

In a competitive equilibrium, explained in detail in Appendix A, it is assumed that each bank earns zero profits. This implies that banks cannot increase profits by offering a different mix of contracts and/or changing the number of loans issued. If any bank made positive profits, some other bank would offer a slightly more attractive set of contracts and therefore attract all of the borrowers previously getting loans elsewhere without changing their choices. As a result in a competitive equilibrium, the rate of interest paid depositors must be equal to the expected return on a loan contract (Stiglitz & Weiss, 1992).

##### 4.2.1.3.1 Equilibrium Conditions

In the Principal – Agent Problem equilibrium, each lender supplies certain types of loans and firms demand certain types of loans to complete their projects. Table 4.2 describes the supply and demand for loan contracts in the market equilibrium.

| Table 4.2: Equilibrium Conditions | |
|---|---|
| Supply of Loans ($L(v_z\{K\})$) | |
| Poor Firm Loan Contract *{F}* | Rich Firm Contract *{G}* |
| Collateral: $C = C_p$<br>the minimum amount of collateral required for poor firms<br><br>Interest rate: $r = r^s{}_p$<br>the highest interest rate consistent with the poor using the safe technique | Collateral: $C = C_r$<br>the minimum amount of collateral required for rich firms<br><br>Interest rate: $r = r^s{}_r$<br>the highest interest rate consistent with the rich using the safe technique |
| Demand for Loans<br>($N = N_p + N_r$) | |
| $N_p$ denotes the poor firms' demand in the credit market. | $N_r$ denotes the rich firms' demand in the credit market. |

The credit market equilibrium is defined as the maximum return obtained on the loan by a lender subject to the expected utility level of the borrowers. For example, if $v_z\{K\}$ defined by either a contract *{F}* or a contract *{G}* is the expected return to the bank on a contract *{K}* obtained by proportion of poor and rich borrowers presented by $z = N_p/N$, the bank's optimization problem is defined as follows:

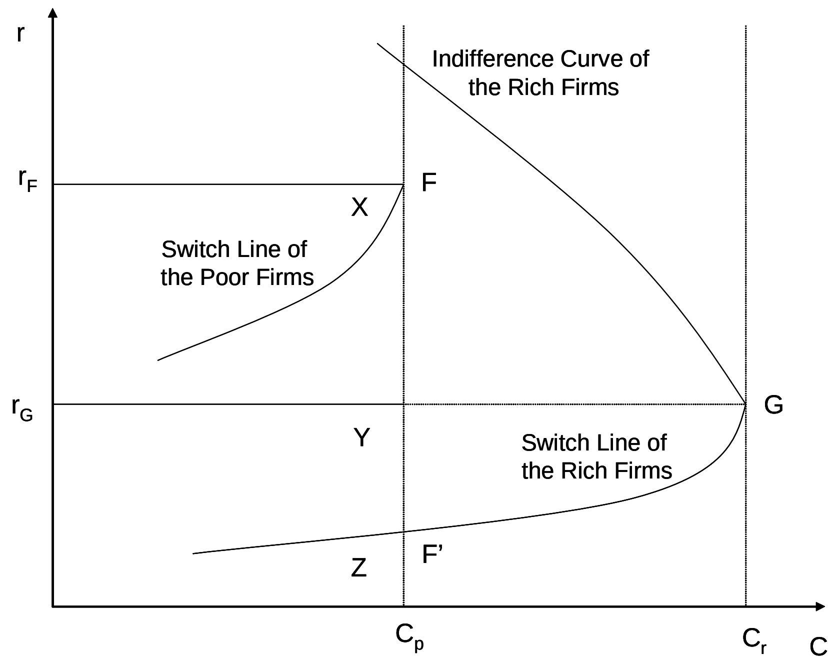
$$Max\ v_z\{K\} = N_p[\ p^s(1 + r^s{}_p) + (1 - p^s)C]\ + N_r[\ p^s(1 + r^s{}_r) + (1 - p^s)C] \qquad (19)$$

$$Subject\ to\ EU^l\{K\} = U_1(W + R - (1 + r_l))p + U_0(W - C))(1\text{-}p). \qquad (20)$$

Two types of loan contracts are offered by lenders and differ in terms of the interest rates as well as collateral amount. The collateral requirement is always higher at *{G}* than at *{F}*, so this type of contract is attainable only by the wealthy firms. A contract *{G}* therefore, increases the bank's expected loan return due to the higher collateral requirement. On the other hand, the interest rate may be lower due to the rich borrower's switch lines lying just below that of the poor line (Figure 4.3). Consequently, increases in rich firms' collateralizable wealth have two effects: the switch line is shifted down decreasing the bank's profits while higher amount of collateral directly increases the bank's profits. The net effect on the bank expected return is ambiguous and depends on the extent to which the borrower's (absolute) risk aversion decreases with wealth. The return also depends on the relative differences between acquired wealth of the rich and poor (Stiglitz & Weiss, 1992).

As noted earlier in this section, each bank can issue only one loan contract, which allows for equilibriums with and without rationing. Either result may be characterized by complete pooling where all rich and poor borrowers receive the same contract or by partial separating where some rich borrowers receive loans at different contract terms from those received by the poor (Stiglitz & Weiss, 1992).

Figure 4.3: Credit Market Equilibrium Conditions




Note 1: *{F}* is the optimal contract to the poor firms; *{G}* is the optimal contract to the rich firms.

Note 2: X-axis represents the amount of collateral required by the banks; Y-axis represents the level of interest rate.

Note 3: Figure 4.3 is based on Stiglitz and Weiss (1992) graphical representation.

#### 4.2.1.1.2 Interior Pooling Equilibrium

Suppose that at contract *{F}*, the contract preferred by the poor firms, the supply of loan contracts exceeds the firms' demand for loanable funds. In this case, the lenders compete for firms willing to borrow. As a result, the banks reduce the collateral requirements as well as the interest rate level such that they continue to provide incentives to the poor firms by offering a loan contract southwest of the contract *{F}* and along the poor firm switch line in Figure 4.3 (Stiglitz & Weiss, 1992).

Assume that *{H}* denotes a contract located along the poor switch line at which the demand for loans (*N*) equals the credit supply ($L(v_z\{H\})$):

$$L(v_z\{H\}) = N. \quad (21)$$

If at a contract *{H}* the loan return is higher than at a contract *{G}*, and if the indifference curve of the poor firms through *{H}* is steeper than the indifference curve of the rich firms, then all firms receive loans at the contract *{H}*. In order to meet the equilibrium criteria, the poor firms' indifference curve through *{H}* is located below the switch line in the region Y. The equilibrium described above represents a pooling equilibrium where all rich and poor firms obtained the same type of loan. Table 4.4 displays the conditions necessary for the interior pooling equilibrium existence (Stiglitz & Weiss, 1992).

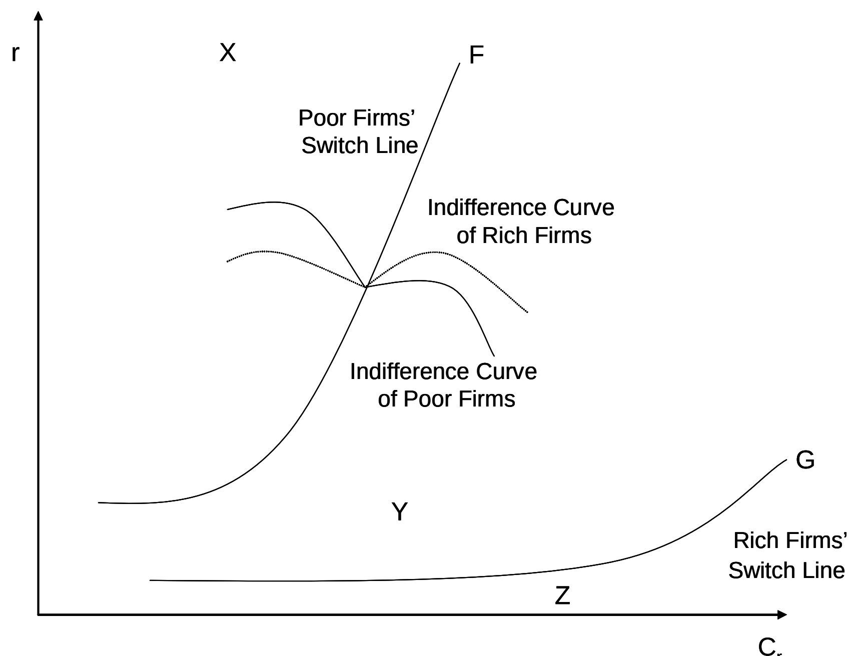
Table 4.3: Interior Pooling Equilibrium Conditions

| | |
|---|---|
| Loan Supply and Demand | Excess Supply:<br>$L(v_z\{F\}) > N$<br>and<br>$L(v_z\{H\}) = N$ |
| For all interest rate and collateral combinations: *{r, C}* | Utility Comparison:<br>$U^p\{r, C\} > U^p\{H\}$ |
| | Infeasibility of profitable separation:<br>Even though at $\{H\}$ $C < C_p$ ,<br>$U^r\{r, C\} > U^r\{H\}$ |
| | Profitability:<br>$v\{H\} > v\{G\}$ |

Under the predetermined conditions concerning the slopes of the indifference curves whenever a contract *{F}*, in the bank's expected return, is preferred to *{H}* by the poor firm, due to its position below the poor's indifference curve through *{H}*, it is also preferred by the rich firms as it lies below their indifference curve through *{H}* (Figure 4.4). Because of decreasing absolute risk aversion, the indifference curves of the rich are flatter than those of the poor in region Y, even though the rich are choosing the riskier technique (Stiglitz & Weiss, 1992).

With full-information, the lender knows exactly the firm type applying for a loan, and for any borrower, the interest rate charged is an increasing function of the amount borrowed. The firm borrows more and therefore, increases the likelihood of the interest rate increase. This phenomenon occurs if the interest rate and loan size combination serve as the self-selection device for firms applying for a credit line (Stiglitz & Weiss, 1992).

Figure 4.4: Interior Pooling Equilibrium




Note 1: X-axis represents the amount of collateral required by the banks; Y-axis represents the level of interest rate.

Note 2: Figure 4.4 is based on Stiglitz and Weiss (1992) graphical representation.

#### 4.2.1.1.3 Pooling Equilibrium with Rationing

Figure 4.3 presents an example of a pooling equilibrium with rationing that exists at a contract *{F}*. This contract requires a collateral amount of at least $C_p$ (the minimum amount of collateral required for poor firms) and offers the highest interest rate at which the poor borrowers invest in the safe project. At the equilibrium contract *{F}*, the supply of loan contract is lower than the demand and therefore, the excess demand for contract *{F}* is observed:

$$L(v_z\{F\}) < N\,. \quad (22)$$

At a contract *{F}*, the bank's profits accounted for by a fraction of the poor population equal to or greater then *z* are higher compared to any other contract employing a risky investment technique (Stiglitz & Weiss, 1992):

$$v_{(1\text{ - }z)}\{F\} = p^r(1 + R^r) + (1 - p^r)C < p^p(1 + R^p) + (1 - p^p)C = v_z\{F\}. \qquad (21)$$

Table 4.4 represents the necessary conditions to obtain the pooling equilibrium with rationing.

Table 4.4: Pooling Equilibrium with Rationing Conditions

| | |
|---|---|
| Loan Supply and Demand | Excess Demand: $L(v_z\{F\}) < N$ |
| For all interest rate and collateral combinations: *{r, C}* | Profitability: $v_p(F) > v_r(G)$ |

Based on the market equilibrium conditions presented in Table 4.4 pooling of rich and poor firms at a contract *{F}* is an equilibrium and the supply of funds is $L(v_z\{F\})$. Hence, rationing at contract *{F}* occurs if, and only if, there is excess demand at a contract *{F}* (Stiglitz & Weiss, 1992).

#### 4.2.1.1.4 Pooling - Separating Equilibrium with Rationing

The rationing equilibriums include those types in which some of the rich and poor firms borrow at the same terms while remaining rich borrowers accept contracts not chosen by any poor borrowers. Suppose the conditions represented in Table 4.5:

| Table 4.5: Pooling - Separating Equilibrium with Rationing Conditions | |
|---|---|
| Loan Supply and Demand | Excess Demand : $L(v\{G\}) < N$ |
| For all interest rate and collateral combinations: $\{r, C\}$ | Utility Comparison: $EU^r\{G\} > EU^r\{F\}$ |
| | Profitability: $v_z\{F\} < v_r\{G\} < v_p\{F\}$ |

Based on the assumption, including the excess demand at a contract *{G}*, there is a rationing equilibrium in which some banks offer contract *{F}* and others offer contract *{G}*. All of the rich borrowers apply for loans at every bank that offers either a contract *{G}* or *{F}*. A rich borrower only accepts a contract *{F}* offer when not offered a *{G}* loan. In equilibrium, the number of rich borrowers receiving *{G}* loans, $N_G$, is such that the proportion of poor borrowers, *x*, among those accepting *{F}* loans satisfies the following equality:

$$v_x\{F\} = v_p\{G\}. \qquad (23)$$

Equation (23) implies that both contract types offered are equally profitable (Stiglitz & Weiss, 1992). By continuity assumption and the observed excess demand for a contract *{G}*, there always exists a value of *x* satisfying equation (23). When the number of *{F}* loans offers made is $N_F$, the total supply of loans is equal to the total loan demand:

$$L(v\{G\}) = N_G + N_F. \qquad (24)$$

Following further analysis, if $L(v\{G\}) > N_G$, $N_F > 0$, then both *{G}* and *{F}* loans are offered and rationed (Stiglitz & Weiss, 1992).

There is also the possibility of establishing the true separating equilibrium with rationing only at a contract *{G}*. In this equilibrium, most but not all of the wealthy firms are offered this type of loan. The separating equilibrium with rationing exists at a contract *{G}*, if $L(v\{G\}) < N_G$ (the demand for a contract *{G}* exceeds the supply of that contract) and therefore only a contract *{G}* is offered. While both rich and poor borrowers are rationed at contract *{G}*, the rich borrowers have a higher chance to obtain loans than are the poor borrowers as they are more likely to meet the loan criteria required by the lenders (Stiglitz & Weiss, 1992).

#### 4.2.1.1.5 Other Types of Separating Equilibrium with Rationing

There may also exist equilibrium types with rationing at one contract but not at another. Suppose there exists a contract *{H}*, lying along the rich firms' switch line. The collateral requirement falls in between the minimum collateral required for receiving a rich firm loan and the poor firm loans: $C_r > C\{H\} > C_p$. This implies that the lender's expected return on a contract *{H}* is equal to the expected return on a contract *{F}*. At the same time, the borrowers prefer a contract *{F}*. Finally, there is excess supply for contract *{F}* (see Table 4.6) (Stiglitz & Weiss, 1992).

Table 4.6: Other Types of Separating Equilibriums with Rationing Conditions

| Loan Supply and Demand | |
|---|---|
| | Excess Demand : $L(v_z\{F\}) > N^r$ |
| For all interest rate and collateral combinations: $\{r, C\}$ | Utility Comparison: $EU^r\{G\} < EU^r\{F\}$ |
| | Profitability: $v\{H\} = v_z\{F\}$ |

Based on the equilibrium conditions listed in Table 4.6, there exists an equilibrium in which all borrowers apply for loans at *{F}* and the rejected rich borrowers borrow at *{H}*. Accordingly, there is rationing at *{F}* but there is no rationing at *{H}* as any rich borrower that gets a loan at *{F}* takes it (Stiglitz and Weiss, 1992).

### 4.2.2 Efficient Market Hypothesis

The second theorem guiding this article is the Efficient Market Hypothesis. It demonstrates that it is impossible to "beat the market" because market efficiency causes existing investment prices to always incorporate and reflect all relevant information with regard to the market conditions. Any investing is always performed at the fair market value. It is impossible to outperform the overall market through market timing, and the only way an investor can obtain higher returns is by purchasing riskier investments (Fama, 1970).

Although the Efficient Market Theory has been widely employed in economics, in finance, several empirical studies including those conducted by Baker et al. (2003), Bosworth (1971), and Malkiel (2004) have found support for profiting from market timing, at least in the short-run. For example, according to the findings of Baker et al. (2003), firms issue long-term debt when bond returns are low and short-term debt when returns are high. On the other hand, debt maturity is related to market conditions such as the interest and inflation rates as found in Bosworth (1971). The authors explain the above evidence of lowering capital costs as managers timing an inefficient capital market using public information to guide their debt maturity decisions (Baker et al.; 2003; Bosworth, 1971).

In this section, the Efficient Market Hypothesis is introduced in which marginal investment return over cost is equal to the prevailing market interest rate. This finding is in an agreement with the General Market Equilibrium Framework described in detail in Appendix A. On the other hand, the Market Timing Hypothesis is addressed as a result of the differential in the optimal investment solutions between the profit-maximizing and utility-maximizing results. The solution discrepancy might be a result of a manager's individual preference system affecting the optimal investment and financing solution. The manager's views might not necessarily reflect the optimal production opportunities faced by the firm for example, resulting in different optimal investment decisions. Often the disparity in solutions is observed in the short-run models and if any positive profits are made, they are minimized to zero in the long-run (Rubinstein, 2002).

#### 4.2.2.1 The Short-Run Efficient Market Hypothesis

##### 4.2.2.1.1 Profit Maximization

Let's consider an example in which a firm makes investment decisions that maximize their expected profits over several time periods or probability states. For simplicity, there are only two time periods, $t = 0$ and $1$, and the investment decision is inter-temporal. Suppose a firm has an initial endowment (earlier referred to as wealth) of inter-temporal output $W$, with output in period 0 being positive ($W_0 > 0$) and output in period 1 being equal to zero ($W_1 = 0$). As a result, there is only an endowment in period 0. The amount of investment involves allocating some amount of period 0 endowment to production for period 1. The excess output in period 0 represents the period 0 consumption ($Y_1$*) and is the amount of initial wealth that investment has not appropriated (Schwert, 2001):

$$Y_0^* = W_0 - I_0^*. \qquad (25)$$

The optimal investment decision ($I_0$*) is obtained by maximizing the firm's profits. The profit maximization concept is introduced in Appendix A. In a world with only two time periods, $t = 0$ and $1$, the firm's profit-maximization problem is defined as follows:

$$Max\ \Pi = R - TC = f(I_0) - (1+r^m_l)\, I_0 \quad (26)$$

where investment $(I_0)$ in period 0 yields output $(Y_1)$ in period 1:

$$Y_1 = f\,(N, I_0) \quad (27).$$

In the above maximization problem, $\Pi$ are the profits, $R$ is the firm's investment return on the project derived by $pY_1$ ($p$ is the output price normalized to one, $p = 1$), $I_0$ is period 0 investment (collateral), $Y_1$ is period 1 output (in this it also represents an investment frontier), $r^m_l$ is the interest rate, $N$ is the amount of labor. Holding $N$ constant implies a concave investment frontier defined as $Y_1 = f\,(N, I_0)$ with increasing marginal product that increases at decreasing rate, $f' > 0$ and $f'' < 0$. With the rate of interest, $r^m_l$, the firm's total costs ($TC$)[40] of investing an amount $I_0$, is equal to $(1+ r^m_l)\, I_0$ (Schwert, 2001).

[40] In the earlier section, total costs were normalized to one. For the description of the Efficient Market Hypothesis, the total costs are not set to one.

The investment decision is optimal where the investment frontier is tangent to the interest rate line represented by - ($1 + r^m{}_l$). The first order conditions for the optimal investment decision are derived by taking the first derivative of the maximization problem with respect to the $I_0$ (see Figure 4.5):

$$f'(I_0) = (1 + r^m{}_l) \qquad (28).$$

Based on the optimal investment decision ($I_0*$) found by equation (28), the inter-temporal allocation of output is defined as (Schwert, 2001):

$$Y_0* = W_0 - I*_0 \qquad (29)$$

$$Y_1* = f(I*_0) \qquad (30)$$

The first order condition referenced by equation (28) implies that

$$f'(I_0) - 1 = r^m{}_l \qquad (31)$$

where $f'(I_0) - 1$ is defined as the marginal rate of return over cost, which is also known as marginal efficiency of investment (*MEI*). Thus, the optimum condition for the firm's investment decision is (Schwert, 2001):

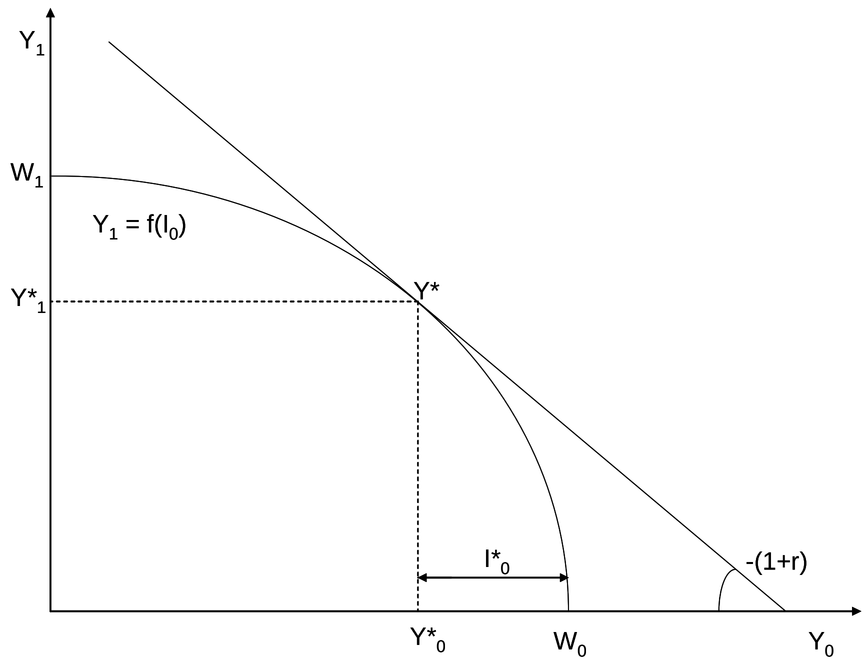
$$MEI = r^m{}_l \qquad (32)$$

implying the *MEI* is equal to the rate of interest. As the rate of interest rises, the investment declines and therefore implies a negative relationship between investment and the interest rate (Schwert, 2001):

$$I = I(r^m{}_l) \text{ where } I_r = dI/d\, r^m{}_l < 0. \qquad (33)$$

This finding represents that the investment return should be always equal to its cost (Schwert, 2001).

Figure 4.5: Firm's Profit Maximization Subject to Investment Frontier




### 4.2.2.1.2 Utility Maximization

Now, suppose that the firm is governed by a firm manager who is endowed with a utility function of $U(Y_{0,} Y_1)$. The investment frontier is defined as $Y_1 = f\,(I_0)$ and the manager maximizes his utility with respect to the inter-temporal investment frontier. The manager is making trade-offs between consumption in period 0 and period 1 that yield the highest utility level. The utility maximization defined in Appendix A is as follows:

$$Max\ U(Y_{0,} Y_1) = U_0(Y_0) + U_1(Y_1)\,/(1 + \rho) \qquad (34)$$

$$Subject\ to\ Y_1 = f\,(I_0) \qquad (35)$$

where $\rho$ is the individual subjective rate of time preference and consumption is defined as $Y_0 = W_0 - I_0$. The first order conditions are derived based on differentiating the utility function with respect to $I_0$:

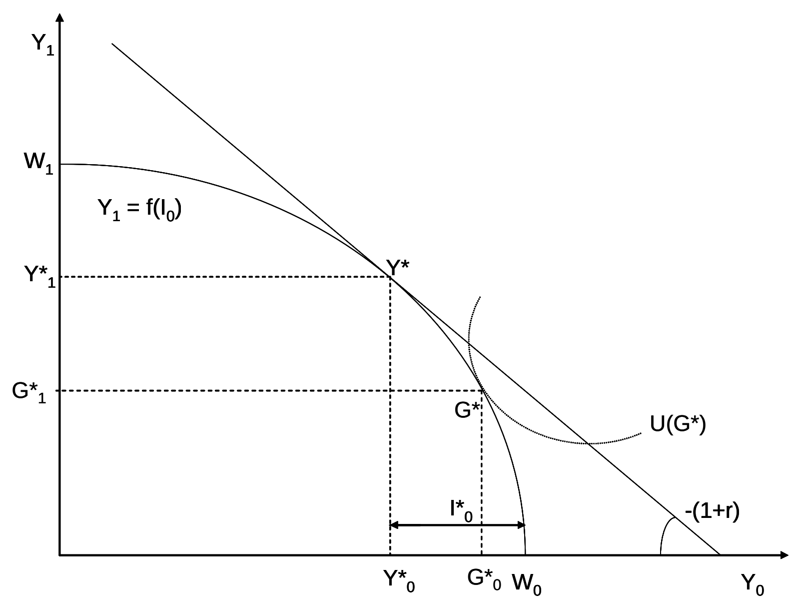
$$-U'_0 + U'_1\, f\,'(I_0)/\,(1+ \rho) = 0 \qquad (36)$$

$$U'_0\,(1 + \rho)\,/\,U'_1 = f\,'(I_0). \qquad (37)$$

Based on the first-order conditions defined in equation (37), the optimal solution is represented by a point $G^* = (G^*_{0,} G^*_1)$ in Figure 4.6. At this point the ratio of marginal inter-temporal utilities is equal to the marginal product of investment. In other words, the Marginal Rate of Substitution of consumption in period 0 and period 1 is equal to the Marginal Rate of Transformation of investment in period 0 and period 1[41] (Rubinstein, 2001).

When the utility maximization solution at point $G^*$ is compared to the profit maximization solution of point $Y^*$, it is obvious that these two solutions are different. In this case, the optimal investment decision of the firm is affected by the manager's time preferences for period 1 investment as manager's utility maximization does not yield the same result as the firm's profit maximization problem (Rubinstein, 2002).

Figure 4.6: Utility Maximization Subject to Investment Frontier




In the final scenario, each firm employs a two-stage budgeting process involving first maximizing the firm's present value of profits, which results in optimal solution at

[41] The definitions of Marginal Rate of Substitution and Marginal Rate of Transformation are provided in Appendix A.

point $Y^*$, and then either borrowing or lending by the firm to obtain the manager's optimal solution presented at point either $C^*$ or $F^*$ displayed in Figure 4.7. The following manager's utility maximization is performed subject to the period 1 budget line:

$$Max\ U(Y_0,\ Y_1) = U_0(Y_0)+U_1(Y_1)/(1+\rho) \qquad (38)$$

$$Subject\ to\ Y_0 = W_0 - I_0\ and\ Y_1 = (1+ r^m{}_l)I_0 \qquad (39)$$

The first-order conditions of the utility maximization problem are derived with respect to $I_0$:

$$-U'_0 + U'_1(1+ r^m{}_l)/(1+\rho) = 0 \qquad (40)$$

The optimal investment decision $I_0$ is derived as follows:

$$U'_0\,(1+\rho)/U'_1 = (1+ r^m{}_l) \qquad (41)$$

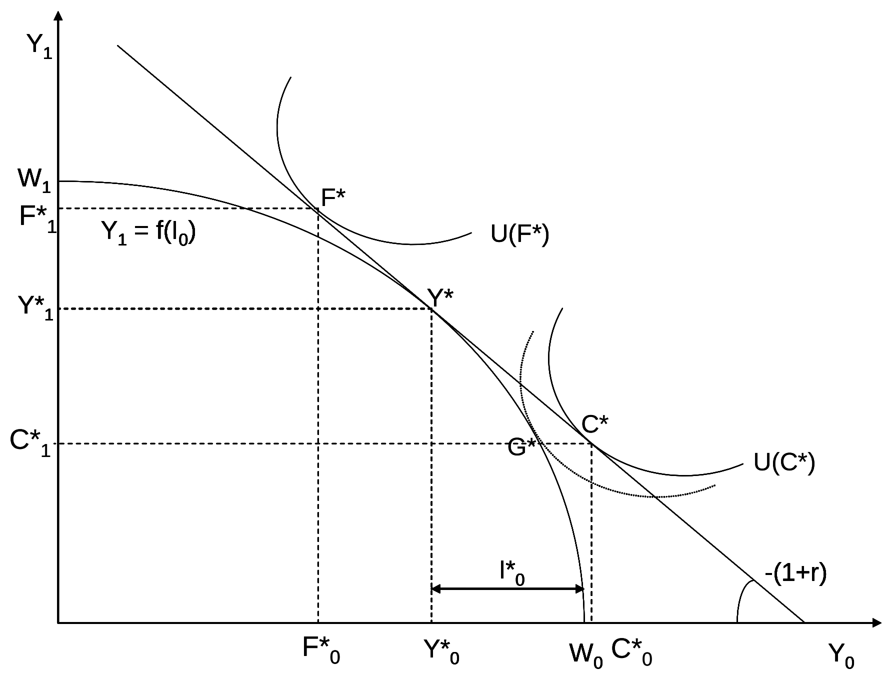
As a result, equation (41) implies that:

$$U'_0(1+\rho)\,/U'_1 - 1= r^m{}_l \qquad (42)$$

The solution in equation (42) implies that the optimal marginal return to investment is equal to the interest rate. The location of the optimal solution depends on the preferences of the manager denoted by the individual subjective rate of time preference (being more or less patient). This implies that the optimal investment decision at point $G^*$ (derived in equation (37)) is not optimal. Furthermore, the optimal investment decision differs when only the "productive opportunities" (point $Y^*$) are taken into account compared to the investment decision that incorporates the "market opportunities" denoted by the market interest rate that leads to optimal solutions at points either $C^*$ or $F^*$ depending on manager's preferences (Rubinstein, 2002).

As mentioned earlier, the above model is not a General Equilibrium Model (long-run model), but a short-run model. Furthermore, the optimal marginal return to investment is equal to or higher than the market determined interest rate. The manager determines the optimal investment by investing in projects with a higher return than the market interest rate. The decision is partly driven by the manager's individual preferences that yield the highest utility. This implies that the project cost of capital must be equal to the market rate. So, if the cost of capital is less than the return of capital, firms will earn excess profits. This is not sustainable in the long-run. In a long-run equilibrium model, the cost of capital has to be equal to the return of capital (Mathiesen, 2008).

Figure 4.7: Utility Maximization Subject to Budget Line in Period 1




### 4.2.2.2 The Long-Run Efficient Market Hypothesis

In the Long-Run Efficient Market Hypothesis, both the manager's utility and firm's profitability is taken into consideration at the same time. This model demonstrates that utility maximizing and perfectly rational mangers pursue the profit-maximizing strategy for firm. Both of the inter-temporal maximizations are taken into account. Based on the firm's maximization problem stated in equations (26) and (27), the firm's inter-temporal profit maximization first-order condition is:

$$f'(I_0) - 1 = r^m{}_l \qquad (43)$$

where the marginal rate of return over cost is equal to the market interest rate. Based on the manager's utility maximization problem stated in equations (38) and (39), the manager's inter-temporal utility maximization first-order condition is:

$$U'_0(1 + \rho)/U'_1 - 1 = r^m{}_l \qquad (44)$$

where the marginal utility of efficient investment is equal to the market interest rate. This implies that in a competitive equilibrium[42], both the firm's profit maximization and the manager's utility maximization result in the same optimal investment decisions, which depend on the market interest rate:

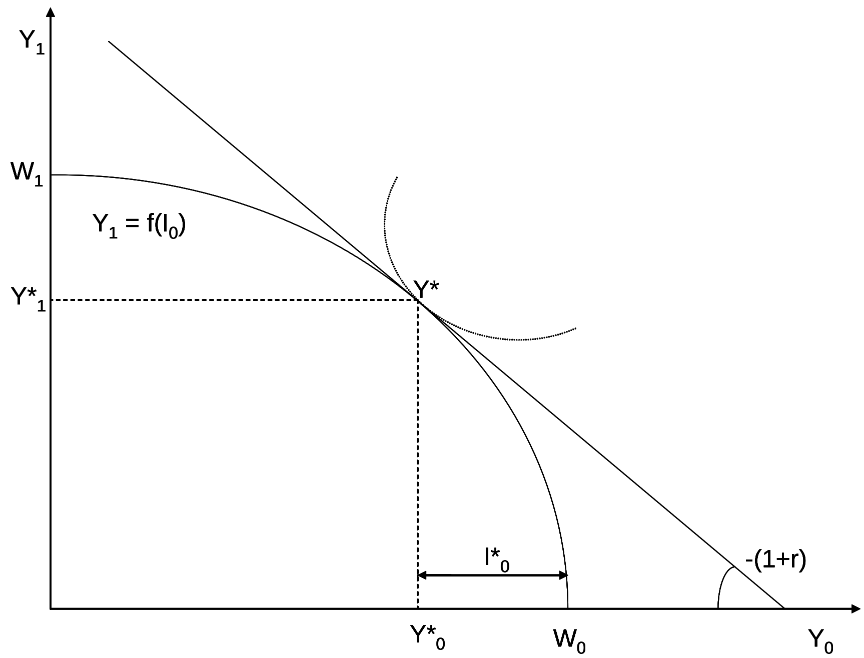
$$f'(I_0) = U'_0(1 + \rho)/U'_1 = r^m{}_l \qquad (45)$$

Based on the above assumption, the optimal investment decision is at point *Y** where the investment frontier, the period 1 budget line, and the manager's indifference curve are tangent to each other. Thus, in the Efficient Market Equilibrium the market interest rate equals the marginal rate of substitution as well as equals the marginal rate of capital. As a result, it is impossible to time the market and make profits as the investment returns are equal to the costs (Rubinstein, 2002).

This model also demonstrates an example of the Principal-Agent Problem where the firm's owners force the manager of the firms they own to pursue the profit-maximizing strategy instead of his own strategy that maximizes his profits. Again the solution of this problem is the equality between the marginal product of investment, marginal inter-temporal utility, and the prevailing market interest rate (Mathiesen, 2008).

[42] The general equilibrium definition is provided in Appendix A.

Figure 4.8: The Efficient Market Hypothesis Equilibrium




## 4.3. Literature Review

The notion of the corporate debt maturity structure is explained by the determinants of debt structure and market conditions. All of the hypotheses tested here are derived based on the Principal-Agent Theory and the Efficient Market Hypothesis. The hypotheses employed in this article include: Signaling and Agency Costs Hypotheses (all based on the Principal-Agent Theory), as well as Market Timing Theory (based on the Efficient Market Hypothesis). The following empirical analysis review represents the empirical evidence either for or against the implications resulting from employment of the proposed theoretical framework when analyzing the financial decisions of corporate companies.

Several terms should be defined before in depth discussion of the empirical issues related to the cost of capital borrowing and corporate debt maturity structure. Table 4.7 defines finance terminology needed for the article understanding.

Table 4.7: Glossary of Finance Terminology

| Terminology | Description |
|---|---|
| Share of Long-Term Debt to Total Debt | The percentage of the firm's total debt that has a maturity of more than three years |
| Excess Bond Returns | An index of investment-grade corporate bonds over commercial paper |
| Abnormal Earnings | The difference between next year's and this year's earnings per share scaled down by this year's share price |
| Market-to-Book Ratio | The estimated market value of assets divided by the book value assets |
| Real Short-Term Rate | The annualized Treasury bill return minus inflation |
| Term Spread | The difference between the Treasury bond yield and the annualized Treasury bill return |
| Inflation | The annual percentage change in the Consumer Price Index |
| Debt Market Timing | Responsiveness to the changes in macroeconomic variables |

According to Barclay and Smith (1995), debt determinants are part of the agency cost and signaling hypotheses derived from the Principal – Agent Theorem. As discussed earlier, lenders are not able to observe directly firms' behavior. As a result, specific loan contract terms help banks to differentiate between different types of borrowers while observing firms' observable characteristics such as size and wealth. Based on the findings of the Principal – Agent Theory, the Agency Cost Hypothesis states that risky debt financing may enlarge the suboptimal investment incentives when a firm's investment opportunity set includes growth options. Managers undertake underinvestment[43] decisions controlled by issuing a short-term debt maturing before the firm's growth opportunity is exercised (Myers, 1977). Consequently, the agency cost hypothesis states that firms with larger growth options in their investment opportunity set issue more short-term debt.

Studies by Antoniou et al. (2002), Jun and Jen (2003), and Chen et al. (1999) find firms with larger growth options issuing more short-term debt. Antoniou et al. (2002) argue that the cost of financial distress of high growth firms is relatively high and therefore mangers are reluctant to raise debt capital (Antoniou et al., 2002). Furthermore,

[43] **Underinvestment** –unwillingness of shareholders to take a positive net present value project when the profits will accrue mainly to the creditors (Foo and Yo, 2005)

according to Jun and Jen (2003), firms with greater financial flexibility and financial strength are more likely to issue more short-term debt. On the other hand, Harvey et al. (2001) also find that closely monitored debt creates shareholder value, because it reduces the agency costs associated with overinvestment. The lower costs of debt financing are a result of a long-term relationship established between the firms and the creditors (Antoniou et al., 2002).

According to Barclay and Smith (1995), firms with relatively large numbers of investment opportunities tend to be smaller and take on short-term debt to avoid the underinvestment decisions and paying higher issuance costs associated with long-term debt. Issuing short-term debt avoids the above problem as the price at which the firm repurchases its debt is fixed. Issuing a short term debt allows the stockholders to capture most of the returns from the new investments (Myers, 1977). Additionally, banks have a comparative advantage in monitoring and maintaining a stronger bargaining position when loaning short-term. Banks' monitoring is especially important for firms with large informational asymmetries as these firms use more bank loans with short-term maturity (James, 1987).

In accordance with the above hypothesis, larger firms issue debt with maturities on average longer by about 6 years (FEMA, 2006). However, Dennis and Sharpe (2005) find that as the borrower size increases, negotiating power with the lender and information transparency increase. The lender is able to spread the fixed costs of loan production across a larger dollar value of the loan and as a result, issue debt to larger firms (Dennis & Sharpe, 2005).

Finally, regulated firms issue longer debt maturity than unregulated firms. This follows from the fact that managers of regulated firms have less discretion over future investment. Lower level of managers' judgment over future investments lowers the adverse impact of the long-term debt. Consequently, a longer maturity period is preferred by the regulated companies (Barclay & Smith, 1995). In summary, it is predicted that small and unregulated firms with more growth options have less long-term debt.

In the Signaling Hypothesis, the pricing of long-term debt is more sensitive to changes in firm value compared to the short-term debt. For example, high-quality firms defined as firms with positive abnormal earnings[44] tend to issue more short-term debt compared to low-quality firms defined as firms with negative abnormal earnings due to lower refinancing and mispricing[45] costs (Barclay & Smith, 1995). Richardson and Sloan (2003) as well as Antonenko et al. (2006) find empirical evidence for low-quality firms obtaining overpriced long-term debt (higher interest rate than usual put on long-term debt). Richardson and Sloan (2003) further find that long-term debt follows a decreasing stock return period. On the other hand, Ozkan (2002) and Antoniou et al. (2002) find no empirical support for the signaling hypothesis. The volatility of earnings is found not to have a significant impact on the capital structure among the studied firms. Although not confirmed by all empirical studies, it is expected that high-quality firms (firms with positive abnormal earnings) issue short-term debt. Consequently, there is a negative relationship between the corporate debt maturity and the quality level of the firm (Barclay & Smith, 1995).

According to Baker et al. (2003) debt maturity is also related to macroeconomic variables such as inflation, interest rates, term spread, and excess bond returns, although the impact of macroeconomic variables on debt maturity has not been examined thoroughly in the literature. As predicted by other finance articles, the macroeconomics variables should be negatively related to the maturity of debt. For example, firms issue long-term debt when the excess bond returns are low (Baker et al., 2003). Furthermore, Stohs and Mauer (1996) find debt maturity to be negatively related to the term spread. Baker et al. (2006) additionally uncover that the term spread is positively related to future excess bond returns, which further implies that short-term debt is preferred when short-term rates are lower compared to long-term. Marsh (1982) as well as Graham and Harvey (2001), who also studied the impact of macroeconomics variables on debt maturity, find the amount of debt to vary with interest rates. The long-term debt is issued when interest rates are particularly low (Graham & Harvey, 2001).

[44] **Abnormal Earnings** - measured as the difference between next year's and this year's earnings per share scaled down by this year's share price (Barclay and Smith, 1995).
[45] **Misspricing –** investor's overestimation of the persistence of accruals and underestimation of the persistence of cash flow (Barclay and Smith, 1995).

The fact that debt maturity is responsive to the changes in macroeconomic variables is called the debt market-timing. Market timing implies the opportunity to obtain higher return on investment by simply observing and responding to the changes in macroeconomic variables. Baker et al. (2003), Baker et al. (2003), and Baker et al. (2006) find evidence for managers timing the debt market by utilizing publicly available information on market conditions as a guide to their debt maturity decisions.  For example, a large fraction of chief financial officers admit to following general debt market conditions while deciding on debt issuance decisions (Baker et. al, 2006).

The phenomenon of market timing is, however, viewed as a short-run model. As presented in the Efficient Market Hypothesis section, in the long-run, it is not possible to profit from market timing as all investments are traded at their fair market value (Fema, 1970). In accordance with the hypothesis, Berlin (2006) believes that managers cannot time the market, because they do not have any information on future interest rate movements and therefore are not able to forecast the interest rates accurately. As a result, if the managers do not have accurate forecasts, the short- and long-term borrowing should lead to exactly the same borrowing costs (Berlin, 2006).

## 4.4. Methodology

### 4.4.1. Empirical Model Description

The debt-maturity decision is modeled based on the Principal – Agent Theory and the Efficient Market Hypothesis.  The lender maximizes its expected returns from providing the investment loans to the borrowers subject to the borrowers' expected utility based on the chosen loan offer. The borrowers maximize their own expected level of utility provided by the chosen risk-level of the investment, the end-of-period wealth level, as well as the associated loan requirements set by the lender: interest rates and loan collateral. Consequently, the theoretical model guiding this empirical analysis is as follows:

**Lender:** $$Max\ v = \sum [p^m(1 + r^m_l) + (1 - p^m)C_l] \qquad (46)$$

**Borrower:** $$Subject\ to \sum EU^l = U_1(W + R - (1 + r^m_l))\, p^m + U_0(W - C_l))(1 - p^m) \qquad (47)$$

where $v$ represent lender's expected returns from providing a loan offer, $p^m$ is the of the borrower's probability of project success based on the chosen investment technique (risky or safe), $r^m{}_l$ represents the loan interest rate variable based on the firms wealth and risk level, $C_l$ is the loan collateral variable based on firms initial wealth, $U$ represents the borrower's utility level, $W$ is the firm's wealth level, $R$ is the firm's investment return on the project.

The optimal condition for deciding on the amount and maturity of debt depends on the probability of project success as well as the borrower's expected utility obtained from the project success:

$$p^m/(1 - p^m) = [U_0{}^{`}(1- p^m)] /[U_1{}^{`}p^m]^{46} \qquad (48)$$

Based on the Efficient Market Hypothesis, the borrower decides on the optimal investment decisions based on the results of the profit and utility maximization. The firm profit maximization is defined as:

$$Max\ \Pi = R - TC = f( I_0) - (1+ r^m{}_l)\ I_0 \qquad (49)$$

where $R$ is the firm's investment return on the project, $I_0$ is the investment level, $TC$ is the total cost of investing in the project, and $r^m{}_l$ represents the loan interest rate. The manager's utility maximization problem is defined as:

$$Max\ U(Y_0, Y_1) = U_0(W_0 - I_0) + U_1((1+ r^m{}_l)I_0) /(1 + \rho) \qquad (50)$$

where $W_0$ is the initial wealth and $\rho$ is the subjective rate of time preference. The first order condition derived by taking the first derivative of both the profit and utility maximization problems with respect to $I_0$ results in the following condition:

[46] The first order conditions with respect to interest rate and collateral are:

$$dl/d\ r^m{}_l = pm\ - \lambda[U_1{}^{`}p^m] = 0$$
$$dl/d\ C = 1- p^m + \lambda[U_0{}^{`}(1- p^m)] = 0$$
$$p^m/[U_1{}^{`}p^m] = 1- p^m/[U_0{}^{`}(1- p^m)]$$

$$f'(I_0) = U'_0(1 + \rho)/U'_1 = r^m_l \qquad (51)$$

Thus, in the Efficient Market Equilibrium the market interest rate offered for the loan ($r^m_l$) is equal to the marginal inter-temporal utility as well as it equals the marginal rate of return on investment.

Based on the maximization problems description, the debt-maturity decision is defined as a series of simultaneous decisions made by the firm and the lender. Equations (46) and (47) present that probability of investment success as well as the prevailing interest rate impact on the investment decision of the firms. As it is difficult to directly predict the firm's probability of project success, the lenders approximate the success probability by observing the firm specific characteristics and based on the information issue a loan contract geared towards the firms' needs. The loan conditions include cost of capital borrowing and debt maturity period. On the other hand, the firms might actively take into account changes in macroeconomics variables such as the interest and inflation rates when making their financing decisions.

As presented above, only a system of simultaneous equations can represent these joint decisions as banks offer a number of loans from which the borrowers choose their desired contract features based on the loans' characteristics as well as their own ability to pay off the debt in the future (Dennis et al., 2000). Financial theory does not provide the appropriate restrictions allowing for such analysis although several studies have attempted to estimate this simultaneous equation framework (Dennis et al., 2000). As a result, the debt-maturity model is a reduced-form regression based on the simultaneous equations system (Baker et al., 2003; Barclay & Smith, 1995).

The corporate debt-maturity model utilized in this article follows the model described by Barclay and Smith (1995) and Baker et al. (2003). As a result, the theoretical debt-maturity model is presented as follows:

$$LTD_{it} = f\,(MB_{it}, LMV_{it}, AE_{it}, EBR_{it}, TS_{it}, INF_{it}, STR_{it}, RD_i) \qquad (52)$$

where $LTD_{it}$ is the share of long-term debt to total debt, $MB_{it}$ denotes the market-to-book ratio, $LMV_{it}$ is the natural logarithm of market value of the firm, $AE_{it}$ is the abnormal earnings, $EBR_{it}$ is the excess bond returns, $TS_{it}$ is the term spread, $INF_{it}$ denotes the inflation, and $STR_{it}$ is the short-term interest rate. The subscript *i* denotes the *i*th firm (*i* = 1,…, 652), and the subscript *t* represents the *t*th year (*t* = 1991,…, 2000). $RD_i$ is the firm regulation dummy variable, which does not vary over time.

### 4.4.2. Empirical Data Tests

Following the model specifications used by Barclay and Smith (1995) as well as Baker et. al (2003), the debt-maturity model estimated in this analysis is as follows:

$$LTD_{it} = \alpha_i + \beta_1 MB_{it} + \beta_2 RD_i + \beta_3 LMV_{it} + \beta_4 AFE_{it} + \beta_5 EBR_{it} + \beta_6 TS_{it} + \beta_7 INF_{it} + \beta_8 STR_{it} + \boldsymbol{D'\lambda}_i + v_{it} \qquad (53)$$

where ***D'*** is a row vector of dummy variables created for each firm's month when the fiscal year ends, excluding the month of December and $\boldsymbol{\lambda}_i$ is a column vector of associated weights. The variable was added to control for the possibility of seasonal time heterogeneity in the model. Furthermore, it is assumed that the disturbance term in equation (2) is specified as a one-way error component model:

$$u_{it} = \alpha_i + v_{it} \qquad (54)$$

where $v_{it} \sim NIID\,(0, \ \delta_v^2)$; $i$ = 1,..., 652; $t$ = 1991,..., 2000; $\alpha_i$ denotes a firm-specific effect, and $v_{it}$ is the idiosyncratic error term (Hsiao, 2002).

In order to adequately specify the equation to be estimated, the debt-maturity model is tested for violation of assumptions such as: normality, heteroskedasticity, autocorrelation, and multicollinearity. The software utilized in this analysis is Stata 8. In order to test the normality assumption, the Bera and Jarque's LM skewness-kurtosis test for normality is used. The test is, however, sensitive to outliers and rejection of the null hypothesis provides no information about the alternative distribution to be used

(Wooldridge, 2002). The problem of heteroskedasticity is examined utilizing the Breusch-Pagan test with the null hypothesis being that the $u_{it}$ are serially uncorrelated. The null hypothesis is rejected for negative values of $\delta_{\alpha}^{2}$ (Wooldridge, 2002). Finally, testing the idiosyncratic errors $v_{it}$ for serial correlation identifies autocorrelation. Testing autocorrelation in panel data uses the time-series errors ($\ddot{v}_{it}$), which are found to be negatively correlated when the $v_{it}$'s are uncorrelated. The null hypothesis for this test is that the time-demeaned[47] errors are serially correlated. If serial correlation is found, the asymptotic variance matrix estimator and test statistics can be adjusted (Wooldridge, 2002).

### 4.4.3 Model Specifications

#### 4.4.3.1 Ordinary Least Square Estimation

There are several ways of estimating the debt maturity model. The Ordinary Least Square (OLS) regression is usually employed as it utilizes the actual values of the dependent variable. In order to allow for panel data estimation, OLS is often employed with either fixed or random-effects models (Berger et al., 2004; Hackethal & Jansen, 2006). The fixed- and random-effects models are discussed in the upcoming sections. On the other hand, other articles employ Tobit regression to estimate the debt maturity. As often the debt variable ranges between values of 0 and 1 (some firms might not have any long term debt), the data is truncated at the 0 point and exclude observations below this threshold from the analysis. Truncation might also occur when the observations are missing from the sample due to other sensitivity limits put on the variable of interest in the data set (Antoniou et al. 2002; Magri, 2006).

Although the Tobit model might seem like an appropriate method to estimate debt maturity data, several studies including Margi (2006) have criticized the utilization of Tobit model for panel data estimation. When Tobit model is employed in combination with random-effects, it may suffer from inconsistent estimates. The estimate inconsistency results from the absence of correlation between regressors and unobserved

[47] **Time-demeaned Data** - panel data where, for each cross-sectional unit, the average over time is subtracted from the data in each time period (Wooldridge, 2002).

individual effects being not satisfied. Although several fixed-effects methods have been proposed to resolve this problem, the estimation process still yields inconsistent results or introduces complicated estimation processes not easily handled by current statistical software (Magri, 2006). Due to the several estimation limitations of different model specifications, different panel data forms of the OLS regression are employed in this article (Antoniou et al. 2002; Magri, 2006).

### 4.4.3.2 Panel OLS Regression Types

According to Dennis et al. (2000), the employment of the reduced form of debt-maturity model estimated using OLS estimation yields unbiased results. There are several OLS estimation methods employed to estimate the relationship between debt maturity and its determinants. First, the pooled OLS estimation is utilized; however, there is a problem of unobserved error term dependence, which results in biased and inconsistent findings (Wooldridge, 2002). In order to correct for this issue, a cross-sectional OLS regression method is utilized in combination with the time-series mean of each variable for an individual firm. However, this method only accounts for the variation across firms without analyzing the time-variation (Barclay & Smith, 1995).

In order to investigate the time-series dispersion, both fixed and random effects regression models are employed. In the fixed effects regression, the cross-sectional variation is arbitrarily correlated with the explanatory variables. In the random effects regression model, it is assumed, on the other hand, that the cross-sectional variation is random and it is a part of the unobserved variation included in the error term. The Hausman specification test is used to test for the sufficiency of random effects estimation. Finally, in order to correct for the existing heteroskedasticity and serial correlation across panels, the generalized least squares estimation is used to obtain efficient estimators (Wooldridge, 2002).

#### 4.4.3.2.1 Ordinary Least Square Regression

The basic technique for estimating the debt-maturity model is pooled OLS estimation. To perform this procedure, the model is rewritten as follows:

$$LTD_{it} = \beta_0 + \beta_1 MB_{it} + \beta_2 RD_i + \beta_3 LMV_{it} + \beta_4 AFE_{it} + \beta_5 EBR_{it} + \beta_6 TS_{it} + \beta_7 INF_{it} + \beta_8 STR_{it} + \boldsymbol{D'\lambda}_i + u_{it} \qquad (55)$$

where $u_{it} = \alpha_i + v_{it}$; $v_{it} \sim NIID\ (0,\ \delta_v^2)$; $i$ = 1,..., 652; $t$ = 1991,..., 2000. The underlying assumptions of this model are: 1) the explanatory variables ($x_{it}$) in each time period are uncorrelated with the idiosyncratic error in each time period: $E(x_{it}'v_{it}) = 0$, $i$ = 1,..., 652, $t$ = 1991,…, 2000; and 2) the explanatory variables are uncorrelated with the unobserved effect in each time period: $E(x_{it}'\alpha_i) = 0$, $i$ = 1,..., 652, $t$ = 1991,…, 2000. The regression estimation provides consistent estimators as long as the underlying assumptions are satisfied (Barclay & Smith, 1995; Wooldridge, 2002).

#### 4.4.3.2.2 Cross-Sectional Regression with Time-Series Mean

Since the OLS assumption of serially uncorrelated composite errors is not satisfied in the pooled OLS specifications, the t-statistics are overstated. To account for the potential error-dependence problem, a single cross-sectional regression with the time-series mean of each variable by firm is used to perform the regression analysis. Running the OLS cross-sectional equations eliminates the problem of serially correlated composite errors. The estimated model is re-specified as follows:

$$LTD_i = \alpha_i + \beta_1 MB_i + \beta_2 RD_i + \beta_3 LMV_i + \beta_4 AFE_i + \beta_5 EBR_i + \beta_6 TS_i + \beta_7 INF_i + \beta_8 STR_i + \boldsymbol{D'\lambda}_i + v_i \qquad (56)$$

where $v_i \sim NIID\ (0,\ \delta_v^2)$; $i$ = 1,..., 652; $t$ = 1991,..., 2000 (Barclay and Smith, 1995; Wooldridge, 2002).

#### 4.4.3.2.3 Fixed and Random Effects Models

Estimating the model by the cross-sectional regressions preserves the dispersion across firms; however, it does not exploit the time-series variation in the observations (Barclay and Smith, 1995). To correct for the serially correlated errors, a random or fixed effects regression model can be utilized. In a random effects model, the $\alpha_i$ is included in the error term, and the model takes the following specification:

$$LTD_{it} = \beta_0 + \beta_1 MB_{it} + \beta_2 RD_i + \beta_3 LMV_{it} + \beta_4 AFE_{it} + \beta_5 EBR_{it} + \beta_6 TS_{it} + \beta_7 INF_{it} + \beta_8 STR_{it} + \boldsymbol{D'}\lambda_i + u_{it} \qquad (57)$$

where $u_{it} = \alpha_i + v_{it}$; $\alpha_i \sim NIID\ (0,\ \delta_\alpha^2)$; $v_{it} \sim NIID\ (0,\ \delta_v^2)$; $i$ = 1,..., 652; $t$ = 1991,..., 2000. In the random effects approach, $\alpha_i$ is incorporated in the composite error term under the assumption that it is orthogonal to the explanatory variables, ($x_{it}$), $E(x_{it}'\alpha_i) = 0$, $i$ = 1,..., 652, $t$ = 1991,..., 2000. Furthermore, the method accounts for the implied serial correlation in the composite error, $u_{it} = \alpha_i + v_{it}$, identical to the generalized least squares (GLS) estimation (Wooldridge, 2002).

In the fixed effects model, the model specification is as follows:

$$LTD_{it} = \alpha_i + \beta_1 MB_{it} + \beta_2 RD_i + \beta_3 LMV_{it} + \beta_4 AFE_{it} + \beta_5 EBR_{it} + \beta_6 TS_{it} + \beta_7 INF_{it} + \beta_8 STR_{it} + \boldsymbol{D'}\lambda_i + v_{it} \qquad (58)$$

where $v_{it} \sim NIID\ (0,\ \delta_v^2)$; $i$ = 1,..., 652; $t$ = 1991,..., 2000. In the fixed effects analysis, $\alpha_i$ is arbitrarily correlated with $x_{it}$, $E(x_{it}'\alpha_i) \neq 0$, $i$ = 1,..., 652, $t$ = 1991,..., 2000 (Wooldridge, 2002).

In order to identify whether fixed or a random effects estimation technique is appropriate for the analysis, the Hausman test is performed to examine the appropriateness of the random effects estimator. A Hausman test compares two estimators. Under the null hypothesis, the fixed and random effects estimators are consistent, but one is more efficient; under the alternative hypothesis, the more efficient of the two becomes inconsistent but the less efficient remains consistent. Thus, if the null is not rejected, the two estimators should be similar; divergence indicates rejection of the null. Rejection of the null further implies the effects are correlated with the individual variances, and the fixed effects should be used for estimation. The Hausman test statistic is as follows:

$$W = (\beta^{\wedge F} - \beta^{\wedge R})'\Sigma^{-1}(\beta^{\wedge F} - \beta^{\wedge R}),\ \ W \sim X^2(k) \qquad (59)$$

where $k$ is the number of estimated coefficients and $\Sigma^{-1}$ is the difference of the estimated covariance matrices from the two estimators. Based on the test results, one of the methods is chosen to perform the econometric analysis (Wooldridge, 2002).

As noted earlier, the problem of heteroskedasticity and autocorrelation is a common issue associated with panel data; GLS estimation can correct for violations of the underlying assumptions. On the other hand, the random effects model accounts for the serial correlation in the composite error term and, therefore, corrects for the serial correlation of errors (Wooldridge, 2002).

#### 4.4.3.4 GLS Model Corrected for Panel Heteroskedasticity and AR(1)

Finally, the last problem associated with panel data is the problem of heteroskesdasticity and serial correlation across panels. The fixed and random effects models overlook these correlations and, therefore, yield inefficient estimators. The heteroskedasticity test compares the maximum likelihood of the model with panel-level heteroskedasticity to the model with homoskedasticity across panels. The autocorrelation across panels is tested based on the assumption that the idiosyncratic errors, $v_{it}$, are serially uncorrelated in the random and fixed effects model specifications (Stata 8 Manual, 2005; Wooldridge, 2002). By correcting for heteroskedasticity and autocorrelation across panels, the estimation procedure yields an error variance-covariance estimator, which is robust to the common problems associated with panel data (Stata 8 Manual, 2005).

### 4.5 Data

In order to investigate the relationship between debt maturity and its determinants, a large sample is constructed, following Barclay and Smith's (1995) and Baker and colleague's (2003) sampling method. The time- and cross-sectional data set merges the COMPUSTAT industrial annual file of debt determinants and the Federal Reserve's file of macroeconomic variables. The sample is restricted to firms with Standard Classification codes from 2000 to 5999 to focus on the industrial corporate sector. Furthermore, firms utilized in the study are present in the sample over the specified time period and have complete data for the explanatory variables. Consequently, the total number of firms included in the empirical analysis is 652. The data span is 1991 through 2000.

The construction method of the sample might cause a survivorship bias, which represents a tendency for some companies to be excluded from performance studies due to the fact that they no longer exist (Investorwords.com, 2005). In the present investigation, only firms with complete data for 1991-2000 are included in the analysis. The bias affects the results towards finding the predicted relationship between the debt maturity and debt determinants, because most of the surviving firms are larger and older with less investment opportunities and, therefore, issue more long-term debt. On the other hand, the bias prevents the analysis from finding the results predicted for the macroeconomic variables due to their high volatility over the time period of the study.

The time- and cross-sectional data series are debt determinants: 1) total assets, common shares outstanding, debt due in one, two, three, four and five years, earnings per share, total equity, and stock price; and 2) macroeconomic variables: twenty-year government bond and commercial paper return, six-month treasury bill yield, inflation, and term spread.

COMPUSTAT reports the amount of long-term debt payable in years one through five from the firm's fiscal year end. Following Barclay and Smith (1995), the maturity structure of a firm's debt is defined as the percentage of the firm's total debt that has a maturity of more than three years. Several firms have less than zero percent or more than 100 percent of their total debt maturing in more than three years. Since these observations reflect data-coding errors, they are deleted.

Excess bond returns are measured by an index of investment-grade corporate bonds over commercial paper. The corporate bond indices track portfolios that are continually redefined to a constant 20-year maturity. Excess government and corporate bond returns are the difference between the long-term corporate bond and commercial paper returns, respectively (Baker et al., 2003).

There are four debt determinant variables used in this study: market-to-book ratio, a regulatory dummy variable, firm size, and abnormal earnings. Consistent with the Agency Cost Hypothesis, the market-to-book ratio is a proxy for the firm's investment opportunity set. The market value of the firm's assets is estimated as the book value of assets minus the book value of equity plus the market value of equity (the price of shares multiplied by the total number of shares outstanding). The market-to-book ratio is the estimated market value of assets divided by the book value of assets (Barclay & Smith, 1995). In accordance with Baker et al. (2006), the variable is hypothesized to carry a negative sign.

To estimate the effect of regulation on debt maturity, a dummy variable is constructed. It takes a value of one for regulated firms, and zero otherwise. Regulated industries include airlines, telecommunications, as well as gas and electric utilities. The firm size is the estimated natural logarithm of the market value of the firm (Barclay & Smith, 1995). As found in the studies by Chen et al. (1999) and Antonenko et al. (2006), the log of firm value is expected to be positively related to the share of long-term debt as a proportion of total debt.

Signaling models assume that managers have better information about the firm's value (or quality) than investors. To estimate the firm's quality empirically, the firm's abnormal earnings are defined as a proxy. The variable is measured as the difference between next year's and this year's earnings per share divided by this year's share price. It is assumed that high-quality firms have positive abnormal earnings, and low-quality firms have negative abnormal earnings (Barclay and Smith, 1995). Following Barclay and Smith (1995), observations with the absolute value of abnormal earnings greater than five were disregarded. The exclusion of the extreme values might affect the coefficient for the abnormal earnings, but should not influence the other variables in the analysis.

Debt market conditions are represented by three variables: inflation, the real short-term rate, and the term spread. Inflation is the annual percentage change in the Consumer Price Index. The real short-term rate is estimated as the annualized Treasury bill return minus inflation. The term spread is the difference between the Treasury bond yield and the annualized Treasury bill return. In order to account for the cross-sectional variation of the debt determinants data, the market variables ending dates are matched with the firm's fiscal year end (Baker et al., 2003). The descriptive statistics for each variable used in the analysis are presented in Table 4.8.

Table 4.8: Data Descriptive Statistics Employed in Corporate Debt Estimation Process

| Variable | | Mean | Standard Deviation | Min Value | Max Value |
|---|---|---|---|---|---|
| Share of Long-Term Debt to Total Debt | Overall | 0.587608 | 0.2489763 | 0 | 1 |
| | Between[a] | | 0.1307427 | 0.085569 | 0.938587 |
| | Within[b] | | 0.2119414 | -0.23633 | 1.29564 |
| Market-to-Book Ratio | Overall | 1.549547 | 0.9116479 | 0.360198 | 22.81232 |
| | Between[a] | | 0.707687 | 0.639052 | 7.222132 |
| | Within[b] | | 0.575302 | -1.81968 | 21.19958 |
| Regulatory Dummy Variable[c] | Overall | 0.21319 | 0.4095923 | 0 | 1 |
| | Between[a] | | 0.4098753 | 0 | 1 |
| | Within[b] | | 0 | 0.21319 | 0.21319 |
| Log of Market Value | Overall | 2.698775 | 0.9371075 | 0.152839 | 5.435521 |
| | Between[a] | | 0.9146363 | 0.550687 | 5.332079 |
| | Within[b] | | 0.2067988 | 0.473086 | 4.102186 |
| Abnormal Earnings | Overall | 0.030876 | 0.4310666 | -5 | 5 |
| | Between[a] | | 0.1028505 | -0.35749 | 0.820588 |
| | Within[b] | | 0.4186344 | -5.67963 | 4.901274 |
| Excess Bond Returns | Overall | 0.010944 | 0.022473 | -0.052 | 0.079 |
| | Between[a] | | 0.004737 | -0.00534 | 0.01641 |
| | Within[b] | | 0.021968 | -0.03572 | 0.080104 |
| Term Spread | Overall | 0.012496 | 0.0120939 | -0.0068 | 0.0351 |
| | Between[a] | | 0.0017139 | 0.01113 | 0.01565 |
| | Within[b] | | 0.011972 | -0.00967 | 0.034366 |
| Inflation | Overall | 0.027276 | 0.0068753 | 0.0137 | 0.0565 |
| | Between[a] | | 0.0009295 | 0.02662 | 0.029 |
| | Within[b] | | 0.0068123 | 0.011976 | 0.055196 |
| Short-Term Rate | Overall | 0.021536 | 0.0127056 | -0.0549 | 0.0501 |
| | Between[a] | | 0.0066697 | -0.01892 | 0.02339 |
| | Within[b] | | 0.0108171 | -0.01444 | 0.090556 |
| January Dummy | Overall | 0.062883 | 0.2427718 | 0 | 1 |
| | Between[a] | | 0.2429396 | 0 | 1 |
| | Within[b] | | 0 | 0.062883 | 0.062883 |
| February Dummy | Overall | 0.019939 | 0.1398002 | 0 | 1 |
| | Between[a] | | 0.1398968 | 0 | 1 |
| | Within[b] | | 0 | 0.019939 | 0.019939 |
| March Dummy | Overall | 0.055215 | 0.2284164 | 0 | 1 |
| | Between[a] | | 0.2285743 | 0 | 1 |
| | Within[b] | | 0 | 0.055215 | 0.055215 |

| Table 4.8: Data Descriptive Statistics, Continued | | | | | |
|---|---|---|---|---|---|
| Variable | | Mean | Standard Deviation | Min Value | Max Value |
| April Dummy | Overall | 0.015337 | 0.1229004 | 0 | 1 |
| | Between[a] | | 0.1229853 | 0 | 1 |
| | Within[b] | | 0 | 0.015337 | 0.015337 |
| May Dummy | Overall | 0.026074 | 0.1593665 | 0 | 1 |
| | Between[a] | | 0.1594766 | 0 | 1 |
| | Within[b] | | 0 | 0.026074 | 0.026074 |
| June Dummy | Overall | 0.07362 | 0.2611709 | 0 | 1 |
| | Between[a] | | 0.2613514 | 0 | 1 |
| | Within[b] | | 0 | 0.07362 | 0.07362 |
| July Dummy | Overall | 0.016871 | 0.1287986 | 0 | 1 |
| | Between[a] | | 0.1288876 | 0 | 1 |
| | Within[b] | | 0 | 0.016871 | 0.016871 |
| August Dummy | Overall | 0.019939 | 0.1398002 | 0 | 1 |
| | Between[a] | | 0.1398968 | 0 | 1 |
| | Within[b] | | 0 | 0.019939 | 0.019939 |
| September Dummy | Overall | 0.075153 | 0.2636589 | 0 | 1 |
| | Between[a] | | 0.2638411 | 0 | 1 |
| | Within[b] | | 0 | 0.075153 | 0.075153 |
| October Dummy | Overall | 0.049233 | 0.216371 | 0 | 1 |
| | Between[a] | | 0.216201 | 0 | 1 |
| | Within[b] | | 0.0117498 | -0.05077 | 0.949233 |
| November Dummy | Overall | 0.016871 | 0.1287986 | 0 | 1 |
| | Between[a] | | 0.1288876 | 0 | 1 |
| | Within[b] | | 0 | 0.016871 | 0.016871 |

[a] Between Group Estimates – variation between mean groups.
[b] Within Group Estimates - deviations from a group mean.
[c] Mean derived based on a dummy variable (i.e. Regulatory Dummy Variable) present the percent of the data falling into the '1' category.

## 4.6 Discussion of the Results

### 4.6.1 Comparison of Alternative OLS Models

Results of tests of the debt-maturity model for violation of assumptions of normality, heteroskedasticity, and autocorrelation are presented in Table 4.9. These results revealed significant heteroskedasticity and serial correlation issues. In order to account for misspecification problems, different model specifications were estimated including pooled and cross-sectional OLS, as well as GLS regressions. The precision and efficiency of the parameter coefficients were found to be sensitive to the chosen form. Finally, the Hausman specification test was utilized to examine the appropriateness of random effects estimation. As presented in Table 4.9, fixed effects estimation appears to be the appropriate model choice as the null hypothesis representing the difference in coefficients not being systematic is rejected ($X^2(8) = 22.24$). For the purpose of proving the debt theory, however, both random and fixed effects are presented for comparison purposes.

Table 4.9: Specification Test of the Corporate Debt Maturity Model Specifications

| Test | Null Hypothesis | Test Statistic | Probability > Test Statistic |
|---|---|---|---|
| Ordinary Least Square Regression | | | |
| Bera and Jarque's LM Skewness-Kurtosis Test for Normality | No Skewness and Kurtosis | $X2(2) = 1.76$ | 0.414 |
| Breusch-Pagan Test for Heteroskedasticity | Constant Variance | $X2(1) = 91.47$ | 0 |
| Test of Residuals | Serial Correlation | t = 20.45 | 0.006 |
| Fixed and Random Effects Regression | | | |
| Breusch and Pagan Lagrange Multiplier Test for Random Effects | Constant Variance | $X2(1) = 844.55$ | 0 |
| Hausman Specification Test | Difference in Coefficients Not Systematic | $X2(8) = 22.24$ | 0.0045 |
| GLS Model Corrected for Panel Heteroskedasticity and AR(1) | | | |
| Likelihood-Ratio Test for Heteroskedasticity Across Panels | No Panel Heteroskedasticity | $X2(651) = 2024.73$ | 0 |
| Wooldridge Test For Autocorrelation Across Panels | No First-Order Autocorrelation | F(1, 651) = 78.425 | 0 |

Overall, the test-statistic and log likelihood values of each model specification indicate that each regression in Table 4.10 is significant at conventional levels. Comparing the models' R-square values suggests that variation in debt maturity structure across firms provides some of the explanatory power in these regressions. The R-square value for the different specifications ranges from 0.0101 to 0.1218, which implies that from 1 percent to 12 percent of the variation in the debt maturity variable, is explained by the explanatory variables in the model specifications.

### 4.6.2 Report of Empirical Results for Different Model Specifications

#### 4.6.2.1 Ordinary Least Square Regression

Results presented in Table 4.10 suggest that debt maturity decisions are related to excess bond returns, debt determinants, and macroeconomic variables. The pooled OLS model specification indicates that excess bond returns are negatively related to debt maturity ($b_6$ = -0.23884) at the 10% significance level. Furthermore, market-to-book ratio and abnormal earnings are inversely associated with the debt maturity variable at the 1% and 5% levels, respectively. Log market value is positively correlated to the dependent variable at a 1% level of significance. There is no regulatory effect found on the debt maturity decisions, although the coefficient on the dummy variable is positive, as predicted. Furthermore, none of the macroeconomic variables is statistically significant, yet all of them are negatively related to the dependent variable, which is in agreement with theoretical predictions. Finally, there is a statistically significant difference for the firms with fiscal year ending in March and November and the amount of long-term debt issued compared to the firms with fiscal year ending in December. For example, firms with fiscal year end in March tend to issue more short-term debt.

#### 4.6.2.2 Cross-Sectional Regression with Time-Series Mean

To account for the error-dependence problem in pooled OLS estimation, a single cross-sectional regression with the time-series mean of each variable by firm was used to estimate the model. The excess bond returns are statistically significant at the 5% level of significance and have a coefficient of $b_6$ = -0.12034. The debt determinants: the market-to-book ratio and abnormal earnings negatively relate to the long-term debt decisions and have higher magnitudes compared to those obtained in the pooled OLS estimation ($b_1$ = -0.01305 vs. $b_1$ = -0.00613; $b_4$ = -0.08694 vs. $b_4$ = -0.01413). Log market value is positively associated with debt decisions and is higher compared to the pooled OLS estimate. Again, no market variables are statistically significant and term spread carries a positive sign. Finally, there is no statistical difference for long-term debt in capital structure for firms with fiscal year end in December and other months.

**4.6.2.3 Fixed and Random Effects Models**

The cross-sectional regressions preserve dispersion across firms, but do not exploit time-series variation in the observations (Barclay & Smith, 1995). The fixed effects regression model was estimated to correct for this problem. Both fixed and random effects results are reported and discussed. For the random and fixed effects regressions, there is a difference between the parameter coefficients and their efficiency levels. In the fixed effects estimation, the excess bond returns are negatively related to debt maturity ($b_6$ = -0.23886). The coefficient is smaller in absolute value than the random effects coefficient by 0.02, yet it is less efficient (the t-statistic associated with the independent variable in fixed effects model is lower compared to the t-statistic in the random effects model specification). Furthermore, the log of firm market value is positively associated with the debt maturity ($b_3$ = 0.11800), and it is the only debt determinant statistically significant in the model. The coefficient is larger than the random effect estimate coefficient ($b_3$ = 0.05250) and has a larger standard error. Parameter estimates for market-to-book ratio and abnormal earnings diverge under the two models; in the random effects equation both variables are statistically significant and inversely correlated with the dependent variable, while in fixed effects equation both variables are statistically insignificant. In the fixed effects model, only term spread is statistically significant among the macroeconomic variables; the positive coefficient contradicts theoretical predictions. Finally, in the case of random effects regression, there is a statistically significant difference between the firms with fiscal year ending in March and November in the level of long-term debt issued compared to the firms with fiscal year end in December. Due to multicolinearity observed in the case of fixed effects, most of the seasonal dummy variables were dropped from the estimation by Stata 8.

#### 4.6.2.4 GLS Model Corrected for Panel Heteroskedasticity and AR(1)

Heteroskedasticity and autocorrelation across panels are corrected by generalized least square estimation applying the correction for heteroskedasticity and first-order autocorrelation across panels with autocorrelation coefficient of $\rho = 0.4070$. The modification improved the precision and efficiency of the parameter coefficients in the regression compared to the previous specifications. The excess bond returns are statistically significant at the 1% level, and negatively correlated to debt maturity decision ($b_6 = -0.06003$). The parameter coefficient is the lowest among all specifications. As before, market-to-book ratio and abnormal earnings are negatively related to the debt maturity decision. Log firm market value is positively related to debt maturity ($b_3 = 0.01366$) at the 1% significance level. In the case of macroeconomic variables, term spread and inflation are negatively associated with debt maturity at the 10% and 5% levels, respectively. These coefficients have the largest magnitudes among all the regression specifications. Finally, there is a statistically significant difference in the debt-maturity decisions for firms with fiscal year ending in January, March, May, June, and November compared to those with fiscal year-ending in December. For example, firms with fiscal year end in January, May, June, and November have a higher share of long-term debt compared to those in December.

### 4.6.3 Results Summary

The fixed effects model as well as model corrected for heteroskedasticity and autocorrelation across panels are the specifications that should be given more weight when interpreting the regression results. The pooled OLS and cross-sectional model with time-series mean display issues with the estimation of panel data. The pooled OLS model does not account for the error-dependence problem while a single cross-sectional regression with the time-series mean of each variable by firm does not employ the time-series variation within each panel. To correct for these estimation issues the fixed effects model is utilized. The model specification corrected for heteroskedasticity and autocorrelation across panels also should carry a higher weight in results comparisons as it improves the precision and efficiency of the parameter coefficients when compared to the other specifications.

Table 4.10: Estimation Results for the Corporate Debt Maturity Model

| Variables | Pooled OLS | Cross-Sectional OLS | Fixed Effects | Random Effects | GLS Model Corrected for Panel Heteroskedasticity and AR(1) |
|---|---|---|---|---|---|
| Dependent Variable | Share of Long-Term Debt to Total Debt | | | | |
| Intercept ($b_1$) | 0.43717*** (17.22) | 0.08306 (0.08) | 0.27612*** (5.06) | 0.47550*** (15.93) | 0.59842*** (39.03) |
| Market-to-Book Ratio ($b_2$) | -0.00951*** ( -2.79) | -0.13057* (1.80) | -0.00474 (-0.99) | -0.00796** (-1.99) | -0.00674** (-2.16) |
| Regulatory Dummy Variable ($b_3$) | 0.01040 (1.29) | 0.01082 (0.83) | Dropped | 0.00809 (0.63) | 0.00656 (0.99) |
| Log of Market Value ($b_4$) | 0.04724*** (13.42) | 0.04365*** (7.40) | 0.11800*** (7.37) | 0.05250*** (9.71) | 0.01366*** (4.48) |
| Abnormal Earnings ($b_5$) | -0.01543** (-2.19) | -0.08694* (-1.79) | -0.00955 (-1.44) | -0.01269* (-1.94) | -0.01212** (-2.18) |
| Excess Bond Returns ($b_6$) | -0.23884* (-1.87) | -0.12034*** (-2.34) | -0.23886* (-1.70) | -0.25854** (-2.03) | -0.06003*** (-2.42) |
| Term Spread ($b_7$) | -0.02786 (-0.08) | 0.02743 (0.34) | 0.70750* (1.89) | 0.029691 (0.09) | -0.29725* (-1.67) |
| Inflation ($b_8$) | -0.34899 (-0.54) | -0.46010 (-0.04) | -0.26200 (-0.44) | -0.34895 (-0.59) | -0.57826** (-2.16) |
| Short-Term Rate ($b_9$) | -0.27210 (-0.72) | -0.19125 (-0.98) | -0.09694 (-0.28) | -0.25854 (-0.75) | -0.11847 (-0.85) |
| January Dummy ($b_{10}$) | 0.01737 (1.29) | -0.04619 (-0.34) | Dropped | 0.01713 (0.8) | 0.03400*** (2.54) |
| February Dummy ($b_{11}$) | -0.00884 (-0.40) | 0.11801 (0.22) | Dropped | -0.00758 (-0.21) | 0.00954 (0.04) |
| March Dummy ($b_{12}$) | -0.0646*** (-4.55) | 0.04413 (0.09) | Dropped | -0.06550*** (-2.96) | -0.05803*** (-5.91) |

| Table 4.10: Estimation Results for the Corporate Debt Maturity Model, Continued | | | | | |
|---|---|---|---|---|---|
| Variables | Pooled OLS | Cross-Sectional OLS | Fixed Effects | Random Effects | GLS Model Corrected for Panel Heteroskedasticity and AR(1) |
| Dependent Variable | | | Share of Long-Term Debt to Total Debt | | |
| April Dummy ($b_{13}$) | -0.00073<br>(-0.03) | 0.00564<br>(0.03) | Dropped | 0.00377<br>(0.01) | -0.02318<br>(-1.06) |
| May Dummy ($b_{14}$) | 0.00897<br>(0.37) | -0.04860<br>(-0.36) | Dropped | 0.00953<br>(0.28) | 0.44310*<br>(1.65) |
| June Dummy ($b_{15}$) | 0.00786<br>(0.64) | -0.09589<br>(-0.38) | Dropped | 0.01010<br>(0.52) | 0.07890***<br>(9.54) |
| July Dummy ($b_{16}$) | 0.00369<br>(0.15) | -0.12577<br>(-0.38) | Dropped | 0.00524<br>(0.36) | 0.02657<br>(1.01) |
| August Dummy ($b_{17}$) | 0.00967<br>(0.44) | 0.00738<br>(0.08) | Dropped | 0.01286<br>(0.14) | -0.00612<br>(-0.33) |
| September Dummy ($b_{18}$) | -0.00963<br>(-0.12) | -0.06278<br>(-0.52) | Dropped | -0.00701<br>(-0.37) | 0.01354<br>(1.16) |
| October Dummy ($b_{19}$) | -0.00170<br>(-0.81) | -0.00627<br>(-0.14) | 0.07469<br>(0.32) | 0.00102<br>(0.04) | 0.00293<br>(0.24) |
| November Dummy ($b_{20}$) | 0.06158***<br>(2.6) | -0.00320<br>(-0.02) | Dropped | 0.06315*<br>(1.66) | 0.11411***<br>(16.49) |

*Note:* The coefficient estimates are represented as the first number for each independent variable. Numbers is parenthesis represents the standard error of each parameter coefficient. An * (**) (***) indicates significance at 10% level (5% level and 1% level). Pooled OLS regression: R-Square = 0.0388; F-stat = F (19, 6500) = 13.79. Cross-sectional OLS regression: R-Square = 0.1218; F-stat = F(19, 632) = 5.17. Fixed effects regression: R-Square = 0.0122; Wald $X^2(19)$ = 137.54. Random effects regression: R-Square = 0.0101; Wald $X^2(19)$ = 596.82. GLS Regression Corrected for Panel Heteroskedasticity and AR(1): Log Likelihood = 1601.881; Wald $X^2(19)$ = 596.82.

Based on the results presented in Table 4.10, there is a negative relationship between debt maturity and excess bond returns in all specifications. Furthermore, market-to-book ratio, abnormal earnings, as well as term spread and inflation are negatively related to the share of long-term debt. The firm size is positively related to debt maturity. Finally, there is a difference in long-term debt issues for firms with fiscal year ending in January, March, May, June, and November. However, no statistically significant relationships were found for the regulatory dummy and short-term interest rate, although the signs for these coefficients are consistent with theoretical predictions. Results here indicate the precision and efficiency of the model were sensitive to the model specification; corrections for violations of underlying assumptions are responsible for the magnitudes of these differences.

**4.7 Discussion and Conclusions**

The debt financing decisions are a set of simultaneous decisions made by lenders and borrowers. Since lenders are unable to observe directly the firms' investment decisions, the banks offer contracts based upon the observable firm characteristics such as wealth and size. The contracts offered to each firm differ with respect to the cost of capital borrowing as well as its maturity. When deciding on the financing decisions, firms might also take into account the changes in macroeconomics variables in order to lower the cost of borrowing. As a result, the goal for this article was to examine the hypothesis of the effect of the debt determinant as well as the macroeconomics variables on the debt maturity structure at the individual firm level data. The research objective was achieved by investigating the relationship between these variables and the corporate debt maturity structure. A reduced form of the simultaneous debt decision model was estimated by employing several OLS estimation methods to analyze the relationship between debt maturity and its determinants.

The examination of the excess bond returns as well as the debt and market determinants of corporate debt maturity supports the hypothesis that excess bond returns are negatively associated with debt maturity. For example, firms tend to issue long-term debt when excess bond returns are low. Furthermore, results here are consistent with Barclay and Smith's (1995) and Antoniou et al. (2002) findings that firms with more growth options in their investment opportunity set issue more short-term debt. As a result, reducing debt maturity helps control the underinvestment problems as presented by Myers (1977). For example, underinvestment occurs when the debt maturity is not appropriately timed and debt refinancing occurs after investment options expire so the gains from new investments accrue to the debt holder (Johnson, 2003).

As discussed by FEMA (2005), there is also evidence for a strong positive association between firm size and debt maturity in which large firms issue a significantly higher share of long-term debt. Finally, there is support for the hypothesis that firms use the maturity of debt to signal information to the market. In accordance with Barclay and Smith's (1995) and Richardson and Sloan's (2003) findings, high-quality firms issue short-term debt, while low-quality firms issue long-term debt. There is no statistically significant evidence found supporting the impact of firms' regulatory status on debt maturity, although the variable does carry the predicted sign, implying that regulated firms issue more long-term debt.

The results obtained for the market variables are consistent with the hypothesis of macroeconomic variables impacting the debt maturity. The evidence is supported by the inverse relationship between the corporate debt maturity and the macroeconomic variables: inflation and term spread. Firms borrow long term when debt market conditions suggest that the relative cost of long-term debt is low. Although there is no evidence found in support of impact of short-term interest rates on debt maturity, the variable carries the predicted negative sign.

The sign and the magnitude of impact of the term spread changes depending on the model specification. These differences can be explained by linking the variable to the impact of taxes on the debt maturity decisions. Due to a firm's default on debt payments, the expected value of the firm's tax liabilities depends on the debt maturity structure as long as it is not flat. For example, if the term structure is upward sloping, issuing long-term debt reduces the firm's expected tax liability and, therefore, increases the firm's market value. As a result, there is a positive relationship between long-term debt and a slope of the term structure (Brick & Ravid, 1985).

The limited impact of the term spread is explained by Lewis (1990) and Terra (2005) as taxes having no effect on the optimal debt maturity when the optimal debt-asset ratios and debt maturity structure are chosen simultaneously. Terra (2005), however, finds that taxes do have a negative relationship with debt maturity in the case of Latin American firms. As many researchers, he struggles to explain why taxes have no impact on debt maturity although the average effective tax rate for the U.S. companies is greater than those of other countries so the negative impact of the term spread should be statistically influential (Terra, 2005).

The finding of the negative relationship between the macroeconomic variables and debt maturity confirms the fact of the chief financial officers timing the market by issuing short-term debt when short-term rates are low or when long-term rates are expected to decline. The finding is especially significant in the case of large firms, which have easy access to financial markets (Berlin, 2006). The findings confirm also the belief that managers watch the debt and therefore lower the borrowing cost. For example, Faulkendler (2005) finds that managers believing in market timing purchase a swap, amplifying the firm's exposure to rising interest rates while undertaking new borrowings instead of hedging against the increase in interest rate risk. This implies that managers are likely to swap fixed interest rate payments for floating interest rates payments at the same time of the debt offering when the term premium is high (Faulkendler, 2005).

On the other hand, not finding the short-term interest rate statistically significant is in agreement with the view shared by Berlin (2006) which states that chief financial officers cannot time the market due to limited information on interest rate movements (Berlin, 2006). As a result, managers might simply be wrong in their beliefs that market timing helps to decrease the cost of borrowing. Even if the cost of borrowing is lowered, it might not be a result of market-timing. Although Baker et al. (2003) find managers to exploit inefficiencies in debt markets and therefore, find lower borrowing cost, Butler et al. (2004) criticizes their econometric techniques and remains unconvinced of the market-timing phenomenon. His opinion as well as Berlin's (2006) opinion is in agreement with the Efficient Market Hypothesis, which states that it is impossible to "beat the market" because market efficiency causes existing investment prices to always incorporate and reflect all relevant information with regard to the market conditions (Fama, 1970).

The analysis here has some potential limitations. First, results obtained in this study are affected by a survivorship bias due to the sample generating process, which only included firms with complete data for the entire 10 years of study. As a result, the firms that are included in the sample as a result are more likely to be larger and older with less investment opportunities. These firm characteristics bias the sample towards finding issuance of long-term debt. On the other hand, the bias works against finding the inverse relationship between debt maturity and excess bond returns as well as macroeconomic variables. High volatility of the explanatory variables over the time of the analysis makes it harder to obtain the predicted results due to lack of persisting trend in the variables. Consequently, based on the bias possibility, the results for firm characteristics have to be taken lightly.

Second, as discussed by Barclay and Smith (1995), the debt-maturity analysis could be better executed if more disaggregated data were used in the study. For example, firms with more growth options in their investment opportunity set issue more short-term debt; however, it is questionable whether the total variation might result in variation among instruments with different maturities such as short-term bank debt or long-term public debt. A more detailed examination of the mix of debt instruments issued by different companies would add to the depth and understanding of firms obtaining the observed debt structure (Barclay & Smith, 1995).

Further research on the debt maturity structure should not only include the mix of debt instruments, but also one, two, and three-year cumulative excess bond returns. By including the cumulative returns in the model, debt maturity sensitivity to the impact of long- and short-term returns could be estimated. Such analysis would provide further information on the cost of debt borrowing at different maturities. Additionally, the sample used in the analysis should include information for as many firms and years as are available, in order to increase the precision and efficiency of the results. Furthermore, a large number of firms included in the sample will decrease the possibility of a survivorship bias and offer a more representative sample of firms, so obtained results could be extrapolated and interpreted for all firms in general. In order to investigate in more detail the relevance of the Market Timing Hypothesis in capital borrowing process, the weekly/monthly data points instead of yearly values should be employed in the model. As mentioned earlier, the market-timing phenomenon is a short-run process, so with more frequent data points finding support for the hypothesis is more likely to be achieved.

Finally, several changes should be made to the functional form of debt maturity and the empirical estimation framework. Since the functional form of the debt maturity function has not been defined by the finance literature at this point, the estimation process of the OLS models should include the specification tests for higher orders and interaction effects between the independent variables. Finding the right functional form of the debt maturity will improve the precision and efficiency of the model estimates and identify the debt maturity drivers. In addition, the empirical analysis framework should be extended to a mixing (averaging) estimator model to improve the efficiency of the estimation process. The idea behind the mixing estimator model is to average the estimators obtained from different OLS models to attain more efficient results compared to those derived from estimating each model individually while controlling for the omitted variable bias. The employment of the mixing estimator model results in fitted estimates that are asymptotically efficient with a minimum value of squared error in the class of discrete model average estimators (Hansen, 2006).

In conclusion, the analysis found evidence in support of impact of debt determinants and macroeconomic variables on the structure of debt maturity. Furthermore, as suggested by Backer et al. (2003), empirical studies of debt maturity need to incorporate the market conditions and the excess bond returns in order to adequately explain patterns in debt maturity data.

## 4.8 Acknowledgements

The author would like to extend thanks to Dr. Jonathan Newitt, Dr. Alexei Ovtchinnikov, Dr. Chetan Dave, Dr. F. Stewart DeBruicker, Mr. Jim Colizzo, the unanimous reviewers of the International Journal of Business and Economics, and the participants of the 2006 Global Conference on Business and Economics at Harvard University, Cambridge, MA for their valuable comments.

## 4.9 References

Antonenko, O., Talavera, O., & Tsapin, A. (2006). *Determinants of Liability Maturity Choice: The Case of Ukraine*. Working Paper. Available online at http://www.eerc.kiev.ua/eroc/anconference/ninth/papers/Antonenko.pdf. Accessed: March 2005.

Antoniou, A., Guney, Y., & Paudyal, K. (2002). The Determinants of Corporate Debt Maturity Structure. *Proceedings of Annual Meeting of the European Financial Management Association 2003*. Helsinki, Finland. Available online at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=302833#PaperDownload. Accessed: March 2007.

Baker, M., Greenwood, R., & Wurgler, J. (2003). The Maturity of Debt Issues and Predictable Variation in Bond Returns. *Journal of Financial Economics 70(2)*, 261-291.

Baker, M., Ruback, R. S., & Wurgler, J. (2006). Behavioral Corporate Finance. In E. Ebcko, (Ed.), *Handbook in Corporate Finance: Empirical Corporate Finance*. Elsevier/North Holland.

Baker, M. & Stein, S., Wurgler, J. (2003). When Does the Market Matter? Stock Prices and Investment of Equity-Dependent Firms. *Quarterly Journal of Economics 118*, 969-1006.

Baker, M. & Wurgler, J. (2002). Market Timing and Capital Structure. *Journal of Finance 57*, 271-372.

Barclay, M. J., & Smith, Jr., C.W. (1995). The Maturity Structure of Corporate Debt. *Journal of Finance 50*, 609-631.

Berger, A. N., Espinosa-Vega, M., Frame, W. S., & Miller, N. H. (2004). *Debt Maturity, Risk, and Asymmetric Information. Finance and Economics Discussion Series* Divisions of Research & Statistics and Monetary Affairs Federal Reserve Board, Washington, D.C. Available online at http://www.federalreserve.gov. Accessed: June 2008.

Berlin, M. (2006). Debt Maturity: What Do Economists Say? What Do CFOs Say? *Business Review*, 3-10. Available online at http://www.phil.frb.org/files/br/Q1_06_DebtMaturity.pdf. Accessed: March 2007.

Bosworth, B. (1971). Patterns of Corporate External Financing. *Brookings Papers on Economic Activity 2*, 253–279.

Brick, I.E., & Ravid, R.A. (1991). Interest Rate Uncertainty and the Optimal Debt Maturity Structure. *Journal of Finance 26*: 1423-1437.

Butler A., Grullon, G., & Westen, J. (2004). *Can Managers Successfully Time the Maturity If their Debt Issues?* Working Paper. Rice University.

Chen, S. S. (1999). The Determinants of Debt Maturity: the Case of Bank Financing in Singapore. *Review of Quantitative Finance and Accounting 12*, 341-350.

*COMPUSTAT Industrial Annual File*. Available online at www.compustat.com. Accessed March 2005, Dennis, S., Nandy, D., & Sharpe, I. G. (2000). The Determinants of Contract Terms in Bank Revolving Credit Agreements. *Journal of Financial and Quantitative Analysis 35*, 87-110.

Dennis, S., & Sharpe, I. (2005). Firm Size Dependency in the Determinants of Bank Term Loan Maturity. *Journal of Business Finance and Accounting 32* 31-64.

Diamond, D. (1993). Seniority and Maturity of Debt Contracts. *Journal of Financial Economics 33*, 341-368.

Diamond, D. (1996). Financial Intermediation as Delegated Monitoring, a Simple Example. *Economic Quarterly* of the Federal Reserve Bank of Richmond *82(3)*, 51-66.

*Federal Reserves File*. Available online at www.federalreserve.gov. Accessed: March 2005.

Fama, E. F. (1970). Efficient Capital Markets: A Review of Theory and Empirical Work. *Journal of Finance 25(2)*, 383–417.

*FMA*. (2006). The Maturity Structure of Private Placements of Debt. Available online at http://www.fma.org/SLC/Papers/MaturityStructureFMA.pdf. Accessed: March 2007.

Faulkender, M. (2005). Hedging or Market Timing? Selecting the Interest Rate Exposure of Corporate Debt. *Journal of Finance* 60, 931-962.

Fisher, I. (1930). *The Theory of Interest: As Determined by Impatience to Spend Income and Opportunity to Invest It*. New York: Macmillan. Reprinted in 1995: New York: Augustus M. Kelley.

Flannery, M. J. (1986). Asymmetric Information and Risky Debt Maturity Choice. *Journal of Finance 41(1)*, 19 -37.

Foo, M., & Yo, W.W. (2005). The Determinants of Corporate Hedging Decision: An Empirical Analysis. *Financial Management Association Annual Meeting*. Available online at: http://www.fma.org/Chicago/Papers/Merlyn_Foo_FMA_2005_submission.pdf. Accessed: June, 2008.

Graham, J. R., & Harvey, C. R. (2001). The Theory and Practice of Corporate Finance: Evidence from the Field. *Journal of Financial Economics 60*, 187-243.

Hackethal, A., & Jansen, C. (2006). New Empirical Evidence on the Debt Maturity Choice and the Role of Credit Risk. China Center for Financial Research. Available online at www.ccfr.org.cn. Accessed: June 2008.

Hansen, B. E. (2006). Least Squares Model Averaging. Department of Economics. Universoty of Winsconsin. Madison, WI. Available online at www.ssc.wisc.edu/~bhansen. Accessed: December 2008.

Harvey, C. R., Lins, K.V., & Roper, A. H. (2001). The Effect of Capital Structure When Expected Agency Costs Are Extreme. NBER Working Papers. Available online at http://nber.org. Accessed: March 2006.

Hsiao, C. (2002). *Analysis of Panel Data (2nd ed.)*. Cambridge, MA: Cambridge University Press.

James, C. (1987). Some Evidence on the Uniqueness of Bank Loans. *Journal of Financial Economics 19(2)*, 217-235.

Kale, J. & Noe, T. (1990). Dividends, Uncertainty, and Underwriting Costs under Asymmetric Information. *The Journal of Financial Research 13(4)*, 265-277.

*Investorwords.com*, Available online at http://www.investorwords.com/5814/survivorship_bias.html. Accessed: April 2005.

Johnson, S. (2003). Debt Maturity and Effects of Growth Opportunities and Liquidity Risk on Leverage. *The Review of Financial Studies 16(1)*, 209-236.

Jun, S.G. & Jen F.C. (2003). Trade-off Model of Debt Maturity Structure. *Review of Quantitative Finance and Accounting 20*, 5-34.

Lewis, C. M. (1990). Multiperiod Theory of Corporate Financial Policy under Taxation. *Journal of Financial and Quantitative Analysis*: 25-43.

Malkiel, B.G. (2004). Can Predictable Patterns in Market Returns Be Exploited Using Real Money? *Journal of Portfolio Management 31 (Special Issue)*, 131-141.

Magri, S. (2006). Debt Maturity of Italian Firms. Bank of Italy Research Paper Nr. 574. Social Science Research Network. Available online at: http://ssrn.com/abstract=895509. Accessed: June 2008.

Marsh, P. (1982). The Choice between Equity and Debt: an Empirical Study. *Journal of Finance*: 121– 144.

Mathiesen, H. (2008). Perfect Capital Market - Fisher Separation Theorem. Version 2.2. Available online at: encycogov.com. Accessed: June 2008.

Modigliani, F. & Miller, M.H. 1958. The Costs of Capital, Corporation Finance, and the Theory of Investment. *American Economic Review 53*, 261-297.

Myers, S.C. (1977). Determinants of Corporate Borrowing. *Journal of Financial Economics 5*, 147-175.

Ozkan, A. (2002). The Determinants of Corporate Debt Maturity: Evidence from UK Firms. *Applied Financial Economics 12*, 19-24.

Richardson, S.A. & Sloan, R.G. (2002). External Financing and Future Stock Returns. University of Pennsylvania Working Paper. Available online at http://finance.wharton.upenn.edu/~rlwctr/papers/0303.pdf. Accessed: March 2007.

Rubinstein, A. (1976). *Comments on the Risk and Time Preference in Economics.* Princeton University Press.

Rubinstein, M. (2003). Great Moments in Financial Economics: Present Value. *Journal of Investment Management.*

Schwert, G. W. (2001). Anomalies and Market Efficiency. *Handbook of the Economics of Finance*, North Holland.

*Stata 8 Manual*. (2005). Available online at http://www.stata.com/sipport/faqxs/stat/panel.html. Accessed: March 2005.

Spence, A. M. (1973). Job Market Signaling. *Quarterly Journal of Economics 87(3)*, 355-374.

Stiglitz, J.E. (1974). On the Irrelevance of Corporate Financial Policy. *American Economic Review 64(6)*, 851-866.

Stiglitz, J. E. & Weiss, A. (1992). Asymmetric Information in Credit Markets and Its Implications for Macro-Economics. *Oxford Economic Papers 44*, 694-724.

Stohs, M.H. & Mauer, D.C. (1996). The Determinants of Corporate Debt Maturity Structure. *Journal of Business 69*, 279-312.

Terra, P.R.S. (2005). Determinants of Corporate Debt Maturity in Latin America. Proceedings of Annual Meeting of the American Accounting Association (AAA) 2005. San Francisco, CA.

Varian, H.R. (1992). *Microeconomics Analysis*. Third Edition. W.W Norton & Company. New York.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, MA: MIT Press.

## 4.10 Appendix A

### 4.10.1 Profit Maximization

The basic assumption of firm behaviour is the goal of maximising firm's profits. There are several fundamental implications resulting from profit maximization. For example, the firm chooses the optimal level of output at which the production of an additional output unit produces marginal revenue equal to marginal cost. Another decision determines the level of each factor in output production (eg. labour). The final implication resulting from profit maximization is the equality of long-run profits. If two firms have identical revenue and cost functions, in the long-run, the two firms must have equal profits due to each firm's imitation of the other's actions (Varian, 1992).

In order to apply implications resulting from profit maximization (optimal output level, optimal input level, and equality of long-run profits), the revenue and cost functions must be broken into more basic parts. For example, revenue is composed of a number of outputs a firm sells and the price of each output. Costs are also composed of a number of inputs a firm employs and the price of each input. The firm's profit maximization problem reduces to determining the optimal output and input price levels as well as production and utilization levels of outputs and inputs. The following equation describes the profit maximization problem faced by a firm (Varian, 1992: page 26):

$$\Pi(p, w) = \max pf(\mathbf{x}) - w\mathbf{x}(w, y) \qquad (60)$$

where $\Pi(p, w)$ is the firm's profit, $pf(\mathbf{x})$ is the revenue, $w\mathbf{x}(w, y)$ is the cost of production, $f(\mathbf{x})$ is the output level, $p$ is the price output, $w$ is the vector of factor prices, and the inputs are measured by the nonnegative vector $\mathbf{x} = (x_1, \ldots, x_n)$. The cost function ($c(w, y) = w\mathbf{x}(w, y)$) measures the minimum cost of producing a given level of output for some fixed factor prices. As such, it summarizes information about the technological choices available to firms.

Firm's profit-maximizing behaviour can be characterized by the following first order condition (Varian, 1992: page 26):

$$p(\partial f(\mathbf{x}^*)/\partial x_i) = w_i \qquad (61)$$

This condition states that the value of the marginal product of each factor must be equal to its price and therefore the marginal revenue must be equal to the marginal cost. The second order condition for profit maximization is represented by the second derivative of the production function with respect to the input. The result of the second order derivative must be nonnegative (Varian, 1992: page 27):

$$\partial^2 f(\mathbf{x}^*)/\partial x_i^2 \leq 0. \qquad (62)$$

Several properties describe the profit function. These properties include: (Varian, 1992: page 41):

1. The profit function is non-decreasing in output prices and non-increasing in input prices. If $p'_i \geq p_i$ for all outputs and $p'_j \leq p_j$ for all inputs, then $\Pi(\mathbf{p}') \geq \Pi(\mathbf{p})$.
2. The profit function is homogenous of degree 1 in $\mathbf{p}$. $\Pi(t\mathbf{p}) = t\,\Pi(\mathbf{p})$ for all $t \geq 0$.
3. The profit function is convex in $\mathbf{p}$. Let $\mathbf{p}'' = t\mathbf{p} + (1-t)\mathbf{p}'$ for $0 \leq t \leq 1$, then $\Pi(\mathbf{p}'') \leq t\,\Pi(\mathbf{p}) + (1-t)\,\Pi(\mathbf{p}')$.
4. The profit function is continuous in $\mathbf{p}$ at least when $\Pi(\mathbf{p})$ is well-defined and $p_i > 0$ for $i = 1, \ldots, n$.

**4.10.2 Utility Maximization**

The second theory guiding the managerial decisions is the utility maximization theory. The basic problem of utility or otherwise known as preference maximization problem represents the set of all bundles (alternatives) satisfying the consumer's budget constraint. Each consumer is assumed to have preferences on each consumption bundles within a consumption set *X*. Consumer preferences are an ordered set of bundles where a bundle $\mathbf{x}$ is at least as good as a bundle $\mathbf{y}$ ($\mathbf{x} \geq \mathbf{y}$). A set of all consumption bundles defined as indifferent to each other is called an indifference curve. The preference maximization problem is defined as maximization of utility function that meets a budget constraint requirement (Varian, 1992: page 98):

$$Max\ u(\mathbf{x}) \qquad (63)$$

$$\text{such that } px \leq m \qquad (64)$$

$$\mathbf{x} \text{ is in } X. \qquad (65)$$

where $u(\mathbf{x})$ represents the utility function and $px \leq m$ represents the budget constraint with $m$ being the fixed amount of money available to a consumer ($\mathbf{B} = \{\mathbf{x} \text{ in } X\text{: } \mathbf{px} \leq m\}$). The theoretical preference maximization can be stated by utilizing the Lagrangian function (Varian, 1992: page 100):

$$\ell = u(\mathbf{x}) - \lambda(\mathbf{px} - m) \quad (66)$$

where $\lambda$ is the Lagrange multiplier. Differentiating the Lagrangian function with respect to $x_i$ gives the following first-order conditions (Varian, 1992: page 100):

$$\partial u(\mathbf{x})/\partial x_i - \lambda p_i = 0, \text{ for } i = 1,\ldots, k. \qquad (67)$$

When the $i^{th}$ first-order condition is divided by the $j^{th}$ first-order conditions the marginal rate of substitution between good $i$ and $j$ is obtained (Varian, 1992: page 100):

$$\partial u(\mathbf{x})/\partial x_i/\partial u(\mathbf{x})/\partial x_i = p_i/p_j, \text{ for } i,j = 1,\ldots, k. \quad (68)$$

The fraction on the right represents the economic rate of substitution between $i$ and $j$. The utility maximization exercise implies that the marginal rate of substitution is equal to the economic rate of substitution (Varian, 1992).

### 4.10.3 General Equilibrium

The economic theory of general equilibrium provides the basis for the efficient market hypothesis. The general equilibrium framework combines the firm's profit maximization problem with the problem faced by the consumer (utility maximization). In this framework, each consumer owns some initial endowment $\boldsymbol{\omega}_i$ of commodities for which he/she receives an income of $\boldsymbol{p\omega}_i$. For any vector of prices of goods and labor, the consumer considers selling off his/her endowment and then buying the desired bundle of goods including leisure. Given an endowment vector $\boldsymbol{\omega}$ and a price $\boldsymbol{p}$, the consumer solves the following problem (Varian, 1992: page 342):

$$Max\ u(\boldsymbol{x}) \qquad (69)$$

$$\text{such that } \boldsymbol{px} = \boldsymbol{p\omega}. \qquad (70)$$

The utility maximization problem involves constraint that includes the total amount of leisure consumed by the consumer being less than 24 hours a day (Varian, 1992).

To complicate the optimization problem, consumers own firms and receive a share of firm's profits ($T_{ij}$). For any firm $j$, share profits sum to one, $\sum_{j=1}^{n} T_{ij} = 1$, such that the firm completely owned by individual consumers. At a price vector $\boldsymbol{p}$ each firm $j$ chooses a production plan that yields a profit $\Pi(\boldsymbol{p}) = \boldsymbol{py}_j(\boldsymbol{p})$, where $\boldsymbol{y}_j(\boldsymbol{p})$ is a net supply function. The total profit income received by a consumer $i$ is the sum of the profits he/she receives from each of the firms: $\sum_{j=1}^{m} T_{ij}\, \boldsymbol{py}_j(\boldsymbol{p})$. As a result, the consumer's budget constraint is defined as follows (Varian, 1992: page 342):

$$\boldsymbol{px}_i = \boldsymbol{p\omega}_i + \sum_{j=1}^{m} T_{ij}\, \boldsymbol{py}_j(\boldsymbol{p}). \qquad (71)$$

The consumer chooses a utility-maximizing bundle that satisfies this newly defined budget constraint. As a result, consumer $i$'s demand function is written as a function of the price vector $\boldsymbol{p}$ (Varian, 1992).

For each consumer, an allocation ($\boldsymbol{x}, \boldsymbol{y}$) is feasible if aggregate holdings are compatible with the aggregate supply (Varian, 1992: page 345):

$$\sum_{i=1}^{n} \boldsymbol{x}_i - \sum_{i=1}^{n} \boldsymbol{y}_j - \sum_{i=1}^{n} \boldsymbol{\omega}_i = 0. \qquad (72)$$

A feasible allocation ($\boldsymbol{x}, \boldsymbol{y}$) is Pareto efficient if there is no other feasible allocation ($\boldsymbol{x'}$, $\boldsymbol{y'}$). Based on the Pareto efficiency definition, the First Theorem of Welfare Economics states that if ($\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{p}$) is Walrasian equilibrium, then ($x$, $y$) is Pareto efficient (Varian, 1992: page 345). The Second Theorem of Welfare Economics states that when ($\boldsymbol{x^*, y^*}$) is a Pareto efficient allocation with each consumer holding strictly positive amounts of each good and presenting convex, continuous, and strongly monotonic preferences. Suppose firm's production possibility sets, $Y_j$ for $j = 1,\ldots, m$ are convex, then there exists a price vector $\boldsymbol{p} \geq \boldsymbol{0}$ such that (Varian, 1992: page 346) :

1. if $\boldsymbol{x^*}_i >_i \ \boldsymbol{x^*}_i$ then $\boldsymbol{px'}_i > \boldsymbol{px^*}_i$ for $i = 1,\ldots,n$.
2. if $\boldsymbol{y^*}_i >_i \ \boldsymbol{y^*}_i$ is in $Y_j$ then $\boldsymbol{py'}_i > \boldsymbol{py^*}_i$ for $i = 1,\ldots,m$.

As a result, the Pareto efficient allocation is achieved by re-appropriation of "wealth". The second way of defining the Second Theorem of Welfare Economics is to assume that ($x^*$, $y^*$) is a Pareto efficient allocation and that preferences are locally non-satiated. When a competitive equilibrium exists from the initial endowments $\boldsymbol{\omega}_i = \boldsymbol{x^*}_i$ with profit share $T_{ij} = 0$ for all $i$ and $j$, and is given by ($p'$, $x'$, $y'$), then ($p'$, $x^*$, $y^*$) is a competitive equilibrium (Varian, 1992: page 348).