Consumer Economic Behavior and the Role of Information: Three Case Studies

Maria V. Vinoles Gomez

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

Doctor of Philosophy
in
Economics, Agriculture

Klaus Moeltner, Chair
Wen You
Kwok Ping Tsang
Steven Blank

September 17, 2014
Blacksburg, Virginia

Keywords: information, nutritional label, social norms, length of residency, community lenders

Consumer Economic Behavior and the Role of Information:
Three Case Studies

Maria V. Vinoles Gomez

(ABSTRACT)

The economics of information is a relatively new and important field of economics. This dissertation analyzes the role of information in three case studies within three different branches of economics: health economics, environmental economics, and finance and banking. First I analyze parental nutritional label usage and its effect of children's dietary outcomes (i.e. Health Eating Index and Body Mass Index). I show that parental usage of nutritional labels is associated with a better quality of their children's diet as well as an overall improvement in their health as measured by their Body Mass Index. Secondly, I study the behavioral effect of length of residency on water demand in the arid cities of Reno and Sparks in Nevada. In this case, I observe that social interaction among households affects their water usage. In particular, newcomers' watering behaviors are influenced by the prevailing social norms among neighbors that have lived in the arid area for a longer period of time. Finally, I compare the performance of local versus larger national and regional lending institutions in the years leading to the 2007 mortgage crisis. I find that local or community lenders have a significantly lower foreclosure rate during these years. Local lenders presumably base their origination decisions on an interpersonal relationship with their customers. This provides them with information that is not contained within the standard risk metrics generally used in loan applications. I discuss the policy implications of these results for each case study.

Acknowledgments

I would like to thank my committee members: Dr. Klaus Moeltner, Dr. Wen You, Dr. Kwok Ping Tsang, and Dr. Steven Blank. This dissertation would not have been possible without their corrections, suggestions and the concepts I learned on their microeconometrics, environmental economics, health economics, and macroeconomics lectures.

I would like to thank my advisor Dr. Klaus Moeltner for encouraging me to continue towards my PhD after I finished my Masters in Reno. I am also grateful to him that I could work under his learned guidance and financial support for the past two years.

I would like to specially thank Dr. Wen You for patiently introducing me to the perks of academic writing and microeconometrics along our research for chapter 1 of this dissertation.

I would like to thank my husband Christopher and my baby daughter Antonella for giving me the necessary motivation to conclude my doctorate degree.

Finally, I would like to say “gracias” to my mother, who carefully spoilt Antonella during the last five months while I completed this dissertation.

Contents

1 Introduction **1**

2 Parental Nutrition Label Usage and Children's Dietary-related Outcomes **4**

2.1 Introduction 4

2.2 Theoretical Model 6

2.2.1 Child's optimization problem 7

2.2.2 Parent's optimization problem 9

2.3 Empirical model 11

2.4 Data set 14

2.5 Results 18

2.6 Conclusions 21

2.A Covariance matrix for the simultaneous 3-equation model 26

2.B Estimates using BMI-percentiles 27

3 Length of Residency and Water Use in an Arid Urban Environment. **28**

3.1 Introduction 28

3.2 Households adaptation to the arid Southwest 32

3.3 Theoretical model 34

3.4 Description of the data 37

3.5 Empirical model and results 40

3.6 Robustness checks 45

3.7 Policy implications 47

3.8 Population growth and LOR 48

3.9 Conclusions 50

3.A Results including all control variables 60

3.B Winter versus summer water consumption. 62

4 National versus local lenders and distressed home sales during the 2007 mortgage crisis. 67

4.1 Introduction 67

4.2 Study objective 71

4.3 Data Set 73

4.4 Results 77

4.4.1 Logit model 77

4.4.2 Matching estimators 80

4.5 Conclusions 85

5 Conclusions 94

Bibliography 98

List of Figures

3.1 Total number of customers by initial year as TMWA customer 64
3.2 Average weekly use by LOR 65
3.3 Percentage of compliant weeks by LOR 66
3.4 Average weekly winter use by LOR 66

4.1 Percentage of foreclosures on total sales by city 93
4.2 Delinquency rate by year of origination 93

List of Tables

2.1 HEI and BMI by label users and non-users 23

2.2 Maximum likelihood estimates for the z-BMI and N (HEI) equations 23

2.3 Marginal Effects Label Usage equation 24

2.4 Average treatment effect on children's outcome: z-BMI, BMI percentiles and HEI 25

2.5 Covariances and their standard errors 26

2.6 Maximum likelihood estimates for the BMI-percentiles and N (HEI) equations . . 27

3.1 Descriptive statistics of water usage and explanatory variables 52

3.2 Mean of property characteristics by LOR 53

3.3 Average number of households per subdivision, within 50, 100 and 200 yards from each property. 53

3.4 Results for the natural log of weekly usage. 54

3.5 Winter weekly use and comparable sample in summer. 55

3.6 Washoe County population increase: historical and projected. 56

3.7 Avgerage weekly usage by LOR (years) 57

3.8 Total New residents (LOR$\leq$ 10 years) in 2008 and 2032 58

3.9 Number of households and aggregate weekly usage: 2008 vs 2032 58

3.10 Estimated difference in aggregate weekly usage accounting for LOR 59

3.11 Summer weekly use model including all regressors. 61

3.12 Winter weekly use and comparable sample in summer. 63

4.1 Number of lenders, number of sales and foreclosure rate by state. 88

4.2 Total and forced sales by year. 88
4.3 Percentage of foreclosure by Local/Non-local lenders. 89
4.4 Local and non-local lenders' average loan characteristics. 90
4.5 Socio-demographic characteristics of borrowers by local and non-local lenders. 90
4.6 Probability of a loan entering foreclosure (Average marginal effect, logit). . . 91
4.7 Matching estimators of ATT. 92

Chapter 1

Introduction

The First Theorem of Welfare Economics assumes perfect information of all agents. Pareto inferior equilibria, however, may be associated with too little or too much information regarding agents and their available choices (Stiglitz, 2002). Asymmetric information, that is when one side has more or better information than the other, can prevent individuals from taking into account all the possible outcomes of their choices and from making optimal decisions. In the real world, it is not rare for economic agents to base their decision on insufficient or incorrect information. For this reason, it is important to understand the consequences of decisions made in this information-poor scenarios, how those sub-optimal decisions can result in economic inefficiencies, and finally how to mitigate the effects of these asymmetries.

One way to improve decision making is by the disclosure of information via labeling (Cason and Gangadharan, 2002). Common examples of labeling include nutritional labeling on packaged food, ingredients lists of commercial products, energy consumption labels on appliances, and side effect lists on medicines. In chapter 2 of this dissertation I analyze parental nutritional label use and its effect on children's dietary outcomes. I show that parental usage of nutritional labels is associated with a better quality of their children's diet as well as an overall improvement in their health as measured by the Body Mass Index.

Alternatively, information can be passed from agent to agent through social interaction among them. For example, coworkers can inform each other about the pros and cons regarding different retirement plans (Duflo and Saez, 2002). In chapter 3 I study how households change their water consumption over time after moving to an arid city in the western USA. In this case, I observe that social interaction among households and between these and the water utility in the city of Reno and Sparks in Nevada, affect their water usage. I found that newcomers' watering behaviors are influenced by the local outdoor watering regulations and by prevailing social norms within their new neighborhood (i.e. watering and landscape norms among households that have lived in the arid area for a longer period of time).

An uninformed agent can also try to exert hidden information through screening mechanisms (Spence, 1973; Stiglitz, 1975). Examples of screening strategies are those used by banks evaluating the risk associated to their prospective borrowers, a firm interviewing candidates for a job position, or an insurance company assessing the optimal premium of its policies. This lead to the third case study of this dissertation. In chapter 4 I analyze the role of local or community lenders as oppose to larger national or regional lending institutions in the period leading to the 2007 mortgage crisis. The results from this chapter emphasize the importance of financial institutions' lending and screening procedures. By comparing the foreclosure rate of local versus non-local lenders, I show that the former have a significantly better performance during these years. Most importantly, I find that non-local lenders relaxed the screening and monitoring of their customers as well as their lending standards more severely than local lenders during this period. Not surprisingly, in the aftermath of the crisis, it was this group of lenders that experienced the highest rates of foreclosure.

I finalize the dissertation in chapter 5 with a discussion of the main conclusions and implications regarding the three empirical studies.

Chapter 2

Parental Nutrition Label Usage and Children's Dietary-related Outcomes

2.1 Introduction

The U.S. childhood obesity prevalence rate has reached 18% in 2010 (Ogden et al., 2012). This surge is more concerning since obese children are more likely to become obese adults. Furthermore, in addition to the onset of chronic diseases at a very young age, obese children are also more likely to be exposed to social discrimination that can cause serious emotional and behavioral problems, such as low self-esteem and depression (Dietz 1998, Guyer et al. 2009).

Policies and programs are being designed and implemented to promote healthy eating. Mandatory nutritional labeling regulation is one of them. The Nutrition Labeling and Educational Act (NLEA) was passed in 1990 and went into effect in 1994. It requires food manufacturers to provide nutritional information and serving sizes on a standardized label. Furthermore, the Health Care Bill, passed in March 2010, requires the presence of nutrition labels in chain restaurants and vending machines nationwide. Researchers have confirmed that the use of nutrition labels can lead to better food choices and overall diet qualities among adults (e.g. Chang and Nayga Jr, 2011; Drichoutis et al., 2009; Guthrie et al., 1995; Kim et al., 2000, 2001; Kreuter et al., 1996; Teisl et al., 2001; Variyam, 2008). However, there is scant literature on the relationship between nutrition label usage and the actual improvement in dietary-related outcomes such as weight. Drichoutis et al. (2009) found no significant linkage between individual's nutrition label usage and Body Mass Index (BMI). Furthermore, with one known exception, no study has explored the effect of parental nutrition label use on children's health. Chang and Nayga Jr (2011) is the only study that examines mother's nutritional label usage effect on children's weight and found a weak positive effect on children's BMI using data from Taiwan. Finally, nutritional label usage has broadly being analyzed with no theoretical model to justify what factors influence the decision to use nutritional information and its effect on dietary and health related outcomes. Teisl et al. (2001) is the only study that utilizes a theoretical framework to assess the impact of nutritional information on social welfare.

This paper addresses the three previously mentioned gaps that still remain in the literature. Firstly, this paper examines the effect of parental nutrition label usage on children's dietary-related outcomes (i.e., Healthy Eating Index (HEI) and BMI) by using U.S. national representative data (the 1994-1996 Continuing Survey of Food Intakes by Individuals (CSFII) and its complementary Diet and Health Knowledge Survey (DHKS)). Secondly, this paper will tackle the identification challenges facing nutrition label usage literature by guiding the empirical model specification and method choices through a theoretical model. The theoretical model is based on the two-stage Stackelberg model developed by You and Davis (2010), in which the parents are the leaders and the children are the followers. Thirdly, in our modified model, we include the parental nutrition label usage decision as one of the parental time allocation variables. This decision enters directly into the child's HEI production function while the child's HEI outcome in turns affects the child's production process of the dietary-related health outcome (i.e., BMI).

In the subsequent sections we present the theoretical model followed by the empirical method of estimation, the description of the data set used, and a summary of the main findings and implications for future research.

2.2 Theoretical Model

The household production theory introduced by Becker (1965) and its following variants by Lancaster (1966) and Grossman (1972), have been widely used and extended to analyze the role of the parents in overall household productivity, and in particular in children's nutritional intake quality and health status (e.g. Pitt and Rosenzweig, 1985; Behrman and Wolfe, 1987; Apps and Rees, 1997; Variyam et al., 1999; You and Nayga Jr, 2005; You and Davis, 2010; Cawley and Liu, 2012).

In the context of the household production model, household member's health (i.e. BMI) is considered a "commodity" produced at home by combining time and market goods subject to the household technology constraints (e.g. human capital and nutritional knowledge). Market goods (e.g. food items) have different characteristics (e.g. taste, nutrient content) and these characteristics are the relevant inputs in the production of the health related outcomes as well as in the utility derived from their consumption. Household members maximize the utility function derived from both market and at home produced goods subject to time, income and technology constraints. Solving this utility maximization problem, individuals determine their optimal demand for inputs and "commodities" production levels. We consider the HEI as a relevant input in the production of the dietary-related health outcome BMI. In turn, the HEI is the resulting demand of nutrients contained in the different food items consumed. This HEI demand function is conditioned on the parent's decision to use the nutritional labels.

Within this framework we adapt the two-stage Stackelberg model developed by You and Davis (2010) where the parents are the leaders and the children are the followers. We consider the parent's decision to read the information contained in the nutritional labels as one of their time allocation variable.

The model assumes continuous and concave utility functions and convex constraint sets as well as perfect information.

2.2.1 Child's optimization problem

By backward induction we start with the child's optimization problem, which can be summarized as:

$$\underset{(x,t)}{\text{Max}}\ u = f(BMI, x, t_f, t_{pa}, t_o, K, E_h, E_s) \tag{2.1}$$

subject to

$$BMI = f(N, t_{pa}, K, E_h, E_s), \tag{2.2}$$

$$N = f(x, X_f, T_f, K, E_h, E_s), \tag{2.3}$$

$$t_f + t_{pa} + t_o = \bar{t}, \tag{2.4}$$

The child maximizes (2.1) subject to (2.2), (2.3) and (2.4). Equation (2.1) corresponds to the utility the child derives from his own body mass index BMI, from the amount he chooses of x (vector of different food items consumed by the child), t_f (time the child allocates to consuming food), t_{pa} (time the child spends doing physical activity), t_o (rest of the child's time) and the three set of exogenous variables: K (the child's biological and genetic factors such as age and gender), E_h and E_s (household and school environmental factors respectively).

Equation (2.2) is the child's health production function (i.e. his BMI) with the different nutrient contents in his daily food intake as inputs (Variyam et al., 1999; Pitt and Rosenzweig, 1985). Specifically, we consider the child's HEI as a compound indicator of his daily nutrient intake. Besides the overall quality of his diet, the second input entering the BMI production function is the child's time allocated to physical activities (t_{pa}). Finally, the three set of exogenous factors that affect the efficiency of the child's BMI production function are: K, E_h and E_s.

Equation (2.3) corresponds to the child's nutrient demand N (i.e. HEI) which is a function of x, X_f (vector of different food items provided by the parent), T_f (parental time allocated to food preparation, including the time spent reading the nutritional information in the food labels) and the same exogenous conditioning factors that affect (2.2): K, E_h and E_s. The time parents spend preparing food has been shown to be positively associated with a better food quality and nutritional content (You and Davis, 2010).

Equation (2.4) is the child's time constraint.

Note that X_f and T_f are conditional variables that the parent determines in the first stage of the game taking into account the final response of the child.

Solving this optimization problem, we obtain the child's best response functions: x^* and t^* with arguments X_f, T_f, K, E_h, E_s and $\bar{t}$.

The child's indirect utility, production and demand functions are:

$$u^* = u(X_f, T_f, K, E_h, E_s) \tag{2.5}$$

$$BMI^* = BMI(X_f, T_f, K, E_h, E_s) \tag{2.6}$$

$$N^* = N(X_f, T_f, K, E_h, E_s) \tag{2.7}$$

2.2.2 Parent's optimization problem

Simplifying the parents' optimization problem, we assume that only one of them (either the father or the mother, the head of the household that is in charge of the grocery shopping) makes the first stage decisions.

The parent's optimization problem can then be summarized as:

$$\operatorname*{Max}_{(X,T)} U = f(BMI^*, u^*, X_f, X_o, T_f, T_w, T_o, \mu, NL, NA, E_h, E_w) \tag{2.8}$$

subject to

$$PX = wT_w + I + Y_s = Y, \tag{2.9}$$

$$T_f + T_w + T_o = \bar{t}, \tag{2.10}$$

The parent maximizes (2.8) subject to (2.9) and (2.10).

In equation (2.8) U is the utility the parent derives indirectly from the child's health status (BMI^*) and the child's utility (u^*), and directly through X_f, X_o (consumption of all other market goods), T_f, T_w (time the parent spends at work), T_o (time the parent spends in other activities), and the exogenous variables: μ (parent's biological and genetic factors such as age, gender and race), NL (parent's perceived benefits of using nutritional labels), NA (parental nutritional and health awareness and knowledge), and the environmental factors E_h and E_w (household and work respectively).

Equation (2.9) is the household budget constraint where P is the vector of market prices, X is the vector of all market goods (food (X_f) and non-food related (X_o)), w is the wage rate, I the unearned household income, Y_s the spouse's income, and Y total household income.

Equation (2.10) is the parent's time constraint.

Substituting (2.5) and (2.6) into (2.8) and solving his optimization problem, the parent optimal input demand functions are:

$$X^* = X(P, w, I, Y_s, K, \mu, NL, NA, E_h, E_w, E_s) \tag{2.11}$$
$$T^* = T(P, w, I, Y_s, K, \mu, NL, NA, E_h, E_w, E_s) \tag{2.12}$$

Where T is a vector with elements T_f, T_w and T_o.

The parent's indirect utility function is:

$$U^* = U(P, w, I, Y_s, K, \mu, NL, NA, E_h, E_w, E_s) \tag{2.13}$$

Substituting (2.11) and (2.12) into equations (2.6) and (2.7) we obtain the child's final reduced BMI^{**} and N^{**} forms:

$$BMI^{**} = BMI(P, w, I, Y_s, K, \mu, NL, NA, E_h, E_w, E_s) \tag{2.14}$$
$$N^{**} = N(P, w, I, Y_s, K, \mu, NL, NA, E_h, E_w, E_s) \tag{2.15}$$

2.3 Empirical model

Identifying the treatment effect of parental label usage T_f on the child's N (i.e. HEI) and BMI outcomes, is a similar problem to that of assessing the effect of training on earnings (Heckman and Robb Jr, 1985; Greene, 2012). The decision to use nutritional labels (treatment) is not random. Variables such as parental age, gender or education, may affect this self-assignment of households into the treated and non treated groups. Failing to account for these factors, and if they happen to affect both the treatment decision and the outcome equations may result in sample selection bias.

We estimate a recursive system of 3 equations: the child's structural (BMI) production function (equation 2.2), his indirect nutrient content demand function (N) conditional on whether the parent reads the information contained in the nutritional labels (equation 2.7) and the reduced form equation corresponding to the parental nutritional label usage (T_f). We use the child's HEI as a surrogate for N. We do not directly observe the time parents spend preparing the food. The variables available are: an indicator variable of whether the parent uses nutritional labels, and the amount of money spent on fast food and on food away from home. We use these last two variables to control for the rest of the time the parent spends in food preparation.

Equations (2.2) and (2.7) are represented by the following linear regression functions:

$$BMI = \boldsymbol{\beta}'_{BMI} \left[N \vdots t_{pa} \vdots \mathbf{K} \vdots \mathbf{E_h} \vdots \mathbf{E_s} \right] + \varepsilon_{BMI} = \boldsymbol{\beta}'_{BMI} \mathbf{Z_{BMI}} + \varepsilon_{BMI} \tag{2.16}$$

$$N = \boldsymbol{\beta}'_{N} \left[X \vdots T_f \vdots \mathbf{K} \vdots \mathbf{E_h} \vdots \mathbf{E_s} \right] + \varepsilon_{N} = \boldsymbol{\beta}'_{N} \mathbf{Z_N} + \varepsilon_{N} \tag{2.17}$$

Where $\boldsymbol{\beta}_{BMI}$ and $\boldsymbol{\beta}_{N}$ are the vectors of coefficients associated with the corresponding variables included in each equation.

In the case of the dummy indicator T_f, we assume the existence of an unobserved latent utility T_f^* such that:

$$T_f^* = \boldsymbol{\beta'_{T_f}} \left[p \vdots w \vdots I \vdots Y \vdots Y_s \vdots \mathbf{K} \vdots \boldsymbol{\mu} \vdots \mathbf{NL} \vdots \mathbf{NA} \vdots \mathbf{E_h} \vdots \mathbf{E_s} \vdots \mathbf{E_w} \right] + \varepsilon_{T_f} = \boldsymbol{\beta'_{T_f}} \mathbf{Z_N} + \varepsilon_{T_f} \tag{2.18}$$

$$T_{fi} = \begin{cases} 1, & \text{if } T_{fi}^* > 0; \\ 0 & \text{otherwise.} \end{cases} \tag{2.19}$$

Where β_{T_f} is the vector of coefficients associated with the corresponding exogenous variables in the nutritional label usage's latent utility equation.

For a given observation i, of matched parent and child, we assume that the error terms follow a trivariate normal distribution with covariance matrix Σ :

$$\boldsymbol{\varepsilon} = \begin{pmatrix} \varepsilon_{BMI} \\ \varepsilon_N \\ \varepsilon_{T_f} \end{pmatrix} \sim \mathcal{N}(0, \boldsymbol{\Sigma})$$

$$\Sigma = \begin{pmatrix} \sigma^2_{BMI} & \sigma_{BMI,N} & \sigma_{BMI,T_f} \\ \sigma_{BMI,N} & \sigma^2_N & \sigma_{N,T_f} \\ \sigma_{BMI,T_f} & \sigma_{N,T_f} & 1 \end{pmatrix}$$

Finally, we assume that the error terms across different pairs of parent and child are independent and identically distributed.

The correlation between the parent's unobservables (ε_{T_f}) and those of the child (ε_N and ε_{BMI}) determine whether the treatment is exogenous or endogenous. In the first case, with no correlation between the parent's and child's unobservables, the probit equation for nutritional label usage can be estimated independently of the child's outcome equations. On the other hand, in the more general case of non-zero correlation between the parent's and child's unobservables, this procedure will lead to sample selection bias and inconsistent estimates.

We estimate the system above by Full Information maximum likelihood (FIML). The likelihood function is given by:

$$L = \prod_{i=1}^{n} \left[\phi(e_{BMIi}, e_{Ni}) * \Phi \left(\frac{\boldsymbol{\beta'_{T_f}} \mathbf{Z_{T_f i}} + \mu_i}{\sigma} \right) \right]^{T_{f_i}} * \left[\phi(e_{BMIi}, e_{Ni}) * \Phi \left(\frac{-\boldsymbol{\beta'_{T_f}} \mathbf{Z_{T_f i}} - \mu_i}{\sigma} \right) \right]^{1-T_{f_i}}$$

Where,

$\mathrm{e}_{BMIi} = \frac{BMI_i - \boldsymbol{\beta'_{BMI}} \mathbf{Z_{BMIi}}}{\sigma_{BMI}}$

$\mathrm{e}_{Ni} = \frac{N_i - \boldsymbol{\beta'_N} \mathbf{Z_{Ni}}}{\sigma_N}$

$$\mu_i = \begin{bmatrix} \sigma_{BMI,T_f} & \sigma_{N,T_f} \end{bmatrix} \begin{bmatrix} \sigma^2_{BMI} & \sigma_{BMI,N} \\ \sigma_{BMI,N} & \sigma^2_N \end{bmatrix}^{-1} \begin{bmatrix} e_{BMIi} \\ e_{Ni} \end{bmatrix}$$

$$\sigma^2 = 1 - \begin{bmatrix} \sigma_{BMI,T_f} & \sigma_{N,T_f} \end{bmatrix} \begin{bmatrix} \sigma^2_{BMI} & \sigma_{BMI,N} \\ \sigma_{BMI,N} & \sigma^2_N \end{bmatrix}^{-1} \begin{bmatrix} \sigma_{BMI,T_f} \\ \sigma_{N,T_f} \end{bmatrix}$$

$\phi(\cdot,\cdot)$ is ε_{BMI} and ε_N's bivariate normal density function.

$\Phi(\cdot)$ is the univariate standard normal cumulative distribution function.

2.4 Data set

The datasets used for this study are the USDA's 1994-1996 Continuing Survey of Food Intakes by Individuals (CSFII) and its complementary Diet and Health Knowledge Survey (DHKS).

The CSFII is a nationally representative dataset containing not only socio-demographic variables but also information on 2 non-consecutive 24-hour recalls of dietary intake. Based on this nutrient intake information the USDA's Center for Nutrition Policy and Promotion (CNPP) calculates the HEI as an overall measure of dietary quality and compliance with federal dietary guidance. The HEI used in the 1994-1996 CSFII takes into account ten different components: component 1-5 measure the degree to which a person's diet conforms to the USDA's Food Guide Pyramid serving recommendations for the five major food groups: grains, vegetables, fruits, milk and meat. Components 6 and 7 measure total fat and saturated fat consumption as a percentage of total food energy intakes. Components 8 and 9 measure total cholesterol and sodium intake, and component 10 takes into account the variety of a person's diet. Each component has a score ranging from 1 to 10, with an overall HEI ranging from 0 to 100. Higher scores indicate closer compliance with recommended values.[2]

The DHKS is a telephone follow-up for adults older than 20 years (usually the head of the household). The DHKS was initially designed to assess compliance with the Dietary Guidelines for Americans in an effort to understand consumers' nutritional knowledge, attitude and behavior, and to identify channels to promote a better nutrition education and overall health quality. The questionnaire's topics range from awareness of dietary related diseases, and knowledge about nutrient content in different food items, to questions such as the main factors that affect food purchase choices (e.g. time restrictions, safety concerns, price). Additionally, the DHKS contains a unique set of questions regarding food label usage (e.g. how often food labels are used, how easy they are to understand, if it takes too much time to read them, and how confident individuals are about the information they contain).

[2] The HEI formula was later revised in 2012 to reflect the 2010 Dietary Guidelines for Americans.

The dependent variables used for our analysis are: the child's BMI (standardized), the child's HEI and the parental nutritional label usage.The child's BMI and HEI were calculated based on self reported weight, height and dietary intake for children older than 12 years old. In the case of children between 2 and 12 years old, this information was provided by the adult in charge. The nutritional label usage is a dichotomous variable for those adults that reported using the Nutrition Facts Panel contained in the labels. Unlike the case of adults, a child is considered overweight or obese according to the distribution of children of the same age and gender. Due to differences in body structures, boys have a higher BMI than girls. Besides, younger children have lower BMI, and children's BMI distribution, as well as the cut-off point after which a child can be classified as overweight or obese have, after an initial decrease at age 2, a consistent positive trend that goes from 18 and 17 in children 3 years old to the standard 25 and 30 in adults 18 years or older (Cole et al, 2000). A child with a BMI above the 85^{th} percentile is considered overweight and he is considered obese if his BMI is above the 95^{th} percentile. For these reasons, we standardized children's BMI and calculate children's BMI z-score and BMI percentiles using the 2000 CDC's (Center for Disease Control and Prevention) national reference growth charts.[1]

The explanatory variables include: the child's biological and genetic factors $\boldsymbol{K}$ (age, gender, health status, being on a special diet, Hispanic origin and race); the parent's biological and genetic factors $\boldsymbol{\mu}$ (age, gender, health status, being on a special diet, parental exercise and smoking habits, the parent's concern with food safety (safe), how important is taste when buying food (taste) and how important is that food be easy to prepare (prepare)); the parent's perceived benefits of using nutritional labels $\boldsymbol{NL}$ (concern about nutritional label's

[1] We use the online SAS program provided by the CDC which corrects for outlier (biological implausible) values.

interpretation (interpret), reading the labels consumes too much time (time), nutritional labels help make better food choices (choices)); the household environmental characteristics $\boldsymbol{E}_h$ (household size, whether the households lives in a central city or in a non-metropolitan area, household's participation in the Food Stamp Program (FSP), and parental BMI and HEI); school environmental factors $\boldsymbol{E}_s$ (whether the child attends school or not); work environmental factors $\boldsymbol{E}_w$ (whether the parent works part time); and proxy variables for the additional time the parent spends on food preparation other than reading the labels $\boldsymbol{T}_f$ (weekly amount of money spent on fast-food and food-away- from-home). Finally, we include the parent's nutritional knowledge, awareness and attitude variables $\boldsymbol{N}A$. Following Variyam et. al. (1999), we calculate a nutrient content knowledge variable NCK and a diet and health awareness variable DHA. The NCK is the count of corrects answers to 15 questions that compare the amount of different nutrients contained in different food items (e.g. "Which has more saturated fat: liver or T-bone steak?"). The DHA is the number of correct answers to 6 questions regarding awareness of the relation between the intake of nutrients (fat, fiber, salt, calcium, cholesterol and sugar) and related health problems (e.g. "Have you heard about any health problems caused by eating too much salt or sodium?"). Last, we separate the response to the question "Have you heard about any health problems caused by being overweight?" into the variable AO to quantify the effect of the awareness of being overweight.

The sample size is 1,085 pairs of parent and child.[2] The children's age ranges from 2 to 18 years old. Over 80% of the parents in charge of doing the grocery shopping are females. About 73% of the parents read the nutritional panel in the labels. The mean BMI of the sampled children is 19.67 and the average HEI of these children is 64.93 which is below the

[2]We consider adults that are the head of the household (or their spouses or partners), that do grocery shopping and only their natural or adopted children.

recommended HEI of 80 (You and Nayga Jr, 2005).

Table 2.1 summarizes the mean HEI and BMI of both parents and children, separating them into two groups: parent label users from parent label non-users. On average, parents in the first group have a 6.39 points higher HEI and a 0.16 point lower BMI than those in the second group. The HEI of children whose parents read the label is also higher than that of children whose parents do not read them but this difference in smaller (3.81). The difference in the BMI is slightly higher in the case of children when comparing the two groups (0.84).

2.5 Results

We first estimate a simultaneous three equation model including equations (2.16) and (2.17) (corresponding to the children's outcome) and the probit equation for parental label usage. There is a significant positive correlation between the error terms in the children's outcome equations ($\sigma_{BMI,N}$). However, the error terms in the BMI and the N equations do not present a significant correlation with the parental label usage's unobservables ($\sigma_{T_f,BMI}$ and $\sigma_{T_f,N}$ are not significantly different from zero). The chi-square with 2 degrees of freedom for the Wald test comparing both models is $\chi^2(2) = 0.87$, and we fail to reject the null hypothesis of both parameters $\sigma_{T_f,BMI}$ and $\sigma_{T_f,N}$ being equal to zero. We conclude that the self selection bias is insignificant in our sample and that it is efficient to estimate the probit label usage equation separately from the children's outcome equations.[3]

[3]Estimates for the covariance matrix of the simultaneous 3-equation model is presented in Appendix A. Table 1.

Results obtained estimating separately the children's outcome equations from the parental label usage probit equation are reported in Tables 2.2 and 2.3. Table 2.2 shows the parameter estimates for the children's dietary outcomes N (i.e. HEI) and BMI (measured by the z-BMI).[4] We start discussing the results on the conditional nutrient demand equation (N). After controlling for self selection, we estimate that nutritional labels have an average treatment effect on the treated (ATT) of 2.405 points on children's HEI. Results for the remaining control variables coincide in general with previous studies on the determinants of US children's HEI (Variyam et al., 1999; You and Nayga Jr, 2005). African American children have on average a 2.65 points lower HEI and children living in metropolitan central city areas have a 2.27 higher HEI. Our findings additional suggest that children attending school (including kindergarten and high school) have on average a 2.125 points lower HEI. The variable *school* was asked only for children 5 to 18 years old. For this reason, we conclude that attending school has a significant negative effect on these relatively older children's HEI. Finally, parental dietary pattern (measured by the parent's HEI) is a significant positive predictor of children's HEI. This result confirms that household environmental factors and familial dietary habits do have an impact on children's overall diet quality (Beydoun and Wang, 2009).

Parameter estimates for the children's BMI equation (measured by the z-BMI) are also presented in Table 2.2. An increase of 1 point on children's HEI, on average, results in a significant lower BMI of 0.039 standard deviations. Taking into account the initial positive effect of nutritional labels on children's HEI, we estimate a significant ATT of nutritional labels on children's z-BMI of -0.09 ($-0.039 * 2.405$) with standard errors equals to 0.04.[5] Within the household environmental variables, parental BMI shows a significant positive

[4]Analogous results are obtained using BMI percentiles. These results are presented in Appendix 2.

[5]Standard error for the ATT on children's z-BMI were calculated using the delta method (Agresti, 2010).

correlation with children's BMI (Strauss and Knight, 1999; Danielzik et al., 2002, 2004).

Table 2.3 reports the parameter estimates of the probit label usage model. Our results are in line with previous studies on US adults' nutritional label usage (Drichoutis et al., 2006; Kim et al., 2000, 2001). The probability of an adult using nutritional labels increases with higher income levels, if the person is of Hispanic origin, if she is a female, on a special diet, that lives in non-urban areas, and if he/she practices sports and does not smoke. Our new findings are that the probability of reading the nutritional labels increases in the case of parents with children that have a health status of poor or fair. Meanwhile, parents with children that attend school (including kindergarten and high-school) have a lower tendency to use the nutritional information in the labels. Furthermore, we specifically include in our model variables directly related to the parents' perceived benefits of nutritional label usage. Parents that use nutritional labels reported that they help them make better choices. Besides, they showed concern about how to interpret the information contained in the labels. On the other hand, parents that do not read the nutritional labels argued that reading them takes too much time. There is reason to believe these variables directly affect the probability of using nutritional labels but do not have a direct impact on children's dietary outcomes.

Summarizing our results, Table 2.4 shows the ATT of parental nutritional label usage on children's dietary related outcomes. Parental nutritional label usage is associated with a 2.405 higher HEI in their children, and a reduction of 2.58 percentile points in their BMI. These ATT are significant at the 1 % and 5 % respectively.

2.6 Conclusions

We analyze US adults' Nutrition Facts Panel usage and how it influences their children's dietary quality (i.e. HEI) and diet-related health status (i.e. BMI). Within the household production theoretical framework we adapt the two-stage Stackelberg model developed by You and Davis (2010) where the parents are the leaders and the children are the followers. In the model, we consider the parental nutritional label usage as one of the parent's time allocation variables. This decision influences directly the child's HEI conditional demand for nutrients. The child's HEI in turn, enters his final production process of the dietary-related health outcome (i.e, BMI). We estimate the resulting 3-equation recursive system by Full Information Maximum likelihood utilizing data from the CSFII 1994-1996 and its companion DHKS. After accounting for self selection, we identify a significant average treatment effect on the treated (ATT) of 2.405 and -0.093 on children's HEI and z-BMI respectively. These ATT's are significant at 1% and 5% respectively. Our analysis then, allows as to identify the sequential mechanism through which parental nutritional label usage affect first and directly the quality of their children's diet and only subsequently and indirectly (through this better dietary quality) their weight status. Furthermore, our results suggest that parents do use nutritional label as a mechanism to choose healthier diets for their children. Parents that reported having children with a poor or fair health status manifested using nutritional labels more often than parents with healthier children. On the other hand, parents with children attending school tend to read nutritional labels less frequently. This result, suggests that parents may delegate some of the care in preparing the food once children enter a formal education outside the household.

Our findings provide direct evidence of the relationship between children's dietary related outcomes and parental nutritional label usage. For this reason, the results emphasize the relevance of programs to educate parents on the proper understanding and usage of nutritional labels. The construction and dissemination of educational programs to teach parents, especially those in lower income areas can help reduce the risk of overweight or obesity among children. Our findings additionally indicate that parents tend to delegate nutritional care once children enter a formal education outside the house. This also suggests the relevance of programs such as the Team Nutrition School Programs that is currently being implemented by the USDA. The program states that good nutrition begins at home and continues in the school. For this reason, the program aims at getting parents involved with the nutritional programs carried out at schools. The program is based on teams formed by parents, teachers, and health educators, and it includes activities that can be later shared with the children's families at home. We believe that these programs can be complemented by a more detailed explanation of what nutritional labels mean, and what information to look at in the food labels.

Table 2.1: HEI and BMI by label users and non-users

	label users	label non users	
Variable	**Mean**	**Mean**	**Difference**
Freq	793	292	
HEI parent	63.34	56.95	6.39***
HEI child	65.96	62.15	3.81***
BMI parent	26.13	26.28	-0.16
BMI child	19.45	20.29	-0.84**

(*), (**) and (***) significant at 10%, 5% and 1 % respectively.

Table 2.2: Maximum likelihood estimates for the z-BMI and N (HEI) equations

	z-BMI	**coeff**	**s.e.**	**(*)**		**HEI**	**coeff**	**s.e.**	**(*)**
	HEI	-0.039	0.016	**	Tf	Label use	2.405	0.681	***
Tpa	TV	-0.001	0.035			Fast food	-0.036	0.022	
	Exercise	0.169	0.120			Food away from home	0.006	0.009	
K	Special diet	0.082	0.279		K	Special diet	-0.059	1.779	
	Child's health	0.141	0.264			Child's health	-4.037	2.136	*
	Non Hispanic	-0.117	0.280			Non Hispanic	0.143	1.187	
	Af. American	-0.146	0.392			Af. American	-2.660	0.905	***
	Asian	0.276	0.244			Asian	1.280	1.938	
	Others	-0.007	0.303			Others	-2.175	1.420	
Eh	HH size	-0.017	0.055		Eh	HH size	0.384	0.257	
	City	0.212	0.178			City	2.204	0.756	***
	Non metro	0.065	0.142			Non metro	-0.127	0.786	
	Food stamp	-0.253	0.217			Food stamp	0.399	0.833	
	BMIp	0.054	0.012	***		HEIp	0.281	0.031	***
Es	School	-0.841	0.166	**	Es	School	-2.125	0.853	**
	Constant	2.248	1.094	***		Constant	45.867	2.803	***

(*), (**) and (***) significant at 10%, 5% and 1 % respectively.

Table 2.3: Marginal Effects Label Usage equation

	Label use	**coefficient**	**robust s.e.**	**(*)**
	Income	0.002	0.001	***
K	Child's age 2 to 4	-0.076	0.056	
	Child's age 5 to 11	0.022	0.025	
	Special diet	0.011	0.076	
	Child's health status (poor or fair)	0.126	0.045	***
	Non Hispanic	-0.108	0.049	**
	Af. American	0.029	0.036	
	Asian	0.026	0.079	
	Others	-0.104	0.059	*
μ	Parent's age	0.002	0.002	
	Parent's gender	-0.190	0.028	***
	Parent on special diet	0.099	0.039	***
	Parent No health	0.028	0.035	
	Grade	0.004	0.006	
	Smoke	-0.091	0.024	***
	Exercise	0.059	0.024	**
	TV	0.005	0.007	
	Safe	-0.031	0.050	
	Prepare	0.029	0.026	
	Taste	0.093	0.080	
NL	Interpret	0.051	0.024	**
	Time	-0.047	0.025	*
	Choices	0.269	0.019	***
NA	DHA	0.005	0.009	
	NCK	0.012	0.005	***
	AO	0.052	0.048	
Eh	HH size	-0.013	0.009	
	City	-0.093	0.028	***
	Non metro	-0.041	0.027	
	Food stamp	0.003	0.035	
Es	School	-0.111	0.046	**
Ew	Employed part time	0.018	0.030	

(*), (**) and (***) significant at 10%, 5% and 1 % respectively.

Table 2.4: Average treatment effect on children's outcome: z-BMI, BMI percentiles and HEI

	ATT	**s.e.**	**(*)**
HEI	2.405	0.681	***
z-BMI	-0.093	0.042†	**
BMI percentile	-2.579	1.102†	**

† Standard errors calculated using the Delta Method.

(**) and (***) significant at 5% and 1% respectively.

2.A Covariance matrix for the simultaneous 3-equation model

Table 2.5: Covariances and their standard errors

	estimate	**s.e.**	**(*)**
σ_{BMI}	4.376	0.215	***
σ_{N}	8.474	0.202	***
$\sigma_{BMI,N}$	10.284	3.614	***
$\sigma_{T_f,BMI}$	-0.138	0.274	
$\sigma_{T_f,N}$	-1.252	0.996	

(***) significant at 1 %.

2.B Estimates using BMI-percentiles

Table 2.6: Maximum likelihood estimates for the BMI-percentiles and N (HEI) equations

	BMI-percentiles	**coeff**	**s.e.**	**(*)**		**HEI**	**coeff**	**s.e.**	**(*)**
	HEI	-0.975	0.344	***	Tf	Label use	2.646	0.637	***
Tpa	TV	0.378	0.549			Fast Food	-0.038	0.021	*
	Exercise	3.719	2.218	*		Food away from home	0.006	0.009	
K	Special diet	1.985	6.762		K	Special diet	-0.056	1.780	
	Child's health	1.632	6.097			No health	-4.037	2.121	*
	Non Hispanic	-7.631	4.862			Non Hispanic	0.180	1.189	
	Af. American	1.909	4.126			Af. American	-2.710	0.907	***
	Asian	9.197	6.428			Asian	1.272	1.932	
	Others	-1.675	5.759			Others	-2.145	1.424	
Eh	HH size	-0.781	0.947		Eh	HH size	0.389	0.257	
	City	2.992	2.937			City	2.243	0.759	***
	Non metro	2.444	2.817			Non metro	-0.138	0.785	
	Food stamp	-3.275	3.429			Food stamp	0.365	0.829	
	BMIp	0.948	0.198	***		HEIp	0.275	0.032	***
Es	School	-13.592	2.968	***	Es	school	-2.114	0.854	**
	Constant	116.868	24.401	***		Constant	46.039	2.805	***

(*), (**) and (***) significant at 10%, 5% and 1 % respectively.

Chapter 3

Length of Residency and Water Use in an Arid Urban Environment.

3.1 Introduction

Southwestern arid urban regions have the highest per capita residential water usage in the United States. While the national average is 98 gallons per day, in the State of Nevada daily per capita consumption amounts to 190 gallons (Kenny et al., 2009). Landscape irrigation accounts for the majority of this higher than average water use (US. Environmental Protection Agency (a)).

A well-known problem affecting the southwestern states is the pressure that the increasing population poses on water supplies (Fort, 2002). During the period from 1993 to 2008, the states of Arizona, Texas, Nevada and Colorado were among the top ten states with the highest net immigration flows from other US states (IRS, U.S. Population Migration Data, 1993-2008). Six of the ten cities with the highest total population increase in 2012 are in an arid or semi-arid western state: Houston, San Antonio, Austin, Phoenix, Dallas and Fort Worth (U.S. Census Bureau, 2013). More importantly, the population in states with the highest per capita residential water usage is expected to continue to rise (U.S. Census Bureau, Population Projection: 2004-2030; U.S. Environmental Protection Agency (b)). This stress on water resources will likely be intensified by the concomitant change in climatic conditions. Future global warming is believed to result in an increase in temperature, changes in precipitation intensity, increased evaporation and more pronounced drought in the region (U.S. Environmental Protection Agency (c)).

Inevitable questions arise within government agencies and water utilities concerned about the efficient provision of future demanded water: What are the water consumption habits of this incoming population? Do household that arrive to the Southwest from other states consume higher volume of water than native residents or households that have lived in the area longer? Do they use water less carefully? How much do they know and how much do they comply with water conservation policies in their new arid urban environment?

Very little research has been conducted comparing new residents' water consumption to that of long-time residents. Remarkably, this scant literature seems to be reaching a rather counterintuitive conclusion: homeowners tend to increase outdoor water consumption the longer they have lived in the arid city. Moreover, the length of residency (LOR, hereafter) appears to have a negative effect on households' attitude towards outdoor water conservation behavior and policies (e.g. Olsen and Highstreet, 1987; Agthe et al., 1988; Martin et al., 2003; Spinti et al., 2004; Yabiku et al., 2008; Harlan et al., 2009; Hilaire et al., 2010; Larson et al., 2011).

Of these studies, Agthe et al. (1988) and Harlan et al. (2009) are the only one that use data on water consumption to quantify the effect that the LOR has on water demand. The former found that an additional year of tenancy in Tucson, Arizona increases households' monthly summer water usage by 35.5 cubic feet. The later estimate that households with more than 5 years of residency in the city of Phoenix, Arizona consume on average 10.5% higher volume of water than newer residents. In spite of this general agreement, the reasons explaining why families increase their water consumption after moving into an arid city are less clear.

Our paper analyzes the effects of LOR on water consumption in the metropolitan area covering the cities of Reno and Sparks in the state of Nevada. We use daily metered water consumption at the household level observed for 63 consecutive days in the summer of 2008. Apart from evaluating the specific Reno-Sparks' case study, the main contributions of our research to the existing literature are threefold.

Firstly, our paper specifically focuses on the LOR and its effect on water demand. In particular, ours is the first study to quantify the behavioral effect of LOR controlling for both unobserved neighborhoods' characteristics, and social norms and regulations (e.g. Agthe et al., 1988; Spinti et al., 2004; Randolph and Troy, 2008; Yabiku et al., 2008; Hilaire et al., 2008; Harlan et al., 2009; Larson et al., 2011). In doing so, we are able to show that the LOR significantly explains variation in water consumption within small areas (subdivisions) that usually share common property and yard characteristics. Additionally, the LOR has a positive effect on water use that can neither be explained by the influence of social norms within close neighborhoods nor by outdoor watering regulations effective in the period of analysis. For this reason, we conclude that with our empirical model, we are able to isolate variations in households consumption due to private changes in preferences as oppose to the effect that social norms or regulations cause on water consumption. We suggest that this changes are due to changes in taste, or the knowledge about the arid environment.

Secondly, we identify that households living within 50 yards from one another significantly influence the volume of water used by each other. While several studies have noted the relevance of community norms on new residents' water consumption, Harlan et al. (2009) are the only to include proxy variables for social norms in their analysis. However, and probably because they are unable to capture actual changes in households' behavior, these regressors (survey responses that summarize: concern about water supply, trust in that others conserve water, and belief that a collective solution is efficient to improve the quality of the neighborhood) were not significant in explaining water consumption. In our study, we calculate the average water consumption of households living within 50 yards of each property to measure how households' outdoor water behavior (reflected in their water usage) influences water used by nearby households. Spatial correlation on residential water demand, however, can also be associated with other structural and sociodemographic factors. For example, a tendency of individuals with similar sociodemographic characteristics to cluster together in similar properties with homogenous yards (Martin et al., 2004). Neighborhoods with varying exposure to climatic variables are also associated with different water consumption patterns (Balling et al., 2008). This geographic heterogeneity may confound the identification of a purely social interaction among nearby households. We remove the effect of this unobserved heterogeneity by including in our empirical model spatial fixed effects at the subdivision level.

Finally, once we isolate both a pure private and social effects, we proceed to investigate the interaction between them. In particular, we analyze how social norms and compliance with effective regulations interact with the LOR along the adaptation process of new households.

The remaining sections of the paper introduce the theoretical model of analysis, followed by a description of the data set, a discussion of the main findings and their policy implication, finalizing with the central conclusions of the study.

3.2 Households adaptation to the arid Southwest

Individuals that move from their original society in order to settle in a new host society undergo an acculturation process that involves the learning of new socio-cultural skills (e.g. making friends, getting used to local foods, complying with local rules and regulations, coping with the climate, understanding the local value system) (Berry, 1997; Ward and Kennedy, 1999). Researchers have found that this process of "cross-cultural adaptation" follows a learning pattern characterized by an initial rapid learning rate ("culture shock") over the first few months followed by a final settling stage once the culture-learning process has been completed (Ward and Kennedy, 1999).

Three opposing theories have been argued to explain the adaptation process of households that move to Southwestern arid cities:

1) Human capital accumulation: in this case, outdoor water demand changes over time as

the result of an increase in knowledge and awareness about local environmental concerns. For instance, long time residents to the Reno-Sparks area may be aware of the pressure that the cities' water demand exerts on the Truckee River's natural inflows to Pyramid Lake (regarded as one of the United States' most beautiful desert lakes, host of endangered fish species, and economic center to the Paiute Tribal Reservation). Over time, households may also learn how to efficiently use water in their new climatic conditions (e.g. plant native vegetation, use water-saving devices). Both the newly acquired knowledge and concern will likely motivate more skillful and prudent watering behaviors, and consequently, result in a more efficient outdoor water usage and an overall reduction or savings in water consumption (Spinti et al., 2004; Yabiku et al., 2008; Larson et al., 2011).

2) Physical capital accumulation: on the other hand, time and resources allocated to property yards may change along the different stages of a household's life cycle (Spinti et al., 2004; Yabiku et al., 2008; Larson et al., 2011). For instance, younger families may not have the resources to invest in rich landscaping, families with children may have an additional incentive to do so, and older people may enjoy gardens in their spare time. New residents may also at first be attracted by native plants and landscape, but these initial preferences may shift over time. In this case, previous investments made in gardening-related goods may result in increasing water consumption over time (e.g. the incorporation of new perennial plants each season, the planting of young trees or bushes whose growing process may take years to complete). On the other hand, the incorporation of more efficient devices (e.g. smart sprinkler controllers, soil moisture sensors) would result in a reduction in water consumption (Hilaire et al., 2008).

3) Social capital accumulation: social interaction among households and between these and

water utilities (i.e. social norms and regulations) affect their water usage (Jorgensen et al., 2009). Previous studies have confirmed that an individual's perception that others are not taking water conservation actions can restrain a person's saving effort (Allcott, 2011; Larson et al., 2009; Miller and Buys, 2008; Athanasiadis et al., 2005; Pretty and Ward, 2001). In a similar way, effective regulations can either promote efficient water usage or they may discourage it. For instance, Castledine et al. (2014) find that compliance with outdoor watering restrictions in the cities of Reno and Sparks is associated with higher level of outdoor water consumption due to wind losses and over-watering on assigned days. Additionally, local social norms and regulations can influence new comers' adaptation process (Pelling and High, 2005). For example, private gardens may help in the process of fitting in with the local neighborhood standards, and new residents may try to conform with the predominant landscape and watering practices from established, surrounding neighbors (Larson et al., 2009, 2011).

For the purposes of our paper we differentiate the human and physical, as private capital from the social capital. The objective of our study is to analyze the interaction between those two complementary sources: private changes in preferences versus those due to social norms and regulations.

3.3 Theoretical model

We describe the household maximization problem with a simplified model of stable preferences (Stigler and Becker, 1977), where instead of a learning by doing process (past decisions influencing current behavior), we express the preference formation process as a direct function of the length of residency in the new arid environment. This simplification will allow us to work with a time-separable instantaneous utility function.

In this context, the household maximization problem can be summarized as:

$$\begin{aligned} &\underset{(w_t, Z_t)}{\text{Max}}\ U_t = U(L_t, Z_t) \\ &\text{subject to} \\ &L_t = L(w_t, P_t, C_t, K_t^s, K_t^p) \\ &p_t w_t + Z_t = Y_t \end{aligned} \tag{3.1}$$

The household chooses the optimal level of outdoor water consumption (w_t) as well as the optimal consumption of the numeraire composite commodity (Z_t), so as to maximizes the utility he derives from a certain overall quality of his landscape at time t (L_t) (greenness, extension, variety of plants, etc.) and from the consumption of all other goods (Z_t). This maximization problem is subject to a technological constraint as well as the usual budget constraints given in the last line of equation (3.1).

(L_t) represents the household's landscape production function. At any given time t, the inputs in this production function are: the amount of outdoor water used (w_t), the property's characteristics (P_t) (e.g. property's age, number of fixtures), the climate related variables (C_t) (e.g. wind, rainfall, temperature), and the household's accumulated outdoor-watering related social and private capital (K_t^s and K_t^p).[1] In our simplified model, we assume that social capital is exogenous to the household and we expressed it as a function of local social

[1]We assume that (L_t) increases with water intensity ($\frac{\partial L_t}{\partial w_t} > 0$).

norms (N_t) and local regulations (S_t):

$$K_t^s = K(N_t, S_t) \tag{3.2}$$

On the contrary, we assume that private capital (K_t^p) implies a gradual accumulation of both human and physical capital by the household. As mentioned before, and due to lack of available data, we made the simplifying assumption that this accumulation process depends solely on the length of residency (i.e. time spent) in the arid city. K_t^p can then be expressed as:

$$K_t^p = K(t) \tag{3.3}$$

The third line in equation (3.1) corresponds to the household's budget constraint. The terms p_t and Y_t are, respectively, the price of residential water and household's income per time period.

Solving this maximization problem, household's water demand can be expressed as:

$$w_t = w(p_t, Y_t, P_t, C_t, K_t^s, K_t^p) \tag{3.4}$$

After substituting (3.2) and (3.3) in (3.4), the final reduced form equation for water demand is given by:

$$w_t = w(p_t, Y_t, P_t, C_t, N_t, S_t, t) \tag{3.5}$$

This relationship forms the foundation for our empirical model, as described in section 4.

3.4 Description of the data

The data set for this study consists of daily metered water consumption at the household level observed for 63 consecutive days from June 22^{nd} till August 23^{rd} in the summer of 2008. It was collected by the Truckee Meadows Water Authority (TMWA), the water utility serving the cities of Reno and Sparks in Northern Nevada. The data set also includes basic property information (e.g. year of construction, lot size, square footage of the construction, property value, number of fixtures, number bedrooms, number of bathrooms, property address, property geocodes, and the subdivision the property is located at). Outdoor water restrictions effective in the period of analysis follow a weekly pattern: watering is officially permitted on Thursday and Sunday for odd addresses, and on Wednesday and Saturday for even addresses. For this reason, and following Castledine et al. (2014), we consider a complete week (from Sunday to Saturday) as the unit of analysis. A week with one or more days of missing metered consumption for a given household is eliminated from the data set. After eliminating observations with missing information on the property's characteristics, customers with less than 3 complete weeks of recorded consumption, and subdivisions with less than 20 properties, we are left with a sample of 6,474 households with a total of 36,804 intact weeks. For a sub sample of 3,953 of this customers we have additional information on monthly water use in the winter months from January to March of 2008.[2]

[2] We also exclude 132 properties whose owners filed for bankruptcy between June 2007 and December 2008, and that were not sold before the period of analysis. Presumably, these household have a disincentive to take care of their gardens.

Available weather variables include: average, minimum, maximum daily temperature, average wind speed, and maximum sustained wind speed. No significant precipitation was observed in the summer of 2008, as it is typical for this high-desert area. The variable LOR is computed as the difference in months between the date the household initiated as a TMWA's customer and the 22^{nd} of June 2008 (first day of the sampled period). The basic descriptive statistics for water usage and explanatory variables are summarized in Table 3.1. The average household has lived in the area 121 months ($\approx$ 10 years), and consumes 5,551 gallons of water per week. The average property is valued at $251,700, has 2 bathrooms, 3 bedrooms, and approximately 20 years old. Table 3.2 outlines the average property features by the customers' LOR. The LOR is measured in months since they became a TMWA's customer. As expected, long time residents live, on average, in relatively older homes. However, these properties are not larger or have more extensive land areas than those belonging to newer residents.

Figure 3.1 shows the total number of households by year they started as TMWA customers. The noticeable increase in the number of new TMWA sampled customers correlates with the growing population in the Reno-Sparks area during the period of analysis. TMWA was founded in 2001 by purchasing the water utility assets from Sierra Pacific Power Company (SPPC). In 1977, SPPC installed a new billing system. The result of this data conversion is reflected in the spike on the number of customers added to the billing database in 1977. These households were not actually new customers; instead, they were current residents included in other billing databases that were then compiled into a single billing database.[3] Figure 3.2 shows the average weekly water use along the summer of 2008 by LOR (Figure 3.2.1 in 1,000 of gallons and Figure 3.2.2 in log of 1,000 gallons). The plot (read from right

[3]We decided to keep these households for our analysis, since their LOR is at least the one we computed. Our results regarding the positive effect of LOR on water consumption should at most be underestimated.

to left) shows an initial positive trend that eventually stabilizes, then decreases. Newcomers to the Reno-Sparks area consume an average of around 4.000 gallons of water per week. This amount increases uniformly until it reaches an average of around 6.000 gallons per week in approximately 10 years ($\approx$ 120 month). The wider variability in the average water usage for customers with more than approximately 20 years of LOR ($\approx$ 240 months) is primarily driven by smaller number of observations in those bins (fewer customers starting with TMWA before 1990), which results in a smaller number of households available to average extreme values.

Since there is no price variation along the short sample span (six weeks), prices are excluded from the analysis. As is customary in studies of residential water demand, the assessed value of the property is used as a surrogate for households' income (e.g. Howe and Linaweaver, 1967; Hewitt and Hanemann, 1995).

We calculate the average water consumption of households living within 50 yards of each property as a proxy variable for the effect of neighbor's water consumption. For robustness, we also calculate the average water consumption of households located 100 and 200 yards from each property.[4] As mentioned before, we include spatial fixed effects at the subdivision level to control for unobserved heterogeneity.[5] For our analysis we eliminate subdivisions with 20 or less sampled properties.[6] Table 3.3 summarizes the number of households per subdivision, and living within 50, 100 and 200 yards of each property. There are 182 subdivisions with more than 20 households in our sample. The average number of sampled properties in

[4]We use ArcMap10. (ESRI, 2010) and the properties' geocodes to identify households within 50, 100 and 200 yards of each property.

[5]Subdivisions in Reno-Sparks are small homogenous areas, that usually share a common residential housing development, school district, style of the constructions, type of landscape, among other characteristics.

[6]We exclude subdivision with less than 20 sampled properties since they may correspond to larger ranches and not to representative households.

these subdivisions is 36. Each household has an average of 4, 13 and 44 neighbors on a 50, 100 and 200 yard radius respectively.[7]

In order to analyze household's compliance with outdoor watering restrictions (OWR, hereafter), Castledine et al. (2014) differentiate days of high usage (outdoor watering) from those of low usage (indoor water consumption) using a series of K-mean clustering algorithms. The authors classify households' weekly watering patterns into: "on schedule" (**S**) (weeks when the household complied exactly with the assigned days, watering on all of and only on the assigned days), "'schedule-plus" (**SP**) (weeks of outdoor watering on all assigned days plus some additional days), and "'off-schedule" (**OS**) (at least one assigned day was skipped). Their study reveals that costumers that adhere to the assigned days (**S**) use water less efficiently than customers with a more flexible schedule. We follow this same weekly classification in our present paper. Approximately 29% of the total observations corresponds to weeks of perfect compliance with the effective outdoor watering regulations (**S**). Over 30% of observed weeks are (**SP**) and the remaining 41% are (**OS**) weeks. Figure 3.3 shows the percentage of perfectly compliant weeks (S) by LOR. While new comers show a more flexible watering pattern, the percentage of compliant weeks increases with the LOR.

3.5 Empirical model and results

Our empirical specification of equation (3.5) is given by:

[7]A total of 148 properties did not have a neighbor within a 50 yard radius. These observations are excluded when using this variable.

$$\begin{aligned} w_{sit} = \lambda Y_i + \boldsymbol{\beta}\mathbf{P_i} + \boldsymbol{\alpha}\mathbf{C_t} + \theta N_{it} + \gamma S_{it} + \delta \ln LOR_i + \mu_s + \nu_{it} \\ \nu \sim iid(0, \sigma_\nu^2) \end{aligned} \tag{3.6}$$

Where, w_{sit} is household i's weekly water consumption (natural log of 1,000 of gallons) at week t. The sub-index s indicates the subdivision the property is located at.

The variable (Y_i) is the household's income, $(\mathbf{P_i})$ is the vector of property characteristics, $(\mathbf{C_t})$ is the vector of climate related variables, (N_{it}) represents local social norms, (S_{it}) denotes compliance with outdoor restrictions, and lnLOR the natural log of LOR.

The parameters λ, $\boldsymbol{\beta}$, $\boldsymbol{\alpha}$, θ, γ and δ are the coefficients associated with the corresponding explanatory variables. We include spatial fixed effect μ_s to control for time-invariant subdivision-specific unobserved heterogeneity. We assume the remaining error term v_{it} is uncorrelated with both the explanatory variables and μ_s.

The value of the property is expressed in 1,000s of 2008 dollars. In the group of property characteristics we include: number of bathrooms, fixtures, bedrooms, lot size (1,000 sqft), square footage of the construction (1,000 sqft), and age of the property (natural log of years). In the set of climate related variables we consider average weekly temperature, maximum temperature registered each week, average weekly wind, and maximum wind speed observed on each week. To identify the effect on water use of social interaction among households, we include the average weekly water used by households located 50 yards within each property. Finally, we include the indicator variable (S) of whether the household complied with the assigned schedule on a given week.

Our baseline model is a log-log specification (to account for the non-linear effect of LOR on weekly water consumption) that includes all control variables, in particular N and S. Results are presented in Table 4 column 1.[8,9] The estimated elasticity between weekly water usage and LOR is 0.033 (significant at the 1% significance level), thus the average household (100 months of LOR, with mean weekly usage of 5,283 gallons) increases his weekly water consumption by 207 gallons after living 10 years in the arid area. This corresponds to an increase of 3.9% in weekly water usage. [10]

Figure 3.2 shows a positive trend that stabilizes at around 120 months of LOR. We replace the log of LOR in the previous model by an indicator variable of households with more than 10 years of LOR. The estimated coefficient of the indicator variables is 0.074 (significant at 1%). On average, this group of households with longer LOR consume 345 more gallons of water per week than customers with less than 10 years of residency.[11]

As discussed earlier, LOR is positively correlated with the age of the properties (See Table 3.2). As an additional control for the effect that older houses may have on overall water usage (e.g. due to older pipes, leakages, older toilets and shower-heads), we estimate the

[8]For exposition purposes, we omit estimated coefficients corresponding to the Y, C, and P variables. Appendix 3.A reports the full model's estimates. Property characteristics and property values exhibit little variability within each subdivision, which may hamper their significance in explaining water use after controlling for subdivision fixed effects. Additionally, with the exception of the property age, there is no clear correlation between LOR and the remaining property characteristics (See Table 3.2). Results for the weather related variables are as expected and consistent with previous studies.

[9]Error terms are robust and clustered by subdivision in all models.

[10]The average households with LOR 100 months consume 5,238 gallons of water per week. An increase of 10 years in LOR (120 month) results in a 0.033 times a 120% increase in weekly water use ($0.033*1.2*5,238$) equals 207 gallons per week (3.9% increase in weekly consumption).

[11]The average weekly use by residents with less than 10 years of LOR is 4,660 gallons. A switch from new to old results in a percentage change in weekly usage equals to $\frac{e^{0.074}-1}{100} = 7.7\%$. This corresponds to a total of 345 ($4,660 * 0.077$) gallons of water per week.

same model but excluding properties constructed before 1988 (more than 20 years old). This sub sample includes a total of 4,784 customers and 26,734 observations (intact weeks). The estimated coefficient for the log of LOR is 0.019 (significant at 5%).[12] Even though this estimated coefficient is smaller than the complete sample's, the positive effect of LOR on households' water demand persists when considering this sub-set of newer houses.

We conclude that the effect of LOR on water consumption cannot be explained by property characteristics, neighborhood unobservables, prevailing social norms, or compliance with effective regulations alone. On the other hand, we argue that the remaining effect of LOR on residential water demand corresponds to a change, over time, in households' private preferences.

The estimated coefficient for the variable N equals 0.010 positive and significant at the 10% significance level. Households that live within 50 yards from each other, show similar water use in the summer months. This similar water consumption cannot be explained by neighborhood unobservables or by households' property characteristics.

What can then explain similar water use during the summer month (mainly outdoor water consumption)? Within a subdivision, properties usually share common exterior features such as landscape extension and design. However, there is certain level of independence in the way households manage these outdoor features. For example, households can choose the composition of plants in their yards (e.g. what percentage of drought tolerant vegetation to include). Also noticeable is the management of sprinklers. It is not unusual for households to delegate the task of outdoor irrigation to automatic systems set at fixed time and days of

[12]Results for this subsample of older properties are presented in Appendix 3.A, column 5.

the week. However, a more proactive behavior such as adjusting the system when it is rainy or windy, or the incorporation of devices like rain and humidity sensors, or drip systems can improve the irrigation efficiency. The adoption of these water-saving actions are subject to social scrutiny since a careless administration of the sprinkler system (e.g. sprinklers running too long, broken sprinkler heads or misaligned), is visible to everyone (e.g. soaked lawn and sidewalks). Resetting the sprinkler system requires household to be home and take notice, and the installation of new devices implies investing additional time and money. However, the perception that others are taking this water conservation actions, and the need to comply with the expected norms by surrounding neighbors, may provide the necessary incentive.

The indicator variable S of customers' strict compliance with OWR on a given week has a positive and significant effect on weekly water consumption. Our result are in line with the findings of Castledine et al. (2014), costumers that comply with TMWA's OWR effective in the summer of 2008 do not adjust their watering schedule to exogenous changes in wind conditions and they over water on assigned days.

The next goal of our study is to analyze how social norms and regulations (measured by the N and S variables) affect the adaptation process of new residents' water consumption, and whether they have a stronger or weaker influence during the first years after the household moved to the new city. For this purpose, in Table 3.4, column 2 we include the interaction terms between the log of LOR and both N and S. The interaction term between log of LOR and N is positive and significant (10% significance level), indicating that the effect of spatially close neighbors on household's water use increases with time spent in the arid city. Initially, new residents' water demand is relatively independent from their neighbors'. As households build social bonds, however, they feel more connected, and they grow involved and familiarized with practices that are common within their neighbors.

The interaction term between log of LOR and S is also positive and significant (10% significance level). This result suggest that compliance with OWR reinforces higher usage of water among residents with longer LOR.

We conclude that in the Reno-Sparks area, where higher volumes of water are consumed by residents with longer LOR, that tend to comply with OWR and cluster together (LOR, N and R are positively correlated), both social norms and OWR have a role in promoting higher water consumption. Our results, then, provide further evidence that residents with longer LOR have grown accustomed to the benefits of using higher volumes of water.

3.6 Robustness checks

We perform two additional robustness checks to verify whether the N variable is capturing a pure interaction effect as oppose to any remaining unobserved similarities among close properties or households (e.g. number of children). First, we estimate our model using the average weekly water consumption of neighbors located 100 and 200 yards from the property. We anticipate the influence of nearby households to decrease the farther the properties are from each other. As expected, the effect of neighboring houses is no longer significant when considering the average water consumption of households located farther than 50 yards (See Appendix I, columns 3 and 4).

Second, we analyze winter water consumption. In this case we predict that, if our identification method is correct, and the positive correlation among close neighbors' water consumption is the result of social interaction among them, it should not be significant in the winter months (mainly indoor use).

We calculate the average weekly water used in the winter months for a sub sample of 3,008 households in our dataset.[13] Table 3.5, column 1 presents estimates for our baseline model using average weekly consumption in winter. As expected, the average weekly consumption of nearby households is no longer significant. Table 3.5, column 2 shows estimates using summer weekly consumption for the same sub-sample of customers. The positive effect of N in the summer months remains significant in this comparable-sample. We conclude that the subdivision-fixed effects are controlling for a spatial dependence among the main determinants of indoor water demand (household's size, the number of children, square footage and age of the construction (Mayer and DeOreo, 1999; Wentz and Gober, 2007)).

Finally, we observe that the effect of LOR on winter use is negative an significant. This result may be associated with the incorporation of indoor water efficient devices, as well as a smaller average household size and fewer children with longer LOR. Additionally, this result suggests that is not households' size or the number of children per se (main determinants of winter water consumption) that explain higher water used with longer LOR in the summer months. On the other hand, the positive effect of LOR on summer water consumption seems to be associated with a change in preference regarding landscape options and outdoor water usage.[14]

[13] We calculate average winter weekly usage based on monthly water consumption from January to March 2008 provided by TMWA.

[14] In Appendix 3.B, figure 3.4 shows the average weekly winter water consumption by LOR in months.

3.7 Policy implications

Some degree of outdoor water restrictions have been proven effective water saving policies, particularly during sever droughts or to manage periods of peak demand (e.g. Kenney et al., 2004; Halich and Stephenson, 2009). However, Castledine et al. (2014) showed that strict compliance with OWRs can produce two unintended consequences on households' water consumption: higher peaks to compensate for foregone watering days, and less flexible behavior to adjust to random weather (wind) changes.

At the same time, researchers have acknowledged the general relevance of non-monetary conservation policies. Among the programs that have received particular attention are the ones implemented by the companies OPOWER (in the case of energy saving) and more recently WaterSmart (promoting water conservation).[15,16] These programs are based in the theory of social psychology and social comparison (Schultz et al., 2007; Allcott, 2011; Ferraro et al., 2011).

The main treatment of these programs consists in sending to each household a consumption report (in the form of a letter, email, or text) with a normative message that compares the household's consumption to the average use of neighbors with analogous characteristics (e.g. property and household's characteristics). Households' awareness of their ranking above or below the median has been proven to induce the desired saving behavior. Usually, a persua-

[15] More information on OPOWER's programs can be found at: htpp://www.opower.com.

[16] More information on WaterSmart's programs can be found at: http://www.watersmartsoftware.com.

sive injunctive message (a smiley face or an encouraging phrase of approval such as "You are doing great") is also included in the letter to avoid a boomerang effect for households with below average consumption. These non-traditional policies have also proven to be a cost-effective treatment with long run lasting success in promoting pro-environmental behaviors (Schultz et al., 2007; Allcott, 2011; Ferraro et al., 2011).

Our results confirm the presence of social interaction among households' water consumption and therefore, the suitability of these non-traditional policies to promote water conservation in the cities of Reno-Sparks. More relevant to the implementation of these programs, is our rather counterintuitive finding: residents with longer LOR consume higher volume of water, and tacit social arrangements are stronger among them. Consequently, these non-traditional policies could also be used to contrast new residents' consumption to that of households with longer LOR. On the one hand, this could help frame the adaptation process of newer residents by rewarding their water conservation practices or warning them about inefficient practices that may be common within their neighborhoods. On the other hand, it would incite a feeling of "hurt pride" among households that are native or have lived in the arid Southwest for a longer period.

3.8 Population growth and LOR

Washoe County (as well as the State of Nevada as a whole) grew at a rapid rate between the years 2000 and 2010. This growth, however, is expected to slow during future decades. Table 3.6 shows Washoe County's historical and projected population growth.

With a fairly constant natural growth rate (birth-deaths) this expected growth is mainly explained by changes in net migration. As a result, the percentage of new residents (roughly approximated by yearly increases over total population) is also expected to decrease. For example, the percentage of population with less than 10 years in the area represented 28% of total population in 2008. This percentage is expected to be of only 13% in 2032 (Nevada Department of Taxation and Nevada State Demographer: Population Estimates 2000 to 2012; Nevada State Demographer: Population Projections 2013 to 2032.).

Table 3.8 summarizes the total number of new residents (LOR$\leq$ 10 years) and their estimated weekly water consumption in both 2008 and 2032. To estimate new residents' consumption we multiply the estimated number of households by the estimated weekly usage for customers with a given LOR (0 to 10 years). Table 3.7 shows average 2008 weekly usage by customers' LOR. We extrapolate this weekly average consumption by LOR to customers with the same LOR in 2032. We calculate the number of households as population divided by average household size in Washoe County. The estimated average Washoe County household size is 2.55 (U.S Census Bureau. Census 2010). We assume this average household size remains constant to calculate the respective number of households in 2032.

Table 3.9 shows the percentage of newer residents (and their consumption) and the total number of households (and total estimated consumption) both in 2008 and 2032. We calculate weekly consumption for residents with more than 10 years of LOR as the average consumptions of customers starting with TMWA between 1976 and 1997 (see Table 6). In 2008, newer residents represent 28% of total households and their consumption represents 24% of total consumption. In 2032 these percentages are estimated to drop to 13% and 11%, respectively.

Table 3.10 summarizes the projected 2032 aggregate weekly usage in two scenarios: a) ignoring the effect of LOR on water consumption by assuming an overall (over all customers) average consumption of 5,948 gallons, b) considering the effect of LOR and the difference in average consumption between relatively newer customers and older residents. Our rough estimate of this difference is a 3.6% higher aggregate demand when considering the change in population's LOR and its effect on water use.

3.9 Conclusions

We estimate the behavioral effect of length of residency on urban water demand in the arid cities of Reno and Sparks in the State of Nevada. Using data provided by TMWA (the water utility serving both cities) on daily metered water consumption at the household levels in the summer of 2008, we found that the length of residency has a positive effect on total water consumption: customers that have lived in the area longer consume higher volumes of water on average. Our study is the first to isolate the effect of length of residency on urban water demand from unobserved neighborhood characteristics and social norms and regulations. In doing so, we were able to show that the increase of water consumed over time is the result of changes in preferences that are inherent to each particular household (i.e. changes in private human and physical capital).

Our study additionally highlights the relevance of prevailing social norm and regulation in shaping households' watering behavior. Against our preliminary expectations, we found that residents with longer length of residency, that tend to comply with both local regulations and established norms within close neighbors (households within 50 yards from their property), are the ones that consume, on average, higher volumes of water.

Interestingly, the influential effect of close neighbors increases with time spent in the Southwest. Additionally, compliance with outdoor water restrictions reinforces the use of higher volumes of water the longer the length of residency. In other words, customers that have lived in the area longer appear to be accustomed to the benefits of higher volumes of water. In a sense, households seem to find in complying with social norms and regulations a means to justify higher water consumption.

For this reason, we believe that non-traditional policies that appeal to a social comparison (comparing households consumption to that of their neighbors') are particularly suitable to promote water conservation among both new comers and households with longer length of residency. In particular, by contrasting their consumption to incite a sense of "hurt pride" among residents already familiar to the area and rewarding new comers' water conservation efforts.

The potential savings in water consumption from implementing these policies are considerable, particularly if we acknowledge future population growth in the area. The smaller projected growth rates in Washoe County are expected to result in a smaller proportion of newer residents by 2030. We estimate that accounting for this relatively "older" population (population with a higher average length of residency) and its relatively higher water consumption increases projected aggregate water demand by 3.6% (when compared to a naive projection that does not account for changes in the average length of residency).

Table 3.1: Descriptive statistics of water usage and explanatory variables

Variable	**Observations**	**Mean**	**Std. Dev.**	**Min**	**Max**
LOR (months)	6,474	121.03	107.46	0	392.00
Water usage:					
Daily use (gallons)	257,628	791	1,112	0	54,280
Summer weekly use (gallons)	36,804	5,551	4,663	0	87,510
Winter weekly use (gallons)	24,767	1,064	774	216	16,148
Property characteristics:					
Bathrooms	6,474	2.34	0.67	1.00	16.00
Fixtures	6,474	11.65	3.23	3.00	64.00
Bedrooms	6,474	3.25	0.81	1.00	23.00
Lot size (1,000 sqtf)	6,474	9.23	6.35	0.04	49.66
Sqft (1,000 sqft)	6,474	1.91	0.72	0.49	15.22
Age (years)	6,474	20.91	17.12	0.00	104.00
Value ($ 1,000)	6,474	251.70	144.71	58.98	2,637.44
Weather variables:					
Avg. temp (F)	63 days	77.71	1.86	75.5	81.13
Min. temp (F)	63 days	59.98	1.54	56.94	62.04
Max. temp (F)	63 days	95.58	1.60	93.24	98.10
Avg. wind (knots)	63 days	5.31	0.86	3.91	6.32
Max. wind (knots)	63 days	16.35	2.36	12.3	19.41
Max. gust (knots)	63 days	23.42	2.88	19.25	26.97

Table 3.2: Mean of property characteristics by LOR

LOR		**# properties**	**mean age**	**mean sqft (1,000 sqft)**	**mean lot size (1,000 sqft)**
0-60 months	($\leq$ 5 years)	2,660	18.06	1.86	8.49
			(17.57)	(0.73)	(5.70)
61-120 months	(5-10 years)	1,250	19.17	1.96	9.62
			(16.31)	(0.75)	(6.78)
121-180 months	(10-15 years)	930	20.56	1.95	9.70
			(15.17)	(0.71)	(6.56)
181-240 months	(15-20 years)	620	22.86	2.00	10.16
			(15.07)	(0.72)	(7.37)
241-300 months	(20-25 years)	409	24.61	1.91	9.73
			(15.03)	(0.71)	(6.21)
301-360 months	(25-30 years)	230	26.72	1.88	9.28
			(14.54)	(0.67)	(6.67)
361-392 months	(30-33 years)	376	36.94	1.79	9.84
			(17.32)	(0.64)	(6.31)
Total		**6,474**	**20.91**	**1.91**	**9.23**
			(17.12)	**(0.72)**	**(6.35)**

* Standard errors in parenthesis

Table 3.3: Average number of households per subdivision, within 50, 100 and 200 yards from each property.

	Average number of households	Std. Dev.	Min	Max
Subdivisions	35.87	15.72	21	123
Within 50 yards	4.08	2.25	1	24
Within 100 yards	13.44	7.51	1	58
Within 200 yards	39.99	22.99	1	122

Table 3.4: Results for the natural log of weekly usage.

	(1) Main effects	(2) Interactions
lnLOR (log months)	0.033***	-0.011
	(0.008)	(0.019)
N	0.010*	-0.016
	(0.005)	(0.017)
S	0.067***	0.042
	(0.020)	(0.075)
N-lnLOR		0.006*
		(0.004)
S-lnLOR		0.028*
		(0.014)
Constant	-1.973**	-1.900**
	(0.951)	(0.943)
Tot weeks	35,275	35,275
Number of subdivisons	180	180
r2_o	0.227	0.227
r2_b	0.648	0.641
r2_w	0.0400	0.0424

All models include property and climate related control variables

Clustered by subdivision and robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.5: Winter weekly use and comparable sample in summer.

	(1)	(2)
	Winter	Summer
lnLOR (log months)	-0.057***	0.038***
	(0.011)	(0.011)
N	-0.011	0.015***
	(0.019)	(0.006)
S		0.036
		(0.025)
Number of households	3,008	3,008
Number of subdivisions	151	151
R-squared	0.045	0.059
r2_o	0.0743	0.244
r2_b	0.168	0.652
r2_w	0.0448	0.0517

Clustered-Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Table 3.6: Washoe County population increase: historical and projected.

Year	Total population	Increase	Percentage	Year	Total population	Increase	Percentage
1980	195,407			2010	417,379	747	0.2%
1981	202,031	6,624	3.4%	2011	421,593	4,214	1.0%
1982	208,590	6,559	3.2%	2012	427,704	6,111	1.4%
1983	211,082	2,492	1.2%	2013	431,035	3,331	0.8%
1984	216,279	5,197	2.5%	2014	434,853	3,818	0.9%
1985	221,644	5,365	2.5%	2015	439,283	4,430	1.0%
1986	228,001	6,357	2.9%	2016	444,495	5,212	1.2%
1987	235,014	7,013	3.1%	2017	450,363	5,868	1.3%
1988	242,125	7,111	3.0%	2018	456,556	6,193	1.4%
1989	248,274	6,149	2.5%	2019	462,924	6,368	1.4%
1990	254,667	6,393	2.6%	2020	469,422	6,498	1.4%
1991	262,455	7,788	3.1%	2021	475,968	6,546	1.4%
1992	268,118	5,663	2.2%	2022	482,563	6,595	1.4%
1993	274,424	6,306	2.4%	2023	489,213	6,650	1.4%
1994	282,816	8,392	3.1%	2024	495,878	6,665	1.4%
1995	290,178	7,362	2.6%	2025	502,559	6,681	1.3%
1996	298,395	8,217	2.8%	2026	509,216	6,657	1.3%
1997	306,297	7,902	2.6%	2027	515,823	6,607	1.3%
1998	313,008	6,711	2.2%	2028	522,349	6,526	1.3%
1999	319,816	6,808	2.2%	2029	528,821	6,472	1.2%
2000	333,566	13,750	4.3%	2030	535,216	6,395	1.2%
2001	353,271	19,705	5.9%	2031	541,541	6,325	1.2%
2002	359,423	6,152	1.7%	2032	547,775	6,234	1.2%
2003	373,233	13,810	3.8%				
2004	383,453	10,220	2.7%				
2005	396,844	13,391	3.5%				
2006	409,085	12,241	3.1%				
2007	418,061	8,976	2.2%				
2008	423,833	5,772	1.4%				
2009	416,632	(7,201)	-1.7%				

(*) 1980-1999: U.S. Census Intercensal County Estimates.

2000-2012: Nevada Department of Taxation and Nevada State Demographer. Population Estimates 2000 to 2012.

2013-2032: Nevada State Demographer. Population Projections 2013 to 2032.

Table 3.7: Avgerage weekly usage by LOR (years)

Year starting as TMWA customer	Avg weekly usage	Std. Dev.	LOR (years)
1976	5.11	2.94	32
1977	6.86	5.39	31
1978	6.64	6.52	30
1979	5.92	3.80	29
1980	6.31	4.40	28
1981	7.34	5.67	27
1982	6.08	5.47	26
1983	6.21	5.15	25
1984	6.22	5.79	24
1985	6.89	4.52	23
1986	6.58	4.48	22
1987	6.79	5.27	21
1988	6.07	4.08	20
1989	6.38	4.66	19
1990	6.32	5.01	18
1991	6.11	4.87	17
1992	6.18	4.93	16
1993	6.65	5.48	15
1994	5.97	4.92	14
1995	5.99	4.89	13
1996	6.01	4.76	12
1997	6.08	4.51	11
1998	5.79	4.26	10
1999	6.00	5.21	9
2000	6.08	5.15	8
2001	5.07	4.30	7
2002	5.49	4.53	6
2003	5.33	4.40	5
2004	5.39	4.33	4
2005	4.93	4.29	3
2006	4.68	4.07	2
2007	4.52	4.27	1
2008	4.31	3.74	0

Table 3.8: Total New residents (LOR≤ 10 years) in 2008 and 2032

LOR (years)	Avg 2008 weekly use by LOR	Year	# HHs increase	Tot use (1,000)	Year	# HHs increase	Tot use (1,000)
10	5,773	1998	2,632	15,192	2022	2,586	14,929
9	5,997	1999	2,670	16,011	2023	2,608	15,639
8	6,051	2000	5,392	32,627	2024	2,614	15,815
7	5,055	2001	7,727	39,065	2025	2,620	13,245
6	5,437	2002	2,413	13,116	2026	2,611	14,193
5	5,315	2003	5,416	28,784	2027	2,591	13,771
4	5,341	2004	4,008	21,406	2028	2,559	13,669
3	4,926	2005	5,251	25,867	2029	2,538	12,502
2	4,665	2006	4,800	22,393	2030	2,508	11,699
1	4,491	2007	3,520	15,809	2031	2,480	11,140
0	4,244	2008	2,264	9,606	2032	2,445	10,375
		Tot 2008:	46,093	239,876	Tot 2032:	28,160	146,977

(*) Washoe County average household size = 2.55 (U.S. Census Bureau. Census 2010.)

Table 3.9: Number of households and aggregate weekly usage: 2008 vs 2032

2008	# Households (*)	%	Weekly usage (1,000)	%
New residents (LOR≤ 10 years)	46,093	28%	239,876	24%
Old time residents (LOR > 10 years)	120,116	72%	757,195	76%
Total	166,209	100%	997,070	100%
2032				
New residents (LOR≤ 10 years)	28,160	13%	146,977	11%
Old time residents (LOR > 10 years)	186,654	87%	1,176,636	89%
Total	214,814	100%	1,323,613	100%

(*) Washoe County average household size = 2.55 (U.S. Census Bureau. Census 2010.)

Table 3.10: Estimated difference in aggregate weekly usage accounting for LOR

Average 2008 usage total households	5.948
Average 2008 usage by households with LOR > 10 years	6.304
Average 2008 usage by households with LOR < 10 years	5.209
Total 2032 weekly usage estimated ignoring LOR:	1,277,668
Total 2032 weekly usage estimated accounting for LOR:	1,323,613
Percentage difference	3.6%

3.A Results including all control variables

Table 3.11: Summer weekly use model including all regressors.

	(1)	(4)	(2)	(3)	(5)
	N (50 yards)		N (100 yards)	N (200 yards)	Properties max 20 years
	Main effects	Interactions	Main eff	Main eff	Main eff
lnValue (log 1,000)	0.140	0.156	0.127	0.129	0.011
	(0.160)	(0.159)	(0.166)	(0.171)	(0.136)
Bathrooms	-0.040	-0.040	-0.040	-0.034	-0.064
	(0.041)	(0.041)	(0.040)	(0.040)	(0.043)
Fixtures	0.030**	0.030**	0.031***	0.030**	0.027**
	(0.012)	(0.012)	(0.012)	(0.012)	(0.013)
Bedrooms	0.010	0.008	0.006	0.006	0.016
	(0.018)	(0.018)	(0.018)	(0.018)	(0.023)
Lot size (1,000 sqft)	0.043***	0.043***	0.039***	0.037***	0.044***
	(0.006)	(0.006)	(0.005)	(0.005)	(0.006)
Sqft (1,000 sqft)	0.025	0.022	0.026	0.030	0.070
	(0.040)	(0.040)	(0.043)	(0.045)	(0.042)
lnAge (log years)	0.105*	0.105*	0.107**	0.106*	0.114*
	(0.055)	(0.054)	(0.054)	(0.055)	(0.064)
Avg. temperature (F)	0.002	0.002	0.001	0.001	-0.007
	(0.007)	(0.007)	(0.007)	(0.007)	(0.008)
Max. temperature (F)	0.013***	0.013***	0.013***	0.013***	0.016***
	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)
Avg. wind (knots)	-0.051***	-0.050***	-0.052***	-0.051***	-0.023
	(0.016)	(0.016)	(0.016)	(0.016)	(0.016)
Max. wind (knots)	0.018**	0.018**	0.018**	0.018**	0.004
	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)
lnLOR (log months)	0.033***	-0.003	0.036***	0.036***	0.019**
	(0.008)	(0.019)	(0.008)	(0.008)	(0.009)
N	0.010*	-0.017	-0.006	-0.027	0.005
	(0.005)	(0.017)	(0.008)	(0.019)	(0.006)
N-lnLOR		0.006*			
		(0.003)			
S	0.067***	-0.054	0.066***	0.066***	0.009
	(0.020)	(0.066)	(0.020)	(0.020)	(0.022)
S-lnLOR		0.028**			
		(0.014)			
constant	-1.973**	-1.897**	-1.815*	-1.692	-0.725
	(0.951)	(0.943)	(0.982)	(1.034)	(0.872)
Tot weeks	35,275	35,275	36,467	36,555	26,734
Number of subdivisions	180	180	181	181	145
R-squared	0.040	0.041	0.039	0.038	0.036
r2_o	0.227	0.223	0.214	0.190	0.258
r2_b	0.648	0.633	0.563	0.514	0.629
r2_w	0.0400	0.0411	0.0390	0.0380	0.0357

Clustered by subdivision and robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

3.B Winter versus summer water consumption.

Table 3.12: Winter weekly use and comparable sample in summer.

	(1)	(2)
	Winter	Summer
lnValue (log 1,000)	0.038	0.145
	(0.184)	(0.211)
Bathrooms	-0.028	-0.058
	(0.044)	(0.048)
Fixtures	0.026**	0.033**
	(0.012)	(0.015)
Bedrooms	0.062***	0.027
	(0.019)	(0.019)
Lot size (1,000 sqft)	-0.001	0.037***
	(0.004)	(0.006)
Sqft (1,000 sqft)	0.137***	0.028
	(0.044)	(0.047)
lnAge (log years)	0.122*	0.131**
	(0.062)	(0.058)
Avg. temperature (F)		0.006
		(0.005)
Max. temperature (F)		0.015***
		(0.003)
Avg. wind (knots)		-0.057***
		(0.014)
Max. wind (knots)		0.024***
		(0.007)
lnLOR (log months)	-0.057***	0.038***
	(0.011)	(0.011)
N	-0.011	0.015***
	(0.019)	(0.006)
S		0.036
		(0.025)
constant	-1.049	-2.612**
	(1.019)	(1.191)
Tot weeks	3,008	3,008
Number of subdivisions	151	151
R-squared	0.045	0.052
r2_o	0.0743	0.244
r2_b	0.168	0.652
r2_w	0.0448	0.0517

Clustered-Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

Figure 3.1: Total number of customers by initial year as TMWA customer

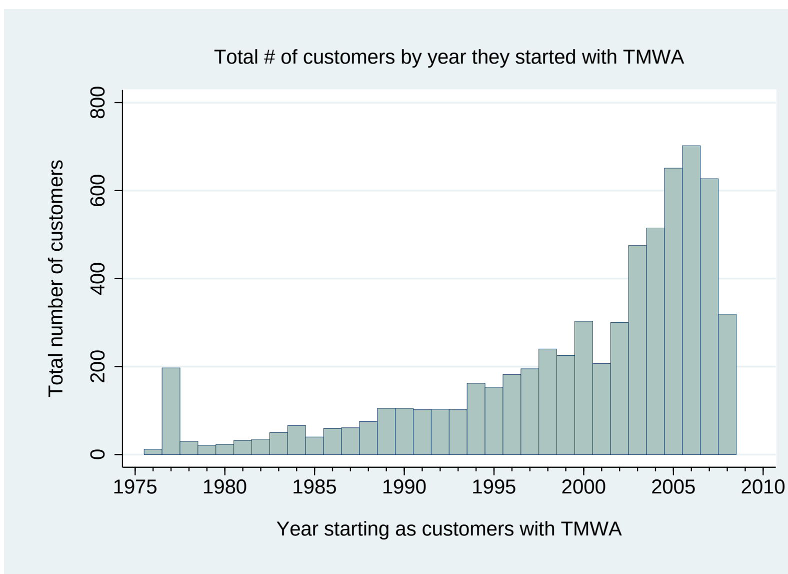


Figure 3.2: Average weekly use by LOR

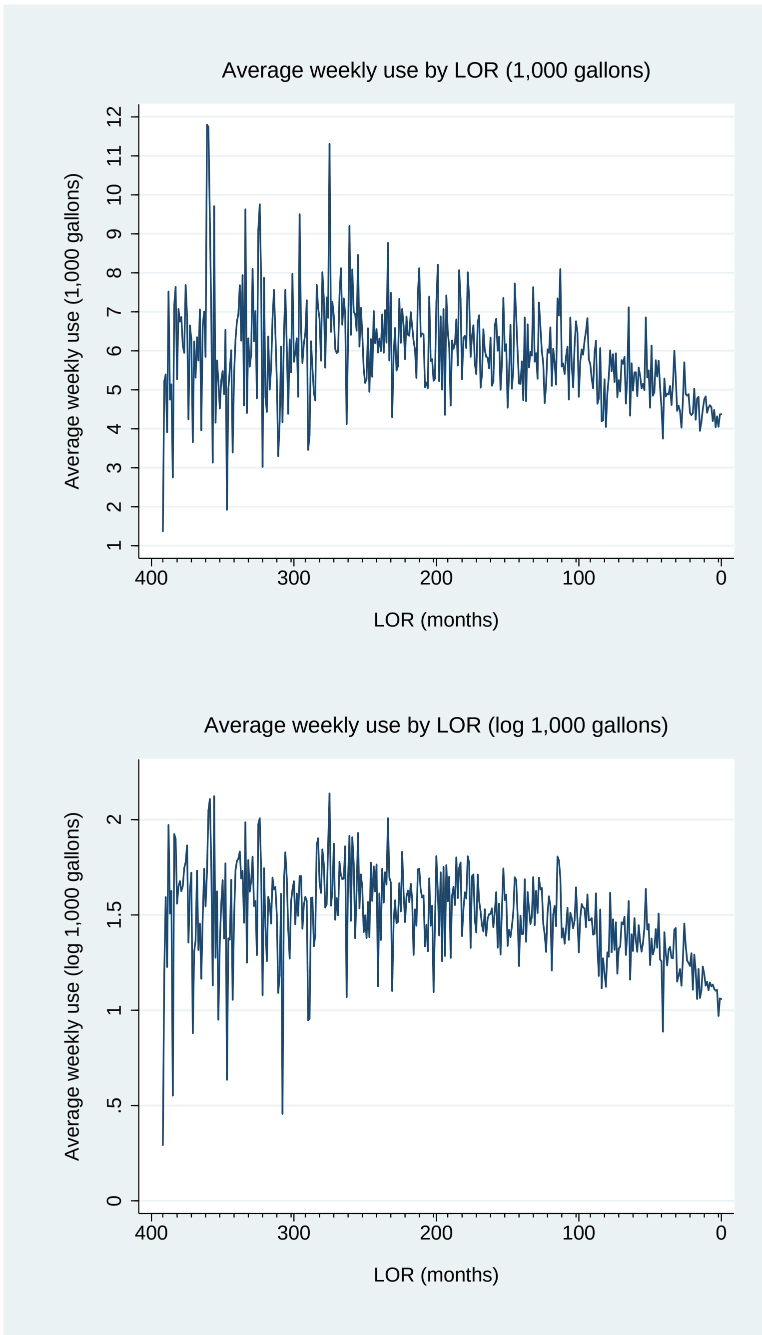


Figure 3.3: Percentage of compliant weeks by LOR

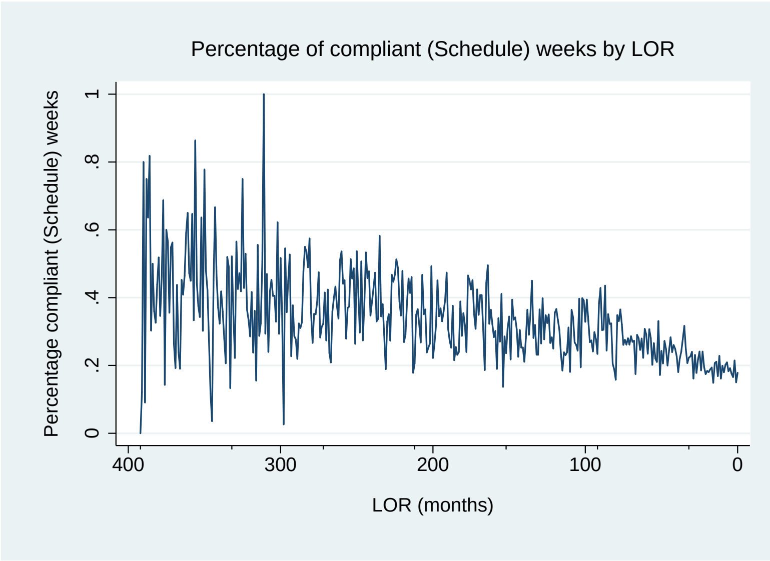
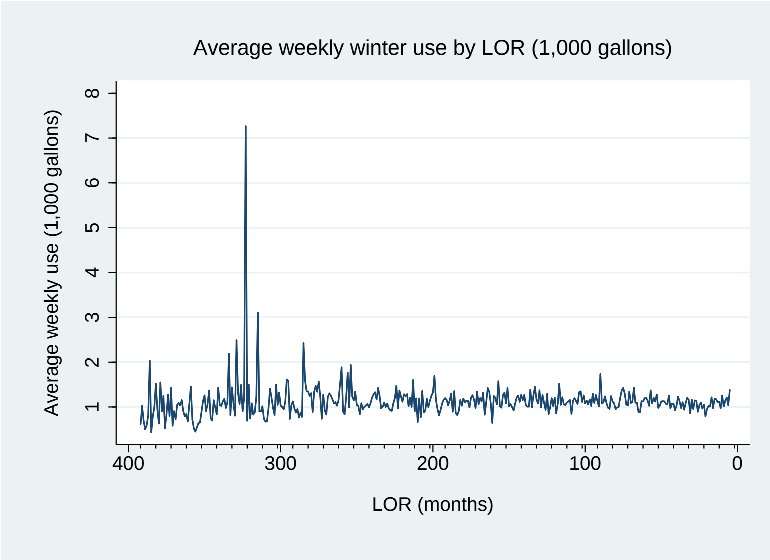


Figure 3.4: Average weekly winter use by LOR



Chapter 4

National versus local lenders and distressed home sales during the 2007 mortgage crisis.

4.1 Introduction

The USA financial system has traditionally been characterized by a dual chartering system, where lending institutions could opt to be regulated by either the state or federal government. Along the decades from 1970 to 1990 a set of regulatory changes were made to encourage the geographic expansion and diversification of the banking industry. One regulatory landmark was the enactment of the Riegle-Neal Interstate Banking and Branching Efficiency Act in 1994, which revoked the majority of previously enforced interstate bank branching restrictions (Carlson and Mitchener, 2005; Dehejia and Lleras-Muney, 2007). As a results of this statewide expansion, the number of small and community lenders across the USA has consistently decreased since 1985 (Critchfield et al., 2004). This decline is particularly relevant in the Midwest of the USA, where a history of unit banking had resulted in a comparatively larger number of small community operators (Critchfield et al., 2004).

Yet, local or community banks play a crucial role in the American economy. Local banks hold the majority of deposits in rural and sub-metropolitan towns in the USA, and more importantly, twenty percent of all US counties lack access to large banks, and are served solely by local lenders (FDIC, 2012). These relatively small banks perform traditional banking activities: collecting deposits to make personal loans, to finance mortgages, and other common financial activities. The scope of their operation is inherently local; profits generated are re-invested within the community, allowing the creation of further loans and mortgages to other local customers (Marsh and Norman, 2013).

Local or community lenders are believed to have an informational advantage over larger lenders regarding their customers and the local housing market, and to be conservative in their lending practices (Critchfield et al., 2004; Dell'Ariccia et al., 2008; Bond et al., 2009; Coulton et al., 2008). Local banks, unlike large banks with widespread locations and constant advertising, are known only to those who live or work within the community served by the bank. Thus local banks must forge and maintain a positive reputation within their community. Local banks are aware that their success is intimately tied to that of their customers' and that they are best served by making mutually beneficial arrangements with them. The foundation of that success is built on these long term relationships, and it is this lending approach based on a close connection to their customers that provide local lenders with a rich set of soft information. For example, a local bank may have a better idea of the ability of a customer to pay back a loan. Additionally, a local bank may know better the conditions of the local real estate market, and originate mortgages accordingly. Local bank thus, ought to have a comparative advantage relatively to non-local operators. Due to their local expertise, local lenders are better able to handle their traditional operations within their local reach (Critchfield et al., 2004).

On the other hand, large banks operating nationally or regionally develop standardized screening procedures in lieu of using their personal relationships with customers to judge risk (Critchfield et al., 2004; Marsh and Norman, 2013). These standardized screening methods rely mainly on hard information (i.e. information that can be summarized by statistics and/or can be documented) such as the borrowers' income and debt history, and FICO score. Moreover, large banks, in contrast to local banks, engage in a wider range of financial activities (FDIC, 2012). Larger banks act in part as investment banks, creating sophisticated financial instruments as opposed to the traditional lending practices. The most notable of these complex structured financial vehicles is the pooling of mortgage loans of varying financial health from across the broad spectrum of nationwide loan holders into mortgage backed securities (MBS) (Marsh and Norman, 2013; Coulton et al., 2008). These securities are later traded and sold in the secondary market, where both investors as well as rating agencies rely again, on the standard information contained within the loan application to judge the risk of the underlying loans (He et al., 2011). A major weakness of the MBS is that it reduces the lenders' incentive to thoroughly collect and process borrowers' information (Keys et al., 2010). Loans are originated with the goal to be sold to secondary parties, rather than to be kept in the lender's own portfolio. For this reason, the originator of a mortgage is no longer the bearer of the burden and losses associated to the mortgage's default (Keys et al., 2010; Agarwal et al., 2012; Purnanandam, 2011).

In the period during the boom of the housing market, the competition to capture new mortgages intensified. Lenders were thus pressured to alter their lending practices, for example by reducing the required FICO score a borrower needed before qualifying for a mortgage. For every reduction in underwriting standards, a new pool of borrowers was created and a new batch of mortgages could be written and securitized. Lenders developed elaborate mortgage products to allow almost anyone to take out a mortgage (Stiglitz, 2010). The development of the "no documentation loan", for example, enabled borrowers with no proof of income to get mortgages. Adjustable rate mortgages, in which the initial interest rate could be close to zero but balloon to fifteen percent or more, were popular products. These predatory lending practices enabled lenders to write larger mortgages, inflate their balance sheets, and increase their leverage (Stiglitz, 2010; Agarwal et al., 2014). Many customers with poor finances who were not interested in home ownership were targeted and encouraged to apply for a mortgage. Many of these targets came from areas with high minority populations and lower incomes, that normally would not qualify for a traditional mortgage (Reiss, 2005; Engel and McCoy, 2006).

These riskiest loans or "subprime mortgages" were originally created in response to government pressure on banks to promote home ownership and increase home ownership rates. The amount of subprime mortgages started small but rapidly expanded during the housing bubble. Subprime mortgages grew from 5% of total originations in 1994 to 20% in 2006 (Federal Reserve, 2008). Subprime lending practices were ubiquitous among lenders. However, the majority of high cost subprime lending was dominated by a small number of big lenders, in particular, national mortgage companies (Coulton et al., 2008).

On the other hand, local lenders participation in MBS was only 0.07 percent of all securitization of mortgages activity during the height of the housing bubble, between 2003 and

2010 (Marsh and Norman, 2013). Thus, it was not the local banks that caused the majority of the crisis.

4.2 Study objective

Our first goal is to compare the different foreclosure rate of loans originated by local versus non local lenders along the years leading to the collapse of the financial system in June of 2007. Our hypothesis is that local operators have an informational advantage over non-local lenders and that they are also more cautious in their lending practices. Non-local lenders, on the other hand, can diversify risk nationally and for this reason, are willing and able to handle higher risk loans. We first estimate a logistic model controlling for all borrowers' and loans' observed characteristics. Secondly, we use matching estimator (nearest neighbor and propensity score) that compare local and non-local lenders that issued similar loans and to observational equivalent borrowers. In both cases, the results confirm that loans issued by local lenders performed better. After controlling for loans and borrowers' characteristics the probability of a loan entering foreclosure is significantly higher if that loan was issued by a non-local lender. We conclude that loans originated by local lenders showed a consistently better performance in the period of analysis.

In our estimations we use information on house sales and their underlying mortgages in the period from 1998 to 2007 in three major western metropolitan areas: Denver, CO, Portland, OR and Seattle, WA. We differentiate the initial years in our dataset 1998-2004 as the pre-boom period, and the years between 2005 and 2007 (years of faster house price appreciation) as the housing market boom period. We observe whether these loans (originated before the start of the crisis) enter foreclosure anytime along the entire sample window that expands until December 2011. We use data on each city separately and compare results across them. Results are robust: local lenders consistently outperformed non-local lenders along the years from 1998 to 2007. This better performance persists after controlling for differences in both the lending conditions and the customers the two groups of lenders served.

The second goal of our study is to compare the foreclosure rates of local and non-local lenders in the period of the housing market boom relatively to the earlier years in our dataset. To evaluate this differential performance we include in our logit model indicator variables in the way of a difference in difference model. Secondly we evaluate the difference between local and non-local lenders' foreclosure rates in the years between 2005-2007, and relatively to the baseline period (1998-2004) using a difference in difference propensity matching estimator. We find that non-local lenders that issued loans observational equivalent to those originated by local lenders, experienced an increase in the foreclosure rate of their portfolio of 3.8% in excess of the increase experienced by local lenders during the years from 2005 to 2007. We conclude that non-local lenders more severely relaxed their lending standards as well as the screening and monitoring of their customers in the height of the housing bubble compared to local lenders.

Our study builds on previous related research. For instance, Keys et al. (2010) analyze the performance of portfolios with different degrees of securitization, Purnanandam (2011) compares foreclosure rates and losses of banks with high level of loans originated to distribute in the secondary market, to the foreclosure rates of banks with low levels of these loans. Dell'Ariccia et al. (2008) examine the decrease in denial rates of incumbents after larger mortgage institutions entered the local market. Coulton et al. (2008) study the probability of default for lenders that mainly issued subprime mortgages versus prime mortgage lenders.

The paper is organized as follows: The next section describes the dataset used for the analysis. Section 4 summarizes the empirical approaches and main results. Section 5 concludes the central findings of the study.

4.3 Data Set

The main data set for our analysis was purchased from a commercial vendor. It contains information on historical real estate transactions in the western metropolitan areas of Denver, CO, Seattle, WA, and Portland, OR in the period from 1998 to 2011. As well as the date and price of the home sales, the data set contains information on the property characteristics, the names of buyer, seller and lender, and information about the loan issued to finance the house.

The group of property characteristics includes, among other variables, the address and assessed value of the property (2011 assessed value by real state local office). The loan information includes the amount of the first and second loan, the type of interest rate (fixed or variable), and whether the transaction was a distressed sales such as a foreclosure. We calculate the loan to value (LTV) as the ratio between the amount of the first loan and the house's sale price. We generate an indicator variable if the loan is secured by a second lien ("piggy back loan"). We also calculate the sale to assessed value (STAV) to quantify how high the house sold relatively to its actualized 2011 assessed value.[1,2]

[1]The STAV value was computed as the amount the house was sold actualized to December 2011 $ over the 2011 assessed value of the property.

[2]To measured house price appreciation we use the house price index (HPI) computed by the Federal

From the publicly available Home Mortgage Disclosure Act (HMDA) dataset, we obtain information on where each lender's main office is located. We consider local lenders to be those whose main office is located within the state the property is located, and we assume non-local lenders those whose main office is located in a different state.

We include in the analysis additional borrowers' socio-demographic variables at the census tract level where the property is located (2010 US Census): unemployment rate, percentage of high school graduate or higher, median household income, and percentage of minority population (black and Hispanic).[3]

After deleting observations with missing property, buyer, lender, or loan information, we are left with a total of 389,353 historic home sales. Table 4.1 reports the number of lenders, number of sales, and the percentage of loans that foreclosure in each of the three metropolitan areas in the entire sample period (1998-2011). The smallest sample corresponds to Denver, CO with 39,620 home sales. This city, however, has the highest foreclosure rate of 17%. The foreclosure rates for Seattle, Wa and Portland, OR are 7% and 9% respectively.

Figure 4.1 shows the percentage of foreclosures on total sales on a given year by each metropolitan area. The city of Denver, CO experienced a much earlier start on its mortgage crisis. Foreclosure rates started to soar as early as 2004 in this city, reaching an early peak in 2008. On the other hand, the cities of Seattle, WA and Portland, OR followed a different pattern. The foreclosure rates in these two cities paralleled each other with a much later

Housing Finance Agency (FHFA).

[3]We use ArcGIS and the properties' address to identify the census tract where each property is located.

increase in foreclosure rates in 2007, which accompanied the general onset of the nationwide mortgage crisis.

Table 4.2 shows the number of sales, number of loans that eventually enter foreclosure by year of origination and the number of foreclosures on a given year for the total sample. Most foreclosures were generated in the immediate pre-crisis years between 2005-2007. Not surprisingly, most of these foreclosures took place after the beginning of the mortgage crisis in 2007.

The quality of the loans originated in the years between 2005-2006 significantly deteriorated compared to previous years. Figure 2 shows the percentage of mortgages that enter foreclosure by year of origination. Loans originated before 2005 have an average accumulated delinquency rate of 2% when they reach a 5 year maturity (60 months). On the other hand, almost 6% and 8% of loans originated in 2005 and 2006 respectively are delinquent five years after origination.

Is there any noticeable difference between local and non-local lender's performance in the pre-crisis years? Table 4.3 shows the percentage of forced sales by local an non-local lenders per year of origination in the total sample. Local lenders issued on average 6% fewer loans that enter foreclosure. Moreover, the difference in foreclosure rates between the two type of lenders continuously increased from 2% in 1998 to 9% in 2006. In the year 2006, for example, this difference represents a 64% higher foreclosure rate of non-local lenders compared to that of local operators.

Could this dissimilar performance be associated with the characteristics of the loans issued by local versus non-local lenders? Table 4.4 summarizes the average loan's characteristics by both group of lenders, for the total sample and for each of the three cities. Non-local lenders originated riskier loans with higher LTV, STAV, and more loans with second lien and with variable interest rate. Noticeable, in the city of Denver, CO lenders (both local and non-local) issued loans with higher average LTV and STAV.

As mentioned in the introduction, in the period before the crisis, some lenders expanded their credit in under-served geographic areas, with lower income, and higher percentage of minority population. Do we observe a systematic difference in socio-demographics of borrowers served by local versus those served by non-local lenders in our dataset? Table 4.5 summarizes borrowers' socio-demographic variables by local and non-local lenders in the total sample and in each of the three cities. Non-local lenders issued loans in census tract with higher percentage of Hispanic home buyers, with slightly higher unemployment rate, lower level of highschool completion, and lower average households' income. The city of Denver, CO, again stands out with a higher percentage of Hispanic and black borrowers, with a lower percentage of highschool completion, and with lower average income.

From our preliminary descriptive analysis we conclude that local lenders did perform better (lower foreclosure rate) particularly in the immediate years before the crisis (2005-2007). The goal in the next section is to identify to what extent this better performance by local lenders can exclusively be explained by the characteristics of the loans they issued and/or the different borrowers the addressed. Or if, on the other hand, local lenders still outperform non-local entities once we control for these observables characteristics.

4.4 Results

4.4.1 Logit model

We start our empirical analysis estimating a logit model. We observe the indicator variable y_{jit} that takes the value one if a mortgage issued by bank i at time t to borrower j enters foreclosure sometime in our dataset. We assume the existence of an underlying latent variable y^* such that:

$$y^*_{jit} = \mathbf{X}'\boldsymbol{\beta} + \varepsilon_{jit} \tag{4.1}$$

Where ε_{jit} follows a logistic distribution. The indicator variable y_{jit} takes the values one and zero according to:

$$y_{jit} = \begin{cases} 1, & \text{if } y^*_{jit} > 0; \\ 0 & \text{otherwise.} \end{cases} \tag{4.2}$$

Our goal is to isolate to what extent the probability that a mortgage enters foreclosure depends on a bank being local. For this reason, we include two additional indicator variables in the index function X'$\boldsymbol{\beta}$ as shown in equation (4.3):

$$\mathbf{X}'\boldsymbol{\beta} = \beta_0 + \beta_l * local_i + \beta_{d05} * d05\_07_t + \beta_{DID} * DID_{it} + \sum_k \boldsymbol{\beta_k}\mathbf{X^k_{jit}} \tag{4.3}$$

These indicator variables are: *local* that takes the value one if the lender has its office in the same state where the property is located, and *d05_07* that equals one if the loan was issued between January 2005 and June 2007. We also include the interaction between these two variables in the way of a difference in difference model (Wooldridge, 2007). We call this interaction term *DID* (*DID=local*d05_07*). The coefficient on the indicator *d05_07* captures the reduction in lending standards, screening and monitoring in the years leading to the mortgage crisis. The financial system as whole experienced this tendency towards lax screening standards during these years (Coulton et al., 2008). For this reason, we expect this coefficient to be positive and significant. On the other hand, and in line with our hypothesis that local lenders issued loans with lower level of risk and that they have an informational advantage regading their customers and the local market, we expect the coefficient of *local* and *DID* to be negative and significant. The coefficient of *local* measures the performance over the entire period of analysis of loans issued by local lenders after controlling for all loans' and borrowers' characteristics in our dataset. Additionally, the coefficient of *DID* captures the differential change over time in lending standards between local versus non-local lenders. We expect this coefficient to be negative and significant, since it was larger banks, operating nationally or regionally, that mainly engaged in the MBS and subprime lending during this period. We control for the set of borrower's and loan's characteristics included in the vector $\mathbf{X^k_{jit}}$.

We estimate equation (4.1) for each of the three metropolitan areas separately, and for the whole sample of loans issued before June 2007. In the later case we include state indicator variables to control for state-specific unobserved housing market characteristics. Results for the four models (average marginal effects) are presented in Table 4.6.

As expected, we found that loans issued in the years between 2005 and 2007 have a higher probability of foreclosure. On average, loans originated in the immediate pre-crisis period have 5.1% higher probability of foreclosing compared to loans originated in previous years. This results confirms the deterioration of lending standards during the boom period of the housing bubble. On the contrary, loans originated by local lenders are 1.83 % less likely to enter foreclosure. This difference represents a nearly 20% higher foreclosure rate of non-local lenders compared to the average foreclosure rate of local lenders in the years from 1998 to 2007. This differential performance persists after controlling for the characteristics of the loans and the borrowers' socio-demographic factors. This result suggests that local lenders issued higher quality loans in the years from 1998 to 2007. Furthermore, foreclosure rates increased significantly more for non-local lenders during the years between 2005-2007. A loan issued by these group of lenders has a average 2.5% higher probability of foreclosure compared to loans originated by local lenders in this period. This suggests that it were non-local lenders the ones that mainly relaxed their lending standards during the height of the housing bubble.

The sign and significance of the borrowers', loans' and state control variables are as expected. Riskier loans, with higher LTV, STAV, with a second lien loan, and with variable interest rate have a higher probability of foreclosure. Analogously, borrowers that live in census tract with higher percentage of minority population (black or Hispanic) with higher unemployment rate and lower average income enter foreclosure more often.

Comparing results across the three cities, we observe that the effect of a lender being local to the community has a similar effect on the three states that ranges from 1.2% to 1.7%. The coefficient of the *DID* indicator, however, is higher in the city of Denver, CO. We conclude that it was in this city, where the poor performance of non-local lenders in the years of the housing market boom was more pronounced. In this city, non-local lenders issued loans with a 4.7% higher probability of foreclose in the years from 2005-2007.

4.4.2 Matching estimators

Our second econometric approach is based on matching estimators. Instead of including the $\mathbf{X^k}$ variables as regressors, these methods evaluate semi or non-parametrically the difference in outcomes between local and non-local lenders that are most similar on this set of observed characteristics.

In order to form a matched sample of lenders we distinguish two periods: 1998-2004 and 2005-2007 and we calculate for each lender, state, and in each of these periods the average of borrowers' and loans' characteristics.[4] The outcome variables is the percentage of loans that foreclose for each lender in each period. The treatment variable is the indicator *local* equals to one if the lender has its office in the same state where the property is located.

Nearest neighbor matching.

Nearest neighbor matching, estimates the average treatment effect on the treated (ATT) non-parametrically, making no explicit functional form assumption for either the treatment

[4]We calculate the average LTV, average STAV, average number of loans with variable interest rate and with second lien loan, average borrower income, average percentage of Hispanic and black borrowers, and average percentage of borrowers with at least high school completion for each lender and in each period (1998-2004 versus 2005-2007).

or the outcome model. To implement this estimator, we match each lender with the lender in the opposite treatment group (local versus non-local lenders) that is most similar on the group of observables characteristics in our dataset (average per lender of borrowers' and loans' characteristics).[5] We use a weighted function of these variables. In particular, we enforce exact matching on the state where the banks originated the mortgages. We implement the nearest neighbor matching suggested by Abadie and Imbens (2006, 2011) that corrects for the bias that results when matching on more than one continuous variable (Abadie et al., 2004). We derive consistent standard errors following Abadie and Imbens (2011). We check that the balancing and overlap assumptions are satisfied.[6] Table 4.7 reports an estimated ATT of -0.035 (significant at 5%) when using the nearest neighbor matching estimation method.[7] This results indicates that local lenders that issued loans with the same level of observable risk as their matched non-local lender, have on average a 3.5% lower foreclosure rate.

Propensity score matching (PSM).

Using the same dataset as before, we estimate the ATT using propensity score matching (PSM). In this case, we use a linear combination of the lenders' averaged covariates to calculate the conditional probabilities of a loan being issued by a local lender (propensity scores) and matched the lenders on these scores to estimate the ATT. After calculating the propensity scores with a logit model, we match each lenders with the first nearest neighbor from the opposite treatment group (closest propensity score) to obtain counterfactuals. As sug-

[5] We use the first nearest neighbor and we allow for more than one match when there are ties.

[6] To check the balancing property we perform a t-test for equality of means in all observed variables between the matched treated and control samples. To check the overlap assumption we control that all observations have at least one nearest neighbor match in the opposite treatment group.

[7] This estimator can be implemented using Stata13 (*teffeccts nnmatch* command).

gested by Abadie and Imbens (2012), we take into account the fact that propensity scores are estimated to calculate the ATT's standard errors. Finally, we check that the overlap and balancing assumption are satisfied.[8] Table 4.7 reports an ATT of -0.046 (significant at 5%) when using the PSM estimation method.[9] On average local lenders' portfolio experienced a 4.6% lower foreclosure rate compared to non-local lenders.

Both matching estimators (nearest neighbor and PSM) confirm a better average performance of non-local lenders along the period of analysis. Local lenders that are observationally equivalent to non-local lenders (e.g. issued loans in the same state, with the same average LTV, average STAV, average percentage of Hispanic and black borrowers) have a statistically significant 3.5-4.6% lower foreclosure rate for the period from 1998 to 2007. This corresponds to a 38-48% higher foreclosure rate compared to the local lenders' average foreclosure rate for loans originated in the years from 1998 to 2007. The better performance of these matched local lenders cannot be explained by observational differences in the quality of the loans they issued, or the socio-demographic characteristics of their borrowers. We conclude that local lenders issued less riskier loans probably because they rely on their better knowledge of both their customers and the local housing market when originating their mortgages.

Does this better performance remain constant along the different stages of the housing bubble from 1998 to 2007, or did non-local lenders have a higher incentive to relax their screening and lending standards in a time of rapidly increasing house prices and profitable origination fees? We answer this question in the following subsection.

[8] To test the balancing property we perform t-tests for the equality of means in the treated and non-treated groups after matching. The overlap assumption is satisfied since all estimated propensity scores are smaller than 0.9.

[9] This estimator can be implemented using Stata13 (*teffects psmatch* command).

Difference in Difference Propensity Score Matching (DID-PSM).

If non-local lenders had an incentive towards lax lending practices and screening of their borrowers, we should have evidence in our dataset that during the immediate pre-crisis period their portfolios significantly deteriorated compared to those of local lenders', and relatively to the initial years in our dataset (1998-2004).

In order to assess this differential performance we use the same sample as before to estimate the ATT following (Heckman et al., 1997, 1998)'s difference in difference propensity score matching estimator (DID-PSM). Our goal is to compare how the difference in foreclosure rate between matched local and non-local lenders changed in the years of the housing market boom relatively to previous years.

In order to do so, we match the difference in the outcomes (before-after) of a local lender to the weighted differences in outcomes (before-after) of the matched non-local lenders. This DID-PSM estimator is superior to the traditional DID model in that it relaxes the linear form specification of the covariates. It also requires weaker assumptions than the matching estimators used before. By calculating the difference in outcomes between the two periods (1998-2004 versus 2005-2007), the DID-PSM estimator eliminates a potential "hidden bias" due to selection on unobservables that are constant over time (e.g. lenders' size). After calculating the propensity scores using a logit model we estimate the average difference in difference for the matched treated and control units using equation (4.2) (Smith and Todd, 2005):[10]

[10] The DID-PSM can be implemented using the Stata user written command *diff* (Villa, 2012).

$$\widehat{ATT} = \frac{1}{N_1} \sum_{i \in I_1 \cap S} \{(Y_{1ti} - Y_{0t'i}) - \sum_{j \in I_0 \cap S} w_{(i,j)}(Y_{0tj} - Y_{0t'j})\}$$
$$w_{i,j} = \frac{G(\frac{P_j - P_i}{a_n})}{\sum_{k \in I_0} G(\frac{P_k - P_i}{a_n})} \tag{4.4}$$

Where t is the period between 2005-2007, $t^{'}$ is the period between 1998-2004. I_1 is the set of local lenders with N_1 elements, I_0 is the set of non-local lenders, S is the region of common support, $w(i,j)$ is the weight given to lender j within all matches to lender i. These weights are calculated as an inverse function of the distance between the estimated propensity scores (i.e. higher weights are given to matched lenders that have the most similar propensity scores). $G(.)$ is a kernel function and a_n is a bandwidth parameter,

Again in this case we test that the balancing and overlap assumption of the covariates are satisfied.[11] We estimate consistent standard errors via bootstrapping. Table 4.7 reports an estimated ATT equals to -0.038 (significant at 5%) when using the DID-PSM estimation method.

Non-local lenders that issued loans observational equivalent to those originated by local lenders, experienced an increase in the foreclosure rate of their portfolio of 3.8% in excess of the increase experienced by local lenders during the years from 2005 to 2007. This difference corresponds to a 42% of the average foreclosure rate in the baseline period from 1998 to 2004. We conclude that not only do local lenders have an average better performance along the period of analysis, but they were also significantly more cautious than non-local lenders in their lending practices in the height of the housing bubble.

[11]To test the balancing property we perform t-tests for the equality of means in the treated and non-treated groups after matching. The overlap assumption is satisfied since all estimated propensity scores are smaller than 0.9.

4.5 Conclusions

We analyze the lending behavior of local or community banks relative to that of larger nationwide or regional lenders during the years leading to the 2007 mortgage crisis. We use information on historic home sales and associated mortgages in three fast growing western metropolitan areas during the period from 1998 to 2011 (Denver, CO, Portland, OR and Seattle, WA).

We found that local lenders experienced consistent lower average foreclosure rates, and this differential performance intensified in the years of the housing market boom from 2005 to 2007. On average, local lenders' foreclosure rate is 1.8% lower than that of national or regional lenders. This difference accentuates an additional 2.5% during the housing market boom years. Taking into account that the average foreclosure rate in the baseline period was 8.6%, both effects combined account for a 50% of this average.

National or regional lenders are more willing and able to better handle higher risk loans through the diversification of their portfolio across states. We show that local lenders' relative efficiency cannot be explained solely by the different quality of the loans issued or differences in socio-demographic characteristics of borrowers served by both group of lenders. Local lenders presumably base their loan origination process on interpersonal relationship with their customers. This relationship approach provides them with a larger set of information beyond the one contained within the basic risk ratios usually found on an otherwise standardized loan application.

Additionally, as the house market experienced a period of expansion, non-local lenders, highly engaged in the securitization of their portfolio, further relaxed their lending and screening procedures. Not surprisingly, in the aftermath of the crisis, it was this group of lenders that experienced the highest rates of foreclosure.

The social cost of foreclosure exceeds the immediate cost borne by the household that is left without a home. After a mortgage forecloses, the property that is abandoned, with lack of maintenance and prone to vandalism, further decreases in price. In the year 2009, for example, foreclosed properties sold at a nationwide average discount of 25% compare to the value a similar non-distressed property. Foreclosed homes exert additional negative externalities to neighboring properties. In a period like the post 2007 economic downturn, once a house entered foreclosure, properties located nearby would experience a contagion effect and a further decrease in their market value (Frame, 2010; Campbell et al., 2011).

Screening and monitoring procedures have significantly tightened after the 2007 financial crisis. In an effort to prevent or minimize the losses of future financial crisis, regulators have institutionalized a series of underwriting standards. One example, emerging from the Dodd-Frank Wall Street Reform and Consumer Protection Act, is the *Ability to repay and qualified mortgage rule (ATR/QM)*. The rule issued by the Consumer Financial Protection Bureau (CFPB), entered in effect in January 2014. It imposes a series of restrictions on what mortgages can by issued as well as how lenders should screen and monitor their customers. Among other restrictions, the rule imposes a cap on total points and fees of no more than 3% of the total mortgage. It requires that mortgages should not be longer than 30 years term, with negative amortization, zero interest or balloon payment features, for example.

The rule also requires lenders to verify consumers' income and asset information, as well as to keep record of this information for three years after the loan is completely paid back. The results in our study indicate the importance of these regulatory measures. In particular, the need to supervise the operation of larger financial institutions, who demonstrated a lack of financial care regarding the expansion of their lending during the years from 1998 to 2007.

Table 4.1: Number of lenders, number of sales and foreclosure rate by state.

State	Number of lenders	Number of sales	Foreclosure rate
WA	310	174,459	7%
OR	328	175,274	9%
CO	416	39,620	17%
Total	631	389,353	9%

Table 4.2: Total and forced sales by year.

Year	Tot. sales	Foreclosures by year of origination		Foreclosures on a given year	
		#	% tot.sales	#	% tot.sales
1998	25,625	1,853	7%	256	1%
1999	25,180	2,199	9%	358	1%
2000	26,405	2,390	9%	605	2%
2001	27,914	2,378	9%	852	3%
2002	31,224	2,554	8%	1,264	4%
2003	35,441	2,960	8%	1,660	5%
2004	36,771	3,682	10%	2,014	5%
2005	41,016	5,074	12%	2,398	6%
2006	36,101	6,174	17%	2,215	6%
2007	30,466	4,250	14%	2,582	8%
2008	19,094	930	5%	3,559	19%
2009	19,542	187	1%	5,566	28%
2010	18,611	58	0%	5,934	32%
2011	15,963	47	0%	5,860	37%
Total	389,353	34,736	9%	35,123	9%

Table 4.3: Percentage of foreclosure by Local/Non-local lenders.

	Local lenders	Non-local lenders	Diff.
1999	8%	10%	2%
2000	8%	11%	3%
2001	7%	11%	4%
2002	7%	11%	4%
2003	7%	12%	5%
2004	10%	16%	6%
2005	12%	19%	8%
2006	14%	23%	9%
2007	12%	18%	6%
2008	4%	7%	3%
2009	1%	2%	0%
2010	0%	0%	0%
2011	0%	0%	0%
Total	7%	13%	6%

Table 4.4: Local and non-local lenders' average loan characteristics.

		LTV*	STAV**	Variable int.rate	Second lien
Tot sample	Non-local lender	78%	89%	43%	53%
	Local lender	76%	88%	36%	41%
	Mean difference (2 sample t-test)	1.6%***	0.3%**	6.8% ***	11.6% ***
WA	Non-local lender	76%	90%	49%	56%
	Local lender	76%	90%	44%	40%
OR	Non-local lender	78%	85%	38%	50%
	Local lender	76%	84%	20%	43%
CO	Non-local lender	81%	100%	48%	56%
	Local lender	80%	92%	41%	49%

* LTV = first loan / sale value, ** STAV = (sale value in dec 2011 $)/ (assessed value)

Table 4.5: Socio-demographic characteristics of borrowers by local and non-local lenders.

		Hispanic	Black	Unemployment	Highschool	Income
Tot sample	Non-local lender	12%	5%	9%	90%	70,717
	Local lender	9%	5%	8%	92%	77,553
	Mean diff (2 sample t-test)	3.3%***	0.1%	1.0%***	-1.8%***	-6,837***
WA	Non-local lender	8%	6%	7%	92%	83,909
	Local lender	7%	5%	7%	93%	84,054
OR	Non-local lender	11%	4%	10%	91%	64,144
	Local lender	10%	4%	9%	91%	65,601
CO	Non-local lender	32%	10%	9%	84%	59,274
	Local lender	24%	9%	8%	89%	63,577

Table 4.6: Probability of a loan entering foreclosure (Average marginal effect, logit).

	WA	OR	CO	All 3
LTV (log)	0.201***	0.271***	0.281***	0.246***
	(0.009)	(0.010)	(0.032)	(0.007)
STAV	0.001***	0.001***	0.002***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Variable int. rate	0.049***	0.073***	0.069***	0.062***
	(0.002)	(0.002)	(0.005)	(0.001)
Second lien	0.032***	0.062***	0.061***	0.048***
	(0.002)	(0.002)	(0.006)	(0.001)
Hispanic	0.001***	-0.000	0.003***	0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Black	0.000	-0.001***	0.002***	0.000***
	(0.000)	(0.000)	(0.000)	(0.000)
Highschool	-0.002***	-0.002***	-0.000	-0.001***
	(0.000)	(0.000)	(0.000)	(0.000)
Unemployment	0.002***	0.002***	-0.001*	0.002***
	(0.000)	(0.000)	(0.001)	(0.000)
Income (log)	0.007**	-0.002	-0.001	0.003
	(0.003)	(0.004)	(0.009)	(0.002)
Local lender	-0.017***	-0.016***	-0.012*	-0.018***
	(0.002)	(0.003)	(0.007)	(0.002)
Loans originated between 2005 and 2007	0.051***	0.055***	0.035***	0.051***
	(0.002)	(0.002)	(0.005)	(0.001)
DID	-0.017***	-0.036***	-0.047***	-0.025***
	(0.003)	(0.004)	(0.015)	(0.003)
OR				0.026***
				(0.001)
CO				0.027***
Observations	131,700	130,602	27,935	290,237

Robust standard errors in parenthesis. *** p< 0.01, ** p< 0.05, * p< 0.1

Table 4.7: Matching estimators of ATT.

	ATT	Std. Err.	z	p-value
Nearest neighbor matching	-0.035	0.017	-2.08	0.038
PSM	-0.046	0.020	-2.31	0.021
DID-PSM	-0.038	0.017	-2.23	0.026

Figure 4.1: Percentage of foreclosures on total sales by city

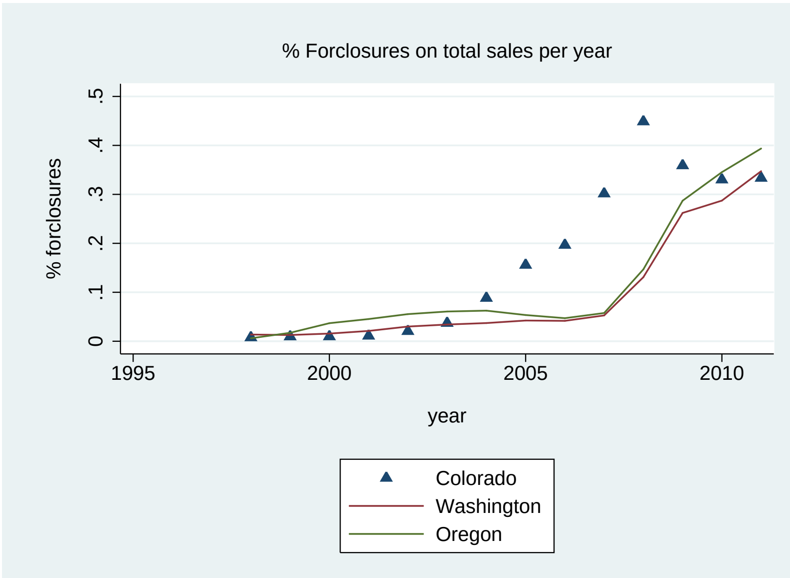
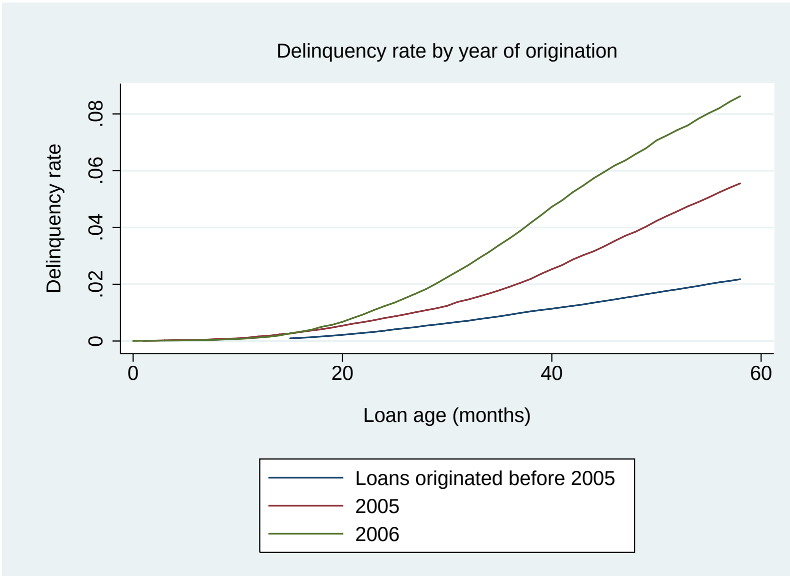


Figure 4.2: Delinquency rate by year of origination



Chapter 5

Conclusions

The economics of information is a relatively new and important field of economics. This dissertation looked at the role of information in three case studies within three different branches of economics: health economics, environmental economics, and finance and banking. We explore how access to, as well as the manipulation of that information, can result in better choices and overall more efficient economic outcomes.

First, we analyze the use of nutritional labels by US adults. We find that parents that read the information contained in the nutritional labels are better able to influence the quality of their children's diet as measured by the Health Eating Index (HEI). Consequently, these children present better health status with significantly lower Body Mass Index (BMI) compared to children whose parents do not read the nutritional labels. We evaluate the effect of nutritional labels within the framework of the household production model. In particular, we adapt the two-stage Stackelberg model developed by You and Davis (2010) where the parents are the leaders and the children are the followers. In the model, we considered the parental nutritional label usage as one of the parent's time allocation variables. This decision influences directly the child's HEI conditional demand for nutrients. The child's HEI in turn, enters his final production process of the dietary-related health outcome (i.e, BMI). We estimate the resulting 3-equation recursive system by Full Information Maximum likelihood using data from the CSFII 1994-1996 and its companion DHKS. After accounting for self selection, we identify a significant average treatment effect on the treated (ATT) of 4.07 and -0.6 on children's HEI and BMI respectively. These ATT's are significant at 1% and 5% respectively. Our analysis then, allows as to identify the sequential mechanism through which parental nutritional label usage affect first and directly the quality of their children's diet and only subsequently and indirectly (through this better dietary quality) their weight status. Furthermore, our results suggest that parents do use nutritional labels as a mechanism to choose healthier diets for their children. Parents that reported having children with a poor or fair health status were observed to use nutritional labels more often than parents with healthier children. On the other hand, parents with children attending school tend to read nutritional labels less frequently.

Next, we estimate the behavioral effect of length of residency on urban water demand in the arid cities of Reno and Sparks in the State of Nevada. We found that the length of residency has a positive effect on total water consumption. Additionally, our study shows the relevance of local social norm and regulation in shaping households' watering behavior. Against our preliminary expectations, we found that residents with longer length of residency, that tend to comply with both local regulations and established norms within close neighbors (households within 50 yards from their property), are the ones that consume, on average, higher volumes of water. This influential effect of close neighbors increases with time spent in the Southwest. Moreover, compliance with outdoor water restrictions reinforces the use of higher volumes of water the longer the length of residency.

Finally, we study the lending behavior of local or community banks relative to that of non-local lenders during the years leading to the 2007 mortgage crisis. We found that local lenders experienced consistently lower average foreclosure rate in the metropolitan areas of Denver, CO, Seattle, WA and Portland, OR. Additionally, this differential performance intensified in the years of the housing market boom from 2005 to 2007. On average, local lenders' foreclosure rate is 1.8% lower than that of non-local lenders. This difference increases by an additional 2.5% during the housing market boom years. National or regional lenders are more willing and able to handle higher risk loans through the diversification of their portfolio across states. We show that local lenders' relative efficiency cannot be explained solely by the different risk properties of the loans they issued or differences in socio-demographic characteristics of their borrowers. Local lenders presumably base their origination decisions on an interpersonal relationship with their customers. This relationship approach provides them with information that is not contained within the standard risk measures used in generic loan applications.

By examining three practical studies, we showed the relevance of information on decision making. Our results also suggest ways to mitigate the consequences associated to the information asymmetries within each case. For example, in chapter 2 we conclude that the disclosure of accurate information (labels) is an effective policy to promote healthy eating habits among US children. In chapter 3 we conclude that the implementation of signals to motivate conservation behavior (e.g. sending reports to households comparing their water consumption to that of their neighbors') can result in residential water savings. Finally the results in chapter 4 suggest that the direct regulation of lending institutions' screening and monitoring procedures can be effective policies to reduce the risk and losses associated to mortgage foreclosure originated in a period like the pre-2007 housing market boom years. We think that the implementation of these measures can help mitigate the information asymmetries in each of the three cases analyzed.

Bibliography

Abadie, A., Drukker, D., Herr, J. L., and Imbens, G. W. (2004). Implementing matching estimators for average treatment effects in Stata. *Stata journal*, 4:290–311.

Abadie, A. and Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1):235–267.

Abadie, A. and Imbens, G. W. (2011). Bias-corrected matching estimators for average treatment effects. *Journal of Business & Economic Statistics*, 29(1).

Abadie, A. and Imbens, G. W. (2012). Matching on the estimated propensity scores. Harvard University and National Bureau of Economic Research. *http://www.hks.harvard.edu/fs/aabadie/pscore.pdf*.

Agarwal, S., Amromin, G., Ben-David, I., Chomsisengphet, S., and Evanoff, D. D. (2014). Predatory lending and the subprime crisis. *Journal of Financial Economics*, 113(1):29–52.

Agarwal, S., Chang, Y., and Yavas, A. (2012). Adverse selection in mortgage securitization. *Journal of Financial Economics*, 105(3):640–660.

Agresti, A. (2010). *Analysis of ordinal categorical data*, volume 656. John Wiley & Sons.

Agthe, D. E., Billings, R. B., and Dworkin, J. M. (1988). Effects of rate structure knowledge on household water use. *JAWRA Journal of the American Water Resources Association*, 24(3):627–630.

Allcott, H. (2011). Social norms and energy conservation. *Journal of Public Economics*, 95(9):1082–1095.

Apps, P. F. and Rees, R. (1997). Collective labor supply and household production. *Journal of political Economy*, pages 178–190.

Athanasiadis, I. N., Mentes, A. K., Mitkas, P. A., and Mylopoulos, Y. A. (2005). A hybrid agent-based model for estimating residential water demand. *Simulation*, 81(3):175–187.

Balling, R. C., Gober, P., and Jones, N. (2008). Sensitivity of residential water consumption to variations in climate: An intraurban analysis of Phoenix, Arizona. *Water Resources Research*, 44(10):n/a–n/a.

Becker, G. S. (1965). A theory of the allocation of time. *The economic journal*, pages 493–517.

Behrman, J. R. and Wolfe, B. L. (1987). How does mother's schooling affect family health, nutrition, medical care usage, and household sanitation? *Journal of econometrics*, 36(1):185–204.

Berry, J. W. (1997). Immigration, acculturation, and adaptation. *Applied Psychology*, 46(1):5–34.

Beydoun, M. A. and Wang, Y. (2009). Parent–child dietary intake resemblance in the United States:Evidence from a large representative survey. *Social science & medicine*, 68(12):2137–2144.

Bond, P., Musto, D. K., and Yilmaz, B. (2009). Predatory mortgage lending. *Journal of Financial Economics*, 94(3):412–427.

Campbell, J. Y., Giglio, S., and Pathak, P. (2011). Forced Sales and House Prices. *American Economic Review*, 101(5):2108–31.

Carlson, M. and Mitchener, K. J. (2005). Branch banking, bank competition, and financial stability. Technical report, National Bureau of Economic Research.

Cason, T. N. and Gangadharan, L. (2002). Environmental labeling and incomplete consumer information in laboratory markets. *Journal of Environmental Economics and Management*, 43(1):113–134.

Castledine, A., Moeltner, K., Price, M., and Stoddard, S. (2014). Free to choose: Promoting conservation by relaxing outdoor watering restrictions. *Journal of Economic Behavior & Organization*.

Cawley, J. and Liu, F. (2012). Maternal employment and childhood obesity: A search for mechanisms in time use data. *Economics & Human Biology*, 10(4):352–364.

Chang, H.-H. and Nayga Jr, R. M. (2011). MotherâĂŹs nutritional label use and childrenâĂŹs body weight. *Food Policy*, 36(2):171–178.

Coulton, C., Chan, T., Schramm, M., and Mikelbank, K. (2008). Pathways to foreclosure: A longitudinal study of mortgage loans, Cleveland and Cuyahoga County, 2005-2008. *Cleveland, Oh.: Case Western Reserve University, Mandel School of Applied Social Sciences, Center on Urban Poverty and Community Development*.

Critchfield, T. S., Samolyk, K., Davison, L., Hanc, G., Gratton, H., and Davis, T. (2004). The Future of Banking in America-Community Banks: Their recent past, current performance, and future prospects. *FDIC Banking Review Series*, 16(3).

Danielzik, S., Czerwinski-Mast, M., Langnäse, K., Dilba, B., and Müller, M. (2004). Parental overweight, socioeconomic status and high birth weight are the major determinants of overweight and obesity in 5–7 y-old children: baseline data of the kiel obesity prevention study (kops). *International Journal of Obesity*, 28(11):1494–1502.

Danielzik, S., Langnäse, K., Mast, M., Spethmann, C., and Müller, M. J. (2002). Impact of parental BMI on the manifestation of overweight 5–7 year old children. *European journal of nutrition*, 41(3):132–138.

Dehejia, R. and Lleras-Muney, A. (2007). Financial development and pathways of growth: state branching and deposit insurance laws in the United States, 1900-1940. *Journal of Law and Economics-Chicago*, 50(2):239.

Dell'Ariccia, G., Igan, D., and Laeven, L. (2008). Credit booms and lending standards: Evidence from the subprime mortgage market.

Drichoutis, A., Lazaridis, P., and Nayga Jr, R. M. (2006). Consumers' use of nutritional labels: a review of research studies and issues. *Academy of Marketing Science Review*, 10(9).

Drichoutis, A. C., Nayga Jr, R. M., and Lazaridis, P. (2009). Can nutritional label use influence body weight outcomes? *Kyklos*, 62(4):500–525.

Duflo, E. and Saez, E. (2002). The role of information and social interactions in retirement plan decisions: Evidence from a randomized experiment. Technical report, National Bureau of Economic Research.

Engel, K. C. and McCoy, P. A. (2006). Turning a blind eye: Wall street finance of predatory lending. *Fordham L. Rev.*, 75:2039.

Ferraro, P. J., Jose Miranda, J., and Price, M. K. (2011). The persistence of treatment effects with norm-based policy instruments: evidence from a randomized environmental policy experiment. *American Economic Review*, 101(3):318.

Frame, W. S. (2010). Estimating the effect of mortgage foreclosures on nearby property values: A critical review of the literature. *Economic Review, Federal Reserve Bank of Atlanta*, 95.

Greene, W. H. (2012). *Econometric analysis*. Prentice Hall. Pearson Education.

Guthrie, J. F., Fox, J. J., Cleveland, L. E., and Welsh, S. (1995). Who uses nutrition labeling, and what effects does label use have on diet quality? *Journal of Nutrition education*, 27(4):163–172.

Halich, G. and Stephenson, K. (2009). Effectiveness of residential water-use restrictions under varying levels of municipal effort. *Land economics*, 85(4):614–626.

Harlan, S. L., Yabiku, S. T., Larsen, L., and Brazel, A. J. (2009). Household water consumption in an arid city: Affluence, affordance, and attitudes. *Society and Natural Resources*, 22(8):691–709.

He, J., Qian, J., and Strahan, P. E. (2011). Credit ratings and the evolution of the mortgage-backed securities market. *The American Economic Review*, 101(3):131–135.

Heckman, J. J., Ichimura, H., and Todd, P. (1998). Matching as an econometric evaluation estimator. *The Review of Economic Studies*, 65(2):261–294.

Heckman, J. J., Ichimura, H., and Todd, P. E. (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *The review of economic studies*, 64(4):605–654.

Heckman, J. J. and Robb Jr, R. (1985). Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics*, 30(1):239–267.

Hewitt, J. A. and Hanemann, W. M. (1995). A discrete/continuous choice approach to residential water demand under block rate pricing. *Land Economics*, pages 173–192.

Hilaire, R. S., Arnold, M. A., Wilkerson, D. C., Devitt, D. A., Hurd, B. H., Lesikar, B. J., Lohr, V. I., Martin, C. A., McDonald, G. V., and Morris, R. L. (2008). Efficient water use in residential urban landscapes. *HortScience*, 43(7):2081–2092.

Hilaire, R. S., VanLeeuwen, D. M., and Torres, P. (2010). Landscape preferences and water conservation choices of residents in a high desert environment. *HortTechnology*, 20(2):308–314.

Howe, C. W. and Linaweaver, F. P. (1967). The impact of price on residential water demand and its relation to system design and price structure. *Water Resources Research*, 3(1):13–32.

Jorgensen, B., Graymore, M., and O'Toole, K. (2009). Household water use behavior: An integrated model. *Journal of Environmental Management*, 91(1):227–236.

Kenney, D. S., Klein, R. A., and Clark, M. P. (2004). Use and effectiveness of municipal water restrictions during drought in colorado. *JAWRA Journal of the American Water Resources Association*, 40(1):77–87.

Kenny, J., Barber, N., Hutson, S., Linsey, K., Lovelace, J., and Maupin, M. (2009). Estimated use of water in the United States in 2005. *U.S. Geological Survey. Circular 1344*.

Keys, B. J., Mukherjee, T., Seru, A., and Vig, V. (2010). Did securitization lead to lax screening? Evidence from subprime loans. *The Quarterly Journal of Economics*, 125(1):307–362.

Kim, S.-Y., Nayga, R. M., and Capps, O. (2001). Food label use, self-selectivity, and diet quality. *Journal of Consumer Affairs*, 35(2):346–363.

Kim, S.-Y., Nayga Jr, R. M., and Capps Jr, O. (2000). The effect of food label use on nutrient intakes: an endogenous switching regression analysis. *Journal of Agricultural and resource Economics*, pages 215–231.

Kreuter, M. W., Brennan, L. K., Scharff, D. P., and Lukwago, S. N. (1996). Do nutrition label readers eat healthier diets? behavioral correlates of adults' use of food labels. *American Journal of Preventive Medicine*, 13(4):277–283.

Larson, K. L., Casagrande, D., Harlan, S. L., and Yabiku, S. T. (2009). Residents' yard choices and rationales in a desert city: social priorities, ecological impacts, and decision tradeoffs. *Environmental Management*, 44(5):921–937.

Larson, K. L., Wutich, A., White, D., Muñoz-Erickson, T. A., and Harlan, S. L. (2011). Multifaceted perspectives on water risks and policies: A cultural domains approach in a southwestern city. *Human Ecology Review*, 18(1):75–87.

Marsh, T. D. and Norman, J. W. (2013). The impact of Dodd-Frank on community banks. *American Enterprise Institute May*, 7.

Martin, C. A., Peterson, K. A., and Stabler, L. B. (2003). Residential landscaping in phoenix, arizona, us: Practices and preferences relative to covenants, codes, and restrictions. *Journal of Arboriculture*, 29(1):9–17.

Martin, C. A., Warren, P. S., and Kinzig, A. P. (2004). Neighborhood socioeconomic status is a useful predictor of perennial landscape vegetation in residential neighborhoods and embedded small parks of Phoenix, AZ. *Landscape and Urban Planning*, 69(4):355–368.

Mayer, P. W. and DeOreo, W. B. (1999). *Residential end uses of water*. American Water Works Association.

Miller, E. and Buys, L. (2008). The impact of social capital on residential water-affecting behaviors in a drought-prone Australian community. *Society and Natural Resources*, 21(3):244–257.

Ogden, C. L., Carroll, M. D., Kit, B. K., and Flegal, K. M. (2012). Prevalence of obesity and trends in body mass index among US children and adolescents, 1999-2010. *Jama*, 307(5):483–490.

Olsen, D. and Highstreet, A. L. (1987). Socioeconomic factors affecting water conservation in Southern Texas. *Journal of the American Water Works Association*, 78(3):59–68.

Pelling, M. and High, C. (2005). Understanding adaptation: what can social capital offer assessments of adaptive capacity? *Global Environmental Change*, 15(4):308–319.

Pitt, M. M. and Rosenzweig, M. R. (1985). Health and nutrient consumption across and within farm households. *The Review of Economics and Statistics*, pages 212–223.

Pretty, J. and Ward, H. (2001). Social capital and the environment. *World development*, 29(2):209–227.

Purnanandam, A. (2011). Originate-to-distribute model and the subprime mortgage crisis. *Review of Financial Studies*, 24(6):1881–1915.

Randolph, B. and Troy, P. (2008). Attitudes to conservation and water consumption. *Environmental Science & Policy*, 11(5):441–455.

Reiss, D. (2005). Subprime standardization: How rating agencies allow predatory lending to flourish in the secondary mortgage market. *Fla. St. UL Rev.*, 33:985.

Schultz, P. W., Nolan, J. M., Cialdini, R. B., Goldstein, N. J., and Griskevicius, V. (2007). The constructive, destructive, and reconstructive power of social norms. *Psychological Science*, 18(5):429–434.

Smith, J. and Todd, P. (2005). Does matching overcome LaLonde's critique of nonexperimental estimators? *Journal of econometrics*, 125(1):305–353.

Spence, M. (1973). Job market signaling. *The quarterly journal of Economics*, pages 355–374.

Spinti, J. E., Hilaire, R. S., and VanLeeuwen, D. (2004). Balancing landscape preferences and water conservation in a desert community. *HortTechnology*, 14(1):72–77.

Stigler, G. J. and Becker, G. S. (1977). De gustibus non est disputandum. *The American Economic Review*, 67(2):76–90.

Stiglitz, J. E. (1975). The theory of" screening," education, and the distribution of income. *The American Economic Review*, pages 283–300.

Stiglitz, J. E. (2002). Information and the change in the paradigm in economics. *American Economic Review*, pages 460–501.

Stiglitz, J. E. (2010). *Freefall: America, free markets, and the sinking of the world economy.* WW Norton & Company.

Strauss, R. S. and Knight, J. (1999). Influence of the home environment on the development of obesity in children. *Pediatrics*, 103(6):e85–e85.

Teisl, M. F., Bockstael, N. E., and Levy, A. (2001). Measuring the welfare effects of nutrition information. *American Journal of Agricultural Economics*, 83(1):133–149.

Variyam, J. N. (2008). Do nutrition labels improve dietary outcomes? *Health Economics*, 17(6):695–708.

Variyam, J. N., Blaylock, J., Lin, B.-H., Ralston, K., and Smallwood, D. (1999). Mother's nutrition knowledge and children's dietary intakes. *American Journal of Agricultural Economics*, 81(2):373–384.

Villa, J. M. (2012). Simplifying the estimation of difference in differences treatment effects with Stata.

Ward, C. and Kennedy, A. (1999). The measurement of sociocultural adaptation. *International Journal of Intercultural Relations*, 23(4):659–677.

Wentz, E. A. and Gober, P. (2007). Determinants of small-area water consumption for the city of Phoenix, Arizona. *Water Resources Management*, 21(11):1849–1863.

Wooldridge, J. (2007). What's New in Econometrics? Lecture 10 difference-in-differences estimation.

Yabiku, S. T., Casagrande, D. G., and Farley-Metzger, E. (2008). Preferences for landscape choice in a Southwestern desert city. *Environment and Behavior*, 40(3):382–400.

You, W. and Davis, G. C. (2010). Household food expenditures, parental time allocation, and childhood overweight: An integrated two-stage collective model with an empirical application and test. *American Journal of Agricultural Economics*, page aap031.

You, W. and Nayga Jr, R. M. (2005). Household fast food expenditures and children's television viewing: Can they really significantly influence children's dietary quality? *Journal of Agricultural and Resource Economics*, pages 302–314.

2008-2012 American Community Survey 5-Year Estimates. Educational Attainment. (`http://factfinder2.census.gov`).

FDIC. (2012). Community Banking Study. (`https://www.fdic.gov/regulations/resources/cbi/report/CBSI-1.pdf`).

ATR/QM. (2014). (http://www.consumerfinance.gov/regulations/ability-to-repay-and-qualified-mortgage-standards-under-the-truth-in-lending-act-regulation-z.

Nevada Department of Taxation and Nevada State Demographer. Population Estimates 2000 to 2012. (`http://nvdemography.org/data-and-publications/estimates/`)

Nevada State Demographer. Population Projections 2013 to 2032.(`http://nvdemography.org/data-and-publications/projections/`)

U.S. Census Bureau, State Interim Population Projections by Age and Sex: 2004-2030.

U.S. Environmental Protection Agency (a). 2008. Water supply and use in the United States. Technical Report EPA 832-F06006.

U.S. Environmental Protection Agency (b). (`http://www.epa.gov/watersense/our_water/tomorrow_beyond.html`)

U.S. Environmental Protection Agency (c). (`http://www.epa.gov/climatechange/impacts-adaptation southwest.html#area`)

U.S. Census Bureau. Population Division. Annual Estimates of the Resident Population for Incorporated Places Over 50,000, Ranked by July 1, 2012 Population: April 1, 2010 to July 1, 2012. Release Date: May 2013.

U.S. Census Bureau. Census of Population and Housing. (http://www.census.gov/prod/www/decennial.html)

U.S. Census Bureau. Intercensal County Estimates.(http://www.census.gov/popest/data/counties/asrh/1980s/PE-02.htmlandhttp://www.census.gov/popest/data/counties/totals/1990s/tables/co-99-08/99C8_32.txt)

Water Research Foundation. (2010). (http://www.waterrf.org/PublicReportLibrary/4031.pdf).

World Health Organization. Fact sheet 311, May 2012. (http://www.who.int/mediacentre/factsheets/fs311/en/index.html).