Evaluating the Sensitivity of Residential Water Demand Estimation to Model Specification and Instrument Choices

A Thesis

Presented in Partial Fulfillment of the Requirements for the

Degree of Master of Science

with a

Major in Applied Economics

in the

College of Graduate Studies

University of Idaho

by

Roshan Puri

Major Professor: Alexander Maas, Ph.D.

Committee Members: Liang Lu, Ph.D.; Jonathan Yoder, Ph.D.

Department Administrator: Christopher McIntosh, Ph.D.

August 2019

Authorization to Submit Thesis

This thesis of Roshan Puri, submitted for the degree of Master of Science with a Major in Applied Economics and titled "Evaluating the Sensitivity of Residential Water Demand Estimation to Model Specification and Instrument Choices" has been reviewed in final form. Permission, as indicated by the signatures and dates below, is now granted to submit final copies to the College of Graduate Studies for approval.

Major Professor: ______________________________ Date: __________
Alexander Maas, Ph.D.

Committee Members: ______________________________ Date: __________
Liang Lu, Ph.D.

______________________________ Date: __________
Jonathan Yoder, Ph.D.

Department Administrator: ______________________________ Date: __________
Christopher McIntosh, Ph.D.

Abstract

Past studies have estimated residential water with different econometric model choices. Inconsistency in the choice of the price signal, its instruments, and appropriate weather variables have offered qualitatively different estimates of price elasticity — both elastic (>1) and inelastic (<1) — in the water demand literature. This distinction is important, not only because elasticity estimates are critical in creating efficient and effective water conservation and management practices, but also because accurate demand responses to price changes help water utilities stabilize and anticipate future revenue. Therefore, an accurate estimate of price elasticity is necessary, which requires specifying the appropriate demand model. However, specifying residential water demand with an appropriate model (under increasing block rates) is challenging for three reasons: 1) little theoretical motivation exists for determining which weather variables affect demand, 2) there is an ongoing debate over the appropriate price signal on residential water bills, and 3) simultaneity issues in estimation require the use of instruments, which vary within the literature.

In this paper, we elucidate the effects of model choices on elasticity estimates by systematically varying the specification of price, instruments, and weather variables across a suite of models. Fixed Effects with Instrumental Variables (FE-IV) is employed in our model to control time-invariant effects and solve the problem of simultaneity, generating consistent and unbiased estimates of the coefficient. First, we conduct a formal Shin Test to investigate the price variables to which the households respond; after selecting the price signal, we fix the price variable and its instrument and re-specify the model across all possible combinations of weather variables (511 different combinations from the use of nine weather variables) to create a distribution of elasticity estimates. Lastly, we hold price signal and weather variables constant and run 31 separate model specifications using a suite of instrumental variables. Thus, this paper systematically evaluates how model choices influence elasticity estimates of residential demand estimation, providing insight into model choices and the validity of previous work.

As found in other work, our results suggest that average price rather than marginal price is the appropriate price signal to which households respond. For the choice of weather, we find 97% of our estimates between -0.83 to -0.57, which implies that choices in weather variables

(average versus maximum daily temperature, for example) have little (or no) impact on elasticity estimates. Notable exceptions to this result are model specifications which only use precipitation variables (excluding temperature information). In these model specifications, elasticity estimates are qualitatively different (>1) and the model fit is worse. These results suggest that, so long as temperature variables are included, parsimonious models can provide consistent and efficient estimates, thereby reducing the need for more in-depth measurements (air pressure, humidity, wind speed, for instances) offered by advanced weather stations. Similarly, the choice of price instruments has a negligible impact on elasticity estimates (estimates clustered between -0.77 to -0.69), so long as the set of instruments are valid.

These results provide guidance to future researchers for if and how to include different weather metrics while modeling residential water demand. Further, these results also help stakeholders and water managers establish appropriate water price and pricing structure, thereby helping utilities stabilize and generate the revenue along with the proper demand-side management of water.

Acknowledgments

First and foremost, I would like to express my deepest appreciation to my advisor, Dr. Alexander Maas, who believed in me and supported me continuously throughout the entire period of this research. He helped me develop and improve my skill sets in different areas of this study and beyond. This thesis would not have been possible without his guidance and persistent help.

Next, I would like to thank my committee members: Dr. Liang Lu and Dr. Jonathan Yoder for their discerning comments, feedback, suggestions, and encouragement in the development process and accomplishment of this research.

I would also like to thank the NSF Sustainability Research Network (SRN) Cooperative Agreement 1444758 for providing financial support for this research.

Last but not least, I would like to thank the faculty and the staff at the Department of Agricultural Economics and Rural Sociology for providing a beautiful and competitive environment to help succeed and building academic careers. I would also like to thank Dr. Raju Pokharel, a postdoctoral fellow in the College of Natural Resources, for helping me on this research and my academic career.

Dedication

Dedicated to my beloved Mom and to the memory of my Dad.

Table of Contents

List of Figures

List of Tables

1. Introduction

Global water demand is expected to increase by 55% from 2000 to 2050, with domestic water use increasing to 130% of current levels by 2050 (OECD, 2012). Accordingly, water utilities have engaged in pricing policies as a viable option for demand-side management, because the associated losses in the economic efficiency with this option are lower than many other conservation methods (Pérez-Urdiales *et al*., 2016; Roibás *et al*., 2007). While water utilities have found block rate structure as an attractive option to encourage conservation (Wichman, 2014), they can obfuscate the direct relationship between price and consumption, making it difficult for policymakers to anticipate the precise consumption and revenue effects of rate changes (Pérez-Urdiales *et al*., 2016). Though most of the literature has reported inelastic estimates (<1) of residential water demand, there are still some cases wherein the estimates are elastic (>1) (see Espey *et al*., 1997; Arbués *et al*., 2003; Worthington & Hoffman, 2008). Even within the inelastic estimates of water, past studies have noted a quantitative difference of price elasticity. For instance, Martínez-Espiñeira (2002) found price elasticity estimates range between -0.12 and -0.17, while Nieswiadomy & Molina (1989) found the price elasticity estimate of -0.86 – both of these studies computed these elasticities to estimate household-level water demand under increasing block rates structure with instrumental variables technique. These distinctions – both qualitative and quantitative – are important since price elasticity of demand is very crucial in finding a way of collecting revenue. For instance, when the demand is price elastic, quantity demanded falls sharply (more than the percentage change in price) even for the small increment in price, thereby reducing the total revenue. However, utilities can significantly encourage conservation through an increase in price since quantity is expected to fall substantially under elastic demand. For an inelastic demand of water, quantity demanded falls only slightly with the increase in price, which increases the amount of revenue. By comparison with elastic demand, if water demand is inelastic, small price increases have limited ability to induce conservation. So, from revenue and conservation perspectives, an accurate estimate of price elasticity is necessary, which requires the choice of an appropriate demand model.

While theory and past work dictate some econometric model choices, there is an ongoing debate over the appropriate price signal (for instance, average vs marginal price where the specification of the model with the former yields significantly higher estimates of price elasticity (Dalhuisen *et al*., 2003; Espey *et al*., 1997)), which is compounded by the need to instrument for a price, because price and quantity are simultaneously determined under increasing block rates structure. To address this problem of simultaneity, different studies use a different set of price instruments which vary across the literature. The choice of weather metrics as control variables is also inconsistent and lacks concrete theoretical motivation for if and how to include these terms in demand estimation. Logic dictates that such choices should not bias price estimates so long as weather is uncorrelated with price, but little work has formally evaluated this relationship empirically. Further, different ranges of price elasticity are observed across different studies. Because the demand for water differs with the geographical location and time, land use patterns, demographic characteristics (household size, income, gender, culture etc.) (Sebri, 2014), it is difficult to determine which differences in elasticity estimates are legitimate (real), and which ones result from model choices, which is why it is essential to realize the outcome of different model choices.[1]

The goal of this paper is to better elucidate the effects of model choices on elasticity estimates by systematically varying specifications of price, instruments, and weather variables across a suite of regression models. This paper estimates water demand for single-family households in the City of Fort Collins, Colorado. Fort Collins Utilities has set the goal of reducing the water use to 130 gallons per capita per day (gpcd) by 2030 (*Water Conservation Annual Report*, 2017).[2] An increasing block rate (IBR) structure has been adopted by Fort Collins Utilities as a management tool to reduce the consumption of water. IBR structure is designed in such a way where the marginal rate of water (price per unit) increases with the volumetric increase in consumption i.e., a specific volume of water is associated with each block which is specified by different (increasing) price levels. Households face lower marginal rates when their consumption is in the lower block group and vice versa. Such block rate structure discourages the consumption of water since households are obliged to pay higher marginal rates when their consumption increases,

---

[1] Model choices in the literature include pooled OLS, Random Effects, Fixed Effects, First Differences, etc. and choice of functional forms include linear, log-linear, log-log, Stone-Geary. Each of these choices may effect the specific point estimate of own-price elasticity (Arbues *et al*., 2003).

[2] This goal was set in 2015 when the water usage was 141 gpcd. The water usage increased by 11.34% (157 gpcd) in 2016 and decreased by 10.19% (141 gpcd) in 2017.

thereby inducing conservation among high users. IBR structures also do not place a burden on low-income/low-users (Wichman, 2014), making it politically attractive.

Results from this work will provide guidance on which choices may be more appropriate based on statistical and economic theory and which choices affect the stability of elasticity estimates. This paper employs the Shin Test for the appropriate choice of price signal in modeling water demand and uses a suite of price instruments and weather variables to evaluate the sensitivity, thereby creating a distribution of price elasticity estimates based on model choices. We provide empirical evidence on how households respond to price and weather variables and how sensitive estimation results are to model specification. Specifically, we examine if:

1. Households respond to average price rather than marginal price;
2. Price elasticity of demand is sensitive to the choice of weather variables; and
3. Price elasticity of demand is sensitive to the choice of instruments.

Results suggest that an average price rather than marginal price is the appropriate price signal to which consumers respond. We find reasonably stable and robust estimates of price elasticity across different weather specification and choice of price instruments, with some notable exceptions. Understanding the implications of model choice can inform better estimation methodologies and ultimately help water managers and stakeholders predict revenue while satisfying households' demand.

The remainder of this paper is comprised of the following sections: Section 2 reviews the existing literature; Section 3 provides some insight about data, methods, and model specification; Section 4 summarizes and discusses results; and Section 5 concludes.

2. Literature Review

The relationship between price and consumption of water under different block rate structures has been reviewed extensively in the literature, and yet the price signal to which consumers respond is controversial. An accurate estimation of water demand requires an appropriate choice of price variable, which is coupled with the need for instruments because price and quantity are simultaneously determined under block rate structure. However, the use of price instruments varies significantly across the literature, which augments the complexity of the estimation process. While water demand is also determined by the weather the households face, there exists little theoretical economic motivation to the choice of weather variables and their inclusion in water demand model. Further, imprecise estimates of elasticity can be found when the model is mis-specified with an inappropriate functional form.

2.1 Choice of Price Variables (Price and its Structure)

In the United States, water rates system – decreasing, constant and increasing block rates – have historically varied across time and location. These rate systems are designed either as a flat rate (constant block rates) or with the combination of fixed fee and variable fee (across blocks; increasing or decreasing block rates structures). Also, within the block rate system, the relative size of the fixed and variable costs varies by city and time period (for instances, City of Santa Fe, New Mexico vs City of Fort Collins, Colorado). Previously, water utilities had favored the decreasing block rates (DBR) structure to stabilize revenue because a greater proportion of consumers can be found in the first block group from where a large portion of revenue can be generated (Griffin, 2016). Under DBR structure, the marginal price of water decreases with an increase in quantity consumed, reflecting economies of scale. Opposingly, under IBR structure, the marginal price of water increases with an increase in consumption, which discourages the consumption of non-basic use of water. Discouraging consumption of water among high-users is important to cope with the increasing demand of water due to the increasing population for sustainable use.

Thus, from conservation perspectives, IBR structure has become more appealing these days – the effectiveness of which can be measured using price elasticity as a tool (Asci *et al*., 2017). Water pricing is considered an economically efficient tool for designing incentives and making the most valuable use of water (Hoyos & Artabe, 2017). Further, price can be used as a conservational tool by water utilities to forecast future revenue. Theoretically, for a given good, an increase in price decreases quantity demanded; but, because water has few substitutes, raising price reduces water demand only slightly – i.e., the demand of water is price inelastic (Worthington & Hoffman, 2008). While water demand is relatively unresponsive, price can still be used as a tool for demand-side management of water (Arbués *et al*., 2003), particularly when IBR are used to promote conservation.

However, under block pricing structure, it is difficult to determine the price to which consumers respond (Wichman, 2014). As many consumers spend only a small portion of their total expenditure on water[3], they are less concerned about the details of the rate structure such that most consumers perceive an average price rather than a marginal price as the cost of water (Foster & Beattie, 1981). The informational costs associated with understanding marginal price are substantial, primarily because it is difficult to keep track of cumulative consumption throughout the billing periods, which advocates the use of average price in lieu of marginal price (Ito, 2014; Pérez-Urdiales *et al*., 2016). Economic theory and some past studies assume that consumers understand the structure of the price and suggest the use of marginal price to equate costs with benefits at the margin (see Nataraj & Hanemann, 2011; Rinaudo *et al*., 2012; Vásquez Lavín *et al*., 2017). In contrast to these assumptions, Nieswiadomy & Molina (1991) commented that it is the pricing structure – IBR or DBR – faced by households which makes them respond to either marginal or average price. They concluded that consumers respond to marginal price under IBR structure and to average price under DBR structure.

Taylor (1975) argued that households may face the same marginal price even if they are in different pricing structures and suggested the inclusion of both marginal and average prices to estimate demand. Taylor *et al*. (2004) used both marginal and average price to estimate residential water demand. They found that the average price specification yields an upward

[3] A notable exception to this statement is low-income households in midwestern United States that block pricing in those cities can be many times what it is in the fixed charge. For instance, households with a 3/4” tap in the City of Fargo, North Dakota pay an average variable charge of $39.6/month (for an average monthly consumption of 9000 gallons) which is two times higher than the minimum charge (fixed charge) of $17.55/month (as of 2019), reflecting the high cost-burden on water consumption across low-income households.

biased (towards unity) estimate of elasticity because of the presence of fixed charge in average price. They also mentioned that the exclusion of fixed charges produces a steep decline in the significance level of average price elasticity estimates. Maas *et al*. (2017) also excluded fixed price from the average price and used average variable price while evaluating the effect of conservational motives on residential water demand. Shin (1985) made an argument in favor of the inclusion of a fixed price while estimating electricity demand. He stated that the exclusion of fixed rates from the average price shifts the perceived budget constraints of the consumers and consequently shifts their indifference curve. Further, the author added that the share of the fixed charges in the total bill is even more for water compared to electricity, which underscores the importance of fixed charges.

While different studies continue to use average price, marginal price or some combination of both as the relevant price, Shin (1985) proposed an econometric test to identify the price signal to which households respond. In his work, he found that households respond to average price rather than marginal price under DBR structure. Nieswiadomy & Molina (1991) used this formal test to estimate the residential water demand and found the results consistent to Shin (1985). Flyr *et al*. (2019) used the same test to explore the price specification in commercial water demand and found firms respond to lagged average price rather than marginal price, contemporaneous average price and the average price over the last 12 months. Most recently, Ito (2014) employed the Encompassing Test, as suggested by Davidson and MacKinnon (1993), to choose among alternative prices (marginal price, expected marginal or average price), though in electricity pricing, and found consumers respond to the average price.

2.2 Choice of Price Instruments

Under linear pricing models, consumers face and react to constant marginal price to determine their consumption level. However, under non-linear pricing (IBR or DBR), consumers face a non-constant marginal price, which cannot be determined without determining households' consumption level, causing simultaneity (Clarke *et al*., 2017). Even though households are assumed to respond to the average price, the simultaneity issue still prevails since water prices increase even with the decrease in water demand because of the associated fixed costs in the average price (Schleich & Hillenbrand, 2009).[4] Even if the water rate structure is known properly, a simultaneous-equation problem theoretically occurs unless there exists an infinitely elastic price (Arbués *et al*., 2003).

Further, it is difficult to include all the relevant explanatory variables in the model due to the limitations of the available data or because the effect of omitted variables from the model is unknown (Kenney *et al*., 2008). If such omitted variables are correlated either with the dependent variable or with both dependent and one or more of the included independent variables, then the estimates from OLS will be biased and inconsistent (Greene, 2003).[5]

Both simultaneity and omitted variables can create endogeneity issues while modeling water demand and require the use of instrument variables (IV). However, the choice of instrument varies across the literature. Clarke *et al*. (2017) used total annual expenditure on water for each household, the price for first block group, and the difference between each successive block price, along with the fixed charge as a set of instrumental variables and regressed them on lagged price and the price-difference variable. Utility determines all these instrument variables, except total annual expenditure, so they are not directly related to the water consumption; however, they are strongly correlated with price variables. Also, total annual expenditure is not substantially correlated with the consumption in a given period, though correlated to some extent with annual consumption. Kenney *et al*. (2008), consistent with Nieswiadomy and Molina (1988), also employed parameters of rate structure i.e., the actual marginal prices that households face at the different consumption level as instrumental variables. They argued that the width of each block i.e., the difference between the consumption between each successive block cannot be the valid instruments because they might be correlated with any unobserved variables during any period when specific water costs are allocated. Wichman *et al*. (2016) followed Olmstead (2009) and used basic service fees and marginal price (full block rate structure i.e., price for each block group) to account

[4] Households might not have the separate information on average fixed and variable price that are associated with average price.

[5] Such omission of variables also affects the standard error of the estimators, the extent of which depends on the correlation between omitted variables and the included independent variables. For example, if omitted variables are highly correlated with the included independent variables, then the omission of variables decreases the standard error of the coefficient; however, if they are also found to be correlated with the dependent variable at the same time, then omission of variables might increase the standard error of the coefficient.

for endogeneity since they are correlated with the water consumption only through the price variables. Similarly, Pérez-Urdiales *et al*. (2016) also used a full set of marginal price in each block as a price instrument. Other literature to use marginal price as a price instrument are Nieswiadomy and Molina (1989), Price *et al*. (2014), and Reynaud *et al*. (2005).

Most recently, the number of billing days was introduced as the possible instrument for the price to solve for endogeneity because it is correlated with the lagged average price but uncorrelated with the current water consumption (Flyr *et al*., 2019). The author also mentioned that using billing days as price instruments reflects the exogenous variation across both households and time, whereas simply using block rates to instrument does not allow for interhousehold variation among instruments because price structures are usually identical across a service area.

2.3 Choice of Weather Variables

Water demand is also determined by the weather that households experience, which necessitates the inclusion of weather variables in regression-based water demand modeling (Gutzler & Nims, 2005; Kenney *et al*., 2008). However, the inclusion of weather variables in the model is challenging for two reasons: first, the weather varies within the region because of the microclimatic variation; and second, the impact of weather on consumption is likely to vary by land-use characteristics (urban density, for instance) (Kenney *et al*., 2008). Moreover, it is unclear how individuals perceive temperature and precipitation events and how those perceptions align with actual evapotranspiration (ET) requirements of the landscape (Kenney *et al*., 2008).

Accordingly, the use of weather variables for estimating water demand is inconsistent in the literature. Vásquez Lavín *et al*. (2017) used monthly average temperature and monthly average precipitation to estimate monthly household water consumption. They found a negative correlation of consumption with both weather variables. In contrast, Hoyos & Artabe (2017) used average annual temperature and found the relation as positive. Espey *et al*. (1997), in their meta-analyses, found significantly lower estimates of elasticity when both ET and rainfall were included in the model. Clarke *et al*. (2017) used potential ET instead of temperature. They also used number of rainy days along with total precipitation and found the former to be a better indicator and more significant in consumer landscape-watering decisions. Martínez-Espiñeira (2002) suggested a similar result – the impact of a number of rainy days on the water demand is larger than that of the total rainfall. Hoffmann *et al*. (2006) used the number of rainy days and the number of warm days in each quarter (those with a daily maximum in the uppermost quartile of all daily temperatures) and found a significantly negative impact on water demand. Romano (2014) found no influence of temperature while estimating the average consumption of drinking water for domestic use; however, they showed that precipitation exerts a significant negative effect on water consumption. Maidment and Miaou (1986) showed the non-linear effect of weather variables in water demand using daily water use data. They specified that the effect of rainfall in water demand diminishes over time; also, there is no effect of daily maximum air temperature between 40 to 70-degree Fahrenheit; beyond 70, water demand increases with the increases in temperature. A consensus for weather variable parameterizations in water demand modeling has not yet been reached (Kenney *et al*., 2008), and since the above studies differ in location, time, and model choice, it is unclear if differences in their results are a function of study setting or model choice.

2.4 Choice of Functional Form

While reviewing water demand literature, Arbués *et al*. (2003) mentioned the choice of functional form as another problem in the estimation process. Vásquez Lavín *et al*. (2017) added that the use of different functional form in water demand estimation yields different estimates of the coefficient and recommended to report the different estimated values for different functional form.

Economic theory states that any specification of functional form while modeling water demand should characterize households' utility function. Two common approaches in the literature for defining the model in the similar outline are Cobb-Douglas utility function (Howe & Linaweaver, 1967) and Stone-Geary utility function (Clarke *et al*., 2017; Gaudin *et al*., 2001; Vásquez Lavín *et al*., 2017). The former approach does not segregate between different uses of water (indoor or outdoor water uses, for example) and models the water demand as a single commodity, while the latter approach separates the demand of water across different uses and models the water demand treating each use of water as a separate entity (Baumann *et al*., 1997). Both Cobb-Douglas and Stone-Geary functional form are inflexible in terms of direct estimation of price elasticity.

Alternatively, simpler and flexible functional form are offered across the literature; among which, the most commonly used functional form is the linear water demand function. Dalhuisen *et al*. (2001) found a linear specification to be much simpler and to have substantial explanatory power as compared to the non-linear specification. This functional form has faced criticism in the literature for two reasons: first, there is a constant change in water demand with respect to the change in price at every price level, i.e., price elasticity decreases along the demand curve under linear price specification (Arbués *et al*., 2003; Gaudin, 2006); and second, the consumption of water is zero at the choke price which contradicts properties of water as an essential good (Arbués *et al*., 2003; Hoyos & Artabe, 2017).

Recent literature has used log-log (double-log) functional form more frequently. Under the log-log functional form, price elasticity coefficients can be directly estimated and can be compared with the previous estimates (Gaudin, 2006; Hewitt & Hanemann, 1995). This functional form assumes elasticities to be constant over the entire domain of variables which actually might differ for the low and high prices (Arbúes *et al.*, 2004). Because of the nonconstant price elasticities, the double-log functional form requires the use of the squared term, which lets the price elasticity to decrease with the price (Gaudin, 2006; Hoyos & Artabe, 2017). Alternately, the semi-log functional form is used because it allows us to measure price elasticity when it is not constant (Arbúes *et al*., 2004).

3. Methods and Model Specification

3.1 Data

3.1.1 Water Billing Data (Water Consumption and Pricing Structure)

Meter-level billing data on monthly water consumption and water rates (pricing structure) are obtained from Fort Collins Utilities for a period of nine years (2006-2014). This raw dataset is an unbalanced panel[6] of single-family households[7] from where we obtain monthly water consumption, number of days in each billing period, amount of revenue collected for corresponding billing periods, bill dates, and a unique identification for households' and water-tap numbers. Differentiation across households and tap accounts allow us to remove accounts which serve many locations, under the assumptions that these are rental homes where utilities are likely included in the lease rate.[8]

Fort Collins Utilities use a combination of base charges (fixed irrespective to the consumption level) and volumetric charges (per 1,000 gallons) to bill the households. Households face different marginal rates (per unit price) based on the block they are in and rates increase across blocks. Three-blocked water rate structure is used to calculate the total bill of the households. These blocks are determined based on the volume of consumption, which breaks at 7,000 and 13,000 gallons (for single-family households) i.e., 0 to 7,000 gallons for the first block, 7,000 to 13,000 gallons for the second block, and over 13,000 gallons for the third block. Although these volumetric breaks of consumption levels for different blocks remained constant throughout the study periods, the price at each block changed five times between 2006 to 2014 (see **Figure 1**).

Figure 1

---

[6] Panel data approaches have comparative advantage over cross-section or time series data as they estimate water demand more accurately, allowing both temporal and subject-based variability to integrate into the coefficient estimates (Polebitski & Palmer, 2009).

[7] About 76% of taps in Fort Collins Utility service areas are for the single-family households.

[8] While these customers may still respond to price increases through higher lease rates or the installation of efficiency appliances, these premises add a complication in single family home estimation and are unlikely to have the same underlying data-generating process as homes who pay their utility bill directly.

3.1.2 Construction of Price Variables

The consumption level of the households determines the block group they face for any given period. Consequently, such block group determines the rate that the households face, which we considered as their marginal price (MP). We create an average price (AP) variable based on total revenue and consumption data as recorded on the water bill.[9] Both MP and AP were lagged by one period since households are aware of the price only from the previous billing period when deciding their current consumption level (Clarke *et al*., 2017).[10] While it is not the primary reason to do so, an additional benefit of using lagged price as the explanatory variable is the possibility of reducing simultaneity between price and consumption (Clarke *et al*., 2017; Garcia-Valiñas *et al*., 2014).

3.1.3 Weather data

Daily weather data from a period 2006 to 2014 are collected from the Colorado Agricultural Meteorological Network (CoAgMet) station located in Fort Collins (CCC, 2018). As shown **Table 1**, we include only those weather variables that are frequently used in the literature which are: maximum temperature (Martínez-Espiñeira & Nauges, 2004; Olmstead *et al*., 2007; Wichman *et al*., 2016), minimum temperature (Guhathakurta & Gober, 2007), average temperature (Klaiber *et al*., 2014; Price *et al*., 2014), average precipitation (Vásquez Lavín *et al*., 2017; Yoo *et al*., 2014), evapotranspiration (Baerenklau *et al*,, 2014; Wichman, 2014), number of cooling degree days (Lyman, 1992; Strong & Smith, 2010), total amount of precipitation (Clarke *et al*., 2017; Kenney *et al*., 2008; Maas *et al*., 2017), number of days with precipitation (Clarke *et al*., 2017; Martínez-Espiñeira, 2002) and average relative humidity (Hung *et al*., 2017). The first five weather variables are obtained directly from CoAgMet station while the other four weather variables are created using maximum temperature, minimum temperature, precipitation, maximum relative humidity, and minimum relative humidity (see **Appendix 1** for construction of weather variables). We aggregate

---

[9] We also computed the actual revenue for each observation using the fixed charge and the block price they face in accordance to their consumption level to calculate actual average price that the households face. Observations were dropped from the dataset if the absolute difference between actual average price and the average price based on the recorded data on water bill is more than $2.

[10] While creating a lagged variables, we ensure that the observations for the given households take the value from the previous bill which is only one month ago.

these daily weather variables across the exact date of each household to match the number of days in the billing cycle.

Table 1

3.1.4 Data Cleaning

Any observations with more than one tap per household, observations with a consumption level of less than 10 gallons and more than 2,160,000 gallons[11], observations with total revenue less than base charge (fixed charge) and observations with days of service (number of days in each billing period) less than 25 days and more than 35 days[12] are dropped from the dataset for analysis. We also drop observations when the difference between the action date (defined as the date when the bill is notified to the households) and the charge date (defined as the date when the bill is actually charged) is more than 30. Further, any observations that do not appear for more than 24 times throughout the period of study are also dropped to assure sufficient within variation in each group of the household. Lastly, we drop observations with any missing information.

3.1.5 Summary Statistics

After cleaning the data, we are left with 1,278,882 unique observations for a total of 21,874 single-family households from the period of Jan 2006 to Dec 2014. The share of these observations across different blocks i.e., first, second and third block group are 55.32%, 22.84%, and 21.84% respectively, indicating that majority of the observations face the lower marginal price.

We begin our analysis plotting mean monthly value for consumption, average price, and marginal price, across the study period (**Figure 2**). As expected, there is a noticeable decline in the average consumption of water over the years. One reason might be the prevalence of IBR structure throughout the study period which induced conservation since households face higher marginal price when their consumption increases. The city of Fort Collins has faced

---

[11] This is the maximum amount of water that a ¾" tap could dispense in a month given the average water pressure in the city. Realistically, this upper bound is likely too high, since it can only be achieved by running water from every faucet continuously for 30 days.

[12] Fort Collins Utility used human meter readers during this study period so the billing length (days) were not same across households. Households might face higher marginal rate if their billing cycles are longer. On that account, we decide to control the number of billing days using this arbitrary number.

the same marginal rate in each block from the period of 2007 to 2009, which means households are consuming either in the same block group or in the lower one, but not in the higher block. This should hold the consumption constant or should decrease it. Not surprisingly, we observe the declining trend of water consumption in those years. The sharpest decline in consumption can be observed during 2013-2014, which coincided with the largest incremental increase in blocked rates. Consistent with economic theory, an increase in price during this period corresponds to a decrease in consumption.

Figure 2

Descriptive statistics of the variables along with their description used in our analysis are listed in **Table 2** and **Table 3** respectively. A large variation can be observed in the water consumption – the dependent variable of our data referring to the total water consumption of a household in a distinct billing period (which is identical to one month) – since standard deviation (8367.87 gallons) is very large and almost close to the mean consumption (9039.83 gallons). We observe a noticeable difference between mean monthly average price ($5.50 per 1000 gallons) and mean monthly marginal price ($2.27 per 1000 gallons) because the fixed charges are embedded only into the average price but not into the marginal price. Except price, other independent variables that have been used for the analysis are different weather parameters as stated in **Section 3.1.3** and number of billing days in each cycle (*DOS*).[13]

Table 2 and **Table 3**

3.2 Model Specification

Pooled Ordinary Least Squares vs Fixed Effects vs Random Effects model

Prior to modeling our residential water demand, we decide to specify our model with the correct form of functional form, which otherwise will lead to imprecise estimates of price elasticity. Valid reasoning for the choice of functional form has often been ignored in the literature. The appropriate functional form in which the data fits better can be known by

[13] We choose not to include income variables in our model, since we expect income of the households to be constant throughout the study period. We also choose not to adjust the price for inflation because adjusting the price without the real adjustment in income will lead to imprecise (upward biased) estimation of price elasticity. While adjusting the nominal price to real one helps to remove the affect of general inflation, the true effect of price changes on consumption cannot be determined without adjusting income.

knowing the data itself. Rather than following many sophisticated approaches of empirically fitting the data to choose different functional form, we choose simply to plot the predicted value of consumption against different possible values of price (average).[14] As shown in **Figure 3,** the relationship between water consumption and price is found to be non-linear. We compare this graph with the graphs of the standard non-linear form (linear-log, log-linear, and log-log) and found it to coincide with the plot of log-log functional form. Also, this functional form preserves the assumptions of zero mean and normality of the residuals for all the specified model in Equation (1), (2) and (4) under both MP and AP specification (see **Figure 4**).[15]

Figure 3 and **Figure 4**

To ensure robustness, we begin our specification with the pooled Ordinary Least Squares (OLS) model:

$$\begin{aligned} ln(Q_{it}) = \beta_1\, ln(price_{it-1}) + \beta_2 tmax_t + \beta_3 tmin_t + \beta_4 tave_t + \beta_5 rhave_t \\ + \beta_6 pp_t + \beta_7 ppdays_t + \beta_8 totalpp_t + \beta_9 ET_t + \beta_{10} cdd_t \\ + \beta_{11} DOS_{it} + \alpha + \varepsilon_{it} \end{aligned} \tag{1}$$

where $ln(Q_{it})$ is the natural log of water consumption for households *i* over billing period *t*; $ln(price_{it-1})$ is the natural log of one period lagged average or marginal price; $tmax_t$ is maximum daily temperature; $tmin_t$ is minimum daily temperature; $tave_t$ is an average of daily mean temperature; $rhave_t$ is average relative humidity; pp_t is average precipitation; $ppdays_t$ is number of days with precipitation; $totalpp_t$ is the total amount of precipitation; ET_t is average evapotranspiration rate; cdd_t is number of cooling degree days; DOS_{it} is number of billing days for each billing period; β_s are the coefficients of parameters to be estimated; α is estimated as the constant term which neither varies with the households nor with the time; and ε_{it} is an error term.

---

[14] We simulate our model with the assumed function i.e., $\hat{Q} = AP^{\beta}$, where $\hat{Q}$= predicted consumption, P is a vector of average prices between $1 to $12 per 1000 gallons, β= price elasticity form preferred model (from **Table 7**), and A = constant generated by using the monthly average (of June 2014) of each of the independent variables.

[15] Before plotting the histogram of the residuals, all three models in Equation (1), (2) and (4) are instrumented with fixed charge, price of block one, successive difference between block groups and days of service as in Equation (5a); and also standard errors estimated are robust to households.

Under zero conditional mean of ε_{it}, homoscedasticity, strict exogeneity of explanatory variables, OLS renders consistent and efficient estimates and needs no further analysis (Greene, 2003). However, our dataset is not rich enough to include housing characteristics like household size, number of bathrooms, income and demographic characteristics like age, sex and social attitudes in our model – which should also be considered as the factors influencing residential water demand as found in different water demand literature (Arbués *et al*., 2003; Klein *et al*., 2007; Maas *et al*., 2017). The estimates from OLS will no longer be consistent and suffer biases when such variables are excluded from the model.

Fortunately, panel data allows us to incorporate such unobserved household heterogeneity, introducing a household-specific fixed or random effect term into the model. If household effects are assumed to be constant over time (i.e., time-invariant) and are correlated with independent variables, then a fixed effects (FE) model (the introduction of fixed-term[16] – household-specific constant effect – into the model) reduces the threat of omitted variable biases and removes the effect of all the variation across households.[17]

The FE model is specified as follows:

$$ln(Q_{it}) = \beta_1\, ln(price_{it-1}) + \beta_2 tmax_t + \beta_3 tmin_t + \beta_4 tave_t + \beta_5 rhave_t + \beta_6 pp_t + \beta_7 ppdays_t + \beta_8 totalpp_t + \beta_9 ET_t + \beta_{10} cdd_t + \beta_{11} DOS_{it} + \alpha_i + \varepsilon_{it} \qquad (2)$$

where α_i is the fixed effects term (time-invariant unobservable) that are averaged out of the model using the within transformation.

Equation (2) can be rewritten as:

$$ln(\overline{Q}_{it}) = \beta_1\, ln(\overline{price}_{it-1}) + \beta_2 \overline{tmax}_t + \beta_3 \overline{tmin}_t + \beta_4 \overline{tave}_t + \beta_5 \overline{rhave}_t + \beta_6 \overline{pp}_t + \beta_7 \overline{ppdays}_t + \beta_8 \overline{totalpp}_t + \beta_9 \overline{ET}_t + \beta_{10} \overline{cdd}_t + \beta_{11} \overline{DOS}_{it} + \overline{\varepsilon}_{it} \qquad (3)$$

[16] If such unobservable heterogeneity of households are time-invariant, then the effects of such variables on households will be same across the study period so they are considered as fixed (Allison, 2009); and such fixed characteristics do not influence any change in the dependent variable (Stock & Watson, 2015).

[17] Here, we assume that any changes over time in unobservable heterogeneity within each household are uncorrelated with changes over time in price and consumption, or else FE model will still suffer from omitted variable bias.

where each of these variables is represented in demeaned form.

However, if we assume such time-invariant characteristics to be uncorrelated with the included variables of all past, present and future time periods, then random effects (RE) model (the introduction of household-specific random effects into the model) produces efficient estimates (smaller standard error than FE model) of the coefficient.[18] Compared to the FE model, the RE model allows us to estimate the effects of time-invariant characteristics – improper control of which in the model may lead to the biased and inconsistent estimates of the coefficient.

The RE model is specified as follows:

$$\begin{aligned} ln(Q_{it}) = \beta_0 + \beta_1\, ln(price_{it-1}) + \beta_2 tmax_t + \beta_3 tmin_t + \beta_4 tave_t + \beta_5 rhave_t \\ + \beta_6 pp_t + \beta_7 ppdays_t + \beta_8 totalpp_t + \beta_9 ET_t + \beta_{10} cdd_t \\ + \beta_{11} DOS_{it} + v_{it} \end{aligned} \tag{4}$$

where $v_{it} = \mu_i + \varepsilon_{it}$ is the composite error term; μ_i refers to the household-specific random effect term which does not vary with the time; ε_{it} is an error term that explains both temporal and spatial variation; and β_0 is the constant term.

We test for heteroscedasticity specifying both FE and RE models with MP and AP specification.[19] For the FE model, we compute modified Wald statistics as suggested by Greene (2003); while for the RE model, we perform the Breusch-Pagan LM test. Regardless of the specified price, heteroskedasticity is present in both models (for FE - with MP: $\chi^2 = 8.7 \times 10^5$, with AP: $\chi^2 = 8.7 \times 10^5$and for RE - with MP: $\chi^2 = 4.2 \times 10^6$, with AP: $\chi^2 = 1.0 \times 10^6$) because the variance in price for the households in each block group differs from each other, which means the effect of price on consumption is different for the households in the different block group. Such variance does not affect the observations individually but affects the households within each group, leading to the groupwise heteroskedasticity.

[18] However, efficiency is rarely a concern in our model because of the very large sample size, which produces very small standard errors.
[19] We do not know the price variables that households respond to, so we conducted test for heteroscedasticity on both average and marginal prices.

We choose between FE and RE models using a Mundlak approach as suggested by Greene (2011), as an alternative to a Hausman Test.[20] The Mundlak approach tests if time-invariant characteristics in the model are correlated with the independent variables in the RE model. This technique computes the average of each of the independent variables at household-level and uses those averages as a regressor (separate variable) in addition to what is already in Equation (4). Then, we test for whether those averages are jointly equal to zero as a null hypothesis. Rejection of the null indicates that the time-invariant characteristics in Equation (4) are correlated with the other independent variables (which is an assumption for the FE model), suggesting the use of the FE model.

This test assumes that the regressors in the RE model are uncorrelated with the household-specific error terms (i.e., μ_i in Equation 4). Under the null, such orthogonality condition is valid, which increases the efficiency of random effects estimators compared to fixed effects; alternatively, the rejection of null suggests the estimates from RE model is not consistent because the condition of orthogonality might not be valid; hence FE model should be preferred.

We employ this Mundlak technique to our model and find the FE model to yield consistent coefficients estimates for both MP ($\chi^2(11) = 5.5 \times 10^4, p = 0.00$) and AP ($\chi^2(11) = 4.0 \times 10^4, p = 0.00$) specification.

Instrumental Variables Approach

Under block rate pricing structure, both price (MP or AP) and consumption are simultaneously determined which biases coefficient estimates (Arbués *et al.*, 2003). The use of instrumental variables to resolve the problem of simultaneity between price and consumption is increasingly common in water demand literature. In the presence of simultaneity, price is endogenous i.e., it is correlated to the error. It appears that much of the variation in consumption is due to the variation in price, though some part of the variation is

[20] Hausman Specification Test relies on the assumption of homoscedasticity so we cannot perform this test to choose between FE and RE model. Unlike Hausman test, this test allows us to cluster the estimates of standard error at households' level. Clustering the error at household level is necessary, since these observations are related, and will lead to more consistent estimates of standard error.

because of the variation in error (caused by the variables outside the equation). Under such condition, OLS suffers from biases and the estimated price coefficient will not be consistent.

An introduction of the instrumental variable, if valid, helps to find the true effect of price on consumption by isolating the exogenous part of the price (the part that is uncorrelated with the error). The validity of the instruments relies on two assumptions: 1) Relevance and 2) Exogeneity (Greene, 2003). The first assumption of relevance states that the instruments should not only be correlated with the endogenous variables (price variable in our model) but also should explain the large variation of that variable (i.e., instruments should not be weak). The second assumption requires that such instruments should not be correlated to the error term. Valid instrumental variables improve the estimates by consistently estimating the coefficients though less precisely than OLS – in the presence invalid instruments, the loss of precision will be severe, and the estimates, compared to corresponding OLS, will be more biased and inconsistent.[21] This highlights the necessity of having valid instruments in the model.

Lagging price variable by one period may help ameliorate the problem of simultaneity; however, we still suspect the presence of simultaneity in our model because households may choose their current consumption level based on the last period of consumption, which is simultaneously determined by the price at that period. To confirm this suspicion, we perform a Davidson-MacKinnon test of exogeneity, which tests for the consistency of the OLS estimates against IV estimates. The null hypothesis for this test states that estimates from OLS estimator are consistent; rejecting the null suggests the presence of endogeneity and hence requires the use of IV estimator (Davidson & MacKinnon, 1993).

While most of the water demand literature rely on the theoretical explanation to justify the validity of the instruments, few – Flyr *et al*. (2019), for instance – have tested the validity of instrument choice empirically. We test a complete set of instruments commonly used in the literature (a fixed charge, price for block one, the successive difference between block groups and days of service).[22] We select two-stage least squares (2SLS) regression as our IV

[21] Further, in the presence of weak instruments, the sampling distribution of 2SLS and its t-stat will not be normal even if the number of observations are very large, leading to possible type I or II errors.

[22] See **Section 2.2** for the theoretical explanation on validity of these instruments.

techniques. First, we regress the lagged price variable on the set of price instruments and weather variables and test the correlation of our instruments with price variables running first stage regression; we also check the Kleibergen-Paap rk Wald F-statistic (hereafter referred as Kleibergen-Paap F-stat) to see whether these instruments are weak.[23]

This first stage of the 2SLS estimator is specified as follows:

$$\begin{aligned} ln(price_{it-1}) = {} & \eta_1 tmax_t + \eta_2 tmin_t + \eta_3 tave_t + \eta_4 rhave_t + \eta_5 pp_t \\ & + \eta_6 ppdays_t + \eta_7 totalpp_t + \eta_8 ET_t + \eta_9 cdd_t + \eta_{10} DOS_{it} \\ & + \delta_1 FC_{it-1} + \delta_2 blockprice1_{it-1} + \delta_3 blockdiff1_{it-1} \\ & + \delta_4 blockdiff2_{it-1} + \delta_5 DOS_{it-1} + \rho_i + \gamma_{it} \end{aligned} \tag{5a}$$

where η_s and δ_s are the coefficients of parameters to be estimated; FC_{it-1} is one period lagged fixed charge set by the utility; $blockprice1_{it-1}$ is one period lagged price in block group one i.e., first block; $blockdiff1_{it-1}$ is one period lagged difference in prices between block group one and block group two; $blockdiff2_{it-1}$ is one period lagged difference in prices between block group two and block group three; DOS_{it-1} is one period lagged days of service; ρ_i is the fixed effect term; and γ_{it} is the error term.

We predict the price from this first stage (using Equation 5a) and then use it as a regressor on the consumption of water in the second stage where we test the exogeneity of the instruments using the Hansen J-statistic.[24]

This second stage of the 2SLS estimator is specified as follows:

---

[23] Since we clustered our error over both households and time to solve the problem of heteroskedasticity in our model, Cragg-Donald Wald F-statistic is no longer valid and hence should not be used as a measure to check the weakness of the instrument; instead Kleibergen-Paap F-statistic can be used (Baum, 2007). However, the rule of thumb for both of these F-statistics is same, that the value of F-statistic (joint significance) of the instrument(s) from first stage regression should be greater than 10 for exactly identified model i.e., when number of instruments equals the number of endogenous regressors and 13.91 when the number of instruments used exceeds the number of endogenous regressors or else those instrument(s) will be considered as weak.

[24] Hansen J-statistic, often regarded as J-test, is not the actual test of exogeneity, however, is the test of overidentifying restrictions, which tests the exogeneity of the overall instruments given at least one of the instruments used in the model is exogenous. Under null, the instruments are exogenous i.e., uncorrelated with the error; rejecting the null suggests at least one of the instruments in the model is not exogenous i.e., not valid. This test does not perform when the model is exactly identified i.e., when number of instruments used in the model are equal to the number of endogenous regressors.

$$ln(Q_{it}) = \beta_1\, ln(\widehat{price_{it-1}}) + \beta_2 tmax_t + \beta_3 tmin_t + \beta_4 tave_t + \beta_5 rhave_t + \beta_6 pp_t + \beta_7 ppdays_t + \beta_8 totalpp_t + \beta_9 ET_t + \beta_{10} cdd_t + \beta_{11} \mathrm{DOS}_{\mathrm{it}} + \alpha_i + \varepsilon_{it} \quad (5b)$$

where $ln(\widehat{price_{it-1}})$ is the predicted price from equation (5a), α_i is the fixed effect term and ε_{it} is the error term.

The Kleibergen-Paap *F*-stats from the first-stage regression for both AP and MP specification are 92.88 and 88.04 respectively (which are greater than the critical value of 13.91) suggesting that the IVs used in Equation (5a) are not weak instruments. The p-values from Hansen *J*-statistic for both AP and MP specification are 0.0018 and 0.30 respectively, indicating that at least one of the instruments is not exogenous when our model is specified with the average price. Since we have valid economic reasoning for exogeneity of each of our instruments, we are not much concerned for the rejection of null with *J*-test and still prefer all five instruments in Equation (5a) to perform Davidson-MacKinnon Test of exogeneity.[25]

By employing Davidson-Mackinnon Test, we reject the null hypothesis (at 5% level of significance) that price variables (both AP and MP) are exogenous, which confirms the presence of endogeneity and requires the use of price instruments to get unbiased estimates of coefficients.

Using the FE-IV model, we proceed with the remainder of this section as follows: first, we employ Shin Test to determine the price variable that households respond to; second, we test the sensitivity of price elasticity to the choice of weather variables; third, we test the sensitivity of price elasticity to the choice of instrumental variables; and fourth we select a model using different measures of goodness-of-fit and cross-validate it.

[25] We recommend that the results from *J*-test should be taken lightly since *J*-test does not strictly determine if instruments are valid. Indeed, the sensitivity tests around the instrument choice suggests, they are valid (in most cases). Even if the instruments used are valid, the test can be rejected for two reasons: i) because of the wrong functional form associated with the model (though not a problem in our case) ii) or because of the heterogeneous effect of the price on consumption. We argue that the rejection of the null hypothesis in our model is because of the second reason (see **Appendix 2** for details).

3.2.1 Shin Test

The very first objective of this research is to determine the price variables – average or marginal price – to which households respond and correctly specify this price signal into the residential water demand model. Shin (1985) developed a formal test called the Shin Test to determine the perceived price while estimating electricity demand. He introduced a perceived price (P^*) as a function of marginal price (MP), average price (AP) and the price perception parameter (k) such that:

$$P^* = MP_{it}\left(\frac{AP_{it-1}}{MP_{it}}\right)^k \tag{6}$$

The price that households perceive depends on the value of k. Households perceive marginal price when $k = 0$; households perceive average price when $k = 1$; perceived price rests between marginal and average price when $0 < k < 1$; and it is above average price when $k > 1$ (see Flyr *et al*., 2019; Nieswiadomy & Molina, 1991; Shin, 1985; Taylor *et al*., 2004 for details).

Because water shares a similar rate structure to electricity, we decide to employ this test, similar to Flyr *et al*. (2019) and Nieswiadomy and Molina (1991), in our water demand model.[26] We lag the average price by one period but use the contemporaneous marginal price to create P^* and substitute the price variable ($\widehat{price_{it-1}}$) in Equation (5b) with P^*.[27] We estimate the following equation to test the price variable that households perceive:

$$\begin{aligned} ln(Q_{it}) = {} & \beta_1\, ln\left(\widehat{MP_{it}}\left(\frac{\widehat{AP_{it-1}}}{\widehat{MP_{it}}}\right)^k\right) + \beta_2 tmax_t + \beta_3 tmin_t + \beta_4 tave_t + \beta_5 rhave_t \\ & + \beta_6 pp_t + \beta_7 ppdays_t + \beta_8 totalpp_t + \beta_9 ET_t + \beta_{10} cdd_t \\ & + \beta_{11} DOS_{it} + \alpha_i + \varepsilon_{it} \end{aligned} \tag{7}$$

where $\widehat{MP_{it}}$ is the instrumented contemporaneous marginal price; $\widehat{AP_{it-1}}$ is one period lagged predicted average price using full set of instruments.

---

[26] Flyr *et al*. (2019) employed this test while modeling commercial water demand and Nieswiadomy & Molina (1991) performed this test while modelling residential water demand.

[27] We first predicted the lagged AP and MP from the first stage (identical to Equation 5a) and then used such predicted prices to construct P^* in the second stage (as an substitute $\widehat{price_{it-1}}$ in Equation 5b) to solve the endogeneity issues in P^*, if any.

Equation (7) can be rewritten as:

$$
\begin{aligned}
ln(Q_{it}) = {} & \beta_1\, ln\big(\widehat{MP_{it}}\big) + \delta\, ln\left(\frac{\widehat{AP_{it-1}}}{\widehat{MP_{it}}}\right) + \beta_2 tmax_t + \beta_3 tmin_t + \beta_4 tave_t \\
& + \beta_5 rhave_t + \ \beta_6 pp_t + \ \beta_7 ppdays_t + \beta_8 totalpp_t + \beta_9 ET_t \\
& + \ \beta_{10} cdd_t + \beta_{11} DOS_{it} + \alpha_i + \varepsilon_{\mathrm{it}}
\end{aligned}
\tag{8}
$$

where $\delta = \beta_1 k$ and δ is not clearly identified.

We use Equation (8) to test the null hypothesis: $\beta_1 = \delta$. Failing to reject this hypothesis means $k = 1$, indicating households are more responsive to average price than marginal price.

3.2.2 Test Sensitivity to the Choice of Weather Variables

The second objective of this paper is to formally evaluate the effect of weather variable choices on own-price elasticity estimates of water demand. To begin, we combine the nine weather variables from our dataset in 511 possible ways.[28] Keeping the instrument variable constant, we remodel the first stage regression Equation (5a) as follows:

$$
\begin{aligned}
ln(price_{it-1}) = {} & \eta_0 + \eta_1 DOS_{it} + \eta_{k+1} X_t + \delta_1 FC_{it-1} + \delta_2 blockprice1_{it-1} \\
& + \delta_3 blockdiff1_{it-1} + \ \delta_4 blockdiff2_{it-1} + \delta_5 DOS_{it-1} + \rho_i + \gamma_{it}
\end{aligned}
\tag{9a}
$$

where X_t is every possible set of weather variable during period *t;* and η_{k+1} are the coefficients of weather variables to be estimated.

Using the predicted price from Equation (9a), we estimate the second stage regression as:

$$
ln(Q_{it}) = \beta_1\, ln(\widehat{price_{it-1}}) + \beta_2 DOS_{it} + \beta_{s+2} X_t + \ \alpha_i + \varepsilon_{it} \tag{9b}
$$

where β_{s+2} are the coefficients of weather variables to be estimated.

Equation (9b) provides us the direct estimation of price elasticity of water demand.

3.2.3 Test Sensitivity to the Choice of Instrumental Variables

The third objective of this paper is to evaluate the effect of the choice of price instrument on own-price elasticity of water demand. We hold the weather variables in Equation (5a) as

---

[28] We determine the number of *n*-combinations for all *n* using formula 2^n -1 where '1' rules out the empty set. In our case, *n* = 9 (total number weather variables), so 2^n -1= 511.

constant and allow the instrumental variables to run across the model. We create 31 different sets of price instruments (using the same technique of 2^n-1 as in **Section 3.2.2**) by combining five different price instruments that we used before and then remodel Equation (5a) as follows:

$$\begin{aligned} ln(price_{it-1}) = \eta_1 tmax_t + \eta_2 tmin_t + \eta_3 tave_t + \eta_4 rhave_t + \eta_5 pp_t \\ + \eta_6 ppdays_t + \eta_7 totalpp_t + \eta_8 ET_t + \eta_9 cdd_t + \eta_{10} DOS_{it} \\ + \delta_K Z_{it-1} + \rho_i + \gamma_{it} \end{aligned} \tag{10a}$$

where Z_{it-1} is every possible set of instrumental variables; and δ_K are the coefficients of instrumental variables to be estimated.

As discussed earlier, the Kleibergen-Paap *F*-stat is reported and compared to the critical value to inspect the weakness of each set of price instrument. We use predicted price from Equation (10a) – ruling out any prediction from a weak set of instruments – and estimate the second stage regression as follows to create the distribution of elasticity.

$$\begin{aligned} ln(Q_{it}) = \beta_1\, ln(\widehat{price_{it-1}}) + \beta_2 tmax_t + \beta_3 tmin_t + \beta_4 tave_t + \beta_5 rhave_t \\ + \beta_6 pp_t + \beta_7 ppdays_t + \beta_8 totalpp_t + \beta_9 ET_t + \beta_{10} cdd_t \\ + \beta_{11} DOS_{it} + \alpha_i + \varepsilon_{it} \end{aligned} \tag{10b}$$

We also report Hansen *J*-statistic (though we recommend this test to be taken lightly) as a test of overidentifying restriction whenever possible.

3.2.4 Model Selection and Cross-Validation

We follow a global search regression technique – an exhaustive search approach – to select the optimal model based on different measures of goodness-of-fit (Gluzmann & Panigo, 2013). Unlike heuristic search approaches with backward and forward looping, this technique assures the optimality of the variable choices while selecting models by completing the regression across every possible combination of independent variables.

We conduct this global search regression technique in two steps: first, we select the best set of instruments based on its strength using the Kleibergen-Paap *F*-stat (as reported from Equation 10a); and second, we use the predicted price from that chosen set of instruments in the second stage (identical to Equation 9b) across all combinations of weather variables. At the same time, we generate a normalized index (*nindex*[29]) based on three different measures of goodness of fit which are adjusted-R^2, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) (weighted on the ratio of 0.3:0.3:0.4 respectively), using *gsreg* command in Stata (Gluzmann & Panigo, 2013). We consider the model with the highest value of *nindex* as the optimal one.

To avoid over-fitting, we perform k-fold cross-validation (CV) techniques on our optimal model –similar to Yang *et al*. (2019). This technique divides the observation randomly into k groups (or folds) of equal size to see whether the model fits out-of-sample data or not (James *et al*., 2013). The model is first fitted on the *k*-1 group and later tested on the remaining one group (*i*). This process is followed iteratively to test each of the *k* different groups, each time fitting on remaining *k*-1 groups. Root mean squared error (RMSE) is estimated for each of the *k* trails which are averaged out to calculate *k*-fold CV estimate.

$$CV_k = \frac{1}{k}\sum_{i=1}^{k} RMSE_i \tag{11}$$

We set the value of *k* as 10 to cross-validate our model.

[29] This is the normalized linear combination of user-selected criterias (goodness-of-fit) which in our case are adjusted-R^2, AIC and BIC.

4. Results and Discussion

This section presents and discusses results from 1) a Shin Test, 2) the sensitivity of estimates to the choice of weather variables, 3) the sensitivity of estimates to the choice of price instruments, and 4) model selection and cross-validation.

4.1 Shin Test

Table 4 presents the results from the Shin Test estimated using Equation (7). The coefficients on both $ln(\widehat{MP_{it}})$ and $ln(\frac{\widehat{AP_{it-1}}}{\widehat{MP_{it}}})$ are tested to see whether they are statistically different from one another or not. Using a Wald Test, we find that the two coefficients – β_1 and δ – are equal (p-value=0.36). At this p-value, we fail to reject the null hypothesis that the two coefficients are statistically different, suggesting that households respond to lagged average price rather than marginal price. This result is different from what Shin (1985) and Nieswiadomy & Molina (1991) proposed that households respond to marginal price under IBR structure and to average price under DBR structure. However, our finding – households respond to average price under IBR structure – is consistent with more recent studies (Flyr *et al.*, 2019; Ito, 2014; Kenney *et al.*, 2008). As addressed in other literature, results from the Shin Test only provide weak evidence on the price to which households responds.[30] So, consistent with Flyr *et al.* (2019), we also do not suggest this test as valid in every case.

Table 4

Lagged average price specification, under IBR structure, can arguably be considered the appropriate and conventional approach for modeling residential water demand because households might not respond to the marginal price since the associated information costs in perceiving marginal price are substantial. Further, during the years of our study, Fort Collins Utility (FCU) provides information only on the payments from the last bill,[31] so households can only respond to the price (average) they face in the previous billing period. The result

[30] For instances, Taylor *et al.* (2004) argued that the estimates of k (Shin's price perception parameter) will be biased towards unity due to presence of fixed charge which might invalidate a Shin Test; Binet *et al*. (2014) also argued that the model specification might be wrong for households responding to marginal price when $k = 0$.
[31] Note that since this time, FCU has moved to a real time metering system, which provides opportunity for future research.

from the Shin Test, in this paper, provides empirical guidance to such theoretical argument, directing us to the use of lagged average price in modeling water demand.

4.2 Test Sensitivity to Choice of Weather Variables

As shown in **Figure 5**, the range of the distribution of price elasticity to the different choice of weather variables is -1.22 to -0.57 with the mean value of -0.71. Though this range falls on a qualitatively different group of elasticity, a large density of estimates is clustered around the mean since the standard deviation is very low (0.09). 97% of these estimates are between the range of -0.83 to -0.57. This range is consistent with the estimates in the meta-analysis done by Espey *et al*. (1997) where they found 90% of their estimates between -0.75 to 0. By comparison, our range is even smaller, and estimates are clustered around the mean because unlike their meta-analysis, our study is time and location-specific. Such high stability in estimates of price elasticity suggests that the residential water demand estimation for the single-family households is not sensitive to the choice of weather variables, which is because most of the weather variables are highly correlated to each other but are very slightly correlated with the price variable (see **Table 5**). Thus, the inclusion of either one or all the weather variables in the model does not strongly affect the price elasticity estimates. For example, the range of price elasticity using only the combinations of temperature-related variables i.e., maximum temperature, minimum temperature, average temperature, evapotranspiration, and cooling degree days – as a set of weather variables is -0.80 to -0.65. This is because these metrics are highly correlated with each other – the highest of 0.99 for maximum temperature, minimum temperature and average temperature among each other and the lowest of 0.73 between evapotranspiration and cooling degree days (see **Table 5**).

Table 5

Since these weather variables (control variables) in our model have little to no correlation with the average price (variable of interest), we observe high stability in the estimates of price elasticity across the suite of weather variables. This result allows us to predict price elasticity efficiently and consistently even without knowing all the weather that the households face since households’ consumption does not change significantly with change in weather conditions.

We find a similar case i.e., narrow range of estimates (-1.21 to -1.18) while using only rainfall-related variables i.e., average precipitation, total precipitation and number of rainy days as a preferred set of weather variables in our model. It is noteworthy that this range is qualitatively different (absolute elasticities greater than 1) from the estimates (-0.80 to -0.65) of only temperature-related variables, which points to different policy formulations (and implications) in terms of collecting revenue and inducing conservation as discussed in **Section 1**. However, so long as one metric for each (temperature and precipitation-related variables) is included in the model, relatively stable (both qualitatively and quantitatively) price elasticity estimates (-0.83 to -0.63) can be obtained. Surprisingly, this range is similar to the estimates of only temperature-related variables, suggesting that a complete exclusion of temperature-related metrics from the model will produce biased estimates of price elasticity whereas doing such for the precipitation (or related) variables may not have any effect on the estimates.

Our mean price elasticity estimate (-0.71) is slightly higher (in terms of absolute value) than what has been reported in previous meta-analysis of residential water demand estimation by Espey *et al*. (1997), Dalhuisen *et al*. (2003) and Sebri (2014), which are -0.51, -0.41 and -0.37 respectively because of four reasons: 1) IBR structure, 2) average price specification, 3) double-log functional form, and 4) panel data. Compared to DBR structure, IBR structure is likely to produce higher estimates of price elasticity; so is the case when the model is specified with the average price rather than marginal price (Dalhuisen *et al*., 2003; Marzano *et al*., 2018). Similarly, the likeliness to get relatively higher estimates of price elasticity is also greater when panel data are used for analysis (compared to time-series and cross-sectional data) and also when the model is specified with double-log functional form (in comparison to semi-log functional form) (Marzano *et al*., 2018).

Figure 5

4.3 Test Sensitivity to Choice of Instrumental Variables

Out of 31 different sets of instruments that are used in Equation (10a), we find the sole use of days of service as a weak instrument since the value of Kleibergen-Paap *F*-stat from first-stage regression is 9.53 (see **Table 6**) which is below the critical value of 10 (for exactly identified model). Though Flyr *et al*. (2019) discussed it as a valid instrument for addressing exogenous variation across both households and time, we find that the sole use of days of service is insufficient without the inclusion of an additional price metric or other instrument.

Table 6

With the use days of service as a sole instrument, demand for water is estimated as almost perfectly inelastic, -0.087 and insignificant (p-value = 0.78). As discussed earlier in **Section 3.2**, the use of weak instruments (invalid instruments) results in the severe loss of precision of the estimates and hence, estimates will be more biased and inconsistent compared to corresponding OLS. Thus, we eliminate the sole use of this instrument from our model.

Across other instruments, we find the range of price elasticity between -0.77 to -0.69 with a standard deviation of 0.01 (see **Table 6** and **Figure 6**). This range is very small and consistent with our expectation that any linear combination of instruments should be valid and should not produce significantly different estimates if each instrument is valid itself. As shown in the fourth column of **Table 6**, each of the instruments in the first five-row (except days of service) pass the weak identification test of instruments, preserving the first properties (i.e., relevance) of valid instruments, because of which all other linear combinations of these five instruments are also strong in explaining the true effect of price on consumption. For the second properties (i.e., exogeneity), we have mixed empirical results – rejecting and failing to reject the null hypothesis – for the p-values of Hansen *J*-statistics (see column 6 in **Table 6**). However, as discussed earlier, we do not rely on this test (see **Section 3.2** and **Appendix 2** for details) and rather prefer to go with the theoretical argument as stated in **Section 2.2**.

Figure 5

The mean price elasticity from the different choice of instrument variables (-0.72) is roughly equal to the mean price estimates from the different choice of weather variables (-0.71). This is because our estimates are consistent across the choice of both weather and instruments variables.[32]

---

[32] To clarify, we run an augmented model allowing both weather and instruments run together across the model. We find consistent estimates of price elasticity, regardless of which weather or instruments are allowed in the model (see **Appendix 3** for details).

4.4 Model Selection and Cross-Validation

We find the sole use of one period lagged price in block one as a strong set of instruments with Kleibergen-Paap *F*-statistic of 315.00, amongst others (see **Table 6** column 4). So $block1_{it-1}$ is our preferred instrument, and there is little reason to include additional instruments. Instrumenting the price variable with this instrument, we conduct a global search technique (identical to Equation 9b) to find the optimal set of weather variables based on the value of *nindex*. We find the highest value of *nindex* as 0.76 when the maximum temperature is excluded from the model, concluding that the model fits better when the full set of weather variables, except *tmax*, are used to predict price elasticity.

Table 7 shows the regression results for our preferred FE-IV model with the inclusion of the optimal set of weather and instruments variables. We find the elasticity of average price as -0.74 at the significance level of 0.01. This elasticity is not significantly different than the mean price elasticity, -0.71 and -0.72, across different choices of weather and price instruments respectively. We also find statistically significant coefficients for each of the weather variables, except for average precipitation (*pp*) and total amount of precipitation (*totalpp*); the high multicollinearity from these explanatory variables makes their interpretation difficult and reduces the statistical significance of each. Multicollinearity increases the standard error of the coefficient and consequently, the null hypothesis cannot be rejected because of the low power of the t-test. However, multicollinearity among control variables does not affect the parameter of our interest (average price) and thus can be ignored with ease, particularly given our large sample.[33]

Table 7

Table 8 shows the predicted RMSE for *k* (=10) different trials. We find similar predicted values of RMSE across each trail, which suggests that our preferred model equally (with accuracy) fits the out-of-sample data and hence can be implemented safely across similar cases.

Table 8

[33] To confirm, we run the regression removing *pp* and *totalpp* from the model and find similar estimates of elasticity and centered R^2 as tabulated in **Table 8**.

5. Conclusion

Overall, results from our in-depth analysis suggest that model choices with different—but reasonable—specification of weather metrics and instrumental variables have a negligible effect on own-price elasticity estimates. This is encouraging and suggests that previous works, despite inconsistency in the choice of weather or price instrument, may be are accurate in the point estimates of price elasticity and suggests that difference in elasticity may be a function of place, not figments of model choice.

Price is an appropriate tool for the demand-side management of residential water, but the effectiveness of price as a conservation tool requires an understanding of the household demand. While there is a different theoretical argument for employing the appropriate price signal into the model, there is no concrete test to provide empirical insight. The result from Shin Test, though it should be taken lightly, accords with our theoretical reasoning for the choice of average price into our model, thereby guiding us to specify our model with the average price.

Further, water demand literature is inconsistent with the choice of weather-related variables in demand modeling. Despite many studies, there is still limited consistency and theoretical motivation for (and how) the inclusion of weather variables into the water demand model. To our knowledge, this study constitutes the first attempt at formally evaluating the sensitivity of price elasticity estimates to weather variable choices. We find consistent estimates of price elasticity – 97% of our estimates between -0.83 to -0.57– across the different combination of weather variables, which suggests that the inconsistency of weather as control variables in past econometric studies may have little effect on their results.

Moreover, this result suggests that low-cost information like average temperature and total precipitation are sufficient to accurately reflect the data-generating-process, reducing costs and resulting in a parsimonious model. However, we do note that qualitatively different elasticity (>1) is estimated when only precipitation variables are used. While the inclusion of at least one variable from each temperature and precipitation related metrics produce relatively more stable (both qualitatively and quantitatively) estimates (-0.83 to -0.63) of price elasticity, similar estimates (-0.80 to -0.65) can be found just with the inclusion of temperature-related metrics. This further reduces the costs of obtaining more information on weather variables.

Similar to the choice of weather variables, we find highly stable estimates of elasticity (-0.77 to -0.69) to the choice of instrumental variables. So long as the instruments variables are valid, this result should hold and need no further analysis. Though days of service introduces enough exogenous variation across both households and time, it fails in explaining the true effect of price on consumption, and hence should not be used alone, which otherwise will lead to the imprecise estimation of price elasticity.

We find the price elasticity of average price as -0.74 for our preferred model i.e., when *tmax* is excluded from the model. This estimate is robust across the choice of both weather and price instruments and has also been cross-validated to fit out-of-sample data estimates with the aim of providing better insight to the stakeholders and water managers in establishing appropriate water pricing and pricing structure to induce conservation.

These findings may have important policy implications from demand-side management perspectives. First, while designing water management policies, it is important to understand that the types of weather variables that the households face will barely affect the estimation of households' demand. Second, since our estimates are consistent with the choice of both weather metrics and price instruments, it is possible to accurately estimate the demand with low-cost information such as only evapotranspiration as a weather variable and only price of the first block as a price instrument;[34] and, with such estimation, policies can be formulated with relatively less risk. Finally, water managers can use price as an effective tool to increase revenue since we find a high elasticity (-0.74) within inelastic range from our preferred model; however, pricing policies alone should not be used as a conservational tool to significantly reduce the water use, implying the need for combination of pricing and other dynamic (non-pricing policies, for example) strategies.

[34] With such choices, we find price elasticity as -0.61, which is roughly equal to what Kenney *et al*. (2008) found (price elasticity = 0.60) for the City of Aurora, Colorado. It is important to note this result because despite having less information in our model, compared to Kenney *et al*. (2008), we are able to precisely estimate the elasticity in a similar environment (i.e., City of Aurora and City of Fort Collins share a similar landscape and weather conditions.).

While we aim to help improve understanding of households' water consumption behavior and contribute to the existing literature by systematically evaluating model specification choices, there are still some methodological considerations in this paper that are debatable. First, we do not consider the possibility of estimating price elasticity at margin i.e., the likeliness of getting different elasticity for a high and low price; instead, we assume a constant elasticity over the entire domain of price variables to avoid the complexity in the estimation process, highlighting the need for a comprehensive modeling approach. Second, the use of the Shin Test to determine price signal can be argued for different reasons as discussed in **Section 4.1** and therefore requires experimental testing. Last, prices are generally more elastic in outdoor water use as evidenced in the literature, which is largely ignored in this research because of the data limitation. We do not know to which water uses (indoor vs outdoor, for instance) households respond when there is an increase in price. This suggests a need for future research to determine how each use type is affected by the increase in price.

References

Allison, P. D. (2009). *Fixed effects regression models* (Vol. 160). SAGE publications.

Arbúes, F., Barberan, R., & Villanua, I. (2004). Price impact on urban residential water demand: A dynamic panel data approach. *Water Resources Research*, *40*(11).

Arbués, F., Garcıa-Valiñas, M. Á., & Martınez-Espiñeira, R. (2003). Estimation of residential water demand: a state-of-the-art review. *The Journal of Socio-Economics*, *32*(1), 81–102.

Asci, S., Borisova, T., & Dukes, M. (2017). Are price strategies effective in managing demand of high residential water users? *Applied Economics*, *49*(1), 66–77.

Baerenklau, K. A., Schwabe, K. A., & Dinar, A. (2014). The residential water demand effect of increasing block rate water budgets. *Land Economics*, *90*(4), 683–699.

Baum, C. F. (2007). ivreg2: Stata module for extended instrumental variables/2SLS, GMM and AC/HAC, LIML and k-class regression. *Http://Ideas. Repec. Org/c/Boc/Bocode/S425401. Html*.

Baumann, D. D., Boland, J. J., & Hanemann, M. W. (1997). *Urban water demand management and planning*.

Binet, M.-E., Carlevaro, F., & Paul, M. (2014). Estimation of residential water demand with imperfect price perception. *Environmental and Resource Economics*, *59*(4), 561–581.

Center, C. C. (n.d.). No Title. Retrieved from https://coagmet.colostate.edu/cgi-bin/dailydata_form.pl

Clarke, A. J., Colby, B. G., & Thompson, G. D. (2017). Household water demand seasonal elasticities: a stone-geary model under an increasing block rate structure. *Land Economics*, *93*(4), 608–630.

Dalhuisen, J. M., Florax, R. J. G. M., de Groot, H. L. F. M., & Nijkamp, P. (2001). *Price and* *income elasticities of residential water demand: Why empirical estimates differ*. Tinbergen Institute Discussion Paper.

Dalhuisen, J. M., Florax, R. J. G. M., De Groot, H. L. F., & Nijkamp, P. (2003). Price and income elasticities of residential water demand: a meta-analysis. *Land Economics*, *79*(2), 292–308.

Davidson, R., & MacKinnon, J. G. (1993). Estimation and inference in econometrics. *OUP Catalogue*.

Espey, M., Espey, J., & Shaw, W. D. (1997). Price elasticity of residential demand for water: A meta-analysis. *Water Resources Research*, *33*(6), 1369–1374.

Flyr, M., Burkhardt, J., Goemans, C., Hans, L., Neel, A., & Maas, A. (2019). Modeling Commercial Demand for Water: Exploring Alternative Prices, Instrumental Variables, and Heterogeneity. *Land Economics*, *95*(2), 211–224.

Foster, H. S., & Beattie, B. R. (1981). On the specification of price in studies of consumer demand under block price scheduling. *Land Economics*, *57*(4), 624–629.

Garcia-Valiñas, M. A., Athukorala, W., Wilson, C., Torgler, B., & Gifford, R. (2014). Nondiscretionary residential water use: the impact of habits and water-efficient technologies. *Australian Journal of Agricultural and Resource Economics*, *58*(2), 185–204.

Gaudin, S. (2006). Effect of price information on residential water demand. *Applied Economics*, *38*(4), 383–393.

Gaudin, S., Griffin, R. C., & Sickles, R. C. (2001). Demand specification for municipal water management: evaluation of the Stone-Geary form. *Land Economics*, *77*(3), 399–422.

Gluzmann, P., & Panigo, D. (2013). *GSREG: Stata module to perform Global Search Regression*.

Greene, W H. (2011). *Econometric Analysis*. Pearson Education.

Greene, William H. (2003). *Econometric analysis*. Pearson Education India.

Griffin, R. C. (2016). *Water resource economics: The analysis of scarcity, policies, and projects*. MIT press.

Guhathakurta, S., & Gober, P. (2007). The impact of the Phoenix urban heat island on residential water use. *Journal of the American Planning Association*, *73*(3), 317–329.

Gutzler, D. S., & Nims, J. S. (2005). Interannual variability of water demand and summer climate in Albuquerque, New Mexico. *Journal of Applied Meteorology*, *44*(12), 1777–1787.

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

Hoffmann, M., Worthington, A., & Higgs, H. (2006). Urban water demand with fixed volumetric charging in a large municipality: the case of Brisbane, Australia. *Australian Journal of Agricultural and Resource Economics*, *50*(3), 347–359.

Howe, C. W., & 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.

Hoyos, D., & Artabe, A. (2017). Regional differences in the price elasticity of residential water demand in Spain. *Water Resources Management*, *31*(3), 847–865.

Hung, M.-F., Chie, B.-T., & Huang, T.-H. (2017). Residential water demand and water waste in Taiwan. *Environmental Economics and Policy Studies*, *19*(2), 249–268.

Ito, K. (2014). Do consumers respond to marginal or average price? Evidence from nonlinear electricity pricing. *American Economic Review*, *104*(2), 537–563.

James, G., Witten, D., Hastie, T., & Tibshirani, R. (2013). *An introduction to statistical learning* (Vol. 112). Springer.

Kenney, D. S., Goemans, C., Klein, R., Lowrey, J., & Reidy, K. (2008). Residential water demand management: lessons from Aurora, Colorado1. *JAWRA Journal of the American Water Resources Association*, *44*(1), 192–207.

Klaiber, H. A., Smith, V. K., Kaminsky, M., & Strong, A. (2014). Measuring price elasticities for residential water demand with limited information. *Land Economics*, *90*(1), 100–113.

Klein, B., Kenney, D., Lowrey, J., & Goemans, C. (2007). Factors influencing residential water demand: a review of the literature. *Unpublished Paper, Colorado State University, January*.

Lyman, R. A. (1992). Peak and off-peak residential water demand. *Water Resources Research*, *28*(9), 2159–2167.

Maas, A., Goemans, C., Manning, D., Kroll, S., Arabi, M., & Rodriguez-McGoffin, M. (2017). Evaluating the effect of conservation motivations on residential water demand. *Journal of Environmental Management*, *196*, 394–401.

Maidment, D. R., & Miaou, S. (1986). Daily water use in nine cities. *Water Resources Research*, *22*(6), 845–851.

Martínez-Espiñeira, R., & Nauges, C. (2004). Is all domestic water consumption sensitive to price control? *Applied Economics*, *36*(15), 1697–1703.

Martínez-Espiñeira, R. (2002). Residential water demand in the Northwest of Spain. *Environmental and Resource Economics*, *21*(2), 161–187.

Marzano, R., Rougé, C., Garrone, P., Grilli, L., Harou, J. J., & Pulido-Velazquez, M. (2018). Determinants of the price response to residential water tariffs: Meta-analysis and beyond. *Environmental Modelling & Software*, *101*, 236–248.

Nataraj, S., & Hanemann, W. M. (2011). Does marginal price matter? A regression discontinuity approach to estimating water demand. *Journal of Environmental Economics and Management*, *61*(2), 198–212.

Nieswiadomy, M. L., & Molina, D. J. (1988). Urban water demand estimates under increasing block rates. *Growth and Change*, *19*(1), 1–12.

Nieswiadomy, M. L., & Molina, D. J. (1989). Comparing residential water demand estimates under decreasing and increasing block rates using household data. *Land Economics*, *65*(3), 280–289.

Nieswiadomy, M. L., & Molina, D. J. (1991). A note on price perception in water demand models. *Land Economics*, *67*(3), 352–359.

OECD. (n.d.). *OECD Environmental Outlook to 2050: The Consequences of Inaction*. Retrieved from www.oecd.org/environment/outlookto2050

Olmstead, S. M. (2009). Reduced-form versus structural models of water demand under nonlinear prices. *Journal of Business & Economic Statistics*, *27*(1), 84–94.

Olmstead, S. M., Hanemann, W. M., & Stavins, R. N. (2007). Water demand under alternative price structures. *Journal of Environmental Economics and Management*, *54*(2), 181–198.

Parente, P. M. D. C., & Silva, J. M. C. S. (2012). A cautionary note on tests of overidentifying restrictions. *Economics Letters*, *115*(2), 314–317.

Pérez-Urdiales, M., García-Valiñas, M. A., & Martínez-Espiñeira, R. (2016). Responses to changes in domestic water tariff structures: a latent class analysis on household-level data from Granada, Spain. *Environmental and Resource Economics*, *63*(1), 167–191.

Polebitski, A. S., & Palmer, R. N. (2009). Seasonal residential water demand forecasting for census tracts. *Journal of Water Resources Planning and Management*, *136*(1), 27–36.

Price, J. I., Chermak, J. M., & Felardo, J. (2014). Low-flow appliances and household water demand: an evaluation of demand-side management policy in Albuquerque, New Mexico. *Journal of Environmental Management*, *133*, 37–44.

Reynaud, A., Renzetti, S., & Villeneuve, M. (2005). Residential water demand with endogenous pricing: the Canadian case. *Water Resources Research*, *41*(11).

Rinaudo, J.-D., Neverre, N., & Montginoul, M. (2012). Simulating the impact of pricing policies on residential water demand: a Southern France case study. *Water Resources Management*, *26*(7), 2057–2068.

Roibás, D., García-Valiñas, M. Á., & Wall, A. (2007). Measuring welfare losses from interruption and pricing as responses to water shortages: an application to the case of Seville. *Environmental and Resource Economics*, *38*(2), 231–243.

Romano, G., Salvati, N., & Guerrini, A. (2014). Estimating the determinants of residential water demand in Italy. *Water*, *6*(10), 2929–2945.

Schleich, J., & Hillenbrand, T. (2009). Determinants of residential water demand in Germany. *Ecological Economics*, *68*(6), 1756–1769.

Sebri, M. (2014). A meta-analysis of residential water demand studies. *Environment, Development and Sustainability*, *16*(3), 499–520.

Shin, J.-S. (1985). Perception of price when price information is costly: evidence from residential electricity demand. *The Review of Economics and Statistics*, 591–598.

Stock, J. H., & Watson, M. W. (2015). *Introduction to econometrics*.

Strong, A., & Smith, V. K. (2010). Reconsidering the economics of demand analysis with kinked budget constraints. *Land Economics*, *86*(1), 173–190.

Taylor, L. D. (1975). The demand for electricity: a survey. *The Bell Journal of Economics*, 74–110.

Taylor, R. G., McKean, J. R., & Young, R. A. (2004). Alternate price specifications for estimating residential water demand with fixed fees. *Land Economics*, *80*(3), 463–475.

Vásquez Lavín, F. A., Hernandez, J. I., Ponce, R. D., & Orrego, S. A. (2017). Functional forms and price elasticities in a discrete continuous choice model of the residential water demand. *Water Resources Research*, *53*(7), 6296–6311.

Water Conservation Annual Report. (2017). Retrieved from https://www.fcgov.com/utilities/img/site_specific/uploads/WC_annual_report.pdf%0D

Wichman, C. J. (2014). Perceived price in residential water demand: Evidence from a natural experiment. *Journal of Economic Behavior & Organization*, *107*, 308–323.

Wichman, C. J., Taylor, L. O., & Von Haefen, R. H. (2016). Conservation policies: Who responds to price and who responds to prescription? *Journal of Environmental Economics and Management*, *79*, 114–134.

Worthington, A. C., & Hoffman, M. (2008). An empirical survey of residential water demand modelling. *Journal of Economic Surveys*, *22*(5), 842–871.

Yang, Q., Yoder, J., & Brady, M. (2019). *The value of including site metrics in meta-analysis for estimate transfer: Reducing generalization error in water demand estimation*.

Yoo, J., Simonit, S., Kinzig, A. P., & Perrings, C. (2014). Estimating the price elasticity of residential water demand: the case of Phoenix, Arizona. *Applied Economic Perspectives and Policy*, *36*(2), 333–350.

Figures

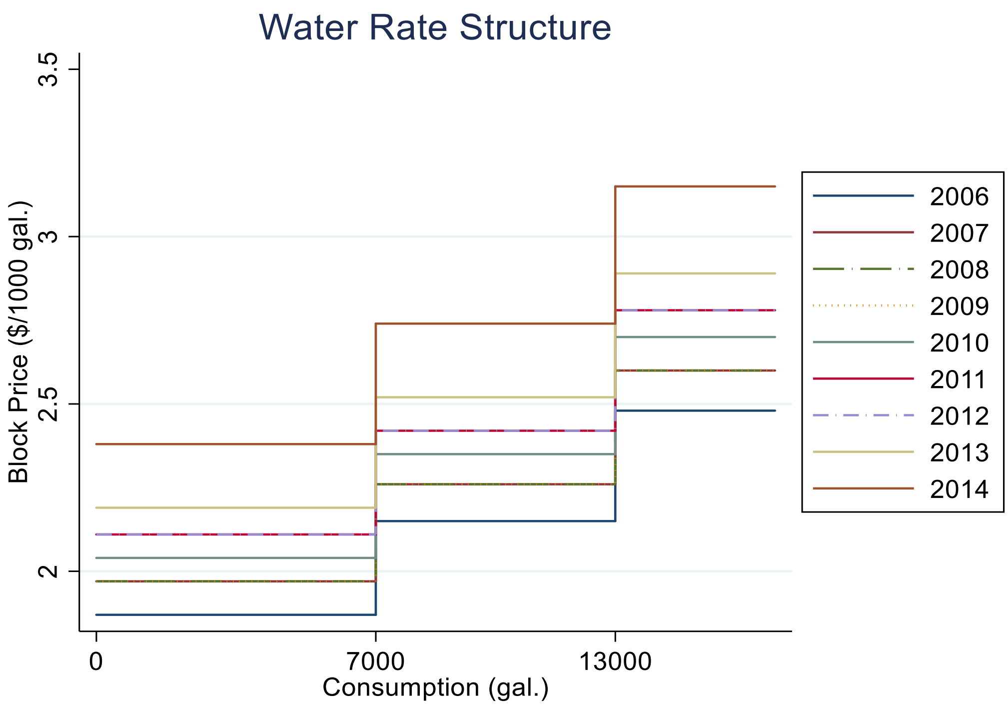

Figure 1: Water Rate Structure (Increasing Block Rate) for the City of Fort Collins

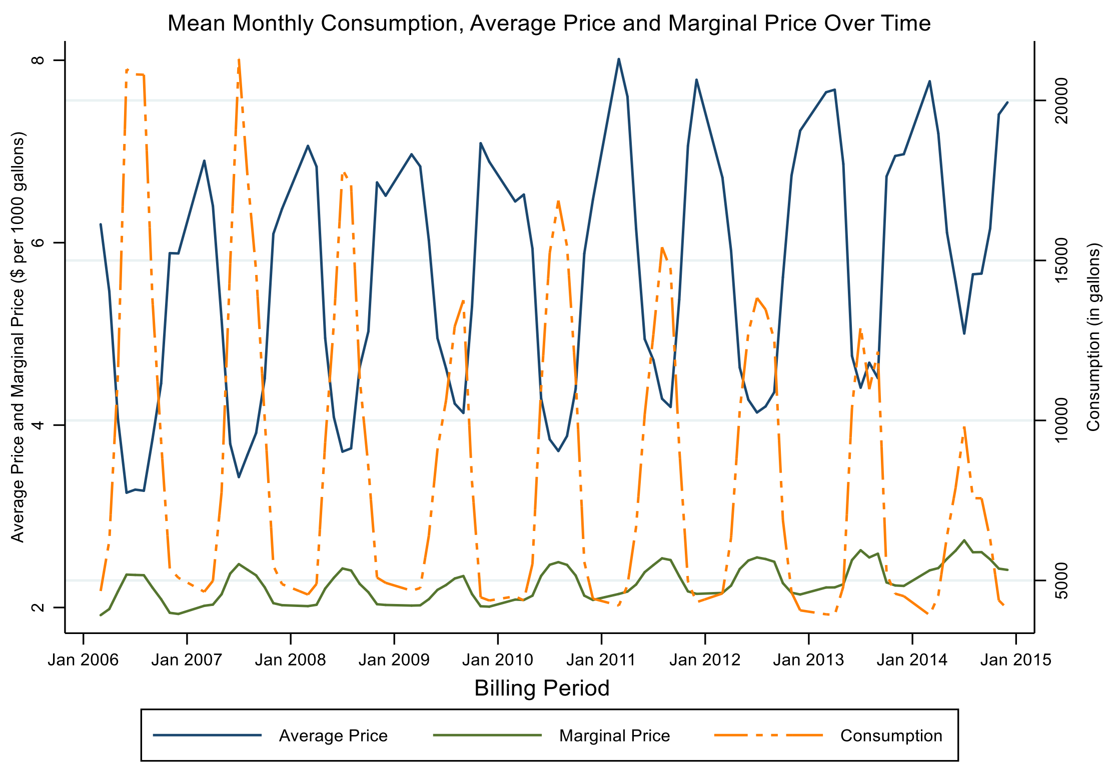

Figure 2: Mean Monthly Water Consumption, Average Price, and Marginal Price of Single-Family Households for the City of Fort Collins over Time

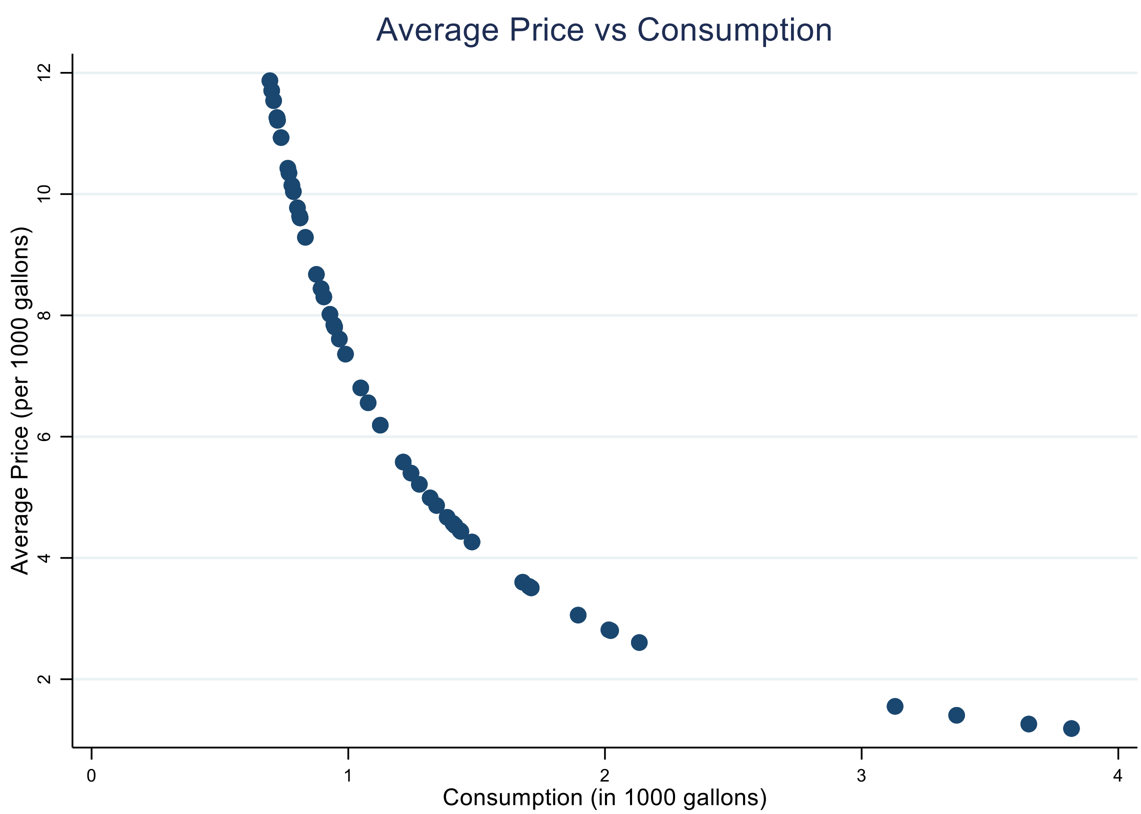

Figure 3: Average Price Vs Consumption

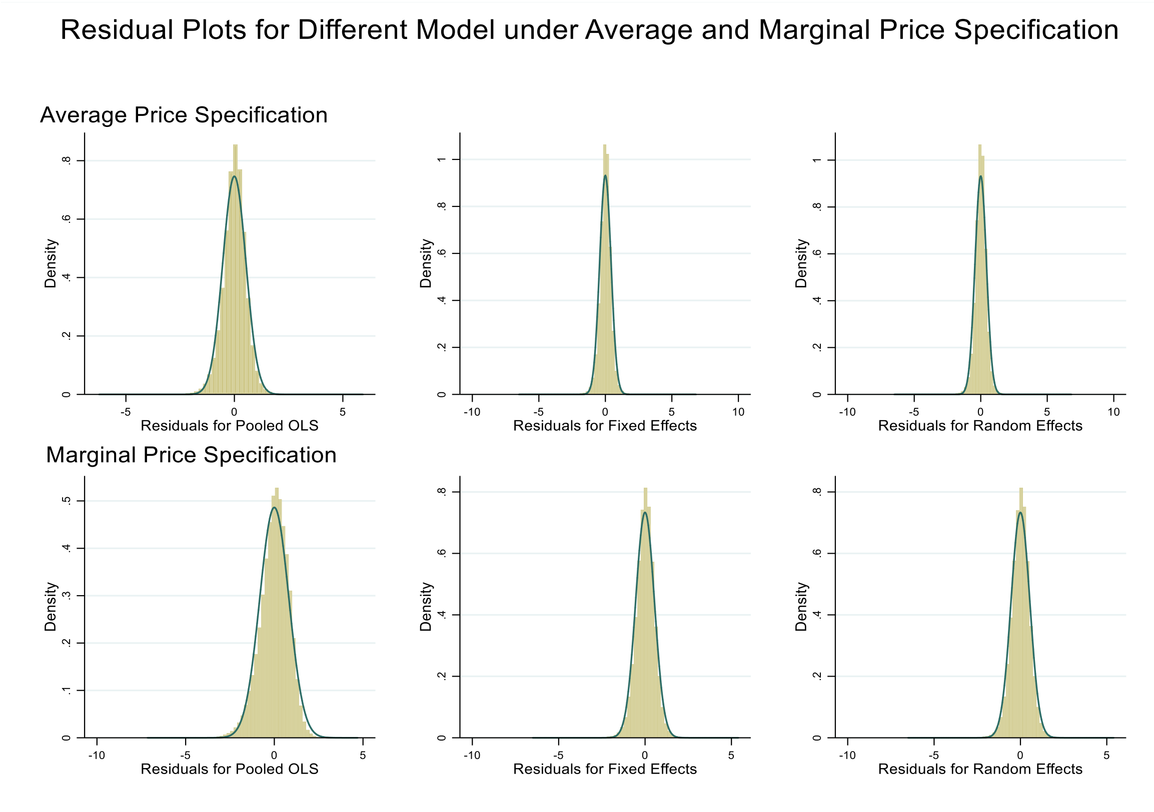

Figure 4: Residuals Plots for Pooled OLS, Fixed Effects, and Random Effects Model under Marginal and Average Price Specification

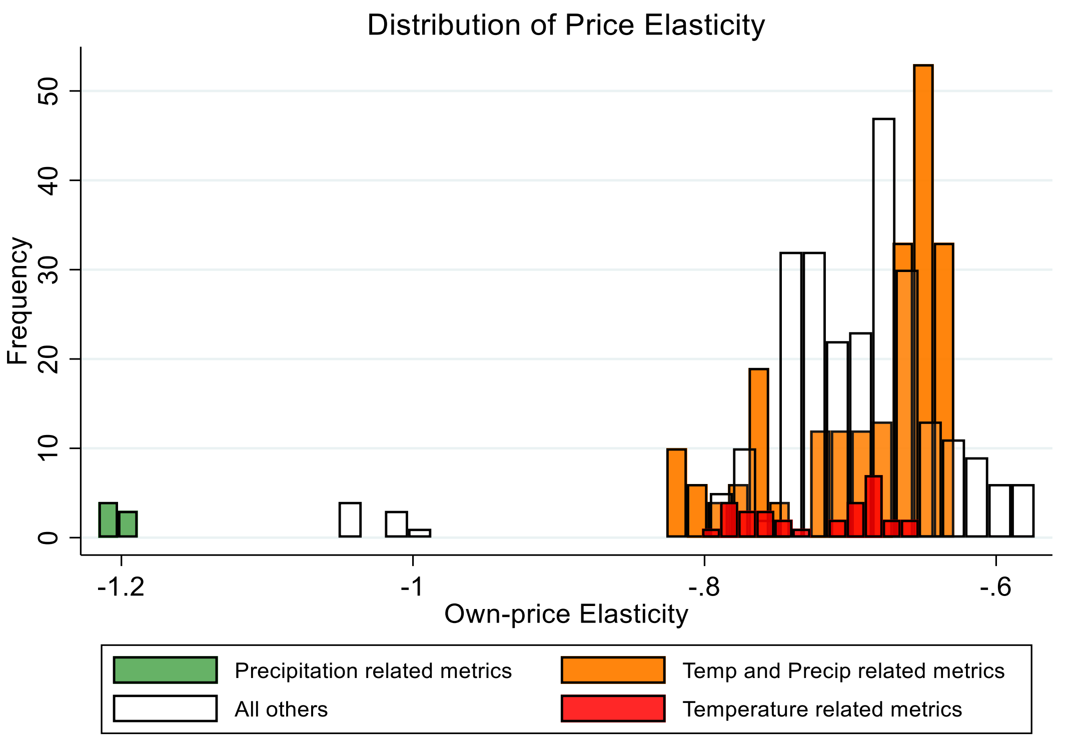

Figure 5: Test Sensitivity to Choice of Weather Variables

Figure 6: Test Sensitivity to Choice of Instruments Variables

Tables

Table 1: Weather Specification across Different Study

Weather Variables	Study
maximum temperature	Martínez-Espiñeira & Nauges (2004)
	Olmstead *et al*. (2007)
	Wichman *et al*. (2016)
minimum temperature	Guhathakurta & Gober (2007)
average temperature	Klaiber *et al*. (2014)
	Price *et al*. (2014)
average precipitation	Vásquez Lavín *et al*. (2017);
	Yoo *et al*., (2014)
number of days with precipitation	Clarke *et al*., (2017)
	Martínez-Espiñeira (2002)
total amount of precipitation	Clarke *et al*. (2017)
	Kenney *et al*. (2008)
	Maas *et al*. (2017)
average relative humidity	Hung *et al*. (2017)
evapotranspiration	Baerenklau *et al.* (2014)
	Wichman (2014)
number of cooling degree days	Lyman (1992)
	Strong & Smith (2010)

Table 2: Descriptive Statistics for Billed Monthly Water Consumption along with the Associated Weather Variables

Variables	Mean	Std. Dev.	Min	Max
Q	9039.83	8367.87	11.00	439820.00
AP	5.50	9.87	2.51	1158.18
MP	2.27	0.28	1.87	3.15
tmax	68.09	14.61	33.38	93.15
tmin	36.83	13.39	8.01	58.90
tave	52.77	14.08	20.97	76.46
rhave	0.55	0.06	0.37	0.69
pp	0.04	0.04	0.00	0.26
ppdays	6.69	3.22	0.00	19.00
totalpp	1.24	1.25	0.00	6.83
cdd	8.87	11.57	0.00	35.00
ET	0.18	0.07	0.03	0.34
FC	13.19	0.59	12.72	14.14
block1	2.06	0.14	1.87	2.38
block2	2.37	0.16	2.15	2.74
block3	2.72	0.18	2.48	3.15
blockdiff1	0.31	0.02	0.28	0.36
blockdiff2	0.35	0.02	0.33	0.41
DOS	30.33	1.95	25.00	35.00
N=1,278,822				
n= 21,874				

Notes: Descriptive statistics are for January to December single-family household water consumption form 2006-2014. Billing records and water rates data were obtained from Fort Collins water utility. Weather variables were obtained from CoAgMet station located at Fort Collins.

Table 3: Description of Variables

Variables	Description	Unit of measurements
Q	monthly household consumption	gallons
AP	Average price	USD per 1000 gallons
MP	marginal price	USD per unit consumption
tmax	maximum daily temperature over a billing period	degree Celsius
tmin	minimum daily temperature over a billing period	degree Celsius
tave	an average of daily mean temperature over a billing period	degree Celsius
rhave	average relative humidity over billing periods	fraction
pp	average precipitation over a billing period	inches
ppdays	number of precipitation days over a billing period	number of days
totalpp	the total amount of precipitation over a billing period	inches
cdd	number of cooling degree days over a billing period	number of days
ET	average evapotranspiration rate over a billing period	inches
FC	fixed charge set by the utility	US Dollars ($)
block1	price in block group one	USD per 1000 gallons
block2	price in block group two	USD per 1000 gallons
block3	price in block group three	USD per 1000 gallons
blockdiff1	the difference in price between block group one and block group two	USD per 1000 gallons
blockdiff2	the difference in price between block group two and block group three	USD per 1000 gallons
DOS	billing length	number of days

Table 4: Results for Shin Test

ln(Q)	Coef.	Robust Std. Err.	z	P>z	[95% Conf. Interval]	
$ln(MP_{it})$	1.13	0.47	-2.40	0.02	-2.05	-0.21
$ln(AP_{it-1}/MP_{it})$	-0.47	0.27	-1.77	0.08	-0.99	0.05
$tmax_t$	0.00	0.02	-0.32	0.75	-0.03	0.03
$tmin_t$	-0.04	0.02	-2.47	0.01	-0.07	-0.01
$tave_t$	0.06	0.03	1.89	0.06	0.00	0.12
$rhave_t$	2.00	0.44	4.50	0.00	1.13	2.88
pp_t	-1.91	2.32	-0.82	0.41	-6.45	2.63
$ppdays_t$	-0.01	0.00	-2.47	0.01	-0.02	0.00
$totalpp_t$	0.03	0.08	0.35	0.72	-0.13	0.19
ET_t	3.94	0.87	4.51	0.00	2.23	5.65
cdd_t	0.01	0.00	2.20	0.03	0.00	0.01
DOS_{it}	0.04	0.00	8.49	0.00	0.03	0.05
Instruments	FC_{it-1}	$block1_{it-1}$	$blockdiff1_{it-1}$	$blockdiff2_{it-1}$	DOS_{it-1}	
Hansen J (p-value)	0.34					
Centered R^2	0.54					
Observations	1,278,882					
Shin Test (p-value)	0.36					

Notes: This test is performed using two stages least squares IV techniques with FE specification. Robust standard errors are produced clustering the errors on both households and time. Hansen *J*-statistic is for the overidentification test of all instruments. Instruments used are fixed charge (*FC*), marginal price in first block (*block1*), difference between each successive block prices *(blockdiff1* and *blockdiff2)* and days of service *(DOS)*, all lagged by one period.

Table 5: Correlation Matrix among Independent Variables

	AP	*tmax*	*tmin*	*tave*	*cdd*	*ET*	*ppdays*	*totalpp*	*pp*	*rhave*
AP	1.00									
tmax	-0.12	1.00								
tmin	-0.12	0.98	1.00							
tave	-0.12	1.00	0.99	1.00						
cdd	-0.10	0.87	0.87	0.87	1.00					
ET	-0.11	0.85	0.82	0.86	0.73	1.00				
ppdays	-0.02	0.20	0.33	0.26	0.15	0.17	1.00			
totalpp	-0.02	0.23	0.35	0.28	0.16	0.18	0.73	1.00		
pp	-0.02	0.22	0.34	0.27	0.14	0.17	0.72	1.00	1.00	
rhave	0.05	-0.36	-0.24	-0.33	-0.27	-0.59	0.51	0.36	0.36	1.00

Table 6: Average Price Elasticity for Different Choice of Price Instruments with Weak Identification *F*-statistics and Hansen *J*-statistics

Instruments	Average Price Elasticity	Z	Weak ID *F*-stat	Hansen *J*-stat	P-value (*J*-test)
DOS_{it-1}	-0.0865	-0.28	9.53	0.00	.
$blockdiff2_{it-1}$	-0.721^{***}	-8.00	255.60	0.00	.
$blockdiff1_{it-1}$	-0.729^{***}	-8.02	244.80	0.00	.
$block1_{it-1}$	-0.734^{***}	-8.45	315.00	0.00	.
FC_{it-1}	-0.769^{***}	-7.18	117.00	0.00	.
$blockdiff2_{it-1}$, DOS_{it-1}	-0.694^{***}	-8.12	150.90	5.17	0.02
$blockdiff1_{it-1}$, DOS_{it-1}	-0.703^{***}	-8.14	141.80	5.28	0.02
$blockdiff1_{it-1}$, $blockdiff2_{it-1}$	-0.729^{***}	-8.02	149.70	0.17	0.68
$block1_{it-1}$, DOS_{it-1}	-0.708^{***}	-8.56	184.30	5.43	0.02
$block1_{it-1}$, $blockdiff2_{it-1}$	-0.747^{***}	-8.75	176.60	0.99	0.32
$block1_{it-1}$, $blockdiff1_{it-1}$	-0.735^{***}	-8.46	163.00	0.10	0.75
FC_{it-1}, DOS_{it-1}	-0.732^{***}	-7.25	71.91	5.82	0.02
FC_{it-1}, $blockdiff2_{it-1}$	-0.730^{***}	-8.03	145.70	0.72	0.40
FC_{it-1}, $blockdiff1_{it-1}$	-0.735^{***}	-8.06	128.20	0.48	0.49
FC_{it-1}, $block1_{it-1}$	-0.734^{***}	-8.45	173.20	0.47	0.49
$blockdiff1_{it-1}$, $blockdiff2_{it-1}$, DOS_{it-1}	-0.703^{***}	-8.14	120.10	5.45	0.07
$block1_{it-1}$, $blockdiff2_{it-1}$, DOS_{it-1}	-0.721^{***}	-8.84	130.90	7.29	0.03
$block1_{it-1}$, $blockdiff1_{it-1}$, DOS_{it-1}	-0.709^{***}	-8.58	127.90	5.67	0.06
$block1_{it-1}$, $blockdiff1_{it-1}$, $blockdiff2_{it-1}$	-0.747^{***}	-8.70	123.80	1.00	0.61
FC_{it-1}, $blockdiff2_{it-1}$, DOS_{it-1}	-0.703^{***}	-8.14	112.40	6.22	0.04
FC_{it-1}, $blockdiff1_{it-1}$, DOS_{it-1}	-0.709^{***}	-8.17	103.10	6.04	0.05
FC_{it-1}, $blockdiff1_{it-1}$, $blockdiff2_{it-1}$	-0.736^{***}	-8.05	108.00	0.85	0.65
FC_{it-1}, $block1_{it-1}$, DOS_{it-1}	-0.708^{***}	-8.56	126.60	5.99	0.05
FC_{it-1}, $block1_{it-1}$, $blockdiff2$	-0.744^{***}	-8.83	134.00	1.12	0.57
FC_{it-1}, $block1_{it-1}$, $blockdiff1$	-0.734^{***}	-8.48	125.70	0.49	0.78
$block1_{it-1}$, $blockdiff1_{it-1}$, $blockdiff2_{it-1}$, DOS_{it-1}	-0.721^{***}	-8.81	104.20	7.29	0.06

FC_{it-1}, $blockdiff1_{it-1}$, $blockdiff2_{it-1}$, DOS_{it-1}	-0.710***	-8.16	94.77	6.45	0.09
FC_{it-1}, $block1_{it-1}$, $blockdiff2_{it-1}$, DOS_{it-1}	-0.719***	-8.92	104.00	7.37	0.06
FC_{it-1}, $block1_{it-1}$, $blockdiff1_{it-1}$, DOS_{it-1}	-0.708***	-8.60	103.30	6.07	0.11
FC_{it-1}, $block1_{it-1}$, $blockdiff1_{it-1}$, $blockdiff2_{it-1}$	-0.745***	-8.78	111.10	1.13	0.77
FC_{it-1}, $block1_{it-1}$, $blockdiff1_{it-1}$, $blockdiff2_{it-1}$, DOS_{it-1}	-0.720***	-8.88	92.88	7.37	0.12

Notes: The dependent variable is the natural log of consumption. Asterisk (***) indicates the statistical significance at 99% levels. Regressions were run clustering the errors on both households and time. Kleibergen-Paap rk Wald *F*-statistic is represented as weak identification (id) *F*-stat and is reported form the first stage regression. Hansen *J*-statistic is for the overidentification test of all instruments, which does not yield any value when the number of endogenous regressors are equal to the number of instruments. P-values from the *J*-test are for the instrument exogeneity, reported from second stage regression.

Table 7: Regression Table for the Preferred Model

ln(Q)	Coef.	Robust Std. Err.	z	P>z	[95% Conf. Interval]	
AP_{it-1}	-0.74***	0.08	-8.80	0.00	-0.91	-0.58
ET_t	4.66***	0.43	10.74	0.00	3.81	5.51
cdd_t	0.00***	0.00	2.07	0.04	0.00	0.01
$rhave_t$	1.97***	0.35	5.65	0.00	1.28	2.65
$tmin_t$	-0.03***	0.01	-3.07	0.00	-0.05	-0.01
$ppdays_t$	-0.01***	0.00	-2.18	0.03	-0.02	0.00
$tave_t$	0.04***	0.01	3.63	0.00	0.02	0.05
pp_t	-2.38	2.05	-1.16	0.25	-6.40	1.64
$totalpp_t$	0.05	0.07	0.70	0.48	-0.09	0.19
DOS_{it}	0.04***	0.00	8.90	0.00	0.03	0.05
Instrument	$block1_{it-}$					
Weak ID *F*-stat	294.52					
Hansen J (p-value)	. (Equation exactly identified)					
Centered R^2	0.62					
Observations	1,278,882					
Number of clusters:	*Across households*: 21,874					
	Across time: 90					

Notes: This test is performed using two stages least squares IV techniques with FE specification. Asterisk (***) denotes the statistical significance at 99% levels. Robust standard errors are produced clustering the errors on both households and time. Kleibergen-Paap rk Wald *F*-statistic is represented as weak identification (id) *F*-stat. Hansen *J*-statistic is for the overidentification test of all instruments. Instrument used is the price in the first block (*block1*), lagged by one period.

Table 8: Model Validation

Estimates	RMSE
est1	0.536
est2	0.537
est3	0.537
est4	0.537
est5	0.536
est6	0.540
est7	0.536
est8	0.536
est9	0.537
est10	0.537

Note: This table represents the *k*-fold (=10) cross-validation of our model predicting the root mean squared error in each fold. RMSE determines the accuracy of the model, out-of-sample in our case.

Appendices

Appendix 1

Construction of Weather variables

Daily weather data on maximum temperature (*tmax*), minimum temperature (*tmin*), average temperature (*tave*), precipitation (*pp*), evapotranspiration (*ET*), maximum relative humidity (*rhmax*) and minimum relative humidity (*rhmin*) for the City of Fort Collins are pulled from Colorado Agricultural Meteorological station. Variables like average relative humidity (*rhave*) and cooling degree days (*cdd*) are constructed in the following ways:

$$rhave = \frac{rhmax + rhmin}{2}$$

$$cdd = \frac{tmax + tmin}{2} - 18^{\circ C}$$

These variables are averaged out for a different number of days of services i.e., 25 to 35 days (which we considered to be standard) and matched to the water consumption data based on these days.

Total amount of precipitation (*totalpp*) is calculated by adding daily precipitation during the given billing cycle, and number of days with precipitation (*ppdays*) is computed by counting the number of days with any amount of precipitation in the given billing period.

Appendix 2

Validity of the Instruments

The households in our sample are divided into three different block groups based on their consumption, and the effects of the price on consumption are different for each of these block group – for example, β_1^i for the first block group, β_1^j for second block group, and β_1^k for third block group, with $i \neq j \neq k$.[35] Considering an unobservable indicator, ζ_i, ζ_j and ζ_k which sums up to one, our demand model can be written as:

$$\begin{aligned} ln(Q_{it}) = {} & \left[\beta_1^i\zeta_i + \beta_1^j\zeta_j + \beta_1^k\zeta_k\right] ln(price_{it-1}) + \beta_2 tmax_t + \beta_3 tmin_t + \beta_4 tave_t \\ & + \beta_5 rhave_t + \beta_6 pp_t + \beta_7 ppdays_t + \beta_8 totalpp_t + \beta_9 ET_t \\ & + \beta_{10} cdd_t + \alpha_i + \varepsilon_{it} \end{aligned} \tag{12}$$

Our set of instruments behave differently to the consumption level for a different group of households. For instances, instrument $blockprice1$ is correlated with the average price mainly for the households in block one, i.e., $cov(block1, \varepsilon_{it}) = 0$, $cov(block_1, \zeta_i price) \neq 0$, instrument $blockdiff1$ is more correlated with the average price for the households in block group one and two i.e., $cov(blockdiff1, \varepsilon_{it}) = 0$, $cov(blockdiff1, \zeta_i price) \neq 0$, $cov\big(blockdiff1, \zeta_j price\big) \neq 0$, while instrument $blockdiff2$ is correlated with the average price for the households in block group two and three i.e., $cov(blockdiff2, \varepsilon_{it}) = 0$, $cov\big(blockdiff2, \zeta_j price\big) \neq 0$, $cov(blockdiff2, \zeta_k price) \neq 0$.

The following conditions hold under statistical regularity:

$$E[\big(ln(Q_{it}) - \alpha_i - \beta_k X_k - \beta_1^i \, ln(price_{it-1})\big)blockprice1] = 0 \tag{13a}$$

$$E[\big(ln(Q_{it}) - \alpha_i - \beta_k X_k - (\beta_1^i + \beta_1^j) \, ln(price_{it-1})\big)blockdiff1] = 0 \tag{13b}$$

$$E[\big(ln(Q_{it}) - \alpha_i - \beta_k X_k - (\beta_1^j + \beta_1^k) \, ln(price_{it-1})\big)blockdiff2] = 0 \tag{13c}$$

These are the conditions which make our instruments valid. However, *J*-test requires error to be orthogonal to each instrument; and when used together, we see no single coefficient of price that allows the condition of orthogonality – rejecting the null hypothesis to make our

---

[35] We make an argument analogous to Parente & Silva (2012), where the authors elucidated the heterogeneous effect of education on wages which makes overidentification restrictions to be invalid even if each instrument for education is valid individually (see Parente & Silva (2012) for details).

instruments invalid. We can raise a similar argument for every other set of instruments used in **Section 4.3**, whenever the p-values for Hansen *J*-statistic for each overidentified model is less than 0.05. Therefore, the test overidentifying restrictions are invalid, regardless of any set (any linear combination) of price instruments.

Appendix 3

Augmented Model

As evidenced in **Section 4**, running the model across the choice of either weather variables or instrumental variables produces stable estimates of price elasticity. To ensure the robustness of our results, we decide to run an augmented model, which runs together across the choice of both weather and price instruments. We build our model combining and redefining the equations in **Sections 3.2.2** and **3.2.3**. We predict the price from first stage regression using each possible set of instruments which picks each possible combination of weather variables as follows:

$$ln(price_{it-1}) = \eta_0 + \delta_l Z_{it-1} + \eta_k X_t + \rho_i + \gamma_{it} \quad (14a)$$

where Z_{it-1} is every possible set of instrumental variables; X_t is every possible set of weather variable during period *t;* δ_K and η_k are the coefficients of instrumental variables and weather variables that need to be estimated.

We rule out any prediction from the sole use of days of service as a price instrument and define the second stage regression as:

$$ln(Q_{it}) = \beta_1\, ln(\widehat{price_{it-1}}) + \beta_s X_t + \ \alpha_i + \varepsilon_{it} \quad (14b)$$

where β_s are the coefficients of weather variables to be estimated

A total of 15,330 point estimates (30 sets of instrumental variables picking 511 different combinations of weather variables) are used to create a distribution of elasticity (**Figure 7**).

We find the mean price elasticity of water as -0.70 with a standard deviation of 0.08, consistent with our previous estimates in **Section 4.2 and 4.3**. Further, we find 92% of our estimates between the range of between -0.80 to -0.52, which ensures the robustness of our previous estimates (**Figure 5** and **Figure 6**) across the choice of weather and price instruments.

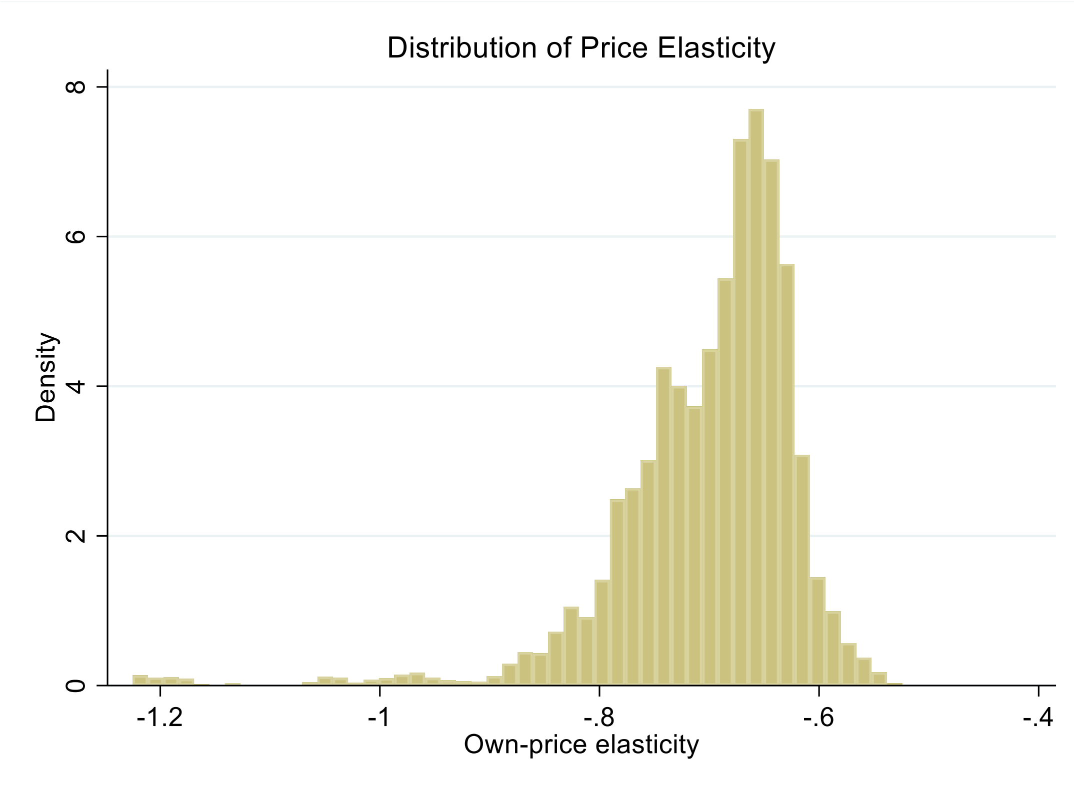

Figure 7: Test Sensitivity to Choice of Both Weather and Instruments Variables Using Augmented Model