Peer Effects in Residential Water Conservation: Evidence from Migration*

Bryan Bollinger
Duke University

Jesse Burkhardt
Colorado State University

Kenneth Gillingham
Yale University

December 14, 2019

Abstract

Social interactions are widely understood to influence consumer decisions in many choice settings. This paper identifies causal peer effects in residential water conservation during the summer using variation from movers. We classify high-resolution remote sensing images to provide evidence that conversions of green landscaping to dry landscaping are a primary determinant of the reductions in water consumption. We also find suggestive evidence that without a price signal, peer effects are muted, indicating a possible complementarity between information and prices. These results inform water use policy in many areas of the world threatened by recurring drought conditions.

Keywords: social interactions; diffusion; information; water policy.
JEL classification codes: R23, Q25, L95.

*Bollinger: Fuqua School of Business, Duke University, 100 Fuqua Drive, Durham, NC 27708, phone: 919-660-7766, e-mail: bryan.bollinger@duke.edu. Burkhardt: Department of Agricultural and Resource Economics, Colorado State University, 1200 Center Ave. Mall, Fort Collins, CO 80523, phone: 503-312-8943, email: jesse.burkhardt@colostate.edu. Gillingham: School of Forestry & Environmental Studies, Department of Economics, School of Management, Yale University, 195 Prospect Street, New Haven, CT 06511, phone: 203-436-5465, e-mail: kenneth.gillingham@yale.edu.

The authors are grateful for conversations and comments from Kerry Smith, Matt Kotchen, Erich Muehlegger, Sheila Olmstead, Michael Hanemann, Soe Myint, and from seminar participants at NBER EEE, Arizona State, Northwestern, and UCLA. All data used and reviewed in this paper is the property of the City of Phoenix and was analyzed in the offices of the City of Phoenix in cooperation with, and for the benefit of, the City

1 Introduction

Social interactions have been shown to play a pivotal role in the diffusion of many new technologies and practices, and undergird classic economic models of technology diffusion (Griliches 1957; Bass 1969; Rogers 1995). The idea that individuals learn from their peers, neighbors, or friends to adopt behaviors or technologies has been explored in settings ranging from agriculture (Foster and Rosenzweig 1995; Conley and Udry 2010) to foreclosures (Towe and Lawley 2013) and schooling (Sacerdote 2001; Graham 2008). In the environmental realm, such 'peer effects' have been shown in the adoption of solar photovoltaic panels (Bollinger and Gillingham 2012; Graziano and Gillingham 2015) and hybrid vehicles (Narayanan and Nair 2013; Heutel and Muehlegger 2015).

This paper is the first to identify causal peer effects in water consumption. Specifically, we use water billing and housing transaction data from over 300,000 households in Phoenix, Arizona to show that a household's water consumption is influenced by the water consumption of nearby households in the previous year. Our identification strategy relies on quasi-experimental variation from movers into the homes surrounding an individual household. We show that a home's water consumption declines when a new household moves into the home, so housing turnover influences the nearby households' average water consumption. At the same time, housing turnover should not otherwise influence an individual household's water consumption after controlling for changes in the housing market and time-varying unobservables. Thus, we use housing turnover in nearby homes as an instrument for the lagged peer water consumption. We find that a one gallon decrease in the mean summer water consumption of households within a 500-foot radius of a home reduces the summer water consumption of that home in the next year by 0.25 gallons.[1]

To better understand this result, we use a machine learning approach on high-resolution remote sensing images to classify the greenness of a household's landscaping. Using these

[1] For simplicity, we define the 'summer' in Phoenix as the six month period from April to September, as this is the period when watering is needed the most.

data, we find evidence that landscape greenness is a primary factor contributing to the peer effects in water consumption. This might be expected if households convert to dry landscaping after observing and/or discussing such a conversion by their neighbors. We further examine the nature of these peer effects by exploiting a natural experiment: some of the Phoenix region has access to extremely low-cost non-potable outdoor water for historical reasons. We find no evidence of peer effects in these areas, while we do find similar peer effects to those in our primary results using a matched set of households in the rest of Phoenix. These results provide suggestive evidence of a complementarity between economic incentives and peer effects that has not yet been noted in the literature.

This work contributes to several literatures. Most directly, it adds new, well-identified evidence to the large and growing literature on peer effects in the diffusion of consumer behaviors. In addition, it adds to the literature on how information transmission, in our case through social learning, can influence consumer decisions about energy and water use. Several papers explore how social norm-based messages aimed at energy conservation can reduce energy use (e.g., Allcott 2011; Allcott and Rogers 2014; Ayres, Raseman and Shih 2012; Costa and Kahn 2013; Dolan and Metcalfe 2015; Gillingham and Tsvetanov 2018; Bollinger and Gillingham 2018) or how prosocial appeals influence energy use relative to economic incentives (e.g., Reiss and White 2008; Ito, Ida and Tanaka 2017; Burkhardt, Gillingham and Kopelle 2017).

Our paper is the first to begin to explore how the effect of social interactions can be directly influenced by economic incentives. Jessoe and Rapson (2014) show that households are three standard deviations more responsive to temporary price increases when provided with high frequency information on electricity usage, and Dolan and Metcalfe (2015) show that the effect of financial incentives can disappear when information on social norms is provided. In the context of residential water demand, Ferraro and Price (2013) show that social comparison messages are the most effective among the least price sensitive households. However, in these studies, information transmission is exogenously manipulated. In contrast to these previous findings, we provide evidence suggesting that social interactions–a phenomenon involving *endogenous* information transmission–may be muted in the absence of a non-negligible price signal.

Our findings can inform active discussions about water policy. Water districts make large upfront infrastructure investments and face continual planning challenges. For example, water is a constrained resource in Phoenix, which is among the fastest-growing and most arid cities in the United States. Between 1980 and 2010, the population of Phoenix increased by approximately 83%, leading to an increase in residential water use of 23%.[2] Yet, even with growing overall water demand in Phoenix, water demand per household has been declining over time, from nearly 230 ccf per year (one ccf is equal to 748 gallons) in 1990 to under 180 in 2014, due to both improved appliance water efficiency and conversion to dry landscaping. This decline in per capita water usage is a positive development for concerns about water availability, but can pose challenges for municipalities that designed water systems for greater demand. For example, in some locations the city of Phoenix has to run clean water through sewer pipes that were overbuilt just to maintain adequate flow.

Thus, understanding the speed and pattern of the diffusion of water conservation activities, such as a transition from green to dry landscaping, is immediately valuable for water planning purposes. An understanding of such diffusion can also enable better targeting of policies. Neighborhoods that begin with some landscape transitions in a given year may be expected to experience contagion from these first transitions, and observe a greater number of transitions in the future relative to other neighborhoods. Policymakers may be interested in one-time targeted information campaigns (Ferraro and Price 2013) or subsidies for dry landscaping for fast-growing areas that are straining the water system (Brelsford and De Bacco 2018), while such efforts may be of less interest for areas that have too much water system capacity.

The remainder of the paper is organized as follows. In the next section we describe our unique dataset, which combines water bills, remote sensing images, and housing transaction data. Section 3 presents the model and discusses our identification strategy. Section

[2]See https://www.biggestuscities.com/city/phoenix-arizona

4 describes the results and mechanisms underlying the peer effect. Section 5 concludes.

2 Data

2.1 Data Sources

The foundation of our analysis is monthly water billing data for all single-family households served by Phoenix Water Services between 2004 and 2012 (308,529 households in the raw data) (Phoenix Water Services 2004-2012*b*). These contain the address of the land parcel and the total monthly water consumption for each household supplied by the Phoenix Water Services. We complement the water billing data with remote sensing images to develop a measure of landscape greenness (Phoenix Water Services 2004-2012*a*). These images, taken once a year in the fall, have resolutions ranging from 0.3 to 0.8 feet. All years from 2004 to 2012 are available at high resolution except 2010, which is only available at a lower resolution. Figure 1 shows a map of Phoenix with the Phoenix Water Services territory, Salt River Project (SRP) territory, and the remote sensing subsample of 71,477 parcels. In SRP territory, households receive potable water from Phoenix Water Services, but can receive nearly-free flood irrigation water from SRP, as we will discuss in section 4.3.

Our next data source is the Maricopa County Assessor's Office, which provided data on housing sales and other physical housing characteristics including pool size, lot size, construction date, home size, and garage size (Maricopa County Assessor 2004-2012). Importantly, we observe the date on which the housing transactions occur and the address of the parcel, allowing us to match these data with the previous data.[3] Our final data source is the U.S. Census, which provides data on a variety of demographic variables (U.S. Census Bureau 2010*b*,*a*).

[3]Summary statistics for physical house characteristics are presented in Table A.1 in the Appendix.

2.2 Data Preparation

2.2.1 Peer Group Definition

A first question in any study on social interactions is how to define the peer group. Defining the peer group membership too broadly could pick up sufficient heterogeneity in the group that it leads to spurious correlations.[4] In our setting, we are interested in how peer effects influence water consumption. While water usage itself may not be visible to peers, landscaping is usually highly visible. This lends itself to a definition of the peer group based on spatial proximity to the household parcel. In a similar setting, Towe and Lawley (2013) define neighbors as the nearest 13 and nearest 25 neighbors by distance. Because there is variation in parcel sizes, we prefer a peer group definition based on geography, as is common in the literature (Topa 2001; Arzaghi and Henderson 2007; Bell and Song 2007; Manchanda, Xie and Youn 2008; McShane, Bradlow and Berger 2012; Narayanan and Nair 2013).

We geocode each address in our data and include all households within a 500-foot radius around each individual household as members of the peer group. This definition includes 25.3 neighbors on average. We also explore different radii in robustness checks. Of course, households can also be expected to have other social groups as well, such as those relating to family, friends, schools, and jobs. So we view our geographic measure as a minimal measure of the social group relevant to water and landscape decisions.

2.2.2 Seasonal Water Consumption Definition

As the focus in this paper is on peer effects in water consumption that may come about through dry landscaping, we focus on water consumption during the hottest six months of the year when irrigation is used the most: April through September. As all of these months have a summer climate, we simplify by referring to these months as "summer months." We refer to the remainder of the year as the "non-summer months."

[4] Indeed, a careful definition of the peer group is central to identification in some studies (e.g., Bertrand, Luttmer and Mullainathan 2000).

2.2.3 Landscaping Data

To convert the remote sensing images to measures of landscape greenness we code each pixel as green landscaping or not. Each image has three color bands (red, green, and blue). There are standard software packages for developing vegetation indices, such as the Normalized Difference Vegetation Index, but our remote sensing images lack the infrared band needed for such standard packages. Therefore, we used a machine learning approach to train the computer to find green pixels through a series of iterations. The approach we used is a supervised maximum likelihood classification routine developed by the company Imagine Software, Inc. (described in more detail in Appendix A.3). To verify the algorithm, we compared the machine learning results to hand-coded results[5] and found that the machine learning approach provided the same coding as the hand-coded pixels in 85-90% of the parcels. The mean greenness based on the machine learning and hand-coding are not statistically different from one another using a simple t-test of differences in means.

Figure 2 provides an example of the output of the classification process. The photo on the left shows several randomly chosen parcels while the photo on the right shows the same parcels with the pixels the computer designates as green landscaping, including tree tops, grasses, and other vegetation, highlighted in green. As is seen in the photo, trees and bright green lawn are coded green. Dry grass is coded (correctly) as not green. It may be hard to see in the photo, but succulents (like cacti) are also not coded as green.

2.3 Summary Statistics

Table 1 presents summary statistics.[6] The average house consumes 16.9 ccf (just under 13,000 gallons) of water per month during the summer months and only 12.3 ccf (about 9,000 gallons) of water per month during the non-summer months. As we are interested

[5]Specifically, a group of interns visually estimated the percentage of turf and greenery in 30,293 parcels and we consider the coding to be accurate if it is within 1-2 percentage points.

[6]Further details on the data cleaning process used to develop the final water and landscaping dataset are included in Appendix A.

in the causal effect of peer water consumption on individual water consumption, we create a variable for the mean water consumption of all neighbors in the peer group (within 500′) in the summer months of the previous year. These summary statistics are very similar to those for the individual households, as would be expected. Roughly 5% of the homes in our sample observe a housing transaction in each year, and not surprisingly, the fraction of houses sold in the peer group (within 500′) is about 0.05. About 40% of our sample is served by SRP.

The landscape sample is a much smaller data sample. The landscaping data summary statistics indicate that the mean fraction of green landscaping is 0.37, with the mean taken over parcels. The summary statistics for the other variables are similar to those in the water data.

Figure 3 shows an important trend in our data: declines in annual water consumption per household over time in Phoenix. This decline may be due in part to reduced water usage from appliances, but may also come about from conversion of green landscaping to dry landscaping.

3 Empirical Strategy

Peer effects are notoriously challenging to identify because the decisions by peers are often endogenous. Put simply, our empirical strategy takes each individual household and uses movers into homes surrounding that household to exogenously shift the average water consumption of all homes surrounding the household (i.e., the peer group). Our identification is facilitated by including a rich set of time-varying fixed effects and housing market characteristics as controls.

In this section, we first present evidence that water consumption (and landscaping) decisions are indeed influenced by movers to motivate our instrumental variables approach for estimating peer effects. Then we present our empirical specification and a discussion of potential identification concerns.

3.1 Evidence on the Effect of Movers on Water Consumption

As our empirical strategy is based on movers influencing the water consumption of peer households, we first examine evidence of the effect of movers into a house on summer water consumption. For all houses that are sold, the mean summer water consumption in the year prior to a move is 15.37 ccf/month, while the mean summer water consumption during the year of a move is 13.41 ccf/month (a difference of 1.96 ccf/month with a t-statistic for a two-sided test of differences in means of 42). For the same group of houses, the mean non-summer water consumption in the year prior to a move is 11.15 ccf/month, while in the year of a move it is 11.35 ccf/month. These statistics from our data suggest that something dramatic is going on in summer water consumption in the year before and after a move.

To examine this more closely, we perform an event study-style analysis, taking the move as the event. We examine how summer water consumption for house i changes in the years proceeding and after a housing transaction, with the following specification:

$$\Delta cons_{it} = \sum_{\tau=-2}^{\tau=2} \beta_\tau 1(transaction_{i\tau}) + \gamma_{tb} + \epsilon_{iy} \tag{1}$$

where $\Delta cons_{it} = cons_{it} - cons_{it-1}$ and $cons_{it}$ is summer water consumption, and $1(transaction_{it})$ is a dummy for whether a housing transaction occurred (i.e., someone new moved into a house). γ_{tb} is a Census block x year fixed effect to capture time-varying unobserved heterogeneity at the Census block level. Figure 4 plots the β_τ coefficients over time (see Appendix B.1 for the regression results).

The results in Figure 4 show that water consumption decreases the year of the sale, and then continues to decrease in the two years after the sale, although by smaller amounts each time, as one would expect. The year prior to the housing transaction, water consumption increases slightly, which is consistent with the owner preparing the house for sale by making sure that plants are well-watered and healthy. In the year of the housing transaction, there is a major decrease in summer water consumption. This may be due in part to the new owner coming in and converting part of a lawn to dry landscaping.[7] We also observe a further decrease in water consumption one year later and then again two years later, which may come about because some households wait a year or two to make the conversion to dry landscaping. By the second year, the change in water consumption from the previous year is much smaller. Taken together, this evidence indicates that having a new household move into a home tends to reduce water consumption and this effect persists.[8]

3.2 Empirical Specification for Peer Effects

Our empirical specification is a classic linear-in-means model, in which the water consumption of household i in year t is given by:

$$cons_{i,t} = \theta \overline{cons}_{i,t-1} + \delta H_{i,t} + \eta_i + \phi_{t,b} + \epsilon_{i,t}, \tag{2}$$

where, $cons_{i,t}$ is the household's water consumption. If we denote household i's peer group (e.g., houses within a 500′ radius) as the set P_i, then $\overline{cons}_{i,t-1} = \frac{1}{|P_i|} \sum_{i' \in P_i} cons_{i',t-1}$ is the average water consumption of the peers, not including household i.[9] $H_{i,t}$ is a vector that includes the average house price in the peer group in t, the change in the average house price in the peer group between t and $t-1$, and the fraction of homes in the peer group that are new construction. η_i contains time-invariant household characteristics, which we model with a separate effect for each parcel x owner combination. $\phi_{t,b}$ captures time-varying factors such as localized economic shocks, gentrification, vegetation shocks such as ash borer infestations, or major new development in a neighborhood, and we model this with Census block x year dummy variables.[10]

We instrument for $\overline{cons}_{i,t-1}$ using the fraction of homes in the peer group that have a

[7] Indeed, when we perform a similar event study-style analysis using the landscaping data, we observe a decrease in greenness in the year of the housing transaction and the following two years (Appendix B.2).

[8] We show the event study results for SRP and non-SRP houses separately in the Appendix.

[9] In the incomplete information framework of Manski (1993), $\overline{cons}_{it}$ would be represented by $\mathbb{E}[cons_{gt}]$, where g refers to the group.

[10] We do not include water prices as a covariate because there is no usable variation in water prices.

housing transaction in year $t-1$, as will be discussed in the next section. We estimate this model *only* on the sample of households that do not move in the current or previous summer (this includes moves during all of year t, and the summer of $t-1$) to avoid any confounding from households moving themselves. Further, we estimate the model in first-differences to difference out η_i, as this requires weaker identification assumptions than demeaning when using a lagged instrumental variable.[11]

3.3 Identification

There are three main categories of concerns in identifying peer effects in a linear-in-means specification (Manski 1993; Brock and Durlaf 2001; Moffitt 2001; Hartmann et al. 2008). The first is 'simultaneity' (sometimes called 'reflection'), which refers to the concern that just as peers may influence a household, the household may influence peers. Our research design addresses simultaneity by using recent, but not contemporaneous, decisions by peers.[12]

The second is 'self-selection' of peers (sometimes described as 'homophily'), which can be an issue if consumers with similar preferences sort into neighborhoods. We address this concern with household fixed effects (e.g., each owner of a house has a separate fixed effect) to capture time-invariant preferences of the household and Census block x year fixed effects for time-varying factors at a fine level of geographic disaggregation. A Census block is an extremely localized area, often covering only a single city block. In our sample, there are 12,485 Census blocks, and each Census block has on average 37 households.[13] These controls are particularly useful because there are frictions in the housing market, such that homebuyers may be able to choose a given broad neighborhood, but

[11]If we demean the data to remove the fixed effects, then we have to assume that all current and future period values of the instrument are uncorrelated with the current period error term (i.e., strong exogeneity), while if we take first-differences, we need only that additionally lagged values of the instruments are uncorrelated with the current period error (i.e., weak exogeneity) (Cameron and Trivedi 2005).

[12]This use of prior peer group decisions to overcome reflection follows several papers in the recent literature (e.g., Towe and Lawley 2013; Bollinger and Gillingham 2012).

[13]We could also use subdivision-by-year fixed effects, which are at a more highly aggregated level. There are 3,930 subdivisions, and each subdivision has on average 135 households.

are very rarely going to be able to choose the exact location of the purchase. For example, there could be some Census blocks where everyone is more liberal and "green," but housing market frictions make it very unlikely that the left side of a street is all environmentalists while the right side is all conservatives. As long as any time-varying sorting into or out of the neighborhood (e.g., due to an ash borer infestation or gentrification on that block) occurs at the Census block level or a greater level of aggregation, then the block-by-year fixed effects nonparametrically control for such sorting.[14]

The third category of concerns is about 'correlated unobservables,' which include the many other factors that may influence both the individual household and peers. For example, if there is an economic downturn facing all households in a neighborhood, their decisions may appear to be aligned, but this alignment is due to the conditions faced by the households, rather than peer effects. Likewise, gentrification may influence whether households change their landscaping to potentially raise the resale value of the home. Our Census block x year and household fixed effects should address most correlated unobservables, but it is possible that some time-varying correlated unobservables work within the Census block group.[15] For example, there could be changes in local amenities, such as the revitalization of a local park, which mean green space is less (or more) important to a small number of houses than it was previously.

To address any correlated unobservables that work within a Census block group, we instrument for lagged peer water consumption using the lagged fraction of movers in the

[14]Bayer, Ross and Topa (2008) make a similar but stronger assumption, arguing that a neighborhood corresponds to a Census block, but that the housing market works at an even higher level of aggregation–at the Census tract level. Our approach allows for sorting at the Census block level, but like Bayer, Ross and Topa (2008), we rely on frictions in the housing market to rule out sorting within our level of geographic aggregation (i.e., within Census block in our case or within Census tract in the case of Bayer, Ross and Topa (2008)).

[15]Using rich fixed effects is a strategy employed by Bayer, Mangum and Roberts (2016) to study housing investment decisions, Towe and Lawley (2013) to study foreclosure decisions in Maryland, and McCartney and Shah (2018) to study the decision to refinance. These papers also use lagged peer group variables. We use more granular fixed effects and controls: Bayer, Mangum and Roberts (2016) uses zip code fixed effects and other controls at larger geographic aggregation, Towe and Lawley (2013) includes controls for housing prices as the Census tract level and county fixed effects, while McCartney and Shah (2018) uses Census block fixed effects.

peer group.[16] For the instrument to be valid, we need the variation in this instrument to be plausibly exogenous. Because housing market frictions make sorting at the block level extremely difficult, the remaining variation in our instrument after including Census block x year fixed effects should be due to individual shocks, such as moves for family reasons or a job. As evidence of these housing market frictions, McCartney and Shah (2018) provide survey evidence indicating that realtors do not field housing requests at the block level, with the exception of new construction of the most expensive homes. Census data from 2014 indicate that roughly 50% of moves are for family or job reasons and we exploit this idiosyncratic variation that should be orthogonal to a neighbor's water consumption.[17]

Of course, differences in annual home sales in Phoenix may be driven by factors such as urbanization, gentrification, housing prices, and new construction. While most of these factors would be expected to be largely picked up by Census block x year fixed effects, we also include peer-group measures of housing prices, changes in housing prices, and new construction in our vector $H_{i,t}$. These additional controls directly address the possibility of a bias from sorting of households due to gentrification or the construction of new dwellings at this finer level of geography.[18]

For there to be a remaining threat to the validity of our instrument, one must believe that people disproportionally move to or from a small radius around a household in a particular year–relative to the rest of that Census block–for factors that *both* directly influence the individual household's water consumption and are not already captured by our controls for gentrification, new home construction, household fixed effects, and Census block-by-year fixed effects.[19]

[16]Our goal is for an exogenous shifter of peer water consumption, consistent with Angrist (2014), who states that this is crucial for a well-identified peer effects study: "Research designs that manipulate peer characteristics in a manner unrelated to individual characteristics provide the most compelling evidence on the nature of social spillovers."

[17]See https://www.census.gov/prod/2014pubs/p20-574.pdf.

[18]Graham (2018) describes the bias that may result if sorting occurs on correlated unobservables that remain even after the extensive use of control variables. In this discussion he also points out that fortunately "sorting into neighborhoods is mediated by the housing market, for which we observe a price."

[19]The common assumption that this does not hold is stated formally in Graham (2018) as the conditions for no sorting or matching on unobservables conditional on predetermined attributes. Graham (2018) further states that such approaches "... have a meaningful role to play in neighborhood-effects research."

A final identification concern in our setting relates to the definition of the peer group as all homes within a 500′ radius around each individual household. This geographic definition allows the radius to cross Census block boundaries, and indeed in just under half of the observations, the peer group does cross the boundaries. This could be a concern because the Census block x year fixed effects may not entirely capture the correlated unobservables that cross boundaries. To address this potential concern, we run a robustness check redefining the peer group by excluding all peers that are in a different Census block than the household.

4 Results

4.1 Peer Effects in Water Consumption

We begin by estimating our primary specification examining peer effects in water consumption (equation (2)). Table 2 presents ordinary least squares (OLS) estimates in columns 1 and 2, and instrumental variables (IV) estimates in columns 3 and 4. Columns 1 and 3 include household and subdivision-by-year fixed effects while columns 2 and 4 include household and Census block-by-year fixed effects. Subdivisions are much larger than Census blocks, with 135 households in a subdivision on average, rather than 37 in a Census block. All specifications include the controls for housing prices and new construction and drop households who moved that year. The IV estimations instrument for the lagged average water consumption of the peers using the lagged fraction of homes in the peer group that have a housing transaction. As one would expect from the evidence in our event study, the instrument is strong, with an F-statistic of over 800 in both columns 3 and 4 (see Appendix C for full first-stage results).

Our preferred specification is in column 4, which show that if the average peer summer water consumption in the previous year decreases by 1 ccf/month, the individual household will decrease summer water use by 0.25 ccf/month (about 187 gallons/month). Recall that the average monthly summer consumption in our data is 16.9 ccf or 12,642 gallons.[20] Our event study provides some further context for this result, by showing that after a move, a household on average reduces water consumption by 1.2 ccf/month in the first year, following by an additional decrease in the second and third years. The sum of the decrease over the three years is approximately 2 ccf/month (about 12% of the average monthly summer consumption). For the sake of comparison, if all peers reduced their water consumption by 2 ccf/month, then our results indicate that the individual household's water consumption over the six-month summer period would decrease by 2,244 gallons. Extrapolated to even 1% of the households in Phoenix, this implies a decrease in water use of nearly 7 million gallons over the summer months. For further comparison, if we use the -0.33 water demand elasticity from Olmstead, Hanemann and Stavins (2007), a 2 ccf/month decrease in water consumption would require an increase in water prices by over 35% to achieve a similar reduction in water consumption.

One question that may arise in interpreting these results is whether we are actually estimating demand-side peer effects. It is possible that there are also supply-side factors that influence the decision to change water use through landscaping decisions. For example, it is possible that landscaping firms undertake focused localized marketing campaigns, such as door-to-door canvassing, that is only in the homes nearby a completed dry landscaping conversion. To provide evidence on whether firm marketing activities may help explain our results, we performed an informal phone survey of landscapers in Phoenix. We did a Google search for "Phoenix landscaper" and called the top 20 landscapers that had a rating of more than three stars. Seven of the landscapers took our call. We learned that none of the companies actively market around an individual installation or use door-to-door canvassing. Four out of the seven firms put signs in the yards of homes during the duration of the landscape conversion, but all remove the signs after the job is completed (See Appendix D for further details on this informal survey). Thus, the signs would not be present a year after the landscape conversion, which means that they should not be a channel explaining our results unless there was a longer-term persistent effect from the short-term display of the signs. We also consider yard signs as another po-

[20] For reference, a typical load of laundry uses 30 gallons of water.

tential contributor of demand-side peer effects, for firms use them to attempt to leverage demand-side peer effects to increase sales by providing information that lowers search costs. Taken together, we view our informal survey findings as evidence that supply-side factors are unlikely to be an explanation for our primary results.

Another interpretation question relates to whether movers affect existing residents' water use through channels other than their own water use. One could imagine movers influencing their neighbors through other types of peer effects (e.g., based on education levels or income levels), in which case the peer effects we capture may be a composite of several types of peer effects. To examine if this is the case, we acquired a cross-section of household-level data on demographics from Acxiom for the landscape subsample in our data (Acxiom 2010). This allows us to examine the demographic characteristics of movers relative to their new peers. We find that on average there is no statistically significant difference between movers and the new peers for the mean of key demographics, such as education, income, housing price, and political affiliation (See Appendix Table A.2). While only suggestive, this evidence is consistent with peer effects occurring due to reductions in water usage by peers, rather than other channels.

4.1.1 Placebo Tests and Robustness Checks

To provide further evidence that our results are well-identified, we run a set of placebo tests. In our first and most important placebo test, we switch the ordering of the timing. In our specifications above, we examined the effect of summer water consumption by peer group households in the previous year ($t-1$) on summer water consumption by the individual household this year (t). In our placebo test, we examine the effect of water consumption by summer peer group households today (t) on summer water consumption by the individual household during the previous year ($t-1$). The only reason that we should find statistically significant results from the effect of peer decisions in t on the household decision in $t-1$ is if there are correlated trends that are influencing both the peers and the individual household. Indeed, it would be physically impossible for such a relationship to be due to peer effects. Thus, if we find a statistically significant effect in our placebo test, that would raise questions about whether there are other unobservable trends influencing our results, rather than actual peer effects.

Table 3 presents the results of our first placebo test. Just as in Table 2, the first two columns present the OLS results, while the second two present the IV results. The first two columns indicate a small but statistically significant relationship between peer group water consumption in t and the household's water consumption in $t-1$. This immediately raises concerns about the identification of peer effects in the OLS specifications, even with the rich set of fixed effects. We view this result as indicating that there is an endogeneity issue, likely due to trends that affect both peer group water consumption and the household's water consumption. On the other hand, the IV results are noisy, but are quite close to zero and show no statistically significant relationship (see Table A.6 for the first-stage results). While this result alone cannot rule out all possible identification concerns, it shows that there is no evidence of unobservable trends confounding identification, further supporting the validity of our primary results. It also highlights the importance of an instrumental variables strategy in identifying peer effects in our setting.[21]

To provide more evidence regarding the validity of our instrument, we run an additional placebo-type test relating to the concern that households may be sorting in a way that might be due to highly localized shocks. If the variation in our instrument–the fraction of households that move in the peer group–is due to highly localized trends that lead to sorting on preferences, we would expect new moves to be clustered. Thus, we use the full sample (including houses that are sold) and regress a dummy for whether the individual house is sold in year t on the fraction of households that moved in the peer group in the previous year, as well as household fixed effects and Census block-by-year dummies. We find a small and statistically insignificant coefficient, which indicates that, after including our controls, the moving process (entry or exit from the neighborhood) does not

[21] We perform further placebo tests by examining the effect of peer group water consumption in t on the household's water consumption in $t-2$ and $t-3$. In both cases, the peer water consumption coefficient is not statistically significant, further confirming our approach.

appear to show clustering (See Appendix Table A.8). We view this as further suggestive evidence that localized trends leading to sorting is unlikely to be a confounding factor in our empirical design.

Table 4 illustrates the robustness of our results to a variety of further checks. In columns 1 and 2, we present the same IV specifications as columns 3 and 4 of Table 2, only we add zip code-specific time trends as further controls to address potentially localized unobserved trends in water consumption. The coefficients are nearly identical to those in Table 2. Columns 3 and 4 are also the same IV specifications only instead of dropping all parcels that had a recent sale, we include the parcels that were sold in the previous year and a half (including the previous summer months) and control for whether there was a transaction with a dummy. Again, the coefficients of interest are identical. These results underscore the robustness of our results to both unobserved trends and the modeling decision we made to exclude parcels that observed a transaction.

In the Appendix, we provide additional robustness checks. We examine specifications that use different definitions of the peer group (Table A.9). We also find no substantial difference in estimated peer effects for larger radii through 700 feet. However, the peer effects become statistically insignificant around 1,000 feet. Next, we limit the peer groups to peers that are both within 500 feet *and* in the same Census block as the household (Table A.10). The reason for this robustness check is that the radius around a household can extend beyond the boundary into another Census block, so the Census block x year fixed effects may not control for sorting and time-vary unobservables. Using this definition of the peer group that does not cross Census block boundaries provides similar results, with slightly larger coefficients, as might be expected when a narrower definition of the peer group is used.

We also explore a set of robustness specifications that are based on a downward 'switch' in water consumption rather than the level of water consumption. Notably, the results are qualitatively similar if we use downward 'switches' in water consumption (Table A.12). We also examine evidence for an asymmetric effect in a specification that is based on an upward 'switch' in water consumption, and we find no evidence of peer effects in increases in water consumption (Table A.14). This finding suggests that the peer effects are strongest for consumer decisions that reduce, rather than increase, water consumption, providing further insight into the mechanisms generating our peer effect results. Finally, we drop all homes that were sold during the sample time period in case the peer effect might be due to the slight increase in water consumption the year before homes are sold. We again find similar results, with just a slightly smaller peer effect coefficient (Table A.15).

Finally, we show that our primary peer effect estimate is robust to including additional lags of peer consumption. The results of including peer consumption in $t-1$ and $t-2$ using OLS and IV are presented in Table A.11. We find that adding peer consumption $t-2$ does not appreciably change the findings for $t-1$. As expected, the effect of peer use in t-2 is also statistically significant, but the magnitude of the effect is much less than the effect in t-1.[22]

4.2 Are the Peer Effects Due to Dry Landscaping?

Our primary results provide strong support for peer effects in water consumption, which is the outcome policymakers care about most. In this section, we examine whether conversion of green landscapes to dry landscaping appears to be a primary driver of these results.

We first examine whether there is evidence of a peer effect in the non-summer months. If outdoor water use is a primary driver of the peer effects in water consumption, we would expect to see little or no effect in the non-summer months. Column 1 of Table 5 presents the same specification as our preferred specification in column 4 of Table 2, only replacing the summer water consumption variables with water consumption variables for the non-summer months. We see that the coefficient on the lagged per water consumption

[22] In column 2, we instrument the peer consumption variables with the fraction of houses sold in the peer group in t-1 and t-2. In column 4, we instrument peer consumption in t-2 with the fraction of houses sold in the peer group in t-2.

in the non-summer months is close to zero and not significant.[23] The evidence of effect in the summer months and not the non-summer months suggests that outdoor water consumption, which is needed much more in the summer months, is a primary driver of the peer effects in water consumption.

We next examine whether the peer effect still holds after controlling for the household's landscape greenness. The idea behind this specification is to isolate the household's outdoor water use from indoor water use. In column 2 of Table 5, we use the landscaping data subsample and run our primary IV specification to confirm that our main result still holds in the subsample.[24] In column 3, we also control for the household's landscape greenness. We find that the effect of landscape greenness is large and significant, as would be expected in Phoenix, where most outdoor plants need substantial watering. The coefficient suggests that increasing landscape greenness by one percentage point increases monthly summer water consumption by 0.62 ccf (463 gallons). Notably, the coefficient on peer consumption is much smaller and is insignificant. This suggests that after controlling for a home's landscape greenness, which determines outdoor water use, we find very little evidence of any remaining peer effects (from non-outdoor use).[25]

Finally, we use the landscape data to examine the trend in landscaping for households that observed a major decrease in summer water consumption between one year and the next (defined as a decrease of at least 2.8 ccf/month that persists at least one more year) and households that did not.[26] Figure 5 shows the trends in the greenness of landscaping over time by these two groups. The mean in the figure is normalized to zero, so the figure can be interpreted as showing relative changes in landscaping over time for the two groups. There is a clear upward trend in greenness residuals for households that did

[23]This finding is robust to our exact specification and even holds for the 'switches' specification we examined for robustness.

[24]See Table A.17 in the Appendix for details on the differences in the landscape subsample and broader water sample.

[25]As a further robustness check, we ran our primary peer effects specification using the landscaping data after correcting for non-classical measurement error in the remote sensing data. The results are reported in A.16 in Section E.3. The results provide further evidence that the peer effect in water consumption is primarily driven by landscaping changes.

[26]2.8 ccf is one half of the average difference in monthly water consumption between the summer months and the non-summer months.

not observe a major decrease in water consumption, and a downward trend in greenness for households that did. This is consistent with the evidence given above and serves as final descriptive evidence that dry landscaping is a primary contributor to the switches in water consumption.[27] It is also useful to note that that after a new household moves into a parcel, the landscape greenness declines as well, consistent with movers converting to dry landscaping, which provided the motivation for our instrument (see Appendix B.2).

4.3 Role of Economic Incentives

We use the Salt River Project's provision of heavily discounted irrigation water for outdoor use as an opportunity to explore whether the price signal for outdoor water influences the strength of the peer effect. Roughly 43% of households in the city of Phoenix Water District are within the boundaries of the SRP. SRP-eligible households pay about $5 per month (which has not changed in recent years) for occasional access to non-potable irrigation water that can be used by households that have constructed berms to direct the water. This is a small fraction of the cost of municipal water. For this setting to be a useful natural experiment, the boundaries of SRP must be plausibly exogenous and we must be confident that there is not sorting into the SRP territory based on water use.

The boundaries of SRP are based on historic water rights boundaries that tend to follow the path of canals built in the late 1800s, which themselves follow the paths of ancient canals built by the Hohokam Indians. These paths were designed to bring Salt River water to irrigated farmland and were based on convenience and historic land rights determined by where settlers created ranches in the 1800s. The areas covered by SRP are of the same general topography, quality, and climate as the rest of the Phoenix basin.[28] Also, it is important to note that today's residential development in Phoenix is not influenced by the

[27]We also regressed the change in landscape greenness on switches in water consumption (and controls such as lot size and year fixed effects and interactions of the two) and find a statistically significant coefficient of 66.5, which indicates that a switch in water consumption is associated with a decrease in irrigated landscaping of 66.5 square feet. This is remarkably close to an estimate by the city of Mesa, AZ of 59 square feet for a reduction of 2.8 ccf/month.

[28]For more on the history of SRP, see https://www.srpnet.com/about/history/StoryofSRP_HistoryBook.pdf.

SRP boundaries. The SRP boundaries cross neighborhoods and it would be difficult to know from looking at a house whether it is covered by SRP without close inspection of the yard and spigots. The distributions of municipal water consumption and landscaping for SRP and non-SRP households are also very similar.[29] SRP irrigation water is also not guaranteed and is not useful for all landscapes, making it unlikely that households sort based on SRP-eligibility rather than more important factors, such as schools, housing quality, and proximity to jobs and amenities.[30]

For our analysis, we matched each SRP-eligible household with a non-eligible household to develop a control group for the SRP-eligible households out of the larger pool of non-eligible households. We explore several matching approaches, but in our preferred specification, we match each SRP household with a non-SRP household using a nearest neighbor approach. We match based on key observables that might influence water consumption: summer water consumption in $t-1$, the lot size, the average house sale price, median household income for the Census tract, the house square footage, the number of bathroom fixtures, and a dummy for whether the household has a pool.[31]

In Table 6, we show the balance of observables between the SRP households and the matched non-SRP households (see Appendix Table A.18 for table of balance comparing SRP households to all unmatched households). The table illustrates that the match is overall very good. While the two groups of houses are not identical, the differences in the observables are small. Several are statistically significant, which is not surprising due

[29]We find that annually, non-SRP houses use approximately 7 ccf (6.6%) more municipal water than on-project houses on average, but this difference is not statistically significant once lot size is controlled for. More broadly, the distributions of water consumption for SRP households and non-SRP households look nearly identical. The distributions of landscaping greenness are also similar, with non-SRP houses having nearly the same landscaping greenness.

[30]SRP water is provided to households based on plot size and is distributed in 45 minute increments. Houses within subdivisions are scheduled to receive water as a group. SRP turns on the water to subdivisions and houses that have been scheduled for service. Houses that are scheduled are required to use the water they have requested. The service comes once a month in the winter months (Oct-March) and twice a month in the summer. There is a "dry up period" for maintenance, which occurs one month out of every year. Subdivisions are affected by the "dry up period," either Dec, Jan, or Feb. Residents could conceivably get SRP water without being scheduled if one of their neighbors had been scheduled to receive water and they left their tap open, but this would be considered water theft.

[31]In Appendix G, we find similar results from matching based on other combinations of variables and using the Mahalanobis distance matching procedure (Rubin 1980).

to the fact that we are matching on so many different variables. In additional robustness checks, we find that regardless of how we match, our results continue to hold. While one can never fully rule out unobservables, this reasonable success in matching and the circumstances of our setting suggest that we are comparing similar households, some of whom are eligible for nearly free water for outdoor use and others who are not.

Our hypothesis is that households that receive heavily discounted outdoor water will appear less susceptible to peer effects in their municipal water consumption. This is because SRP households receive essentially free water several days a month during the summer and at least one day a month during the winter. This means that their marginal price for outdoor water use on these days is basically zero. On all other days, non-SRP and SRP houses pay the same marginal price for outdoor water. Thus, on average, the marginal price (and average price) for SRP houses is lower than for non-SRP houses.

There are a few limitations to this analysis. One important limitation is that we are exploiting a natural experiment and using a matching approach rather than actually randomly varying SRP assignment in an RCT, so we cannot fully rule out sorting on unobservables. Another major limitation of this analysis is that we do not observe SRP-provided outdoor water consumption, but it is important to note that this consumption is limited both in quantity and in the times it is available (which are not always predictable). Thus, to keep green plants alive, SRP-eligible households almost always use a combination of municipal water and SRP-provided water. We observe this in our data, as municipal water consumption is higher in the summer than in the non-summer months for SRP households (17.9 ccf versus 12.4 ccf), just as it is for non-SRP households. Thus, examining how SRP-eligible households respond differently than non-SRP households in their municipal water consumption can provide at least suggestive evidence on the role of economic incentives on peer effects in municipal water use.

Table 7 presents the results of our matching analysis. Each column runs our preferred IV specification from column 4 in Table 2. Column 1 presents the results using only the sample of SRP-eligible households, while column 2 presents the results using the matched sample of households. The coefficient in column 1 suggests no statistically significant peer effects for the SRP-eligible households, while the coefficient for the matched non-SRP households in column 2 is significant and larger than our primary results using the entire sample. Importantly, the peer consumption coefficient in column 1 is smaller (although not significant) than coefficients for the matched non-SRP households, despite a similar sample size in each of the regressions.[32]

While there are limitations to this analysis, these results provide the first evidence we are aware of that suggests economic incentives to reduce outdoor water use are important for the operation of peer effects in water consumption. In neighborhoods where green landscaping is more costly due to a lack of heavily-discounted irrigation water, neighbors may be more likely to discuss xeriscaping (landscaping with slow-growing, drought-tolerant plants) as a money-saving tool. Economic incentives could also influence the peer effects because households are more susceptible to peer effects when they are looking to save money on their water bill, and it is possible that households are responding to the average price of water provision, which could amplify the effect. Neighbors may also be more likely to discuss and share information about dry landscaping when there is an obvious monetary benefit. It is likely that a combination of these reasons can explain our result.

4.4 Implications for Policy and Targeted Interventions

Our primary results provide strong evidence that water consumption decisions made by peers influence a household's water consumption. We provided further evidence that conversions to dry landscaping appear to be a primary driver for the observed peer effects and that economic incentives appear to influence the strength of the peer effects. There are several possible channels for the peer effects we find, including information channels (e.g., social learning that reduces search costs), social norm channels, and combinations

[32]Note that it is possible that SRP-eligible households exhibit peer effects in the SRP irrigation water, although the fact that both municipal water is necessary to keep plants alive during the times when SRP water is unavailable makes this unlikely.

of these. For example, word-of-mouth could provide information that helps households learn the value of dry landscaping and such word-of-mouth could be facilitated by a changing social norm.

From a water district policymaker perspective, it is useful to just know that there are causal peer effects. Regardless of whether peer effects stem entirely from word-of-mouth or a social norm channel, such causal peer effects can lead to the diffusion of lower outdoor water consumption that occurs in a spatial pattern that has ramifications for water provision. Of course, knowing the exact channel may be useful for the design of the optimal policy to facilitate such spatial spillovers. For instance, if the key channel driving the peer effects is an informational channel, providing information may be more cost-effective than simply subsidizing dry landscaping. But subsidies for dry landscaping could still provide a benefit either way.

A key take-away from our results is that policies to promote dry landscaping may have broader effects than might be expected. Consider a dry landscaping subsidy that leads a household to reduce water consumption by 1 ccf/month. Our primary result suggests that all of the peers would together reduce consumption by 0.25 ccf/month–a sizeable spillover.[33] So if a dry landscaping subsidy leads a single household to reduce water consumption by 4.6 ccf/month (i.e., the difference between the average summer consumption and non-summer consumption), then all of that household's peers would reduce water consumption by 1.15 ccf/month or 864.6 gallons/month. This may even be an underestimate of the spillover, for our robustness checks show that if the peer group is defined more broadly to include additional households, we still find an effect on the additional households.

These findings can be put into context by comparing them to the effects of other approaches. For instance, Ferraro and Price (2013) find that information provision reduces water consumption among treated houses by 1% relative to control houses, while the re-

[33]Put differently, the spillover from the single household making a conversion on one of their peers is a 1% water consumption decrease on average (i.e., 0.25 divided by 25.3, which is the average number of houses in a peer group).

duction for social comparisons is 4.8%. Our results imply that if there is a reduction of 4.8% for an individual household, the spillover to neighboring peers would reduce their water consumption on the order of 1.2%, which is remarkably close to the effect of the Ferraro and Price (2013) information provision intervention.

Of course, these estimates deserve several caveats. Most notably, because we are using an IV estimator, our results should be interpreted as a local average treatment effect, and may not apply to the broader population. Similarly, they are an estimate of the average effect, but as we saw in the section above, economic incentives appear to matter. Policymakers may not want to target a dry landscaping program to SRP households, as our results suggest no peer effects in municipal water use for those households. It is also possible that targeting households that have greater economic incentives may lead to even larger peer effects. Another caveat is that our study provides the magnitude of causal peer effects in equilibrium over our sample period. Brock and Durlaf (2001) show that multiple locally stable equilibria may exist when the social interaction effects are sufficiently large and decisionmaking is noncooperative. In addition, if firms or consumers respond differently in a different empirical setting, the estimated effects would be different. However, Phoenix is not only a large desert city itself, but it also has similar water concerns and a similar diffusion of dry landscaping as many other desert cities, such as San Diego, Las Vegas, Tucson, Albuquerque, etc.

5 Conclusions

In this paper, we estimate causal peer effects in residential water consumption using a unique IV strategy that leverages consumer migration into the peer group. Specifically, we study reductions in summer water usage consistent with (often observable) dry landscape transitions. To identify the effects of interest, we exploit within-household and within-Census block-by-year variation after controlling for housing prices and new construction; our key identifying assumption is that the remaining variation in peer housing transactions serves as an exogenous shock to the individual's peer group. We further perform a series of placebo tests and robustness checks that uniformly support the contention that our IV strategy allows us to identify causal effects.

Our primary result is that a 1 ccf/month decrease in the average water consumption in neighboring households within 500′ of a given household reduces the given household's water consumption by 0.25 ccf/month. We then use machine learning techniques on high resolution remote sensing data to develop a robust measure of landscape greenness. Using this measure and several other diagnostics, we provide evidence that suggests a close relationship between outdoor water use, dry landscape adoption, and water consumption. This evidence supports dry landscape adoption as a primary factor that generates our water peer effects results.

One suggestive new finding of this study is that economic incentives appear to complement the effect of social interactions. By exploiting a natural experiment offered by the existence of the Salt River Project's provision of heavily discounted irrigation water and a matching approach, we provide suggestive evidence that the peer effect in municipal water is close to zero and not statistically significant for households eligible for the discounted water. While not dispositive, this finding suggests of the power of economic incentives on the influence of peers.

Our study is relevant because access to water is a major issue in many locations around the world. While the municipal sector uses less water than the agricultural sector in Arizona, municipal water use was approximately 1.47 million acre feet in 2017, which is more than the annual consumption in New York City and approximately 6% of the total volume of Lake Powell at full capacity.[34] Thus, municipal water use in Arizona is not trivial. Moreover, droughts and water scarcity can have serious implications. For example, the California droughts of 2012 and 2015 led to combined economic losses of approximately $5.2 billion, and only a few years earlier, droughts in the Southwest and Midwest led to losses of $20 billion.[35] The U.S. Environmental Protection Agency predicts

[34] Municipal water use was 21% of total water use in Arizona in 2017, with agricultural water use making up most of the remainder. See: `http://www.arizonawaterfacts.com/water-your-facts`.

[35] See `https://www.ama.org/publications/eNewsletters/Marketing-News-Weekly/Pages/economic-loss-us-droughts.aspx` and `http://www.cnbc.com/2015/03/03/`

up to a 40% decrease in snow runoff and soil moisture in parts of the Western United States by 2050, further exacerbating concerns about droughts.[36]

Our results have clear policy implications. The presence of peer effects suggests that policies influencing outdoor water use may have sizable indirect effects. Our results further suggest that targeted policies may be more effective than a uniform policy. For example, the intervention could avoid SRP households and focus on households in areas with a greater economic incentive to reduce water use. Of course, policymakers may be most interested in such targeted policies when there is a strain on parts of the water system from increasing demand or reduced supply. Thus, optimal policy design will inherently involve a consideration of both water district constraints and the potential indirect effects possible in the target audience.

california-drought-seen-having-worsening-3-billion-economic-impact-in-2015.html

[36]See https://19january2017snapshot.epa.gov/climate-impacts/climate-impacts-southwest_.html

References

Acxiom. 2010. “Infobase: Proprietary Dataset of Household Characteristics: Dataset.” Data can be purchased from Acxiom: https://www.acxiom.com/what-we-do/infobase/ (accessed Spring 2014).

Allcott, Hunt, and Todd Rogers. 2014. “The short-run and long-run effects of behavioral interventions: Experimental evidence from energy conservation.” *The American Economic Review*, 104(10): 3003–3037.

Allcott, Huntt. 2011. “Social norms and energy conservation.” *Journal of Public Economics*, 95(9-10): 1082–1095.

Angrist, Joshua D. 2014. “The perils of peer effects.” *Labour Economics*, 30: 98–108.

Arzaghi, M., and J. V. Henderson. 2007. “Networking off Madison Avenue.” *Review of Economic Studies*, 75(4): 1011–1038.

Ayres, Ian, Sophie Raseman, and Alice Shih. 2012. “Evidence from Two Large Field Experiments that Peer Comparison Feedback Can Reduce Residential Energy Usage.” *Journal of Law, Economics, and Organization*, 29(5): 992–1022.

Bass, Frank M. 1969. “A new Product Growth Model for Consumer Durables.” *Management Science*, 15: 215–227.

Bayer, Patrick, Kyle Mangum, and James W. Roberts. 2016. “Speculative Fever: Investor Contagion in the Housing Bubble.”

Bayer, Patrick, Stephen Ross, and Giorgio Topa. 2008. “Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes.” *Journal of Political Economy*, 116(6): 1150–1196.

Bell, David R., and Sangyoung Song. 2007. “Neighborhood effects and trial on the Internet: Evidence from online grocery retailing.” *Quantitative Marketing and Economics*, 5: 361–400.

Bertrand, Marianne, Erzo F. P. Luttmer, and Sendhil Mullainathan. 2000. "Network Effects and Welfare Cultures." *Quarterly Journal of Economics*, 115(3): 1019–1055.

Bollinger, Bryan, and Kenneth Gillingham. 2012. "Peer Effects in the Diffusion of Solar Photovoltaic Panels." *Marketing Science*, 31(6): 900–912.

Bollinger, Bryan, and Kenneth Gillingham. 2018. "Making Pro-Social Social: The Effectiveness of Online and Offline Social Communication."

Brelsford, Christa, and Caterina De Bacco. 2018. "Are 'Water Smart Landscapes' Contagious? An Epidemic Approach on Networks to Study Peer Effects." *Working Paper*.

Brock, William, and Steven Durlaf. 2001. "Discrete Choice with Social Interactions." *Review of Economic Studies*, 68(2): 235–260.

Burkhardt, Jesse, Kenneth Gillingham, and Praveen Kopelle. 2017. "How Do Households Respond to Critical Peak Pricing? Experimental Evidence on the Role of Information and Incentives." *Yale University Working Paper*.

Burkhardt, J., N. Chan, K. Gillingham, and B. Bollinger. 2019. "What is the Value of Conformity? Evidence from Home Landscaping and Water Conservation." *CSU Working Paper*.

Cameron, A. Colin, and Pravin Trivedi. 2005. *Microeconometrics: Methods and Applications.* Cambridge University Press.

Conley, Timothy, and Christopher Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review*, 100(1): 35–69.

Costa, Dora L., and Matthew E. Kahn. 2013. "Energy Conservation "Nudges" and Environmentalist Ideology: Evidence from a Randomized Residential Electricity Field Experiment." *Journal of the European Economic Association*, 11: 680–702.

Dolan, Paul, and Robert Metcalfe. 2015. "Neighbors, knowledge, and nuggets: Two natural field experiments on the role of incentives on energy conservation."

Ferraro, Paul J., and Michael K. Price. 2013. "Using Nonpecuniary Strategies to Influence Behavior: Evidence from a Large-Scale Field Experiment." *The Review of Economics and Statistics*, 95(1): 64–73.

Foster, Andrew, and Mark Rosenzweig. 1995. "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture." *Journal of Political Economy*, 103(6): 1176–1209.

Gillingham, Kenneth, and Tsvetan Tsvetanov. 2018. "Nudging Energy Efficiency Audits: Evidence from a Field Experiment." *Journal of Environmental Economics & Management*, 90: 303–316.

Graham, Bryan. 2008. "Identifying Social Interactions Through Conditional Variance Restrictions." *Econometrica*, 76(3): 643–660.

Graham, Bryan S. 2018. "Identifying and Estimating Neighborhood Effects." *Journal of Economic Literature*, 56(2): 450500.

Graziano, Marcello, and Kenneth Gillingham. 2015. "Spatial Patterns of Solar Photovoltaic System Adoption: The Influence of Neighbors and the Built Environment." *Journal of Economic Geography*, 15(4): 815–839.

Griliches, Zvi. 1957. "Hybrid Corn: An Exploration in the Economics of Technological Change." *Econometrica*, 25(4): 501–522.

Hartmann, Wesley R., Puneet Manchanda, Harikesh Nair, Matthew Bothner, Peter Dodds, David Godes, Kartik Hosanagar, and Catherine Tucker. 2008. "Modeling Social Interactions: Identification, Empirical Methods and Policy Implications." *Marketing Letters*, 19: 287–304.

Heutel, Garth, and Erich Muehlegger. 2015. "Consumer learning and hybrid vehicle adoption." *Environmental and resource economics*, 62(1): 125–161.

Ito, Koichiro, Takanori Ida, and Makoto Tanaka. 2017. "Moral Suasion and Economic Incentives: Field Experimental Evidence from Energy Demand." *American Economic Journal: Economic Policy*, forthcoming.

Jessoe, Katrina, and David Rapson. 2014. "Knowledge is (less) power: Experimental evidence from residential energy use." *The American Economic Review*, 104(4): 1417–1438.

Manchanda, Puneet, Ying Xie, and Nara Youn. 2008. "The Role of Targeted Communication and Contagion in Product Adoption." *Marketing Science*, 27(6): 961–976.

Manski, Charles. 1993. "Identification of Endogenous Social Effects: The Reflection Problem." *Review of Economic Studies*, 60: 531–542.

Maricopa County Assessor. 2004-2012. "Assessor Data for Maricopa County: Dataset." Data can be accessed on site in Phoenix Water Services Building or from Maricopa County Assessor Office. https://www.mcassessor.maricopa.gov/ (accessed Spring 2014).

McCartney, W. Benedict, and Avni Shah. 2018. "I'll Have What Shes Having: Identifying Social Influence in Household Mortgage Decisions."

McShane, Blake, Eric T. Bradlow, and Jonah Berger. 2012. "Visual Influence and Social Groups." *Journal of Marketing Research*, 49(6): 854–871.

Moffitt, Robert. 2001. "Social Dynamics." , ed. Steven Durlauf and Peyton Young, Chapter 3, Policy Interventions, Low-level Equilibria, and Social Interactions, 45–82. MIT Press.

Narayanan, Sridhar, and Harikesh Nair. 2013. "Estimating Causal Installed-Base Effects: A Bias-Correction Approach." *Journal of Marketing Research*, 50(1): 70–94.

Olmstead, Sheila M, W Michael Hanemann, and Robert N Stavins. 2007. "Water demand under alternative price structures." *Journal of Environmental Economics and Management*, 54(2): 181–198.

Phoenix Water Services. 2004-2012*a*. "Remote Sensing Images of Maricopa County: Dataset." Data can be accessed on site in Phoenix Water Services Building (accessed Spring 2014).

Phoenix Water Services. 2004-2012*b*. "Residential Water Consumption for Maricopa County: Dataset." Data can be accessed on site in Phoenix Water Services Building (accessed Spring 2014).

Reiss, Peter C, and Matthew W White. 2008. "What changes energy consumption? Prices and public pressures." *The RAND Journal of Economics*, 39(3): 636–663.

Rogers, Everett. 1995. *Diffusion of Innovations.* The Free Press.

Rubin, Donald. 1980. "Bias Reduction Using Mahalanobis-Metric Matching." *Biometrics*, 36(2): 293–298.

Sacerdote, Bruce. 2001. "Peer Effects with Random Assignment: Results for Dartmouth Roommates." *Quarterly Journal of Economics*, 116: 681–704.

Topa, G. 2001. "Social Interactions, Local Spillovers and Unemployment." *Review of Economic Studies*, 68(2): 261–295.

Towe, Charles, and Chad Lawley. 2013. "The contagion effect of neighboring foreclosures." *American Economic Journal: Economic Policy*, 5(2): 313–335.

U.S. Census Bureau. 2010*a*. "American Community Survey: Income In The Past 12 Months (In 2010 Inflation-Adjusted Dollars): Dataset." Data accessed via American FactFinder. https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml (accessed Spring 2014).

U.S. Census Bureau. 2010*b*. "Decennial Census: Profile of General Population and Housing Characteristics: Dataset." Data accessed via American FactFinder. https://factfinder.census.gov/faces/nav/jsf/pages/index.xhtml (accessed Spring 2014).

Tables & Figures

Table 1: Summary Statistics

Variable	Mean	Std. Dev.	Min.	Max.
Panel A: Water Consumption (N=1,535,545; 260,307 households)				
summer water consumption (ccf/month)	16.9	11.3	0	100
non-summer water consumption (ccf/month)	12.3	8.1	0	99.3
peer summer water consumption in $t-1$ (ccf/month)	16.7	6.2	0.5	82.6
1(housing transaction)	0.05	0.21	0	1
fraction of peer houses sold in $t-1$	0.05	0.05	0	0.8
1(SRP-eligible)	0.40	0.49	0	1
Panel B: Landscape (N=531,650; 71,477 households)				
lot size (ft^2)	9,706	570	1,537	299,200
fraction of green landscape	0.37	0.10	0.19	0.68
1(housing transaction)	0.04	0.20	0	1
fraction of peer households sold in $t-1$	0.04	0.05	0	1
1(SRP-eligible)	0.38	0.49	0	1

Notes: An observation is a household-year. The peer group is defined as households within 500′ of each house. Sample from preferred specification shown here. SRP refers to the Salt River Project. The dummy for housing sales for parcel i is zero because we drop all houses sold in t for our primary specification.

Table 2: Peer Effects in Summer Water Consumption

	(1) OLS	(2) OLS	(3) IV	(4) IV
mean peer consumption in $t-1$	0.19***	0.33***	0.14**	0.25***
	(0.01)	(0.01)	(0.06)	(0.06)
Housing Market Controls	Y	Y	Y	Y
Household Fixed Effects	Y	Y	Y	Y
Subdivision x Year Dummies	Y	N	Y	N
Census Block x Year Dummies	N	Y	N	Y
First Stage F-statistic	N/A	N/A	796	880
R-squared	0.04	0.08	0.04	0.08
N	1,537,435	1,535,545	1,537,435	1,535,545

Notes: The dependent variable in each specification is summer water consumption in t (in ccf/month). An observation is a household parcel-year. The peer group is defined as all houses within a 500′ radius of the household and on average, there are 25.3 houses within a 500′ radius of any household in our study. The 'mean peer consumption in $t-1$' refers to the average peer summer water consumption. Column 1 and 2 present OLS peer effect results. Columns 3 and 4 instrument for peer consumption using the fraction of parcels with housing transactions within 500′ in the previous year. All models are estimated in first differences to difference out the household fixed effects. Housing market controls include the average sales price of homes in the peer group, the change in the price of homes in the peer group, and the fraction of parcels in the peer group that had new construction. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

Table 3: Placebo Tests

	(1) OLS	(2) OLS	(3) IV	(4) IV
mean peer consumption in t	0.20***	0.34***	0.02	0.01
	(0.01)	(0.01)	(0.06)	(0.09)
Housing Market Controls	Y	Y	Y	Y
Household Fixed Effects	Y	Y	Y	Y
Subdivision x Year Dummies	Y	N	Y	N
Census Block x Year Dummies	N	Y	N	Y
First Stage F-statistic	N/A	N/A	191	109
R-squared	0.04	0.08	0.04	0.08
N	1,500,611	1,498,693	1,500,611	1,498,693

Notes: The dependent variable in each specification is summer water consumption in $t-1$ (in ccf/month). An observation is a household parcel-year. The peer group is defined as all houses within a 500′ radius of the household and on average, there are 25.3 houses within a 500′ radius of any household in our study. The 'mean peer consumption in t' refers to the average peer summer water consumption in year t. Column 1 and 2 present OLS peer effect results. Columns 3 and 4 instrument for peer consumption using the fraction of parcels with housing transactions within 500′ in year t. All models are estimated in first differences to difference out the household fixed effects. Housing market controls include the average sales price of homes in the peer group, the change in the price of homes in the peer group, and the fraction of parcels in the peer group that had new construction. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

Table 4: Robustness Checks

	(1)	(2)	(3)	(4)
	Time Trends		Include Sold Homes	
mean peer consumption in $t-1$	0.14**	0.25***	0.14**	0.25***
	(0.06)	(0.06)	(0.06)	(0.06)
Housing Market Controls	Y	Y	Y	Y
Housing Transaction Dummy	N	N	Y	Y
Household Fixed Effects	Y	Y	Y	Y
Subdivision x Year Dummies	Y	N	Y	N
Census Block x Year Dummies	N	Y	N	Y
Zip code-specific Time Trends	Y	Y	N	N
First Stage F-statistic	1,150	1,342	788	893
R-squared	0.03	0.08	0.04	0.08
N	1,537,435	1,535,545	1,624,823	1,623,031

Notes: The dependent variable in each specification is summer water consumption in t (in ccf/month). All specifications instrument for peer consumption using the fraction of parcels with housing transactions within 500′ in the previous year. The specifications that include sold homes include a dummy for whether the parcel had a transaction in the last year. An observation is a household parcel-year. All variable definitions are the same as in Table 2. All models are estimated in first differences to difference out the household fixed effects. Columns 3 and 4 include homes that were sold in the current and previous summer months. Housing market controls include the average sales price of homes in the peer group, the change in the price of homes in the peer group, and the fraction of parcels in the peer group that had new construction. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

Table 5: Is Dry Landscaping a Driver of Peer Effects in Water Consumption?

	(1) Non-Summer Months	(2) Landscape subsample	(3) Landscape added
mean peer consumption in $t-1$	0.02	0.37*	0.16
	(0.18)	(0.19)	(0.15)
landscape greenness			0.62***
			(0.13)
Housing Market Controls	Y	Y	Y
Household Fixed Effects	Y	Y	Y
Census Block x Year Dummies	Y	Y	Y
First Stage F-statistic	1,177	1,165	1,049
R-squared	0.23	0.09	0.09
N	1,545,060	395,048	395,048

Notes: Column 1 uses water consumption in the non-summer months for both the dependent variable and the peer group variable. Column 3 is identical to Column 4 in Table 2, only with the new covariate, which is household i's landscape greenness. Column 2 uses the sample from Column 3 but does not include landscape greenness. All specifications instrument for peer consumption using the fraction of parcels with housing transactions within 500′ in the previous year. An observation is a household parcel-year. All models are estimated in first differences to difference out the household fixed effects. Housing market controls include the average sales price of homes in the peer group, the change in the price of homes in the peer group, and the fraction of parcels in the peer group that had new construction. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

Table 6: Table of Balance for Matched Households

	(1) SRP-eligible	(2) Matched non-SRP	(3) Diff	(4) t-stat
median household income (1,000s $)	45.7	46.1	0.4	6.8
lot size (1000s ft^2)	8.0	8.1	0.19	11.0
average house sale price (1,000s $)	143.6	153.1	9.5	31.1
water consumption (ccf)	15.59	15.99	0.40	10.74
house size (1000s ft^2)	1.56	1.60	0.04	16.99
# bath fixtures	6.05	6.14	0.08	10.13
1(has pool)	0.19	0.18	-0.01	-3.75

Notes: Column 1 reports means for SRP households in the water consumption data with standard deviations in parentheses. Column 2 reports means for the matched non-SRP households, using nearest neighbor matching, in the water consumption data with standard deviations in parentheses. Column 3 reports the difference in means, while column 4 shows the t-statistic for a two-sided test of differences in means. Median HH income refers to the median household income at the Census tract level. There are 133,496 on SRP-eligible houses in the sample and 131,355 matched with frequency weights non-SRP houses in the sample. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

Table 7: Effect of Receiving Discounted SRP Water

	(1) SRP-eligible	(2) Matched non-SRP
mean peer consumption in $t-1$	0.16	0.42***
	(0.10)	(0.11)
Housing Market Controls	Y	Y
Household Fixed Effects	Y	Y
Census Block x Year Dummies	Y	Y
First Stage F-statistic	3026	1956
R-squared	0.15	0.09
N	604,023	689,099

Notes: The dependent variable in each specification is summer water consumption in t. An observation is a household parcel-year. The peer group is defined as all houses within a 500′ radius of the household. The 'mean peer consumption in $t-1$' refers to the average peer summer water consumption in period $t-1$. All specifications instrument for peer consumption using the fraction of parcels with housing transactions within 500′ in the previous year. The 'matched non-SRP' estimations in columns 3 and 4 include only the subsample of matched homes, identified using a nearest neighbor matching routine. All models are estimated in first differences to difference out the household fixed effects. Housing market controls include the average sales price of homes in the peer group, the change in the price of homes in the peer group, and the fraction of parcels in the peer group that had new construction. Standard errors clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

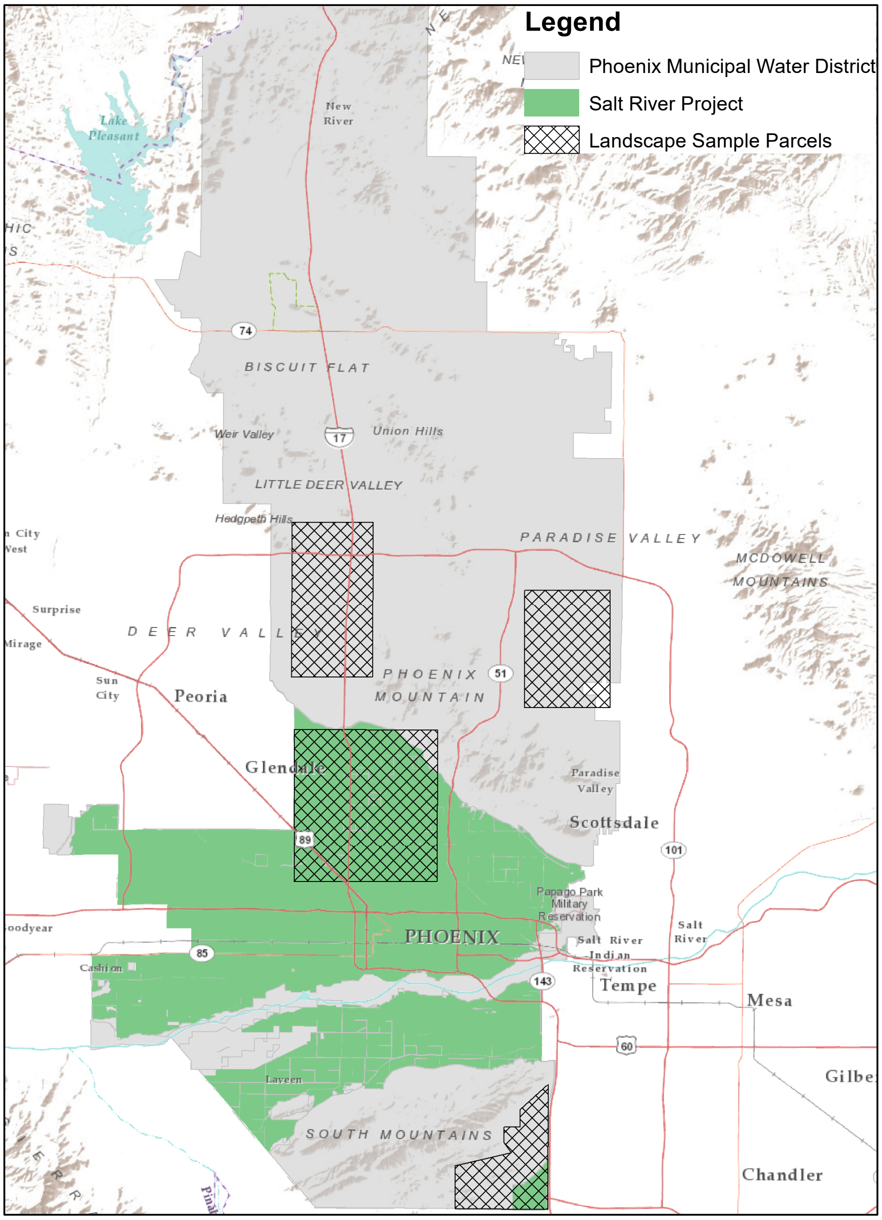

Figure 1: Phoenix water district boundary, along with identification of areas under the Salt River Project and landscape sample parcels.

Figure 2: Illustrative remote sensing images demonstrating the classification of green space. Panel A on the right shows what our remote sensing images look like, while Panel B on the left shows how the machine learning algorithm codes the pixels of green space.

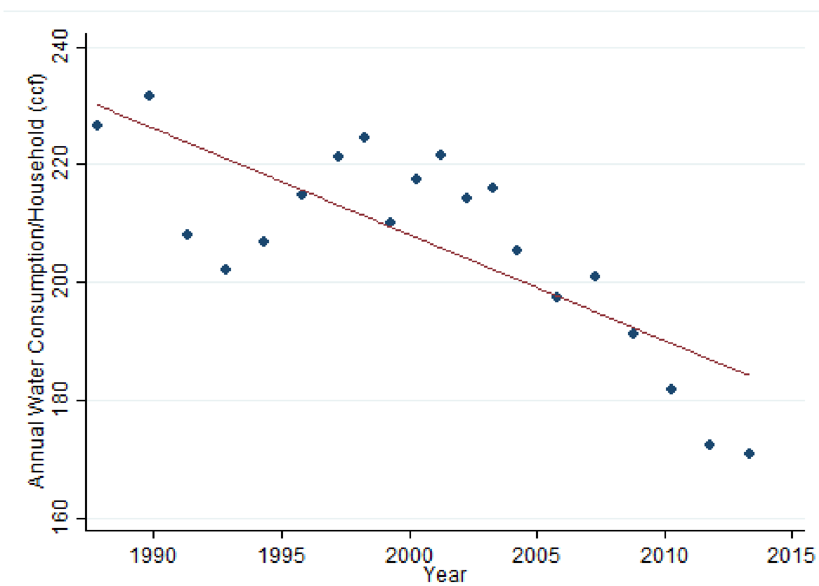

Figure 3: Binscatter plot of the annual water consumption per household served by Phoenix Water Services along with a linear best-fit trendline.

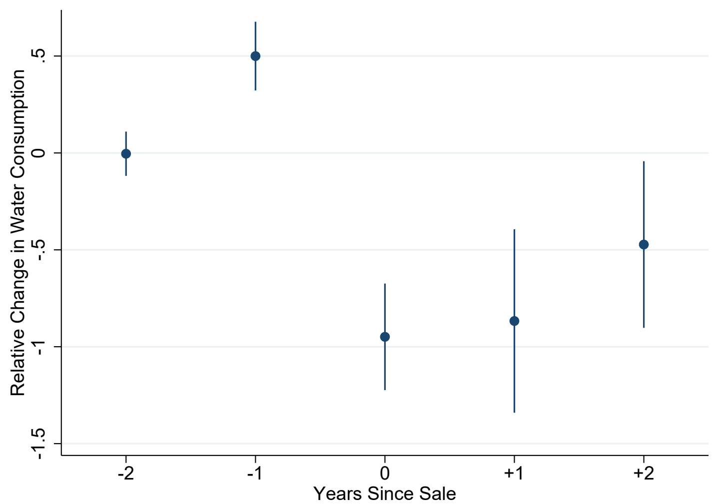

Figure 4: Change in water consumption by year since a housing transaction occurs (Year 0 = year of transaction). Only homes that have a transaction are included. All changes are relative to year -3. See Table A.3 for the full regression results.

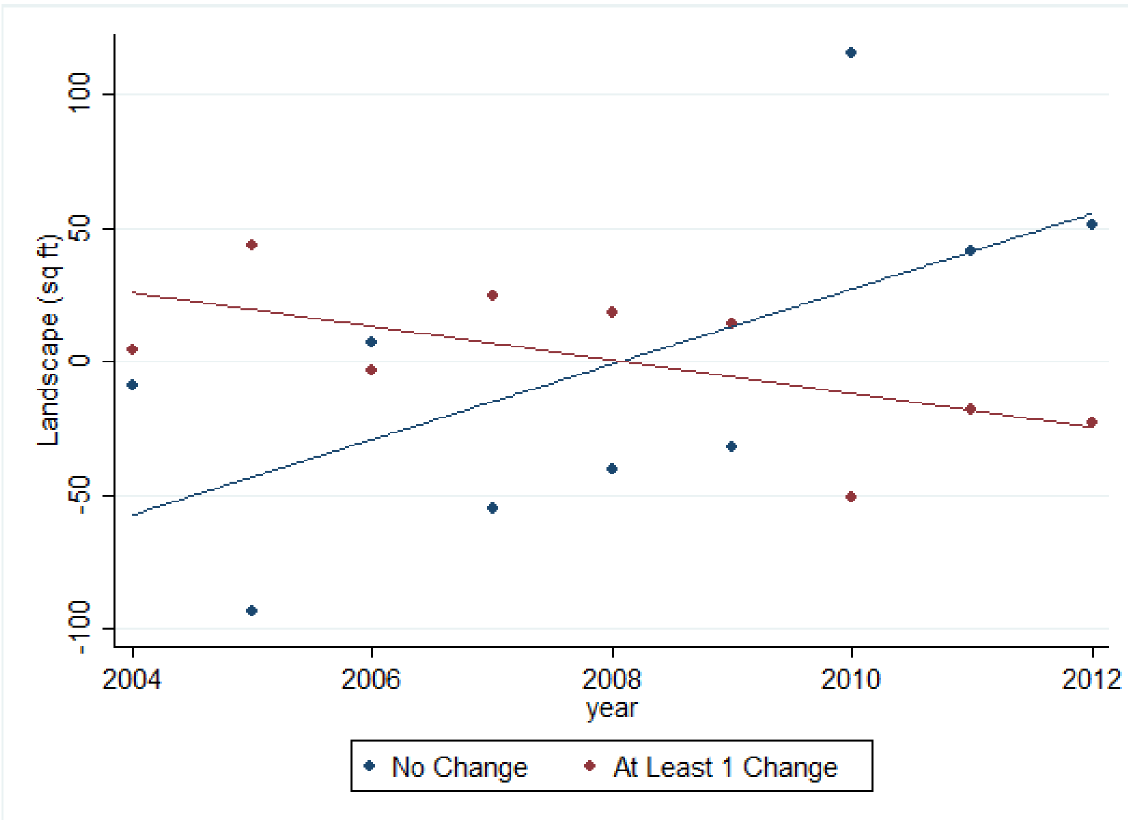

Figure 5: Trends in landscape over time using landscape subsample stratified by whether the parcel exhibits at least one switch in water consumption, where a switch is defined a a decrease of at least 2.8 ccf/month that persists for at least one additional year. To address remote sensing measurement error and allow for comparisons across years, we first regressed the raw data of square footage of green landscaping on a house dummy and year dummy. Each point is a residual, so the mean is normalized to zero. The two lines are best linear fits to the data.

Online Appendix

Peer Effects in Residential Water Conservation: Evidence from Migration

Bryan Bollinger, Jesse Burkhardt, Kenneth Gillingham

A Further Data Details

A.1 Details on Data Cleaning

The raw water consumption data set contains 2,599,862 observations and 308,529 households. Often the utility will make billing adjustments by manually changing a household's water consumption in a particular month to either credit or charge an account. Hence, negative and extremely large consumption values are either errors or billing adjustments and do not reflect actual consumption. Accordingly, we drop annual consumption below the 1st percentile of consumption and above the 99th percentile of consumption, as these are very likely to be outliers that do not reflect actual consumption. For our main specification, we also drop houses that were sold anytime during the current summer months, the current non-summer months, or the lagged summer months (196,925 observations). We also do not have sales price information for all peer groups. For example, if no houses were sold within the peer group for a particular time period. This drop an additional (589,383 observations). Finally, including household-level fixed effects leads to 27,997 singleton observations. The remaining data set, after dropping parcels that had a transaction in the last year, has 1,535,545 observations and 260,307 households.

The raw fitted landscaping data set contains 544,882 observations and 74,112 households. Again, we drop observations below the 1st percentile of consumption and above the 99th percentile of consumption for similar reasons as above. The remaining data set contains 540,451 observations and 72,007 households. Finally, including household-level fixed effects leads to 7,313 singleton observations, resulting in a final data set of 531,650 observations and 71,477 households.

A.2 Further Summary Statistics

This appendix section contains summary statistics for some of the physical household characteristics.

Table A.1: Summary Statistics for Household Characteristics

Variable	Mean	Std. Dev.	Min.	Max.
# bathroom fixtures	6.92	2.32	2	31
house size (ft^2)	1,716	607	288	11,564
pool dummy	0.344	0.475	0	1
lot size (ft^2)	8,148	3,612	531	131,979

Notes: Means are taken over the 130,382 households in the water data.

Table A.2: Means of Movers and Non-moving Peers

	(1) sold house	(2) peer mean	(3) t-statistic
household income (10 brackets)	6.09	6.07	0.38
education level (6 brackets)	1.69	1.68	0.23
1(Democrat)	0.29	0.31	1.17
house price (15 brackets)	10.63	10.58	0.77

Notes: These data on demographics were acquired from Acxiom at the address level. The t-statistic is the statistic for a two-sided test of difference in means.

A.3 Further Details on the Remote Sensing Data Processing

The remote sensing data is three-inch resolution resulting in up to seven terabytes of remote sensing data for all eight years of the sample. To reduce computational burden, we chose image samples from the city, resulting in approximately one terabyte of remote sensing data for the sample period.

To process the images, we used Erdas Imagine software's Supervised Classification routine. This is a common routine used by many in the remote sensing community, including many graduate students in geography. The routine proceeds as follows: the user selects pixels that represent patterns in the images, which are then placed into classes or categories by the user. In our case, we selected pixels that represented green landscaping including grass, trees, or shrubbery as our primary class of interest. The second class, by default, is all other types of land cover. The program uses the mean and covariance matrix of the values of the image bands (e.g., red, blue, green) of the selected pixels to produce a "parametric signature" for the specified classes of pixels. This process produces a training sample that the program then uses to classify pixels in a selected out-of-sample subsample of data. The program uses maximum likelihood to determine the probability that a particular pixel in the out-of-sample data belongs in either class. The user then evaluates the out of sample classification to determine if the training sample effectively classifies green landscaping out-of-sample. If not, then the initial classes are updated and the process is repeated. In this way, the process is iterative. Once, the training sample is deemed appropriate, then the program uses the training sample to classify the remaining pixels in the data. For further details on the approach and for the exact equations used, see `http://geography.middlebury.edu/data/gg1002/Readings/Extras/ERDAS_FieldGuideClassification.pdf`.

B Event Study of the Effect of Housing Transactions on Water Use and Landscaping

This Appendix presents further evidence on how home sales are correlated with changes in summer water consumption and landscape choices, which motivates the first stage of our regression.

B.1 Water Consumption and Movers

We begin by performing an event study-stye analysis on summer water consumption, following equation (1). The key point of this estimation is to see how water consumption changes for households before and after a move. The results from estimating the event study are presented in Table A.3. Column 1 includes data from the year of the move and the year prior to the move. Column 2 includes data from the year of the move, 1 year after the move, and 2 years prior to the move. The omitted category is 2 years prior to the move in column 2. Column 3 includes data from the year of the move, 2 years after the move, and 3 years prior to the move. The omitted category is 3 years prior to the move in column 3.

The results in Table A.3 show that summer water consumption after a home sale decreases relative to a baseline prior to the sale. This decrease continues in the years after the sale. We also see a small increases in summer water consumption the year prior to sale, which is consistent with home owners watering lawns and greenery to increase the visual aesthetic of a home on the market. The results in column 3 of Table A.3 are used to create Figure 4.

B.2 Landscaping and Movers

For landscaping, we first regress the square footage of green landscaping on a year fixed effect to at least partly remove intertemporal measurement error in the landscape data. We use the residuals from this regression as our measure of green landscaping for our

Table A.3: Change in Water Consumption Before and After a Move

	(1)	(2)	(3)
2 years prior to sale			-0.00
			(0.06)
1 year prior to sale		0.46***	0.50***
		(0.09)	(0.09)
year of sale	-2.93***	-1.28***	-0.95***
	(0.12)	(0.15)	(0.14)
1 year after sale		-1.06***	-0.87***
		(0.30)	(0.24)
2 years after sale			-0.47**
			(0.22)
R-squared	0.69	0.64	0.48
N	43,715	66,282	148,630

Notes: This table presents the coefficients showing the change in summer water consumption from estimating the model in (1). The omitted year is always one year less than the earliest year presented. For example, the omitted year in column 2 is 2 years prior to the home sale and the omitted year in column 3 is 3 years prior to the home sale. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

event study, which we call $\Delta landscape_{it}$. We then estimate Equation (1) using $\Delta landscape_{it}$ as the dependent variable. The results of the landscaping event study are presented in Table A.4. The coefficients indicate statistically significant declines in green landscaping the year of the home sale and for up to 2 years after the home sale. However, unlike Table A.3, we do not find statistically significant increases in green landscaping the year prior to the move. While only suggestive, this may indicate that some of the increase in water consumption prior to the move may be caused by non-landscaping uses of water consumption, such as increasing the number of watered indoor plants.

Table A.4: Change in Green Landscaping Before and After a Move

	(1)	(2)	(3)
2 years prior to sale			-15.05
			(18.83)
1 year prior to sale		7.66	-12.49
		(18.69)	(19.81)
year of sale	-49.21**	-50.90**	-47.57**
	(21.96)	(19.58)	(19.29)
1 year after sale		-37.37**	-55.92***
		(17.27)	(19.33)
2 years after sale			-48.40***
			(13.70)
R-squared	0.11	0.11	0.11
N	14,508	29,699	37,300

Notes: This table presents the coefficients showing the change in summer water consumption from estimating the model in (1), only with water consumption replaced by green landscaping. The omitted year is always one year less than the earliest year presented. For example, the omitted year in column 2 is 2 years prior to the home sale and the omitted year in column 3 is 3 years prior to the home sale. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

We can quickly visualize the results in column 3 of Table A.4 in Figure A.1.

Finally, we show the water event study figures for SPR houses and non-SRP houses in Figures A.2 and A.3 respectively. These two figures provide evidence that the first stage is strong for both SRP and non-SRP houses.

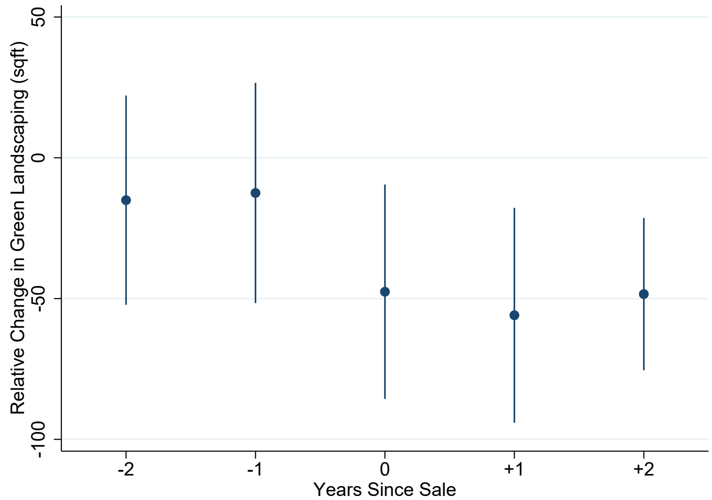

Figure A.1: Change in landscaping by year since a housing transaction occurs (Year 0 = year of transaction). Only homes that have a transaction are included. All changes are relative to year -3. See Table A.4 for the full regression results.

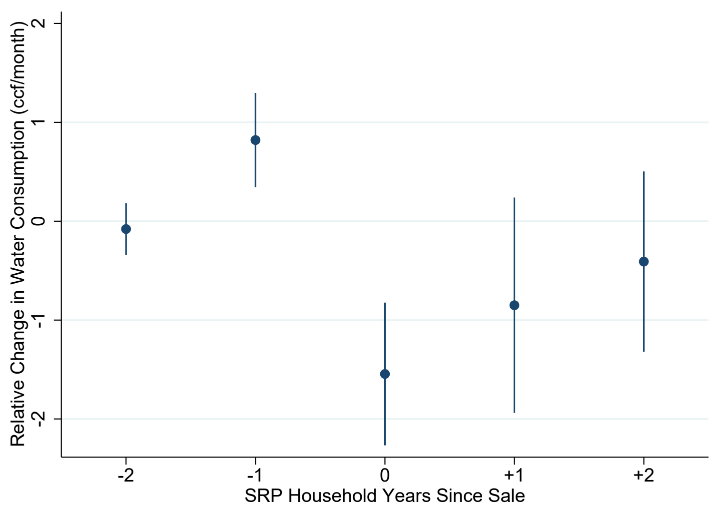

Figure A.2: Change in water consumption for SRP houses only, by year since a housing transaction occurs (Year 0 = year of transaction). Only homes that have a transaction are included. All changes are relative to year -3.

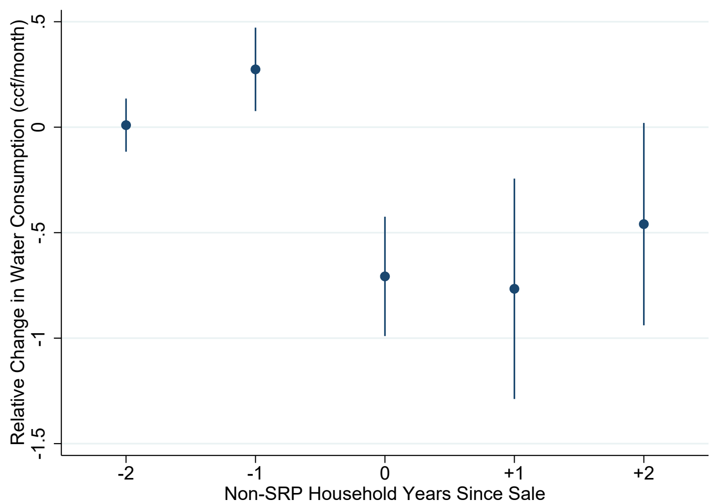

Figure A.3: Change in water consumption for non-SRP houses only, by year since a housing transaction occurs (Year 0 = year of transaction). Only homes that have a transaction are included. All changes are relative to year -3.

C First Stage Results

This appendix provides the results from the first stage of the IV specifications in our results.

Table A.5: First Stage for Preferred IV Specification

	(1)	(2)
mean peer sales in $t-1$	-1.72***	-1.75***
	(0.06)	(0.08)
peer sales prices (1000's \$ /1000ft^2)	39.10***	29.30***
	(5.60)	(5.63)
Δ peer sales prices (1000's \$ /1000ft^2)	-41.71***	-34.88***
	(6.05)	(5.91)
fraction peer new construction	-2.85***	-3.54**
	(1.06)	(1.38)
Household Fixed Effects	Y	Y
Year by Subdivision Dummies	Y	N
Year by Census Block Dummies	N	Y
R-squared	0.54	0.63
N	1,537,435	1,535,545

Notes: This table presents the first stage results from our preferred IV specifications in Table 2. The dependent variable is the mean peer consumption in $t-1$. The independent variable of interest is the fraction of houses sold in the peer group (mean peer sales in $t-1$). Standard errors are clustered at the Census block level. All models estimated in first differences. *** denotes significance at the 1 percent level, ** at the 5 percent level, * at the 10 percent level.

Table A.6: First Stage for Placebo Checks

	(1)	(2)
mean peer sales in t	-2.62***	-2.44***
	(0.09)	(0.12)
peer sales prices (1000's \$ /1000ft^2)	-2.32	8.37
	(5.90)	(6.94)
Δ peer sales prices (1000's \$ /1000ft^2)	14.57***	3.12
	(5.55)	(6.22)
fraction peer new construction	-4.88***	-7.71***
	(1.12)	(1.54)
Household Fixed Effects	Y	Y
Year by Subdivision Dummies	Y	N
Year by Census Block Dummies	N	Y
R-squared	0.53	0.62
N	1,500,611	1,498,693

Notes: This table presents the first stage results from our preferred IV specifications in Table 3. The dependent variable is the mean peer consumption in t. The independent variable of interest is the fraction of houses sold in the peer group (mean peer sales in t). Standard errors are clustered at the Census block level. *** denotes significance at the 1 percent level, ** at the 5 percent level, * at the 10 percent level.

Table A.7: First Stage for Role of Economic Incentives

	(1) SRP-eligible	(2) Matched non-SRP
mean peer sales in $t-1$	-2.19***	-1.51***
	(0.11)	(0.10)
peer sales prices (1000's $ /1000ft^2)	19.03**	39.90***
	(7.80)	(12.48)
Δ peer sales prices (1000's $ /1000ft^2)	-21.53***	-48.15***
	(7.61)	(13.76)
fraction peer new construction	-4.87**	1.98
	(2.34)	(2.18)
Household Fixed Effects	Y	Y
Year by Census Block Dummies	Y	Y
R-squared	0.70	0.62
N	604,244	689,195

Notes: This table presents the first stage results from our preferred IV specifications in Table 7. The dependent variable is the mean peer consumption in $t-1$. The independent variable of interest is the fraction of houses sold in the peer group (mean peer sales in $t-1$). Standard errors are clustered at the Census block level. All models estimated in first differences. *** denotes significance at the 1 percent level, ** at the 5 percent level, * at the 10 percent level.

D Informal Survey of Landscapers

This appendix provides additional details on our informal survey of landscapers in Phoenix. This survey was performed on June 11-13, 2019.

The goal of the phone survey is to determine whether landscapers do actually perform highly localized marketing. Note that conversations with family and friends who live in Phoenix suggested to us that they do not, but the survey is intended to provide clearer evidence of this.

We performed a Google search for "Phoenix landscaper." We called the first 20 of the landscapers that received star ratings above 3 stars. We were able to get 7 out of the 20 landscaping companies on the phone and willing to answer our questions (a response rate of 35%). The landscapers who we were able to speak to were the following: Crystal Green Landscaping, Outside Living Concepts, HMI, Landscaping Services Phoenix, Hawkeye Landscaping Inc, Master Azscapes, and Landscaping Contractors.

We asked the following three questions:

1. When your company performs a landscape conversion on residential properties, not simply maintenance, do you actively market to the nearest neighbors by knocking on doors?

2. When your company performs a major landscape conversion on residential properties, not simply maintenance, do you put up marketing signs in the yard that you are converting?

    - If so, how long do you leave the signs up?

3. Does your company market to people who recently moved into a neighborhood by sending promotional material in the mail?

Our results are the following. None of the firms actively marketed by sending information in the mail or knocking on doors. 4 out of the 7 companies put signs in the yards of the homes they are working on, and the signs remain in place for the duration of the job, and then they are taken down. Note, that if these signs influence neighbors to change their landscaping too, then this would be considered a demand-side factor, as it is a peer making a decision to landscape (and thus have the sign put up) that influences the individual to make the decision. The companies stated they spend most of their marketing funds on websites and online advertising.

Taken together, we see these findings as strongly suggesting that our results are unlikely to be driven by supply-side effects.

E Robustness Checks

This section presents a set of additional robustness checks. The first, in Table A.8, examines whether there is clustering in peer housing transactions, which could suggest sorting affecting our instrument. We regress a dummy for whether a house is sold on the fraction of housing transactions. The coefficient on the fraction of housing transactions is not statistically significant and is close to zero, suggesting that this type of clustering is not a concern.

Table A.8: Clustering in Peer Housing Transactions?

	1(sold)
fraction of housing transactions $t-1$	0.01
	(0.01)
Household Fixed Effects	Y
Census Block x Year Dummies	Y
R-squared	0.06
N	1,535,871

Notes: This table reports the results of regressing a dummy variable for whether house i was sold in period t on the fraction of parcels with housing transactions within 500′ in the previous year, which is our instrumental variable in the primary regressions. The model is estimated in first differences to difference out the household effects. The number of observations is not the same as our primary specification because in our primary specification, we drop houses sold in year t. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

The next robustness check, in Table A.9, examines our primary IV specification as in column 4 of Table 2, only using different definitions of the peer group. Specifically, we use 400′, 600′, and 700′ radii. The results show similar peer effects out to 700′, but the effects slightly weaken as the radii become larger, as one would expect. The peer effect coefficient does not become statistically insignificant until we go our to 1000′.

In Table A.10, we redefined the peer group to drop peers in other Census blocks. After this redefinition, the average peer group has 13 households, rather than 25.3 households. We find a similar peer effect coefficient, although the coefficient is even larger. This might be expected because this peer group definition is capturing closer peers who are more likely to interact with the household.

Table A.9: Robustness: Different Radii

	(1) 400′	(2) 600′	(3) 700′
mean peer consumption in $t-1$	0.26**	0.22***	0.21***
	(0.06)	(0.06)	(0.06)
Housing price controls	Y	Y	Y
New construction controls	Y	Y	Y
Household Fixed Effects	Y	Y	Y
Census Block x Year Dummies	Y	Y	Y
First Stage F-stat	4792	5602	5734
R-squared	0.07	0.07	0.07
N	1,571,039	1,672,005	1,686,514

Notes: All columns run our preferred IV specification using the fraction of movers in the peer group as the instrument for the peer group variable. Each column uses a different radius for the peer group definition. On average, there are respectively 22.9, 26.25, and 26.7 houses within a 400, 600, and 700 foot radius of any household in our study. The models are estimated in first differences to difference out the household effects. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

Table A.10: Redefined Peer Group to Drop Peers in Other Blocks

	(1) OLS	(2) IV
mean peer consumption in $t-1$	0.35***	0.33***
	(0.01)	(0.05)
Household Fixed Effects	Y	Y
Year by Census Block Dummies	Y	Y
First Stage F-statistic	N/A	4762
R-squared	0.08	0.08
N	1534843	1534843

Notes: This table replicates columns 2 and 4 of Table 2, only we redefined the peer group to drop peers in other Census blocks. Standard errors are clustered at the Census block level. The models are estimated in first differences to difference out the household effects. *** denotes significance at the 1 percent level, ** at the 5 percent level, * at the 10 percent level.

Table A.11: Lagged Peer Effects

	OLS	IV	OLS	IV
mean peer Δ consumption in $t-1$	0.38***	0.35***		
	(0.01)	(0.06)		
mean peer Δ consumption in $t-2$	0.17***	0.17***	0.06***	0.11*
	(0.01)	(0.06)	(0.01)	(0.07)
Household Fixed Effects	Y	Y	Y	Y
Census Block x Year Dummies	Y	Y	Y	Y
Housing price controls	Y	Y	Y	Y
New construction controls	Y	Y	Y	Y
First Stage F	N/A	2133.258	N/A	3902.078
R-squared	0.077	0.077	0.074	0.074
N	1514990	1514990	1514990	1514990

Notes: This table reports the results of estimating our primary peer effects specification using OLS and IV, but including mean peer Δ consumption in $t-1$ and mean peer Δ consumption in $t-2$. The models are estimated in first differences to difference out the household effects. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

E.1 Robustness: Discrete Switches Specification

We next we examine a robustness check that uses discrete downward switches in water consumption, rather than the level of water consumption. The idea behind the approach is that the peer effects in water consumption are most likely driven by conversions of green landscaping to dry landscaping, so a specification that looks at the relationship between large persistent reductions in summer water consumption by peers and large persistent reductions by an individual household can provide further useful evidence.

An approach based on persistent decreases in summer water consumption raises the question of what threshold to use for the change in water consumption that would be consistent with a change in landscaping. If we use a very high threshold–such as the difference between the average summer and non-summer water consumption–we will miss more modest landscape changes and may also miss changes that occur in parcels with small lot sizes. If we use too low of a threshold, we risk simply picking up noise, rather than real changes in water consumption. Further, if we use a threshold based on a percentage of water consumption or greenness, translating our results to gallons saved would be much more difficult.

These considerations suggest a threshold that is roughly one half the average difference between the summer and non-summer season water consumption, which comes out to be about 2.8 ccf. For reference, a widely used irrigation calculator for Mesa, AZ indicates that 36 gallons per square foot is required in a summer month,[37] so if all of the decrease in water usage is from dry landscaping, 2.8 ccf would imply a switch of just over 50 square feet of irrigated landscaping to dry landscaping. This threshold captures a noticeable change, but still leaves open the possibility for an average household to have multiple switches in the time frame of our data if the household phases in dry landscaping over time. For some large parcels it is even possible to have more than two switches over different seasons.

We should emphasize that it is important that the decrease in water consumption is

[37] http://apps.mesaaz.gov/watercalculator/

persistent. We would not want to classify households as making a switch if they decrease their water consumption in one season and then increase it the next season, as might be expected when there is mean reversion. Thus, we define a switch in water consumption as a decrease of at least 2.83 ccf in monthly water consumption during the summer months that is persistent through at least the following year. For example, the dummy variable for a switch is equal to one for a household if the household exhibited a decrease in consumption between summers $t-1$ and t that was greater than 2.8 ccf and persisted for at least one more season. An alternative definition we considered requires persistence for multiple seasons, or even until the end of our time frame. This alternative approach is problematic because it treats households making a switch early in our sample differently than households making a switch later in our sample. Thus, our preferred definition considers persistence for one additional season ($t+1$), but we also explore results using the alternative definition to show that mean reversion does not appear to be an issue.

In our sample, 37,098 parcels make more than one switch, 51,581 make a single switch, and the remaining 41,703 make no switches. The largest number of switches in our sample is four, which happens to be for a very small number of parcels with large lot sizes. For households that make one switch, the average difference between the summer and non-summer water consumption per month is 6.6 ccf before the switch and 3.9 ccf after the switch (for the full remainder of the time period in our sample). Similarly, for households that make more than one switch, the average difference between the summer and non-summer water consumption per month is 9.3 ccf before the switch and 5.4 ccf after the switch (for the full remainder of the time period). These statistics indicate that the switches we are modeling are indeed persistent switches and are not simply capturing random variation.

We similarly define our peer decision variable as the average fraction of houses in $i's$ peer group that made a switch between the $t-2$ and $t-1$ summers that persists for at least one more season. Finally, we create a variable for the fraction of houses in $i's$ peer group that were sold between the growing and non-summer months of $t-2$ and $t-1$

(including the non-summer months of $t-1$).

In this 'switches' specification, we model a persistent switch in water consumption during the summer months by household i in year t as a function of the peer group's aggregate choices in $t-1$, peer group housing attributes, time-invariant household characteristics, and time-varying characteristics of the local neighborhood or Census block b:

$$1(\Delta w_{i,t}) = \theta \overline{\Delta w}_{i,t-1} + \delta H_{i,t} + \eta_i + \phi_{t,b} + \epsilon_{i,t}. \tag{3}$$

The term $1(\Delta w_{i,t})$ is a dummy for a persistent switch in summer water consumption.

If we denote household i's peer group as the set P_i, then $\overline{\Delta w}_{i,t-1} = \frac{1}{|P_i|}\sum_{i' \in P_i} 1(\Delta w_{i',t-1})$ is the fraction of household i's peers that complete a major transition in the previous summer, not including household i. $H_{i,t}$ is a vector that includes the average house price in the peer group in t, the change in the average house price in the peer group between t and $t-1$, and the fraction of homes in the peer group that are new construction. η_i contains time-invariant household characteristics, which we model as a household fixed effect (i.e., a fixed effect for each parcel x owner combination, so that there is a different fixed effect after a sale). $\phi_{t,b}$ captures time-varying factors such as localized economic shocks, gentrification, vegetation shocks such as ash borer infestations, or major new development in a neighborhood, and we model this with Census block x year fixed effects.

The results are in Table A.12 and they show clear evidence of a peer effect in terms of downward switches. If there is a larger fraction of peers that make a downward switch, there is also a higher probability of a household making a downward switch. The instruments are strong in this specification, with F-statistics above 700. The placebo tests also hold with this specification, and we find that the specification is robust to the exact choice of the threshold (shown in Table A.13).

One might be concerned that the IV estimates are larger than the OLS estimates in the downward "switches" specification. There are several possible reasons for why this might happen in this particular specification. For instance, consider the possibility of attenuation bias from classical measurement error. Recall that we have very rich fixed effects in our specification, and one often worries more about measurement error with highly disaggregated fixed effects. We expect measurement error to be more problematic when using the downward "switches" specification because this specification converts a continuous variable to a dummy variable. On the other hand, a continuous specification estimated in first differences allows us to use more of the variation in the data, so measurement error would be expected to be less of an issue.

Table A.12: Robustness Check Using Downward Switches in Water Consumption

	(1) OLS	(2) OLS	(3) IV	(4) IV
fraction of peer switches in $t-1$	0.03***	0.06***	0.36**	0.35**
	(0.005)	(0.007)	(0.14)	(0.15)
Housing Market Controls	Y	Y	Y	Y
Household Fixed Effects	Y	Y	Y	Y
Subdivision x Year Dummies	Y	N	Y	N
Census Block x Year Dummies	N	Y	N	Y
First Stage F-statistic	N/A	N/A	796	880
R-squared	0.19	0.22	0.19	0.22
N	1,546,584	1,545,060	1,546,584	1,545,060

Notes: The dependent variable is 1(household persistent switch in water consumption in *t*), where a switch is defined as an average reduction during summer months of at least half the difference between the summer and non-summer consumption that is persistent in the next season. An observation is a household parcel-year. The peer group is defined as all houses within a 500′ radius of the household and on average, there are 25.3 houses within a 500′ radius of any household in our study. The 'fraction of peer switches' refers to the fraction of households in the peer group that make a switch in water consumption in the previous summer. Column 1 and 2 present OLS peer effect results. Columns 3 and 4 instrument for the fraction of peer switches using the fraction of parcels with housing transactions within 500′ in the previous year. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

In this 'switches' specification, we can also examine specifications to help better understand the nature of the peer effects. For example, we can perform the same 'switches' regressions only using non-summer water consumption. Similarly, we can look at upward switches instead of downward switches to see if the peer effects are asymmetric. Finally, we can use the landscaping subsample and add the landscape greenness, as we do in Table 5 to see whether any peer effects remain after landscape greenness is controlled for.

Table A.14 provides this further evidence using the switches specification. In column 1, we see that there is no evidence of a peer effect in non-summer water consumption (indicative of landscaping being a primary driver). In column 2, we see that the peer effect seems to be asymmetric and only applies for decreases in water consumption. This finding is consistent with conversion to dry landscaping being a primary force. In column 3, we see that when landscape greenness is added, the peer effect coefficient becomes statistically insignificant, providing further evidence in support of the findings in Table 5.

Table A.13: Robustness: Alternative Thresholds for Defining a Switch

	(1) 35th	(2) 25th	(3) 10th
fraction of peer switches in $t-1$	0.64***	0.35**	0.26**
	(0.24)	(0.15)	(0.13)
Housing Market Controls	Y	Y	Y
Household Fixed Effects	Y	Y	Y
Census Block x Year Dummies	Y	Y	Y
First Stage F-statistic	430.723	879.3	1514.49
R-squared	0.22	0.21	0.22
N	1,545,060	1,545,060	1,532,692

Notes: The dependent variable is 1(household switch in water consumption in t). Each column presents the results from a specification that is the same as in Table A.12, only with a different threshold for defining a "switch." All specifications instrument for the fraction of peer switches using the fraction of parcels with housing transactions within 500′ in the previous year. An observation is a household parcel-year. All variable definitions are the same as in Table 2. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

Table A.14: Further Evidence on Peer Effects in Switches

	(1) Non-summer usage	(2) Increase in use	(3) Landscape added
fraction of peer switches in $t-1$	0.02	0.08	-0.35
	(0.18)	(0.42)	(0.47)
household landscape greenness			-0.01***
			(0.003)
Housing Marking Controls	Y	Y	Y
Household Fixed Effects	Y	Y	Y
Census Block x Year Dummies	Y	Y	Y
First Stage F-statistic	1177.3	143.6	31.4
R-squared	0.23	0.25	0.14
N	1,545,060	1,545,060	306,480

Notes: Column 1 uses downward persistent switches in water consumption in the non-summer for both the dependent variable and the peer group variable. Column 2 uses increases in water consumption (upward switches) for both the dependent variable and the peer group variable. Column 3 is identical to Column 4 in Table 2, only with the new covariate, which is household i's landscape greenness. All specifications instrument for the fraction of peer switches using the fraction of parcels with housing transactions within 500′ in the previous year. An observation is a household parcel-year. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

E.2 Robustness: Dropping All Sold Homes

In the next robustness check, we address the concern that the changes in summer water consumption prior to the year of sale might suggest that some of the peer effect is due to increases in water consumption prior to the year of sale. To address this concern, we re-estimate our primary model with Census block x year fixed effects using OLS and IV but dropping all houses that were sold during the entire sample period. For reference, in our primary specification, we drop homes that were sold in year t. The results are presented in Table A.15. We find the results do not qualitatively change, although the coefficient is a bit smaller. This might not be surprising because we are looking at an unusually selected sample in this robustness check.

Table A.15: Robustness check dropping all sold homes

	(1) OLS	(2) IV
mean peer Δ consumption in $t-1$	0.31***	0.14**
	(0.01)	(0.07)
Housing Marking Controls	Y	Y
Household Fixed Effects	Y	Y
Census Block x Year Dummies	Y	Y
First Stage F-stat	NA	3659
R-squared	0.09	0.09
N	1,158,216	1,158,216

Notes: This table replicates columns 2 and 4 of Table 2 but dropping homes that were sold at any time during our sample period. The models are estimated in first differences to difference out the household effects. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

E.3 Robustness: Peer Effects in Landscaping

In this short appendix, we present the results directly estimating peer effects in landscaping using our remote sensing data. As mentioned in the main text, we are concerned about the intertemporal measurement error in the remote sensing data, which is a non-classical measurement error because the magnitude of the error is likely correlated with the size of the dependent variable. For instance, larger lots are likely to have more green landscaping simply because they have more space to do so due to the fact that homes only cover so much space. However, larger lots are also likely to contain more measurement error in the remote sensing data because they also have more shadows or tree coverage, which the remote sensing algorithm cannot control for. For this reason, a simple IV approach would not address the non-classical measurement error present in the remote sensing images.

To address the intertemporal measurement error, we use a data correction procedure outlined in section 3.2 of the working paper Burkhardt et al. (2019).[38] Note this working paper is on an entirely different topic (the value of conformity in home prices) and does not include these peer effects results.

In Table A.16 below, we present our preferred peer effects specifications (columns 2 and 4 of Table 2) using the corrected landscaping data. We have far fewer observations than in our main sample because we do not have remote sensing images of all houses in the sample. However, the results provide evidence that there are peer effects in landscaping itself, which further supports our contention that the water consumption peer effects we find (which are the policy-relevant peer effects) can be at least primarily attributed to changes in landscapes.

If we use the uncorrected landscape data, we see positive coefficients similar to the ones in A.16, but our instrument is weak and we do not find statistically significant coefficients in our IV specification.

[38]This paper can be accessed at: https://drive.google.com/file/d/1c6sBbsD5Z3Sb1ZU0TlAVuuD9OXwOk44j/view.

Table A.16: Peer Effects in Landscaping

	(1) OLS	(2) IV
mean peer landscaping in $t-1$	0.32***	0.28**
	(0.01)	(0.12)
Housing market controls	Y	Y
Household Fixed Effects	Y	Y
Census Block x Year Dummies	Y	Y
First Stage F-statistic	N/A	202
R-squared	0.08	0.06
N	1,109,674	1,109,674

Notes: The dependent variable in each specification is corrected landscaping greenness in t. An observation is a household parcel-year. The peer group is defined as all houses within a 500′ radius of the household and on average, there are 25.3 houses with a 500′ radius of any household in our study. The ′mean peer landscaping in $t-1$′ refers to the average peer corrected landscaping. Column 1 presents OLS peer effect results of our preferred specification. Column 2 instruments for peer landscaping using the fraction of parcels with housing transactions within 500′ in the previous year. All models are estimated in first differences. Standard errors are clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

F Additional Tables of Balance

The following table displays comparisons of summary statistics of important observable household and demographic characteristics for key subsets of our data. Table A.17 compares the means of key variables between the water consumption data and the landscaping data. We performed two-sided t-tests of differences in means for each variable and report the standard errors of the differences in means in parentheses below the differences in means in column 3 of the table. Table A.18 presents further summary statistics of key variables for the matched and unmatched samples used in Table 7 in the primary text.

Table A.17: Summary Statistics by Water and Landscape Data Sets

	Water Data	Landscape Data	Difference
median household income (1000s $)	60.0	62.9	-2.9
	(24.1)	(24.0)	(37.2)
lot size (1000s ft^2)	8.1	8.7	-0.6
	(3.6)	(4.1)	(6.0)
house sqft (1000s ft^2)	1.7	1.8	-0.06
	(0.61)	(0.60)	(0.94)
# bath fixtures	6.9	7.1	-0.14
	(2.3)	(2.3)	(0.004)
% white	72.1	75.5	-3.4
	(16.8)	(14.3)	(0.03)
% black	4.8	4.4	0.36
	(4.5)	(3.5)	(0.007)
% latino	33.1	26.3	6.7
	(26.8)	(22.2)	(0.04)
N	260,307	71,477	N/A

Notes: Column 1 reports means for households in the water consumption data with standard deviations in parentheses. Column 2 reports means for households in the landscape data with standard deviations in parentheses. Column 3 reports the difference in means with standard errors of differences in means in parentheses.

Table A.18: Summary Statistics by SRP Status

	(1)	(2)	(3)
		Matched	
	SRP	non-SRP	non-SRP
median household income (1000s $)	45.7	46.1	70.4
average house sales price (1000s $)	143.6	153.1	225.1
water consumption (ccf)	15.59	15.99	17.92
lot size (1000s ft^2)	8.0	8.1	9.5
house size (1000s ft^2)	1.6	1.6	2.0
# bath fixtures	6.05	6.14	7.85
1(has pool)	0.19	0.18	0.43

Notes: Table reports means of variables. There are 133,496 SRP-eligible houses in the sample and 131,355 matched non-SRP houses in the sample.

G Robustness of the SRP Matching

This section performs several robustness checks using different approaches to matching the SRP households to non-SRP households. Recall that our primary table uses nearest-neighbor matching in which each SRP household is matched to a single non-SRP household. We present four additional approaches. The first is the same as our primary specification but using Mahalanobis matching on all the same variables (see Table A.19 for the table of balance). The second uses nearest neighbor matching to match only on Census variables (see Table A.20 for the table of balance). The third uses Mahalanobis matching to match only on variables that vary at the household level (see Table A.21 for the table of balance). The fourth uses nearest neighbor matching in which we also match on the probability of home sales in the subdivision in addition to the other variables (see Table A.22 for the table of balance). The take-away is that each of these matching approaches has a slightly different trade-off in terms of the observables included and the balance of observables.

Table A.19: Table of Balance for Matched Households: Mahalanobis Matching

	(1) SRP-eligible	(2) Matched non-SRP	(3) p-value
lot size (ft^2)	7950	7932	0.29
average house sale price	143553	144876	0.00
median household income	45662	46134	0.00
water consumption (ccf)	15.59	15.57	0.54
house size (ft^2)	1556	1557	0.69
# bath fixtures	6.05	6.06	0.35
1(has pool)	0.19	0.19	0.95

Notes: Column 1 reports means for SRP households in the water consumption data. Column 2 reports means for the matched non-SRP households, using Mahalanobis matching, in the water consumption data. Column 3 reports the p-value for a two-sided test of differences in means. Median HH income refers to the median household income at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

The results from running each of these matching estimations are given in Table A.23. In general, we see clear evidence that the matched households exhibit a peer effect in

Table A.20: Table of Balance for Matched Households: Census Variable Matching

	(1) SRP-eligible	(2) Matched non-SRP	(3) p-value
median household income	45662	45663	0.97
median age	30.45	30.45	0.98
percentage white	56.41	56.42	0.98
percentage Latino	57.17	57.17	0.99

Notes: Column 1 reports means for SRP households in the water consumption data. Column 2 reports means for the matched non-SRP households, using nearest neighbor matching, in the water consumption data. In this table, we only match on variables that vary at the census block level. Column 3 reports the p-value for a two-sided test of differences in means. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

Table A.21: Table of Balance for Matched Households: Household Variable Matching

	(1) SRP-eligible	(2) Matched non-SRP	(3) p-value
lot size (ft^2)	7950	7951	0.99
water consumption (ccf)	15.59	15.03	0.00
house size (ft^2)	1556	1556	0.99
# bath fixtures	6.05	6.06	0.98
1(has pool)	0.19	0.19	0.99

Notes: Column 1 reports means for SRP households in the water consumption data. Column 2 reports means for the matched non-SRP households, using Mahalanobis matching, in the water consumption data. This table matches only on variables that vary at the household level. Column 3 reports the p-value for a two-sided test of differences in means. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

water consumption. We lose some statistical significance in some of the specifications, but all of the results are in the same order of magnitude. We view this as revealing that our main finding about the non-SRP households being different than the SRP households is robust to the exact matching approach that we use.

Table A.22: Table of Balance for Matched Households: Nearest Neighbor with Migration

	(1) SRP-eligible	(2) Matched non-SRP	(3) p-value
lot size (ft^2)	7950	8012	0.21
average house sale price	143553	143602	0.16
median household income	45662	45537	0.00
water consumption (ccf)	15.59	16.68	0.17
house size (ft^2)	1556	1525	0.38
# bath fixtures	6.05	5.92	0.07
1(has pool)	0.19	0.19	0.67
probability of home sale	0.022	0.022	0.39

Notes: Column 1 reports means for SRP households in the water consumption data. Column 2 reports means for the matched non-SRP households, using nearest neighbor matching, in the water consumption data. Column 3 reports the p-value for a two-sided test of differences in means. Median HH income refers to the median household income at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.

Table A.23: Robustness: Role of Economic Incentives

	(1) Mahalanobis Match	(2) Census Match	(3) Household Match	(4) Unmatched non-SRP	(5) Movers Match
mean peer consumption in $t-1$	0.48*	0.32***	0.48**	0.35***	0.24*
	(0.27)	(0.11)	(0.23)	(0.09)	(0.13)
Housing Market Controls	Y	Y	Y	Y	Y
Household Fixed Effects	Y	Y	Y	Y	Y
Census Block x Year Dummies	Y	Y	Y	Y	Y
First Stage F-statistic	2,429	1,686	2,096	2177	2595
R-squared	0.29	0.09	0.21	0.08	0.26
N	517,575	700,681	652,687	909,563	545,243

Notes: This table replicates column 2 of Table 6 in the main text using alternative matching routines and matching on different groups of observables. Column 1 uses the same set of variables as our primary specification in Table 6 but uses Mahalanobis matching. Column two uses nearest neighbor matching but matches only on variables that vary at the census block level. Column 3 uses Mahalanobis matching but matches only on variables that vary at the household level. Column 4 presents the results on the unmatched sample of non-SRP eligible houses. The tables of balance for each specification are presented in the three tables preceding this one and Table A.18. The dependent variable in each specification is growing season consumption in t. An observation is a household parcel-year. The peer group is defined as all houses within a 500′ radius of the household. The 'mean peer consumption in $t-1$' refers to the average peer growing season consumption in period $t-1$. All specifications instrument for peer consumption using the fraction of parcels with housing transactions within 500′ in the previous year. All models are estimated in first differences to difference out the household effects. Standard errors clustered at the Census block level. *** denotes significance at 1% level, ** at 5% level, * at 10% level.