Title
Land Markets and the Value of Water: Hedonic Analysis Using Repeat Sales of Farmland

Journal
American Journal of Agricultural Economics, 96(4)

ISSN
0002-9092

Authors
Buck, Steven
Auffhammer, Maximilian
Sunding, David

Publication Date
2014-07-01

DOI
10.1093/ajae/aau013

Peer reviewed

Land Markets and the Value of Water: Hedonic Analysis Using Repeat Sales of Farmland

Steven Buck, Maximilian Auffhammer, and David Sunding

The lack of robust water markets makes it difficult to value irrigation water. Because water rights are appurtenant to land, it is possible to infer the value of water from observed differences in the market price of land. We use panel data on repeat farmland sales in California's San Joaquin Valley to estimate a hedonic regression equation with parcel fixed effects. This controls for sources of omitted variables bias and allows us to recover the value of irrigation water to landowners in our sample. We show that a more traditional cross-sectional regression results in an artificially low value of irrigation water.

JEL codes: Q15, Q24, Q25.

Determining the value of surface water used for crop irrigation is an important economic policy question. The California Department of Water Resources estimates that surface water accounts for two-thirds of the state's dedicated water supplies and that irrigated agriculture in the state uses approximately 34 million acre-feet of water, most of it surface water, annually. Social policies to protect aquatic habitat impact the availability of irrigation water supplies to farmers (see, for example, Moore, Mulville, and Weinberg). Reliable estimates of the economic damages of expected reductions in irrigation water supplies are of crucial importance to the design of efficient water policy as well as the calculation of the benefits from infrastructure projects such as those under consideration in California. Economists can help assess the impacts of such changes in water availability by developing sound measures of the marginal value of water in agriculture.

Because of the general lack of robust, competitive water markets, it can be difficult to directly observe the marginal value of surface water used in agriculture. A common approach to measuring the value of irrigation water has been through the use of programming models that identify the profit-maximizing input mix given resource availability constraints, including constraints on available surface water. The marginal value of water is determined as a shadow value of the programming problem. Examples of this approach include Marques, Lund, and Howitt and Schaible, McCarl, and Lacewell, among others.

This article takes a different approach to the problem of determining the marginal value of surface water in agriculture. We estimate a hedonic model of farmland values with surface water availability as a covariable. The estimated parameter on the surface water availability variable is the capitalized value of a marginal unit of water availability. A novel feature of our article is that we estimate a hedonic model with parcel fixed effects to control for unobservable factors influencing land value. The use of repeat sales, combined with a significant cross-section and time-series variation in surface water availability across the fields in our sample, allows us to consistently estimate the capitalized value of a marginal unit of surface water.

The hedonic approach is commonly used for evaluating environmental policies; a few prominent examples include the Clean Air Act (Chay and Greenstone 2005; Chong, Phipps, and Anselin 2003) and the Superfund

Steven Buck is an associate economist at The Brattle Group; Maximilian Auffhammer is an associate professor in the Department of Agricultural & Resource Economics, U.C., Berkeley; and David Sunding is a full professor in the Department of Agricultural & Resource Economics, U.C., Berkeley.

The research was supported by the Giannini Foundation of U.C., Berkeley, and U.C., Davis, and the Hewlett Foundation. The authors thank participants in seminars at U.C., Berkeley, and The Ohio State University, as well the editor, Brian Roe, and two anonymous referees for invaluable feedback. All errors are our own.

Amer. J. Agr. Econ. 96(4): 953–969; doi: 10.1093/ajae/aau013
Published online April 10, 2014

program for cleanup of hazardous waste sites (Greenstone and Gallagher 2008). The approach is also used for valuing natural resources such as water quality (Leggett and Bockstael 2000) and climate (Mendelsohn, Nordhaus, and Shaw 1994; Schlenker, Hanemann, and Fisher, 2006). In recent years there has also been a surge of methodological work on hedonic methods in environmental and resource economics. Kuminoff and Pope (2012) present a solution to recover the marginal willingness-to-pay (MWTP) function from the hedonic model; specifically, they describe an instrumental variables strategy to identify the MWTP from the second stage of Rosen's original model. Bishop and Timmins (2011) tackle the same issue to recover the MWTP function for air quality without using instrumental variables. Elsewhere, Gamper-Rabindran and Timmins (2013) evaluate how the scale at which localized amenties are measured affects hedonic estimates of those amenities. In related work, Abbott and Klaiber (2011) investigate how the spatial scale of fixed effects can affect the capitalization of amenities such as "open space" in the housing market, which vary spatially. Following Banzhaf and Walsh (2008), who estimate the effect of air pollution on demographic composition, the work of Bishop and that of Bayer, Keohane, and Timmins (2009) relaxes the assumption that people can move freely to consider the costs of migration in a hedonic model used to recover the MWTP for air quality. Also, Klaiber and Smith (2009) develop a framework for assessing quasi-experimental estimates of the willingness to pay for changes in housing property amenities and apply their framework to study the value of converting land cover from xeric to wetland landscape and the value of cleaning up hazardous waste sites. Relatedly, Kuminoff, Parmeter, and Pope (2010) point out that the use of econometric models using spatial fixed effects or quasi-experimental research designs are now a common way to control for time-constant omitted variables and evaluate the performance of such design specifications at removing omitted variables bias from hedonic price estimates.

In this vein, we show that, in the presence of omitted variables, the hedonic price function for agricultural land can produce severely biased estimates of the value of irrigation water. To demonstrate this bias, we evaluate the quality of hedonic valuations of irrigation water supplies treating repeat sales data as a pooled cross-section compared with exploiting the repeat sales data to estimate a model with parcel fixed effects. We base our analysis on a dataset of farmland transactions in California's San Joaquin Valley.[1] The region is the largest agricultural user of surface water in California and is one of the most productive agricultural regions in the world.

Starting with Selby (1945) , who examines how irrigated land values averaged at the county level covary with the cost of irrigation, a number of empirical studies have analyzed how access to irrigation water is capitalized into farmland values. Other examples include Hartman and Anderson (1962), Faux and Perry (1999), Petrie and Taylor (2007), Shultz and Schmitz (2010), and Schlenker, Hanemann, and Fisher (2007). Studies such as these pursue a hedonic valuation of irrigation water because water in the agricultural setting may not be allocated by a price mechanism, and even when it is, competitive water prices may not be easily observable. There is, however, evidence of a competitive market for farmland (Just and Miranowski 1993), and information on the value of farmland is often available, which makes the hedonic approach attractive. Hartman and Anderson (1962) run a pooled regression using data on agricultural land sales from 1954 to 1960 in Colorado and shares of irrigation company stock appurtenant to the land to estimate a linear specification of the hedonic regression equation, which they use to recover a value irrigation water. Faux and Perry (1999) use cross-section data on agricultural land sales from 1991 to 1995 in Oregon and permit access to different sources of irrigation water; they also estimate a linear specification of the hedonic regression equation. Shultz and Schmitz (2010) run a pooled regression using data on agricultural land sales from 2000 to 2008 in Nebraska and irrigation access to recover the value of access on a per-acre basis. Perhaps the most defensible valuation of irrigation water is the work of Petrie and Taylor (2007), who use a difference-in-differences estimation to recover the value of irrigation water from a policy change in Georgia. They use data on agricultural land sales, irrigation permits, and the timing of a

[1] The San Joaquin Valley consists of eight counties: Fresno, Kern, Kings, Madera, Merced, San Joaquin, Stanislaus, and Tulare.

regional policy affecting irrigation permits to estimate a log-linear specification of the hedonic regression equation. Closest to our own work is that of Schlenker, Hanemann, and Fisher (2007), who combine a pooled cross-section of agricultural land values with measures of surface water deliveries, climate, groundwater availability, soil quality, and population density to recover a value of irrigation water over a time period (1998–2003) and location (39 counties in California) that is similar to that in our own study. However, there are important differences between our study and that of Schlenker, Hanemann, and Fisher—for example, the study area, sample selection criteria, and the measures of land values and surface water availability.

Despite these differences, when we run Schlenker, Hanemann, and Fisher's (2007) cross-sectional linear specification of the hedonic regression equation on our sample of repeat sales, we obtain point estimates for the value of surface water that are statistically indistinguishable from, although smaller than, values obtained by Schlenker, Hanemann, and Fisher (2007). However, as we include spatial fixed effects (hydrological unit, city, school district, and parcel) at an increasingly finer scale, we obtain larger estimates of the value of water. The finest scale of spatial fixed effects, parcel fixed effects, produces estimates of the value of irrigation water in California's San Joaquin Valley that are four times the estimates obtained from specifications using the most aggregate level spatial fixed effects.

In summary, this article contributes two economically significant findings to the literature: First, we find that the value of irrigation water in the San Joaquin Valley may be significantly higher than that suggested by previous work in California. Second, we show evidence that the source of bias, if it exists, in past empirical estimates may be due to the impossibility of accounting for unobserved heterogeneity in cross-sectional/pooled hedonic analyses. These findings both provide information on the value of irrigation water of immediate consequence to public policy makers in California and identify a methodological issue worth considering by those using hedonic analyses to recover water values in empirical settings.

In this article, we present a simple economic model to motivate the empirical analysis. We follow with a description of the data sources. We then discuss our empirical research design and present the main estimating equations. Next we share the estimation results along with a discussion. Finally, we provide a conclusion and indicate areas for future work.

Economic Model

Rosen's theory of hedonic price function allows one to infer the market value of a good by examining the prices of a composite good, which includes the good of interest. Applying the theory, we infer that there is an implicit market for irrigation water deliveries that works through the explicit farmland market. The main economic concept exploited in the hedonic price analysis is that the price of farmland equals the net present value of economic rents expected from the farmland, whereas the price of the differentiated attribute is the shadow value of the attribute in terms of net present value.

A potential buyer of a parcel of farmland observes that the land comes with an expected quantity of available water. This land and water can be combined with variable inputs to produce output in year t. We define output in year t in terms of a production function, $f(L, W, v_t)$ where L is land, W is water available to the parcel, and v_t is some optimal quantity of a variable input in time period t. We suppress the subscript t on L and W because we assume these quantities are fixed inputs that come with a farm parcel and do not change over time. The buyer assumes that the price of the output is p_t and the cost of the variable input is c_t in period t; the cost of the land with its associated water supply is $\theta(L, W)$, which is the hedonic price or bid function. Based on these factors, the potential buyer considers the economic rents that may be derived in a year by choosing the optimal level of v_t given available land and water and facing prices p_t, c_t, and $\theta(L, W)$. Said differently, the potential buyer solves the following profit maximization problem:

$$\max_{(v_t)} \Pi = \sum_{t=0}^{\infty} \delta^t \cdot (p_t \cdot f(L, W, v_t) - c_t \cdot v_t) - \theta(L, W). \quad (1)$$

The resulting Π^* equals the economic rent aggregated across all time periods to be obtained from the land. Next we consider how to assess the shadow value of a permanent change in yearly water availability. First, we note that a change in W will affect the economic rent in year t, Π_t^*; this change in Π_t^* is the expected shadow value of additional water in a year of production, λ. The shadow value of a permanent change in yearly water availability is the sum of discounted shadow values of additional water for each year. The net present value of a permanent change in annual water supply can be written as: $\lambda_W = \sum_{t=0}^{\infty} \delta^t \cdot \lambda(L, W)$. In this simple framework, λ_W is the value an additional unit of irrigation water deliveries in perpetuity.

To complete this analysis, we tie λ_W to changes in land prices associated with a unit change in irrigation water deliveries. Based on Just and Miranowski (1993), we assume that land markets are competitive. Therefore, all economic rents will be bid away so that the price of land will reflect the net present value of production on the land, which may be inferred from productive qualities of the land such as soil quality and water availability. Information on federal surface water deliveries to a parcel of farmland in the San Joaquin Valley can be estimated based on publicly available information on (*a*) how much surface water a parcel's water district receives from the Bureau of Reclamation Central Valley Project[2] and (*b*) irrigated farmland in the water district to which a parcel of land belongs.[3] Given the availability of this information, we assume both sellers and buyers are aware of how much water is available to farmland parcels. Per Rosen, implicit water prices are revealed to buyers and sellers by observing the prices of farmland parcels differentiated by varying levels of water deliveries. As a consequence, the partial derivative of the hedonic price function with respect to W is equivalent to the shadow value of a permanent increase in water availability:[4]

$$(2) \qquad \frac{\partial \theta(L, W)}{\partial W} = \lambda_W = \sum_{t=0}^{\infty} \delta^t \cdot \lambda(L, W).$$

The shadow value of an acre-foot of water to an acre of land for one year is denoted with $\lambda(L, W)$, which will be a function of L and W unless L is separable from W in θ and θ is linear in W. The aim of the empirical analysis is to recover λ_W.

Data

We collected data on sale prices[5] of all land transactions in eight California counties between 2001 and 2008, water deliveries and land acreage for irrigation districts, groundwater depth measurements, historical temperature and precipitation, soil quality measures, land classification codes, and measures of population density. All data resources are combined into a single dataset for analysis. This final dataset of repeat farmland sales represents eight counties, twenty-seven hydrological unit areas, thirty irrigation districts, sixty-three cities, sixty-three school districts,[6] and 140 distinct parcels of farmland.

Outcome Variable: Farmland Sale Price per Acre

The farmland price data were purchased from DataQuick, a private firm that collects data on land sales, mostly from county courthouses. Each observation is geo-referenced with latitude and longitude, and the street address is observed, which allows for clear identification of repeat observations. In addition to land prices and the aforementioned variables, the transaction data include information on transaction characteristics, such as whether the sale was a foreclosed property. The dataset contains characteristics related to structures on the property, including whether there is a building on the property, total square footage of the building, the number of bedrooms, and the number of bathrooms, as well as an estimate of the percentage improvements made to the property. Other property characteristics include a measure of the primary use of the land according to a

[2] Report of Operations Monthly Delivery Tables. Central Valley Operations Office, United States Bureau of Reclamation. Available online at: http://www.usbr.gov/mp/cvo/deliv.html.

[3] California Department of Conservation. Farmland Mapping & Monitoring Program (FMMP). Available online at http://www.conservation.ca.gov/DLRP/FMMP/Pages/Index.aspx.

[4] We direct the reader to the orginal work by Rosen for the full theoretical development of hedonic prices and implicit markets.

[5] This is distinct from Schlenker, Hanemann, and Fisher (2007) who use self-reported land values from the June Agricultural Survey of the U.S. Department of Agriculture.

[6] Cities and school districts are not equivalent; fourteen of the school districts contain land sale observations from multiple cities; ten of the cities contain land sale observations from multiple school districts.

county administrator and lot size, which we use to construct the price-per-acre variable. All prices are converted to year 2012 real prices.

Explanatory Variable of Interest: Surface Water Delivery Right per Acre

In 1933 the state of California passed the Central Valley Project Act. This act authorized the government to begin fundraising for the construction of water infrastructure such as reservoirs, dams, and canals to support irrigated agriculture. Because of the Great Depression, the federal government made several financial transfers to the state of California to complete the project. Today, the United States Bureau of Reclamation is responsible for the administration of the Central Valley Project, which annually delivers approximately 5 million acre-feet of surface water to irrigation districts in California's Central Valley. Farmland within an irrigation district has a contractual right to buy a fixed amount of water in a given year from the Bureau of Reclamation. Like Schlenker, Hanemann, and Fisher (2007), we assume each irrigable acre within an irrigation district receives the irrigation district's average federal surface water deliveries per irrigable acre.[7] Division of water according to irrigable acres is an assumption used elsewhere in the literature, most notably the Statewide Agricultural Production model (for an example, see Medellin Azuara et al. 2008); the most current version of this dynamic programming model was developed in collaboration with the California Department of Water Resources.

A separate issue in constructing the explanatory variable of interest is how one measures expectations about future surface water deliveries. Schlenker, Hanemann, and Fisher (2007) use historical mean deliveries of federal plus private surface water based on data from 1992 to 2002, which does not permit changing expectations. However, water received from the Central Valley Project is not fixed. Each year farmers may receive a different allotment of surface water from the project, which we assume is observed and used to update expectations about future deliveries. Table 1 shows temporal variation in federal deliveries across counties. Although these depict within-county, as opposed to within-parcel variation, the general pattern is emblematic of what we observe on the parcel level. In our empirical analysis, we address the issue of expectation formation by considering moving averages of deliveries of different window lengths, ranging from a contemporaneous specification up to a five-year moving average. Naturally, the measures based on shorter windows of time contain more information about recent history, whereas the measures based on longer windows contain more information of historical persistence of deliveries; accordingly, these attributes may produce differences in implicit prices.

Control Variables

All together surface water may account for as much as 20 million acre-feet in California's irrigated agriculture. Besides the Central Valley Project, other sources of surface water deliveries in the San Joaquin Valley are private (local) projects, which are usually administered by local governments, and to a lesser degree in our dataset, the State Water Project; access to both sources of surface water are observed in our dataset. The last major source of water for agriculture is groundwater, which varies significantly across the state and within our sample; on the whole, groundwater may account for a little less than one-third of irrigation water in the state. We gather data on groundwater availability based on well measurements of groundwater levels obtained from the California Department of Water Resources. Many wells are not measured regularly or were not measured during the sample period, and the wells used for groundwater measurements are not evenly distributed across agricultural land. Therefore, we use regional averages of groundwater availability to create parcel measures of groundwater availability; these are likely measured with error. Data on groundwater quality and other hydrological characteristics such as groundwater flow direction are not easily collected and so are unobserved in this analysis. However, hydrologists have surveyed California and defined areas with similar hydrological characteristics; we reference these as "hydrological units." These are contiguous areas smaller

[7] We collected data on annual federal surface water deliveries from the Bureau of Reclamation. Data on irrigable acres within an irrigation district were recovered from the Farmland Mapping & Monitoring Program, which is managed by the California Department of Conservation.

Table 1. The Distribution of Average Deliveries (AF/Acre) across Counties and Years

County	2001	2002	2003	2004	2005	2006	2007	2008	Total
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Fresno	0.02	0.22	0.005	0.24	0.21	0.03	0.004	0.14	0.12
	(8)	(8)	(7)	(17)	(13)	(14)	(7)	(7)	(81)
Kern	0.00	0.00	0.83	1.14	1.06	1.84	0.00	0.32	0.80
	(1)	(1)	(2)	(3)	(1)	(2)	(1)	(2)	(13)
Kings	0.00	NA	0.00	0.01	0.01	0.00	0.00	NA	0.003
	(1)	(.)	(1)	(2)	(4)	(2)	(1)	(.)	(11)
Madera	0.50	0.96	0.72	1.13	0.41	0.58	0.53	0.78	0.74
	(2)	(3)	(3)	(4)	(3)	(2)	(2)	(3)	(22)
Merced	0.00	0.90	1.05	0.57	0.54	NA	NA	0.52	0.69
	(1)	(4)	(7)	(7)	(6)	(.)	(.)	(6)	(31)
San Joaquin	NA	0.05	0.00	0.47	0.07	0.57	0.04	0.00	0.22
	(.)	(3)	(3)	(4)	(2)	(4)	(4)	(1)	(21)
Stanislaus	0.00	0.00	NA	0.00	0.00	0.00	NA	0.00	0.00
	(1)	(3)	(.)	(2)	(1)	(1)	(.)	(2)	(10)
Tulare	0.55	0.74	0.48	0.78	1.27	0.62	0.60	0.67	0.72
	(10)	(19)	(13)	(21)	(11)	(12)	(7)	(10)	(103)
Total	0.28	0.55	0.48	0.57	0.55	0.40	0.25	0.44	0.47
	(24)	(41)	(36)	(60)	(41)	(37)	(22)	(31)	(292)

Note: Data used to construct values reported are from the U.S. Bureau of Reclamation and from the Farmland Mapping and Monitoring Program, which is managed by the California Department of Conservation. Number of parcels used in the calculation of each average is reported in parentheses beneath each average.

than a county; the land transactions in the sample are spread across 27 hydrological units.

We collect data on soil quality from the United States Department of Agriculture's Natural Resources Conservation Service, which maintains both STATSGO2[8] and SSURGO[9] soil databases; our measure of soil quality is the Storie index. Neither of these soil databases are ideal for parcel cross-sectional analysis because, as with the groundwater measures, they are a weighted average of soil type over large swaths of land; this implies significant unmeasured variation within each soil survey unit and measurement error. This poses a problem if the unmeasured component of soil quality is correlated with both deliveries and land price. For example, there may be a particularly high-quality piece of land within a low average soil quality area. If the high-quality land requires less irrigation water to effectively water plants, then this source of measurement error would cause a downward bias in our point estimate on deliveries.

In addition to irrigation water availability and soil quality, climate is likely to be another important determinant of farmland value. We use the same high-resolution temperature and precipitation climate data that has been used by others (Schlenker, Hanemann, and Fisher 2006). These climate data were organized by the Spatial Climate Analysis Service at Oregon State University for the National Oceanic and Atmospheric Administration; parcel measures of climate are interpolated using the PRISM model also developed by researchers at Oregon State.[10] We use thirty-year historical annual rainfall and both maximum and minimum temperatures to control for climate in our analysis. However, we only use climate measures in the cross-sectional analysis—the impact of climate on farmland values will ultimately be parsed out in the panel data analysis because climate does not vary in our short sample period.

Another important factor to control for in any analysis of farmland values is urban development potential. Our approach to this problem is to create a one-mile buffer

[8] Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. U.S. General Soil Map (STATSGO2). Available online at http://soildatamart.nrcs.usda.gov.

[9] Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database. Available online at http://soildatamart.nrcs.usda.gov.

[10] PRISM Climate Group, Oregon State University. Data are available at http://www.prism.oregonstate.edu/.

zone around population centers, where population centers are defined to Census block groups with population densities greater than 1,000 people per square mile and Census blocks with population densities great than 500 people per square mile. Finally, using information reported in the data file received from DataQuick, we construct measures of whether there is a building on the property, total square footage of the building, distance to closest freeway, the primary use of the land, and lot size.

Sample Selection

For inclusion in the final sample, we only use observations that satisfy several criteria. We begin with a sample of 9,683 property sales received from DataQuick for the period covering the years 2001 through 2008. Keeping only parcels for which there are repeat sales reduces our sample to 1,431 parcels (3,604 land sale observations). We are concerned that property with higher turnover may be correlated with unobservable characteristics correlated with price and water deliveries; keeping parcels that sold less then four times in our eight-year window and never within the same year reduces our sample to 803 parcels (1,696 land sale observations). We then drop all parcels for which the transaction had an unusual characteristic (e.g., foreclosure) affecting price.[11] Removing these parcels reduces the sample to 583 parcels (1,211 land sale observations). Of these, 405 parcels (844 land sale observations) are designated as irrigable farmland. Schlenker, Hanemann, and Fisher (2007) indicate that the most expensive land parcels in terms of price per acre are located close to urban areas. They hypothesize that these land values are driven by urban land markets rather than the local agricultural economy. Therefore, in their analysis they consider three samples of land sales, which are limited to parcels of land with a price per acre of less than \$10,000/acre, \$15,000/acre, and \$20,000/acre. For similar reasons and for advantage of comparability, we limit the sample to parcels of land with a price per acre of less than \$20,000/acre (and more than \$200/acre-foot);[12] this reduces our sample to 263 parcels (546 land sale observations). Another concern is that parcels with houses will greatly affect the price per acre. Similar to urban development potential, parcel attributes such as housing may dwarf variation in prices due to surface water deliveries. To address this issue, we drop all observations with bedrooms, which leaves 142 parcels (296 land sale observations). Finally, we remove all parcels for which the production activity is timber, poultry, dairy, livestock, or vacant land; this drops one dairy parcel and one vacant land parcel.[13] The final result is an admittedly small panel sample that has 292 observations representing 140 parcels of farmland. To evaluate the external validity of our panel estimates, we compare these parcels of farmland for which we have repeated observations (sample A) to all parcels satisfying the described selection criteria (sample B) other than the first criteria of having repeat sale observations. Table 2 reports the means and sample standard deviations for observable covariables by sample A and sample B. The explanatory variables in these samples are comparable, and we do not observe any statistically significant differences between them. There is a significant difference in the dependent variable, price per acre, between sample A and sample B. This is the result of sample composition over time. In sample A, we only keep the first transaction for each parcel so that sample A is weighted more toward the beginning of our study period when land prices were lower in real terms. However, the sample mean of price per acre for all transactions in the sample of repeat sales is \$11,625 (S.E. \$6,617), which is not significantly different than the sample mean in sample B.

Empirical Research Design

The first empirical modeling decision we make is the selection of a functional form for the hedonic price equation. From a theoretical perspective, the functional form

[11] These unusual characteristics include if one of the parcel's recorded sales had a price that included additional property not reported in the sale record (multiproperty sale), or if a recorded sale price was for only part of the property recorded in the sale record, or if the property was a foreclosure sale.

[12] To follow the Schlenker, Hanemann, and Fisher (2007) criteria, this is \$20,000/acre in Y2000 real prices.

[13] Our results are invariant to this final selection criteria because these types of parcels are already absent. We explicitly maintain the criteria because we want to drop these types of parcels to construct a cross-section dataset that is comparable with our panel dataset.

Table 2. Comparison of Parcels Sold Repeatedly (Sample A) to All Parcels (Sample B)

	Sample A Mean (S.D.)	Sample B Mean (S.D.)	p Value	Sample A Minimum / Sample B [Minimum]	Sample A Maximum / Sample B [Maximum]
	(1)	(2)	(3)	(4)	(5)
Price per acre (Y2012 USD)	9,809[a]	12,223	0.000	757	31,249
	(5,783)	(7,070)		[319]	[31,346]
2 year mean moving average of deliveries (AF/acre)	0.491	0.481	0.898	0.000	3.687
	(0.809)	(0.832)		[0.000]	[3.966]
Lot size (acres)	45.26	45.84	0.917	0.99	632.37
	(66.08)	(61.87)		[0.99]	[640.00]
Building structure/storage shed (d)	0.079	0.094	0.549	0.000	1.000
	(0.270)	(0.292)		[0.000]	[1.000]
Square footage of building structure	223	223	0.998	0	14,192
	(1,278)	(948)		[0.00]	[14,210]
Private water deliveries access (d)	0.54	0.50	0.372	0.00	1.00
	(0.50)	(0.50)		[0.00]	[1.00]
No groundwater (d)	0.407	0.412	0.909	0.000	1.000
	(0.493)	(0.492)		[0.000]	[1.000]
Mean depth to groundwater (ft)	35.49	37.71	0.604	0.00	317.11
	(47.20)	(48.31)		[0.00]	[317.11]
Elevation (meters)	88.47	84.23	0.204	0.00	211.10
	(38.86)	(37.51)		[0.00]	[228.27]
Historical mean annual rainfall (mm)	283.5	292.3	0.138	172.8	447.6
	(47.0)	(52.3)		[160.2]	[478.0]
Historical mean max temp (°C)	24.90	24.75	0.199	23.31	26.05
	(0.57)	(0.67)		[22.04]	[26.05]
Historical mean min temp (°C)	9.27	9.20	0.120	8.41	11.59
	(0.35)	(0.36)		[8.23]	[11.95]
Storie Index for soil quality	0.59	0.61	0.368	0.000	0.98
	(0.26)	(0.27)		[0.000]	[0.98]
Orchards (d)	0.364	0.360	0.922	0.000	1.000
	(0.483)	(0.480)		[0.000]	[1.000]
Vineyards (d)	0.150	0.210	0.095	0.000	1.000
	(0.358)	(0.407)		[0.000]	[1.000]
Distance to freeway (meters)	12,542	11,944	0.413	119	30,479
	(8,514)	(8,201)		[5]	[38,690]
1 mile urban buffer (d)	0.26	0.26	0.906	0.00	1.00
	(0.44)	(0.44)		[0.00]	[1.00]
Total observations	140	1383			

[a] Difference in price is the result of sample composition across time. For sample A, we have eliminated more than half the observations because we only keep the first observation per parcel, which by definition occurs earlier in the study period during which land prices were lower in real terms. Sample A minimums and maximums are reported to the right of the means, and Sample B minimums and maximums are reported in square brackets.

depends on the structure imposed on the demand equation, which, as Rosen describes, generally indicates a nonlinear hedonic price function except in special cases. Cropper, Deck, and McConnell (1988) find that the flexible functional form of the quadratic Box-Cox performs best among several forms when there are no omitted variables. However, they also find evidence that suggests simple regression forms, including linear, semi-log, log-log, and linear Box-Cox, perform well in terms of average percentage bias and maximum percentage bias at estimating the hedonic price function whenever there are omitted variables. As a result of their finding and because of concern over omitted variables, the majority of hedonic analyses since Cropper, Deck, and McConnell (1988) have relied on one of these simple forms (see Kuminoff, Parmeter, and Pope (2010) for a thorough review). Among the studies we reviewed that use hedonic analyses of agricultural land values to recover the value of irrigation water, the majority have used a linear specification of the hedonic price function, including the study by Schlenker, Hanemann, and Fisher (2007), who estimate a value of irrigation water most directly comparable with our own estimates. From the scientific perspective of incrementing on what has already been done, we consider the linear specification as the relevant base specification of functional form. Another reason to use the linear specification is because it is an approximation of the average first derivative of the maximum value function, in this case $\theta(L, W)$ differentiated with respect to water deliveries (W). Despite its inadequacies, the linear specification is a reasonable start to assess the average behavior of a nonlinear hedonic price function without a closed form solution. We suppress the arguments of λ in equation (2) under the assumption that the hedonic price function is linear in W, which implies λ is constant. To test the sensitivity of our results to this functional form assumption, we also estimate a quadratic Box-Cox model and recover very similar mean and median values of surface water in agriculture.

Putting aside the discussion of functional form, our main empirical concern is omitted variables bias—particularly that an unobservable parcel characteristic is correlated with both water deliveries and land prices.

Empirical Challenge: Omitted Variables Bias

To begin, we assume land and water are the only two factors of production and that land quality is homogeneous. Therefore, the hedonic price equation may be estimated using

$$(3) \qquad \text{Price/Acre}_{it} = \beta_0 + \beta_1 \cdot \text{Deliveries}_{it} + \epsilon_{it}.$$

In this equation, i denotes the parcel of land, the outcome variable is the price per acre of the land sale, and "deliveries" captures the expected stream of annual surface water to be delivered in perpetuity to parcel i. The parameter β_1 can be interpreted as the shadow value of a permanent shock to water deliveries supply (λ_W); β_1 is expected to be positive. If one observes land price and volumetric water deliveries, then one can estimate β_1 consistently if the following identifying assumption holds:

$$(4) \qquad \mathbb{E}[\epsilon_{it} | \text{Deliveries}_{it}] = 0.$$

Assuming equation (4) holds, then β_1 can be estimated using ordinary least squares to recover the capitalized value of surface water deliveries. However, the identifying condition can be violated in many ways. The most basic way is that there are unobservable environmental quality characteristics of the parcel that are correlated with deliveries so that

$$(5) \qquad \epsilon_{it} = \gamma \cdot EQ_i + \eta_{it}$$

where EQ_i is the underlying environmental quality of parcel i. If EQ_i is observable, then one can consistently estimate the coefficient of interest if the following condition holds:

$$(6) \qquad \mathbb{E}[\eta_{it} \mid \text{Deliveries}_{it}, EQ_i] = 0.$$

This is the assumption maintained in pooled cross-sectional analyses such as in the work of Schlenker, Hanemann, and Fisher (2007). This identifying assumption is violated if there are parcel-specific unobservable characteristics that are correlated with either the other environmental quality measures or deliveries. That is, the error term from equation (3) may take the form

$$(7) \qquad \epsilon_{it} = \xi_i + \gamma \cdot EQ_i + \nu_{it}$$

where,

$$(8) \qquad \mathbb{E}[\xi_i \mid EQ_i, \text{Deliveries}_{it}] \neq 0.$$

Relaxing the assumption of homogeneous land quality and assuming the expressions in equations (7) and (8) hold, then the only way to estimate β_1 consistently is using a fixed-effects estimator, which requires repeat sales of the same parcel.

Estimating Equations

We begin with a specification that treats our sample as pooled cross-section data and, like Schlenker, Hanemann, and Fisher (2007), we use a random effects estimator. Schlenker, Hanemann, and Fisher (2007) cluster based on square mile as defined in the June Agricultural Survey. This level of cluster does not exist in our dataset, so we cluster on hydrological units as defined by the California Department of Water Resources; these are regions with common hydrological characteristics. The estimating equation is

$$(9) \quad \text{Price/Acre}_{ijt} = \beta_1 \cdot \text{Deliveries}_{it} + \beta_2 \cdot X_i + \tau_t + \eta_{ijt} \quad \text{with} \quad \eta_{ijt} = \mu_j + \epsilon_{ijt}$$

where i indicates the parcel; j indicates the hydrological unit; t indicates the year; deliveries$_{it}$ represents water deliveries; X_i represents time-invariant observables including lot size, whether there is building on the property, the building size, alternative surface water supply availability, groundwater depth, historical rainfall and temperature, soil quality, crop production (orchards, vineyards and row crops/pasture land as the excluded group), and population density;[14] τ_t is the year fixed effect; and η_{ijt} is a composite error term where μ_j is the hydrological unit intercept and ϵ_{ijt} represents all unobservable factors affecting the outcome variable. It is worth noting that the interpretation of β_1 is that of λ_W from equation (1), the capitalized value of an acre-foot of water in perpetuity.

We later estimate models with spatial fixed effects at increasingly fine spatial level: hydrological unit fixed effects, city/municipality fixed effects, and school district fixed effects. The problem with these spatial fixed estimators is that they all ultimately rely on cross-sectional variation to identify the effect of irrigation water availability on land price. If there still remains unobserved cross-sectional variation correlated with both water deliveries and land values, then the point estimate on water will be biased. Controlling for soil quality using STATSGO2 measures or through the use of spatial fixed effects may not be sufficient because soil quality measures are based on weighted averages over relatively large areas and thus, as with spatial fixed effects, assume homogeneous soil quality for large regions of land. Hydrological unit, city/municipality, and school district units may encompass areas that still exhibit significant variation in environmental quality. Thus, the use of more aggregate-level spatial fixed effects does not satisfy the homogeneity assumption. As already described, to overcome this, one could collect parcel measures of soil quality, although if there are remaining time-invariant omitted variables then the estimates may still be biased. For this reason, an estimator using panel data is attractive.

We argue environmental quality is slow to change over time so one may consider it time-invariant over the sample period; a parcel fixed-effects estimator will parse out all variation due to time-invariant parcel characteristics. The estimating equation for the parcel fixed effects analysis is given by:

$$(10) \quad \text{Price/Acre}_{ijt} = \delta_1 \cdot \text{Deliveries}_{it} + \xi_i + \tau_t + \epsilon_{ijt}$$

where i, j and t are as before; ξ_i is the parcel fixed effect; and δ_1 is interpreted as λ_W from equation (1), the capitalized value of an acre-foot of water in perpetuity. This specification will control for variation in the outcome variable due to time-invariant characteristics, including underlying environmental quality of the land.

Standard Error and Pivotal Statistic Adjustments

Statistical inference is complicated because of the clustered structure of the data. To address clustering and within-cluster heteroskedasticity, we compute robust standard errors clustered at the county level (eight clusters). Bertrand, Duflo, and Mullainathan (2004) show that when there are a small number of clusters (ten or less), the performance of statistical inference using cluster-robust standard errors is unreliable. Cameron, Gelbach, and Miller (2008) find the same result and consider a variety of standard error adjustments for clustered data and then evaluate which adjustments offer reliable performance for statistical inference. They suggest using the wild cluster-bootstrap method to obtain pivotal test statistics because this method offers asymptotic refinement. We follow their advice and compute a wild cluster-bootstrap pivotal test statistic for each regression coefficient; then the analytically computed cluster-robust standard error is used with the adjusted pivotal test statistic to perform hypothesis testing.

Reconsideration of Functional Form

As suggested earlier, Kuminoff, Parmeter, and Pope (2010) investigate whether the reliance on simple functional forms for the hedonic price function is appropiate given advances in

[14] For summary statistics of the full list of control variables, see table 2.

data availability and methods for addressing omitted variables over the past two decades since the Cropper, Deck, and McConnell (1988) study. Regarding data availability, half the studies Kuminoff, Parmeter, and Pope (2010) review use more than five years of sale data, which presents the concern of whether implicit prices should be regarded as time invariant. In response, we also consider more general forms in which we interact time-period dummies with water deliveries and then test for differences in the capitalized value of water over time.

With respect to methods, Kuminoff, Parmeter, and Pope (2010) find evidence that, when using spatial fixed effects, using a quadratic Box-Cox functional form performs better than any of the simpler functional forms.[15] The result is not surprising because Cropper, Deck, and McConnell (1988) find that without omitted variables the quadratic Box-Cox model produces the lowest percentage bias of all the forms they consider. If spatial fixed effects successfully account for variables omitted from the hedonic price equation, then the simpler functional forms may not be justified on a performance basis. Further, the Box-Cox functional form does not rule out the simpler functional forms; indeed, a Box-Cox test allows us to formally test whether the parameters of the Box-Cox transformation produce a model that is distinguishable from the linear or log-log specifications. Crouter (1987) discusses the issue at length and finds that the linear functional form is justified on theoretical grounds only when water is not appurtenant to land; otherwise, they recommend using the Box-Cox method to identify the functional form of the hedonic price equation. Because water may be considered appurtenant to land in our setting, we do not satisfy Crouter's theoretical condition, so have additional justification to evaluate the quadratic Box-Cox model Kuminoff, Parmeter, and Pope (2010) and Cropper, Deck, and McConnell (1988) recommend. The regression equation for this model is given by:

$$\mathrm{y}_{ijt} = \delta_1^{BC} \cdot \mathrm{x}_{it} + \delta_2^{BC} \cdot \mathrm{x}_{it}^2 + \xi_i^{BC} + \tau_t^{BC} + \epsilon_{ijt}^{BC} \tag{11}$$

[15] Regarding specifications exploiting panel data, Kuminoff, Parmeter, and Pope (2010) point out that in the presence of time-varying omitted variables, Cropper, Deck, and McConnell's (1988) original advice to use simpler functional forms still applies.

where, $y_{ijt} = (\text{Price/Acre}_{ijt}^{\phi} - 1)/\phi$ if $\phi \neq 0$ and $y_{ijt} = \ln(\text{Price/Acre}_{ijt})$ if $\phi = 0$. Similarly, $x_{it} = (\text{Deliveries}_{it}^{\phi} - 1)/\phi$ if $\phi \neq 0$ and $x_{it} = \ln(\text{Deliveries}_{it})$ if $\phi = 0$; all else is as in equation (10).

Results and Discussion

The main results of the analysis are summarized in table 3, which presents the implied capitalized value of an acre-foot of water from linear regressions that treat the data as a cross-section and then as a panel. We begin with a random-effects estimator, and then, as we move across columns, introduce estimators using spatial fixed effects at an increasingly finer scale. The dependent variable in all specifications is the real price of land, whereas the focal independent variable is the two-year moving average of federal surface water deliveries. For each estimated regression, we present the coefficient on our water deliveries measures, its cluster-robust standard error (based on eight county clusters), and the pivotal statistic we obtain from the wild cluster bootstrap procedure. The latter should be used in lieu of the usual critical t values for hypothesis testing and constructing confidence intervals. In column 1, we present the point estimate from the results of a random-effects specification, which assumes clusters at the hydrological unit level. The implied capitalized value of an acre-foot of water per acre of land is \$137 and indistinguishable from zero. In the second specification, we present the estimate from the fixed-effects model clustered at the hydrological unit level ($m_{hu} = 27$). The estimate is \$826 per acre-foot, and a Hausman test shows that the estimates in columns 1 and 2 are significantly different at the 1% level. In the third and fourth columns, we present results from specifications that, as with the specification in column 2, treat the data as pooled cross-section data. The specification using city fixed effects ($m_{city} = 63$) estimates the value of water to be \$1,857/acre-foot, whereas the specification using school district fixed effects ($m_{sd} = 63$) indicates the capitalized value of water is \$2,416/acre-foot. In the final column, we present a specification that treats the data as a panel and thus includes parcel fixed effects ($m_{plot} = 140$); the point estimate on the two-year moving average of federal surface water

Table 3. Ordinary Least Squares Hedonic Regression Results

Dependent variable: Price per acre of farmland (mean: \$11,625, S.D.: \$6,617)
Water deliveries variable: 2-year moving average (acre-feet/acre)

	(1)	(2)	(3)	(4)	(5)
Capitalized value of water					
	137	826	1,857	2,416	3,723***
	(537)	(950)	(767)	(1,095)	(1,392)
	[2.99]	[3.18]	[3.24]	[3.98]	[1.85]
Observations R^2	292	292	292	292	292
	0.316	0.416	0.591	0.593	0.742
Parcel level controls	Yes	Yes	Yes	Yes	No
Year fixed effects ($m_{year} = 9$)	Yes	Yes	Yes	Yes	Yes
Hydrological unit random effects	Yes	No	No	No	No
Hydrological unit fixed effects ($m_{hu} = 27$)	No	Yes	No	No	No
City fixed effects ($m_{city} = 63$)	No	No	Yes	No	No
School District fixed effects ($m_{sd} = 63$)	No	No	No	Yes	No
Parcel fixed effects ($m_{parcel} = 140$)	No	No	No	No	Yes

Note: Cluster-robust standard errors (eight county clusters) are presented in parentheses beneath the coefficient estimates. Hypothesis testing performed using pivotal statistics obtained from the wild cluster bootstrap procedure; pivotal statistics corresponding to two-sided tests of significance with level $\alpha = 0.05$ are presented in square brackets beneath the standard errors. The t statistic (the ratio of the coefficient to the S.E.) must be larger than the pivotal statistic (the value in the square bracket) to demonstrate statistical significance at the level of $\alpha = 0.05$. A triple asterisk indicates $p < 0.01$; a double asterisk indicates $p < 0.05$; a single asterisk indicates $p < 0.10$.

deliveries is \$3,723/acre-foot and is significant at the 1% level. Using our bootstrapped pivotal statistic, the 95% confidence interval for this estimate is \$1,146–\$6,300/acre-foot. The pattern of point estimates as we move from spatial fixed effects that encompass larger swaths of land to spatial fixed fixed effects at the parcel level makes sense—the finer spatial fixed effects account for more omitted variables than those encompassing larger swaths of land. We have evidence that the estimated capitalized value of water apparently suffers from downward bias if fine-scale spatial fixed effects are not taken into account.

In figure 1, we show how the point estimates from the parcel fixed-effects estimation changes when we use moving averages of different window length to measure water rights. We observe that the point estimates increase the larger the moving average window. Further, the value is rising but at a decreasing rate in window size. One interpretation of these results is that as water reliablity, measured per window length of the moving average, increases, then so does the value of water; a more stable longer term supply is valued more highly. Adopting a water rights measure based on a moving average of greater window length clearly has a cost in econometric terms. Given our already small sample size, if we use longer moving averages, we remove variation from our main covariable and the coefficients are less precisely estimated the larger the moving average window.

In table 4, we present the results of a series of robustness checks. In column 1, we drop observations that list a buyer and seller with the same last name; the goal is to limit our sample to arms-length sales. The remaining sample consists of 259 observations; the parcel fixed effects estimate is \$4,623/acre-foot and is marginally significant. In column 2, we drop observations that have access to surface water from alternative private local projects. The remaining sample consists of 132 observations; the parcel fixed-effects estimate is \$4,189/acre-foot and is significant at the conventional level. Our main result is robust to these alternative samples and to the designation of other subsamples not shown. Regarding functional form, Kuminoff, Parmeter, and Pope (2010) examine two stylized facts: (*a*) amenities have time-invariant implicit prices and (*b*) simple functional forms such as the linear or log-log models are less susceptible to omitted variables bias than more flexible functional forms such as the quadratic Box-Cox. Based on evidence from simulation results, they advise analysts to allow for time-variant implicit prices. In column 2, we allow for time-variant implicit prices by interacting our water measure with an indicator for whether the land sale occurred after 2004; doing so permits one

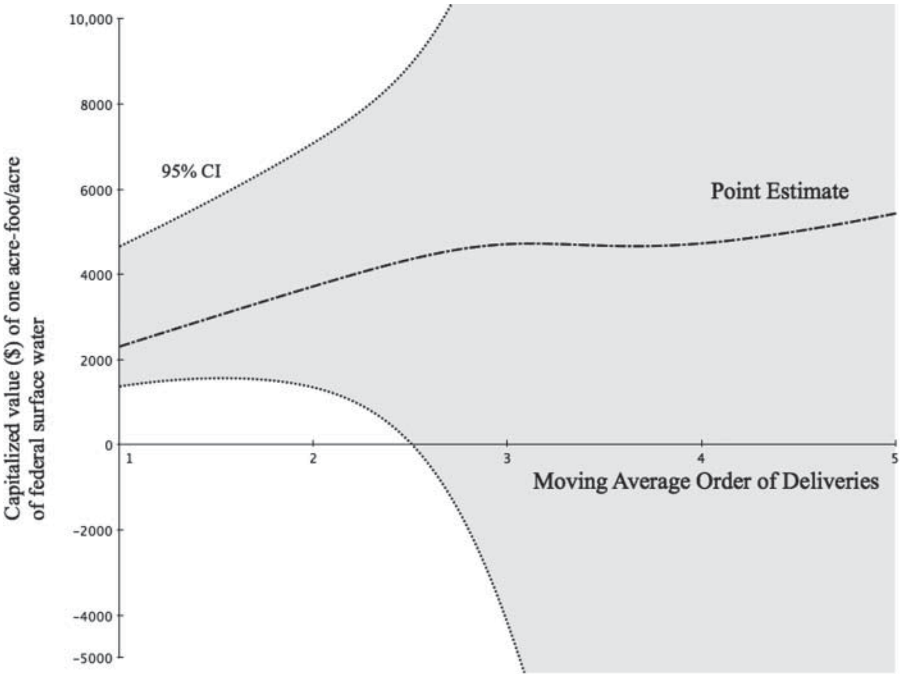

Figure 1. Depiction of how the estimated capitalized value of water changes with moving average order of deliveries

Note: This figure is based on the results of five separate regressions. Each regression is identical to the one associated with the results present in column 5 of table 3 except that we vary the measure of water deliveries. For example, 1 on the x-axis denotes the regression using the 1-year moving average as the relevant measure of deliveries; the corresponding point on the y-axis represents the estimated capitalized value of an acre-foot of water from regression output. We also present the 95% confidence interval, which is calculated using the cluster-robust standard error and the relevant pivotal statistic [$\alpha = 0.05$] obtained from the wild cluster-bootstrap procedure.

implicit price during the first four years of our sample and a different implicit price during the final four years of our sample. The results in column 3 show that the implicit price for surface water was slightly less in the latter years of our sample, although the difference is statistically indistinguishable. This result is robust to alternative functional forms in which we allow implicit prices to vary over different periods of time. For example, we estimate separate regressions that allow the structural break in implicit prices to occur in every year between 2002 and 2008; we also estimate a model in which we interact our water measure with year dummy variables. These alternative specifications do not suggest time-variant price equilibria, although a larger dataset could test this more precisely.

Next we estimate a model using the quadratic Box-Cox form under the assumption that our parcel fixed effects adequately control for omitted variables. We choose ϕ to maximize the likelihood function, which results in the value $\phi = 0.229$ with a standard error of 0.113. We run likelihood ratio tests to evaluate the null hypotheses that $\phi = 0$ and then $\phi = 1$, which correspond to the log-log and linear models, respectively. We reject the null hypothesis in both instances, and estimate the model in equation (11)—a parcel fixed-effects specification for which we assume a quadratic Box-Cox functional form. In column 4 of table 4, we see the mean predicted value of an acre-foot of water using the quadratic Box-Cox model—$3,840/acre-foot; the median predicted value is $3,482/acre-foot. Further, the range of the predicted values of an acre-foot of water in our sample of repeat sales is $2,626/acre-foot to $7,000/acre-foot, which is a relatively tight distribution, and the model itself is highly predictive ($R^2 = 0.846$) when compared with the other regressions in our analysis.

One disadvantage of our estimates from the parcel fixed-effects specifications is that we can only consider parcels that are sold repeatedly, thus leaving us with a small

Table 4. Ordinary Least Squares Hedonic Regression Results, Robustness Checks

Dependent variable: price per acre of farmland
Water deliveries variable: 2-year moving average

	Alternative samples		Alternative functional forms	
	Remove if non-arms-length sale	Remove if receives other surface water	Time-variant price equilibria	Quadratic Box-Cox
	(1)	(2)	(3)	(4)
Capitalized value of water	4,623*	4,189**	3,924*	3,840
	(2,085)	(2,176)	(2,106)	*3,482*
	[2.61]	[1.69]	[1.64]	
Additional value in year 2004 and beyond			−165	
			(877)	
			[2.53]	
Observations	259	132	292	292
R^2	0.757	0.780	0.742	0.846
Number of parcels	124	64	140	140
Year fixed effects	Yes	Yes	Yes	Yes
Parcel fixed effects	Yes	Yes	Yes	Yes

Note: The fourth column reports the mean predicted value of water from the quadratic Box-Cox model; the median predicted value is presented in italics. Cluster-robust standard errors (eight county clusters) are presented in parentheses beneath the coefficient estimates. Hypothesis testing performed using pivotal statistics obtained from the wild cluster bootstrap procedure; pivotal statistics corresponding to two-sided tests of significance with level $\alpha = 0.05$ are presented in square brackets beneath the standard errors. A triple asterisk indicates $p < 0.01$; a double asterisk indicates $p < 0.05$; a single asterisk indicates $p < 0.10$.

sample that is potentially different from the general population of land parcels in San Joaquin Valley. To investigate whether our sample of repeat observations is similar to the rest of agricultural land in the region, we compare the parcels in our sample of repeat sales (sample A) to parcels in a sample that is expanded to those land parcels only sold once (sample B). Table 2 shows the results of this exercise, and we see that the parcels in each sample are similar on average. Although there are no significantly different explanatory variables, it is noteworthy that our sample A of repeat sales is slightly hotter and drier than sample B, which could make irrigation water more valuable to irrigable farmland in sample A than in sample B; this is also true when we compare sample A to all 7,510 parcels in the original DataQuick dataset.

If the parcels in sample A are a random sampling of all agricultural land parcels sold during the study period, then we would expect the estimated value of water to be similar when we estimate identical regressions using sample A and then sample B. The results of this exercise are presented in table 5. In the first two columns, we compare the coefficients from regression specifications with hydrological unit fixed effects; in the second two columns, we compare the coefficients from regression specifications with city fixed effects; and in the third set, we compare the coefficients from regression specifications with school district fixed effects. For each pair of regressions, we run a z test to compare whether the regression coefficients are significantly different from each other. Although the point estimates are larger from the regressions using sample A, which is comprised of repeat sales, than the point estimates from sample B, we fail to reject the null hypothesis that the coefficients are equal ($\alpha = 0.05$) in all three cases. These findings are robust to basing sample A on keeping the most recent transaction for each parcel among the repeat-sale parcels rather than on keeping the first transaction. Although caution should be taken because of the large standard errors on the coefficient estimates, the results in table 2 and table 5 provide some evidence that the $3,723/acre-foot estimate from the parcel fixed-effects specification using the subsample of repeat sales is generalizable to agricultural land in the San Joaquin Valley. This is significant because the estimate from the school fixed-effects specification using the full sample of cross-section data, the basis for our strongest strawman estimate, is $768/acre-foot, which

Table 5. Ordinary Least Squares Hedonic Regression Results, Alternative Cross-Sectional Samples

Dependent variable: price per acre of farmland
Water deliveries variable: 2-year moving average

Sample	A	B	A	B	A	B	B
	(1)	(2)	(3)	(4)	(5)	(6)	(7)[a]
Capitalized value of water	359 (877)	99 (331)	500 (1,691)	356 (414)	1,623 (3,117)	768 (448)	275 (373)
z Statistic from test comparing value of water	0.257		0.066		0.218		
Observations[b]	131	1,376	109	1,364	109	1,359	1,359
Count of spatial clusters	18	34	32	81	32	108	108
R^2	0.473	0.289	0.651	0.334	0.669	0.341	0.255
Time-invariant controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Year fixed effects ($m_{year} = 7$)	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Hydrological unit fixed effects	Yes	Yes	No	No	No	No	No
City fixed effects	No	No	Yes	Yes	No	No	No
School district fixed effects	No	No	No	No	Yes	Yes	No
School district random effects	No	No	No	No	No	No	Yes

[a]Water deliveries is measured using the historical mean deliveries from 1993 to 2002, not a moving average.
[b]In each regression, we drop all observations that are singletons within their cluster. Cluster-robust standard errors (seven county clusters) are presented in parentheses beneath the coefficient estimates. Hypothesis testing performed using pivotal statistics obtained from the wild cluster bootstrap procedure. $p < 0.01, p < 0.05, p < 0.1$.

is, in both statistical (z test with $\alpha = 0.05$) and economic terms, significantly less than the \$3,723/acre-foot estimate. Whether or not our estimate applies generally to California agriculture, we cannot say; however, our finding is economically significant for the state given that the San Joaquin Valley is by far the largest agricultural user of surface water in California.

In related work, Schlenker, Hanemann, and Fisher (2007) find the value of surface water in California agriculture to be \$1,046/acre-foot[16] using pooled cross-section data without the benefit of repeat observations on parcels because these were not available in their dataset. Interestingly, our cross-sectional estimates of the capitalized value of surface water using cross-section data are between \$99/acre-foot and \$1,623/acre-foot. To make a cleaner comparison, we run a linear specification on the pooled cross-section with school district random effects using historical mean deliveries as our measure of deliveries—this is the data and specification most similar to that implemented by Schlenker, Hanemann, and Fisher (2007). The resulting point estimate of \$275/acre-foot (S.E. \$373/acre-foot) is indistinguishable from their estimate of \$1,046/AF (S.E. \$226/acre-foot). However, because of differences between this study and that of Schlenker, Hanemann, and Fisher—most notably, the study area, sample selection criteria, and the measures of land values and water availability—we do not have evidence that Schlenker, Hanemann, and Fisher (2007) would recover our estimate if only they had repeat observations on land values. Nonetheless, we have provided evidence that accounting for parcel-specific unobservable time-invariant characteristics may be a requisite to recover the implied value of agricultural surface water from hedonic analyses of farmland sales.

In terms of magnitude, our estimate of the capitalized value of an acre-foot of surface water, \$3,723,[17] is one and a half to four times the size of the cross-sectional estimates using the same data. Conditioned on receiving federal water deliveries, the average parcel of farmland in our sample has a price of \$11,254/acre and receives 1.09 acre-feet of federal deliveries per acre of land. Our estimate of \$3,723/acre-foot suggests that, conditioned on receiving federal surface

[16] Converted from Y2000 real prices to Y2012 real prices.

[17] As an approximation, if we assume 5% discount rate then the implied value of an acre-foot of water in a given year is close to \$190, which is net of any annual costs the land owner must pay for the surface water received.

water, 1.09 acre-feet in irrigation water rights would account for approximately $4,033 of the $11,254/acre sale price (36% of the average sale price with the 95% confidence interval of 11%–61%). Taken together, our results suggest (*a*) that farmers in the San Joaquin Valley likely have a higher willingness to pay for surface water than previously thought and (*b*) that this result plays a sizable role in the determination of irrigated farmland prices.

Conclusions

Using a small sample of repeat sales of agricultural land in California's San Joaquin Valley, we find evidence that existing cross-sectional estimates of the value of surface water deliveries in agriculture are vulnerable to significant downward bias. Using a parcel fixed-effects estimation, we account for unobserved factors correlated with deliveries and farmland prices and obtain novel estimates of the value of surface water deliveries in California agriculture. The estimated capitalized value of one acre-foot of water is one and a half to four times larger than the estimate obtained in the cross-sectional analyses. Although we have limited evidence on whether the average capitalization value of $3,723/acre-foot of surface water extends to all agricultural land in the San Joaquin Valley, the result that studies relying on cross-sectional data and methods may underestimate the capitalized value of irrigation water is empirically demonstrated. Based on this result, changes in water availability are likely to induce larger changes in producer welfare than previously thought. Finally, these findings inform policy analysis on issues related to water infrastructure projects, the protection of habitat for endangered species, and climate change. A natural way to extend this work is to explicitly incorporate groundwater availability and crop choice into an agro-economic model to make predictions about how the shadow value of surface water varies with these other sources of heterogeneity.

References

Abbott, J. K., and H. A. Klaiber. 2011. An Embarrassment of Riches: Confronting Omitted Variable Bias and Multi-Scale Capitalization in Hedonic Price Models. *Review of Economics and Statistics* 93 (4): 1331–42.

Banzhaf, S. H., and R. P. Walsh. 2008. Do People Vote with Their Feet? An Empirical Test of Tiebout's Mechanism. *American Economic Review* 98 (3): 843–63.

Bayer, P., N. Keohane, and C. Timmins. 2009. Migration and Hedonic Valuation: The Case of Air Quality. *Journal of Environmental Economics and Management* 58 (1): 1–14.

Bertrand, M., E. Duflo, and S. Mullainathan. 2004. How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics* 119 (1): 249–75.

Bishop, K. C. 2010. A Dynamic Model of Location Choice and Hedonic Valuation. Washington University in St. Louis.

Bishop, K. C., and C. Timmins. 2011. Hedonic Prices and Implicit Markets: Estimating Marginal Willingness to Pay for Differentiated Products Without Instrumental Variables.

California Department of Water Resources. 2005. California Water Plan Update 2005, Bulletin 160-05. Working Paper, California Department of Water Resources.

Cameron, C., J. Gelbach, and D. Miller. 2008. Bootstrap-Based Improvements for Inference with Clustered Errors. *Review of Economics and Statistics* 90 (3): 414–27.

Chay, K., and M. Greenstone. 2005. Does Air Quality Matter? Evidence from the Housing Market. *Journal of Political Economy* 113 (2): 49.

Chong, W., T. Phipps, and L. Anselin. 2003. Measuring the Benefits of Air Quality Improvements: A Spatial Hedonic Approach. *Journal of Environmental Economics and Management* 45 (1): 24–39.

Cropper, M. L., L. B. Deck, and K. E. McConnell. 1988. On the Choice of Funtional Form for Hedonic Price Functions. *Review of Economics and Statistics* 70(4): 668–75.

Crouter, J. 1987. Hedonic Estimation Applied to a Water Rights Market." *Land Economics* 63 (3): 259–71.

Faux, J., and G. Perry. 1999. Estimating Irrigation Water Value Using Hedonic Price Analysis: A Case Study in Malhuer

County, Oregon. *Land Economics* 75 (3): 440–52.

Gamper-Rabindran, S., and C. Timmins. 2013. Does Cleanup Of Hazardous Waste Sites Raise Housing Values? Evidence of Spatially Localized Benefits. *Journal of Environmental Economics and Management* 65 (3): 345–60.

Greenstone, M., and J. Gallagher. 2008. Does Hazardous Waste Matter? Evidence from the Housing Market and the Superfund Program. *Quarterly Journal of Economics* 123 (3): 951–1003.

Hartman, L., and R. Anderson. 1962. Estimating the Value of Irrigation Water from Farm Sales Data in Northeastern Colorado. *Journal of Farm Economics* 44 (1): 207–13.

Just, R., and J. Miranowski. 1993. Understanding Farmland Price Changes. *American Journal of Agricultural Economics* 75 (1): 156–68.

Klaiber, H. A., and V. K. Smith. 2009. Evaluating Rubin's Causal Model for Measuring the Capitalization of Environmental Amenities. Working paper, National Bureau of Economic Research.

Kuminoff, N. V., C. F. Parmeter, and J. C. Pope. 2010. Which Hedonic Models Can We Trust to Recover the Marginal Willingness to Pay for Environmental Amenities? *Journal of Environmental Economics and Management* 60 (3): 145–60.

Kuminoff, N. V., and J. C. Pope. 2012. A Novel Approach to Identifying Hedonic Demand Parameters. *Economics Letters* 116 (3): 374–6.

Leggett, C., and N. Bockstael. 2000. Evidence of the Effects of Water Quality on Residential Land Prices. *Journal of Environmental Economics and Management* 39(2): 121–44.

Marques, G. F., J. R. Lund, and R. E. Howitt. 2009. Modeling Conjunctive Use Operations and Farm Decisions with Two-Stage Stochastic Quadratic Programming. *Journal of Water Resources Planning and Management* 136(3): 386–94.

Medellin Azuara, J., J. J. Harou, M. A. Olivares, K. Madani, J. R. Lund, R. E. Howitt, S. K. Tanaka, M. W. Jenkins, and T. Zhu. 2008. Adaptability and Adaptations of California's Water Supply System to Dry Climate Warming. *Climatic Change* 87 (1): 75–90.

Mendelsohn, R., W. D. Nordhaus, and D. Shaw. 1994. The Impact of global Warming on Agriculture: A Ricardian Analysis. *American Economic Review* 84 (4): 753–71.

Moore, M., A. Mulville, and M. Weinberg. 1996. Water Allocation in the American West: Endangered Fish versus Irrigated Agriculture. *Natural Resources Journal* 36 (Spring 1996): 319.

Petrie, R., and L. Taylor. 2007. Estimating the Value of Water Use Permits: A Hedonic Approach Applied to Farmland in the Southeastern United States. *Land Economics* 83 (3): 302–18.

Rosen, S. 1974. Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. The Journal of Political Economy 82(1): 34–55.

Schaible, G. D., B. A. McCarl, and R. D. Lacewell. 1999. The Edwards Aquifers Water Resource Conflict: USDA Farm Program Increase Irrigation Water-Use? *Water Resources Research* 35(10): 3171–83.

Schlenker, W., W. M. Hanemann, and A. C. Fisher. 2006. The Impact of Global Warming on U.S. agriculture: An Econometric Analysis of Optimal Growing Conditions. *Review of Economics and Statistics* 88 (1): 113–25.

———. 2007. Water Availability, Degree Days, and the Potential Impact of Climate Change on Irrigated Agriculture in California. *Climatic Change* 81 (1): 19–38.

Selby, H. 1945. Factors Affecting Value of Land and Water in Irrigated Land. *Journal of Land and Public Utility Economics* 21(3): 250–8.

Shultz, S., and N. Schmitz. 2010. The Implicit Value of Irrigation through Parcel Level Hedonic Price Modeling. Paper presented at the Agricultural & Applied Economics Association's 2010 AAEA, CAES & WAEA Joint Annual Meeting, Denver, CO.