Liquid Constrained in California:
Estimating the Potential Gains from Water Markets

[Current version here]

Nick Hagerty*

October 29, 2019

Abstract

Water markets may help societies adapt to rising water scarcity and variability, but their setup costs can be substantial and their benefits uncertain. I estimate the gains available from strengthening the wholesale surface water market in California, where conveyance infrastructure is well-developed yet transaction volume remains low. To do so, I develop a new empirical framework to analyze welfare in water markets that uses transactions data. First, I recover marginal valuations of water in the presence of unobserved transaction costs, by using particular price comparisons to find the incidence of both known and unknown cost determinants. Second, I estimate demand using yearly water endowments, which have rich variation driven by weather and amplified by historical rules. Then, I combine this demand model with a hydrological network model to simulate counterfactual outcomes. I find that efficient trading across regions and sectors would achieve benefits of only $86 to $278 million per year, without accounting for any environmental costs. These results suggest that promoting large-scale water markets may not achieve large gains without also reforming the policies and institutions that govern local water allocation.

*University of California, Berkeley, nhagerty@berkeley.edu. I am extremely grateful to Esther Duflo, Ben Olken, and Chris Knittel for guidance and support. I thank Clay Landry and WestWater Research, LLC for generously sharing the data that made this project possible, as well as Kavish Gandhi for excellent research assistance. Many people contributed helpful feedback that improved this paper; I particularly thank David Atkin, Sydnee Caldwell, Josh Dean, Dave Donaldson, Michael Greenstone, Peter Hull, Donghee Jo, Jack Liebersohn, Matt Lowe, Rachael Meager, Arianna Ornaghi, Doug Parker, Gary Sawyers, Matt Zaragoza-Watkins, Ariel Zucker, and seminar participants at MIT, Harvard, EDF, UC Davis, UC Berkeley, PPIC, and the Tinbergen Institute. Funding for this research was provided by the Abdul Latif Jameel Water and Food Systems Lab (J-WAFS) at MIT.

1 Introduction

Water supplies are becoming scarcer and more variable in many parts of the world (UNDP 2006). Fueled by population pressure and climate change, water scarcity can increase poverty and conflict (Sekhri 2014; Burke et al. 2015) and is set to worsen in the coming decades (World Bank Group 2016). To help societies adapt to water scarcity, many observers advocate for greater use of water markets. Like other markets, water markets may yield benefits by allocating scarce resources to the greatest social need and allowing participants to flexibly respond to changing conditions. However, the costs of water markets can be substantial, often involving new infrastructure for conveyance and monitoring, overhauls to legal and regulatory institutions, and potential ecological damages from reallocating water. Meanwhile, the benefits are uncertain. Historical experience is thin, since robust, large-scale water markets are rare even in wealthy industrialized countries. Relatively little quantitative research is available to estimate these prospective benefits and help policymakers weigh them against the costs.

In this paper, I estimate the potential gains from a more efficient water market in California. I take a revealed-preference approach, in which welfare calculations depend on parameters that I estimate from observed trading behavior in the existing water market. I begin by modeling California's thin water market as an exchange economy with transaction costs, rationalizing observed transaction prices with an unobserved set of transaction costs and demand curves. As in any exchange economy, these demand curves plus initial endowments are sufficient to find the efficient allocation without transaction costs. I derive an empirical procedure to construct demand curves, and I apply it to comprehensive new data on water transactions and annual entitlements in California. Lastly, I combine the demand curves to simulate the result of a fully efficient market and calculate the resulting gains from trade.

There are two key reasons why California is a useful setting in which to study the potential benefits of water markets. First, there are potentially large economic gains available from water reallocation. California has a large and diverse economy, but most of it depends on water supplies that are imported over great distances and prone to droughts. This water is not allocated by a price mechanism but rather according to historical rules and formulas. A secondary market exists, but it is highly regulated, transaction volume is low, and price dispersion is high. Perhaps as a result, retail prices can vary over more than two orders of magnitude: in 2017, according to their websites, commercial and industrial customers in the city of San Diego paid $2,491 per acre-foot[1], while agricultural customers less than two hours away in the Imperial Valley paid just $20 per acre-foot. Second, substantial reallocation may be possible through policy reforms alone. California has already built most of the physical infrastructure necessary to support a robust water market, and it has spare capacity. Canals, pipelines, and rivers together form a nearly complete hydrological network connecting the vast majority of water users in the state, such that it is technologically feasible to transfer water between nearly any two consumers.

[1] An acre-foot, the standard unit of volume for water in California, is the amount of water that would cover one acre of land with twelve inches of water.

Two salient features of California's statewide water market are high price dispersion and low transaction volume, relative to the number of water districts and independent consumers who might transact. To rationalize these facts as equilibrium outcomes, I model California's water market as an exchange economy in which consumers trade endowments of a single homogeneous good via intermediaries. Consumers incur transaction costs that may be location pair-specific and directionally asymmetric. Transaction costs, which I define broadly, may arise from a range of cost determinants, some of which are observable (such as regulatory reviews or conveyance distance) and others which are not (such as search or contracting difficulty). The results of this model rationalize a set of observed trading outcomes with a set of demand curves and transaction costs, and they provide an equilibrium condition that can be used to empirically recover these objects.

My goal is to calculate the possible gains from trade from an efficient market without transaction costs. (This scenario could be achieved through a combination of deregulation – i.e., streamlining the policies governing water transfers – and the creation of market-supporting institutions – e.g., setting up a central exchange – that together eliminate the wedge between buying or selling and *in situ* use.) I note that the efficient allocation is fully determined by initial endowments (which can be read directly from the data) and demand curves. To construct these demand curves and estimate the potential gains from trade, I follow a four-step empirical procedure:

Step 1: **Estimate transaction costs arising from observable cost determinants, by comparing prices across different transactions.** Many water districts sell to (or buy from) more than one other district in a given year. Sometimes, one of these transactions is subject to an observable cost determinant (such as an extra regulatory review) and another is not. Under an assumption of perfect competition, a district transacting with multiple counterparties is indifferent between them in equilibrium, so any difference in prices can be interpreted as marginal transaction costs. By averaging over all such cases for both buyers and sellers, I estimate the marginal transaction costs associated with specific, observable regulatory and physical cost determinants.

Step 2: **Recover market participants' marginal valuations of water at the observed equilibrium, using revealed-preference conditions on prices.** I estimate marginal valuations of water for all market participants using simple revealed-preference conditions: a buying district's marginal valuation must be at least as high as its highest price paid, and a selling district's marginal valuation must be at least as low as its lowest price accepted, after adjusting prices for observed transaction costs (as estimated in Step 1). If there are remaining unobserved transaction costs, these estimates will understate the true dispersion in marginal valuations, in which case my final results would likely provide a lower bound on the potential gains from trade.

Step 3: **Estimate price elasticities of demand, exploiting supply shocks driven by weather and amplified by historical allocation rules.** Step 2 gives an equilibrium point on each district's demand curve; to extrapolate away from equilibrium, I also need an elasticity. To identify price elasticities of demand, I exploit California's historically-determined allocation rules, which turn shared precipitation shocks into vastly different supply shocks for different regions. Because transaction costs lead to persistence of these initial allocations, I can measure how steeply marginal valuations change in response to exogenous changes in quantity consumed.

Step 4: **Simulate the efficient allocation and calculate the resulting gains in consumer surplus, by combining demand curves in an optimization problem.** I first construct demand curves by combining my estimates of equilibrium marginal valuations and demand elasticities with data on equilibrium quantities. Then, I combine demand curves in a constrained optimization problem to solve the social planner's problem, finding the efficient allocation and calculating the resulting gain in consumer surplus. An ideal market could implement the efficient allocation, so this gain represents the value of an efficient market. Since true physical transportation costs can never be eliminated, I include them (as estimated in Step 1) in the objective function.

To conduct this empirical analysis, I construct what may be the most comprehensive dataset yet compiled on California's water economy. First, I assemble for the first time the universe of yearly surface water entitlements in California, including federal and state water project allocations and surface water rights. Second, I use a proprietary dataset on open-market water transactions that to my knowledge is the most complete in existence; crucially, it provides a mostly complete record of prices. I build a large crosswalk file to link users across datasets and years and a geospatial dataset on user locations and boundaries. For supplementary analysis in an appendix, I also incorporate two uniquely high-resolution datasets on land use and farm-level agricultural finances.

In Step 1, I document large price gaps resulting from specific, observable cost determinants. For example, I find that transactions that must cross the Sacramento–San Joaquin Delta (an environmentally sensitive juncture triggering additional regulatory reviews) are associated with a price premium for sellers of \$76 per acre-foot, and a price discount for buyers of \$31. I interpret these as marginal transaction costs, totaling \$107 per acre-foot. These are large as compared with the mean price in my data, \$221. Several other observable factors also each result in similarly large marginal transaction costs.

In Step 2, I find that even though observed prices of surface water transactions are quite dispersed, the estimated marginal valuations have even greater dispersion. This suggests that substantial welfare gains are available from reducing transaction costs and enabling more transactions to occur. Consistent with conventional wisdom in California, marginal valuations tend to be low in the northern Sacramento Valley and high in the southern San Joaquin Valley.

In Step 3, I find that water agencies in this wholesale market have fairly inelastic demand. I estimate that the price elasticity of demand is -0.10 for the urban sector and -0.23 for the agricultural sector, magnitudes that are smaller than previous studies using household- or farm-level data. However, this single agricultural estimate masks considerable heterogeneity across regions; estimated elasticities are just -0.07 for the Sacramento Valley and -0.81 for the Tulare Lake Basin.

My central results come from Step 4. First, I find that observed trading in the existing market achieves welfare gains of \$10 to \$88 million per year, depending on water availability conditions. Then, I simulate an efficient market, in which trading increases until marginal valuations are equalized up to physical transportation costs. I estimate that this scenario would result in additional gains of \$86 to \$278 million per year – figures that are meaningful but tiny when compared with overall water-related expenditures in California.

These results carry three important limitations. First, I cannot identify transaction costs that are both unobserved and constant within user, such as search or contracting costs. This may lead my estimates to understate the true dispersion in marginal valuations. Because greater dispersion in marginal valuations results in greater gains from trade, my simulations likely are lower bounds on the true potential gains. Second, I treat water districts as the key economic agents, leading me to miss any gains from reallocation among the retail customers within these water districts.

Third, my approach accounts for only the potential benefits of water markets. It omits the benefits from existing market-restricting regulations (such as achieving ecological goals and avoiding hydrological externalities) as well as the costs of setting up market-supporting institutions (such as expanded water-use monitoring systems or a centralized trading exchange) that may be required to achieve the full potential benefits. Many stakeholders in California believe there is scope for reforms to dramatically simplify regulatory review processes while accomplishing the same environmental goals; regardless, a complete policy analysis needs to account for both the benefits and costs of reallocating water.

Finally, in an appendix, I investigate whether the price gaps I document in Step 1 might partly be explained by market power rather than marginal transaction costs. I relax the assumption of perfect competition and allow both buyers and sellers to exercise market power, extending my model in ways that follow Atkin and Donaldson (2015). I then derive a further two-step empirical procedure to adjust prices for possible markups and markdowns and to re-estimate marginal transaction costs, net of market power. The results do not suggest that market power explains the large marginal transaction costs I estimate in Step 1, and so I interpret the gains from trade estimated in Step 4 as true deadweight loss rather than transfers between buyers and sellers.

This paper makes several contributions. First, I provide a new approach to estimating the prospective gains from trade in water markets, and I assemble new data that enables this approach to overcome previous limitations. My approach differs from the prior literature in two important ways: (1) it is based on a small number of parameters that are econometrically estimated within the model, and (2) these parameters are estimated from data on observed transactions. A large literature uses calibrated optimization models to estimate the prospective gains from water markets, for California (Howitt et al. 1999; Sunding et al. 2002; Jenkins et al. 2003; Medellín-Azuara et al. 2007) as well as for Australia (Peterson et al. 2005; Qureshi et al. 2009) and Chile (Rosegrant et al. 2000). While these models incorporate rich institutional and scientific knowledge, their economic components rely on large numbers of imputed parameters and functional form assumptions (Mérel and Howitt 2014). In contrast, my approach is more parsimonious and estimates parameters with particular attention to causal identification. In addition, by inferring the preferences of market participants directly from observed transactions, I sidestep the need to directly model agricultural production functions or other fundamental determinants of water demand. Working directly with the implied objective functions of water districts also makes my approach more policy-relevant, since these districts will continue to be the primary market participants under most proposals aiming to strengthen water markets. There may be additional frictions between the districts and their retail customers, but eliminating them would likely be more difficult than simply easing trading among districts.

More generally, this paper proposes a method to analyze the welfare impacts of transaction costs in thin asset markets. This may be particularly relevant to other settings in environmental economics, such as pollution permits or individual transferable quotas. There is a large literature in financial economics on thin markets and liquidity, but it is typically focused on strategic trading rather than transaction costs (Kyle 1989; Rostek and Weretka 2012, 2015). The literature on pollution permit markets covers the theoretical effects of market power (Hahn 1984; Malueg and Yates 2009; Liski and Montero 2011) and transaction costs (Stavins 1995; Liski 2001), with some empirical analysis of transaction costs (Gangadharan 2000; Cason and Gangadharan 2003), but there are few empirical studies of the welfare impacts of transaction costs. In the context of water, Carey et al. (2002) and Regnacq et al. (2016) study the effects of transaction costs on trading quantities in water markets, while Ayres et al. (2017) study transaction costs in groundwater management decisions.

This paper also relates to a literature in international trade that estimates trade costs from price gaps (Donaldson 2012; Atkin and Donaldson 2015; Bergquist 2016). Furthermore, it contributes to a broad and growing literature on the costs of misallocation, in settings such as housing (Glaeser and Luttmer 2003), capital (Hsieh and Klenow 2009), energy (Davis and Kilian 2011), labor (Bryan and Morten 2015; Adamopoulos et al. 2017), and land (Restuccia and Santaeulalia-Llopis 2017).

Finally, my results contribute to a literature in agricultural economics on the value of water in irrigated agriculture (Schlenker et al. 2007; Buck et al. 2014; Mukherjee and Schwabe 2014). I find using a revealed-preference analysis that the marginal value of water is small. Prior studies find much larger estimates, but in using land values, they measure a long-run marginal value. My approach instead measures within-year marginal values, in which farmers likely have greater ability to substitute toward groundwater.

2 Background on Water in California

Water is scarce in most of California, the largest economy and most populous state in the United States. A majority of the state's population lives in Southern California, where there is little rainfall and no major rivers. Farms in the Central Valley, a major agricultural region, receive little rainfall during the summer growing season and instead rely on irrigation. Most precipitation in the state falls during the winter in mountain ranges in the north and east.

Moving water throughout the state is technologically feasible, thanks to an interconnected system of water infrastructure that is the world's most complex. Federal, state, and local authorities operate canals and pipelines that, together with the river system, form a fully connected hydrological network among the vast majority of water users in the state of California. Although there are capacity constraints, at the margin it is possible to transfer water between nearly any two consumers in the state.

California can be thought of as a closed hydrological system. Essentially none of its precipitation flows to other states or countries. The only major water source it shares with other states is the Colorado River, but the amounts that each state receives are governed by long-term interstate compacts that I treat as fixed.

2.1 Water is initially allocated by fixed rules and environmental conditions

Property rights to water in California are distributed not according to private or social value but instead following a complex system of historical precedent. Some districts and consumers hold entitlements that are almost never curtailed, while others are rationed according to precipitation and runoff during the previous winter. In my empirical exercises, I exploit this rationing to estimate demand elasticities and the marginal value of water in agriculture.

I summarize California's hierarchy of water entitlements in Figure 1. All lawful surface water use in the state derives from a legal framework of appropriative and riparian rights. Some independent consumers (such as rural households or isolated farmers) hold their own water rights; others obtain water from federal or state water projects. More commonly, retail consumers (including farms, households, and other consumers) obtain water from their local water district, which in turn either holds its own water rights or long-term contracts with the federal or state water projects. Water districts go by many legal classifications, such as irrigation district, water agency, or mutual water company, and may be public, private, or a blend. They may also have multiple layers, in which a wholesale district sells to retail districts.

The different sources of water entitlements are governed by different allocation rules:

1. **Appropriative and riparian rights.** Rights follow a seniority rule determined by the date of first use; in droughts, senior rights-holders are entitled to their full claim before junior rights-holders are entitled to any. However, it is rare for this seniority system to substantially affect water diversions in major rivers, since the residual claimants are generally the high-volume federal and state water projects.

2. **Central Valley Project (CVP).** Operated by the U.S. Bureau of Reclamation (USBR), the CVP stores and delivers water to irrigation districts, municipal water districts, and individual farms throughout the Central Valley. Contractors do not buy water at market-clearing rates; instead they are entitled to a certain volume of water each year. Contractors each have a fixed maximum volume, specified in multi-decade contracts. Actual yearly allocations vary from year to year, mostly on the basis of weather in the mountains during the previous winter, as well as environmental regulations. Allocations are announced as percentages of maximum contract volumes, determined separately for each of 14 contract categories based on history,

geography, and sector. Some categories tend to have priority over others, but the ordering is neither constant (due to regional differences in water availability) nor lexicographic (like appropriative rights) (Stene 1995).

3. **State Water Project (SWP).** Operated by the California Department of Water Resources (DWR), the SWP stores and distributes water to users throughout the state. Contractors do not buy water at market-clearing rates but rather are entitled to a certain volume of water each year. SWP entitlements also vary from year to year on the basis of weather during the previous winter as well as environmental regulations. Contractors each have a fixed maximum annual volume, specified in multi-decade contracts. In droughts, all contractors receive cutbacks in equal proportion, except before 1992 when there were different proportions for urban and agricultural contractors. These proportions have varied from 100 percent as recently as 2006 to 5 percent of contract maximums in 2014 (California Department of Water Resources 2015).

4. **Lower Colorado Project.** Also operated by the USBR, the Lower Colorado Project distributes California's share of Colorado River water (fixed in compacts dating back to 1922) to contractors in Southern California. To date, California's Lower Colorado contactors have always received their full entitlements.

5. **Groundwater.** Local groundwater is another major source of water for both farms and cities, but its use is generally unmonitored. Availability and pumping costs vary considerably across regions. In this paper, I treat local groundwater supplies as fixed.

2.2 Secondary markets are inhibited by transaction costs

Because water is not allocated according to a price mechanism, a robust secondary market might be expected. In fact, California's statewide water market is thin. Figure 2 plots the total volume of market transactions over time, as compared with total water supply, using data described in Section 4.

I focus on the statewide water market, which I define as transactions directly among districts and other independent consumers at freely negotiated prices. This definition excludes transactions between wholesale districts and retail districts, and between retail districts and retail consumers. Such transactions take place within fixed, long-term relationships in which neither prices nor quantities are always flexible. It also excludes transactions involving retail customers within a water district. Intra-district transactions between consumers are rare in urban water districts but common in some irrigation districts. Unfortunately, data is scarce: even when districts keep records of these transactions, they rarely record prices. Retail consumers are usually not allowed or able to negotiate their own transactions with external districts; instead they must rely on their own district to represent their interests on the statewide water market.

Many factors may make transactions in this market costly. Next, I outline a typology of transaction costs, building on Regnacq et al. (2016), Scheer (2016), and others.

Administrative transaction costs To trade water, a buyer or seller must first search for a potential trading partner. Without a central exchange, this happens mostly by word of mouth in social networks; sometimes a professional broker helps with matchmaking. There is no single standard contract for water transactions. The buyer and seller must negotiate over the quantity, duration, price, payment terms, delivery date, point of delivery, and delivery pathway. Transaction durations fall into two basic types: (1) permanent sales of water rights or contract entitlements, and (2) intra-year leases, in which the seller transfers a certain quantity of water while retaining the underlying entitlement. Together, search and contracting processes may create considerable administrative transaction costs, both explicit (i.e., attorney fees) and implicit (e.g., hassle costs), for both buyers and sellers.

Physical transaction costs Water is heavy; moving it from one place to another is costly. Water is lost in conveyance to evaporation and percolation. Pumping water uphill into canals requires energy to run turbines. Not all transactions incur these costs: upstream transfers may not incur any conveyance losses, while downstream transfers on a river may not incur any pumping costs. However, these pure physical costs may differ from the costs directly incurred by buyers or sellers. Transactors pay "wheeling" charges to the owners of the intermediate conveyance facilities along the delivery pathway, including canals, pumping stations, and reservoirs. Wheeling charges are not generally equal to the true physical marginal cost of conveyance and pumping; some stakeholders believe they are often substantially higher (Western Water Company 2000).

Regulatory transaction costs Proposed transactions can be subject to regulatory review by three main agencies: California's State Water Resources Control Board (SWRCB), California's Department of Water Resources (DWR), and the U.S. Bureau of Reclamation (USBR). Some transactions are reviewed by counties. Depending on the proposed source and destination, transactions may be reviewed by more than one of these agencies, or none of them.

In these reviews, agencies (1) require sellers to compile records proving they have legal entitlement to the water and physical ability to transfer it; (2) carefully estimate consumptive use (the amount not returned to the water system); (3) set up monitoring systems to verify sellers do not continue using water once transferred; (4) conduct extensive environmental impact analyses to meet requirements of the National Environmental Policy Act (NEPA) and the California Environmental Quality Act (CEQA); (5) estimate impacts on the local economy; and (6) schedule delivery for a time with available capacity (California State Water Resources Control Board 1999; California Department of Water Resources and U.S. Bureau of Reclamation 2015).

Transactions that move water across the Sacramento–San Joaquin Delta, the biggest bottleneck in the system, must meet additional environmental regulations concerning outflow volumes, salinity levels, and endangered fish. These transactions can be risky; because the DWR and USBR are sometimes not allowed to pump water into their canals, there is no guarantee that the seller's water will actually reach the buyer. In addition, these transactions are assessed "carriage losses" to satisfy environmental goals and regulatory constraints.

In sum, regulatory reviews may create substantial policy-induced transaction costs, which again may be explicit or implicit. Explicit costs include agency review fees and the engineering fees to prepare documents. Implicit costs include the hassle or general disutility from the process, the time costs of the review plus its public notice and public comment periods, and the risks of disapproval or delivery failure. Both buyers and sellers may bear these costs.

Political-economy transaction costs Implicit transaction costs may also arise from political economy effects. First, water districts imperfectly represent the interests of their retail customers, driving a wedge between statewide water prices and an individual farmer's or household's willingness to pay or to accept. This is especially relevant in districts controlled by popular vote rather than property value or land area. Second, farmers may be reluctant to sell water because they fear voters will think they don't need it and may take away their property rights in the future (Carey and Sunding 2001).

Many types of transaction costs may be pure loss to society, such as administrative costs. Some types may represent transfers, in the case of attorney fees. Others may have important benefits in preventing negative externalities. Environmental reviews protect public goods like wildlife and ecosystem services, and determining consumptive use protects the property rights of downstream users. Political economy constraints may help prevent pecuniary externalities to the origin community.

However, the evidentiary standards for regulatory approval of water transactions are much higher than for the use of the water in the first place. Many stakeholders have proposed reforms that could accomplish many of the same regulatory goals while streamlining the process (Western Water Company 2000; Culp et al. 2014; Gray et al. 2015; Association of California Water Agencies 2016). In my analysis, I focus on the costs of these regulations, which then can be weighed against their benefits. In addition to the costs incurred for transactions that do take place, transaction costs also may prevent many mutually beneficial trades from occurring.

3 Theoretical Framework

In this section I lay out a model of California's water market to guide my empirical analysis. This model serves two purposes. First, it allows me to rationalize transactions in the data as equilibrium outcomes that are consistent with a conceptually precise set of demand curves and transaction costs. This can explain the high price dispersion and low transaction volume seen in California, and it also clarifies the relationship between the observed equilibrium and the efficient allocation. Second, the model yields equilibrium conditions that allow me to empirically identify these demand curves and find the efficient allocation.

I first present a simplified graphical model to build intuition. I then model an exchange economy in which consumers trade endowments of a homogeneous good through intermediaries, which incur transaction costs. Finally, I derive my four-step empirical procedure.

3.1 Simplified graphical model

Consider two districts, d and o (for destination and origin), which might be thought of as an urban water district d and agricultural irrigation district o. They have initial endowments of water E_o and E_d and inverse demand curves $V_d(Q_d)$ and $V_o(Q_o)$, which give their consumers' marginal valuations of water as a function of quantity demanded. Figure 3 plots these two inverse demand curves, with the axis reversed for district o and lined up such that the overall width of the graph is equal to the overall resource constraint – the sum of the endowments – in the style of an Edgeworth box.

If consumers in these two districts were allowed to costlessly trade with each other, we would expect them to arrive at the competitive equilibrium, resulting in allocation (Q_d^*, Q_o^*) and price P^*. Suppose they instead arrive at allocation (Q_d', Q_o') and transaction price P_{od}. Also suppose that each district is internally efficient such that at this new allocation they reach internally district-clearing prices $P_d = V_d(Q_d')$ and $P_o = V_o(Q_o')$, where $P_d > P_{od} > P_o$. How can we make sense of this scenario, in which three different prices are observed in equilibrium?

Observation 1: Transaction costs can rationalize price gaps. Different prices can be observed in equilibrium within and across districts if trade across districts incurs a marginal (per-unit) transaction cost. In Figure 3, buyers in d must be paying per-unit transaction cost $\tau^b = P_d - P_{od}$, while sellers in o must be paying transaction cost $\tau^s = P_{od} - P_o$. These transaction costs shift down the buyer's marginal willingness-to-pay for cross-district trade, $V_d - \tau^b$, relative to its marginal valuation: the buyer requires a price discount equal to the transaction costs in order to be indifferent between origins. Similarly, the seller's marginal willingness-to-accept for cross-district trade, $V_o + \tau^s$, is higher than its marginal valuation: the seller requires a price premium in order to be indifferent between destinations. Under an assumption of perfect competition, transaction costs must be exactly equal to the price gap. If they were lower, the buyer and seller would trade more, and if they were higher, the buyer and seller would trade less (or not at all).

Observation 2: Potential gains from trade can be estimated with knowledge of demand curves and initial endowments. The welfare gains from reducing transaction costs can be calculated as the change in consumer surplus of the two districts. This equals $\Delta CS_d + \Delta CS_o$, as shown in Figure 3. Buyers' consumer surplus is standard welfare analysis; sellers' consumer surplus here is analogous to producer surplus. Four vectors of information suffice to approximate this gain: initial endowments (E_d, E_o), baseline allocation (Q_d', Q_o'), internal prices at this allocation (P_d, P_o), and demand elasticities (η_d, η_o).

Unfortunately, P_d and P_o are unobserved in my data, which contains mostly cross-district transactions. To overcome this obstacle, below I develop an empirical approach leveraging the fact that districts often trade with more than one other district.

3.2 Model of an exchange economy with transaction costs

I now develop a more formal model, in order to more precisely clarify the insights above, microfound them with explicit assumptions, and extend them to more than two consumers. In this model, two layers of intermediaries engage in spatial arbitrage of a homogeneous good across districts. Intermediaries, which might be thought of as brokers representing each district, incur transaction costs that may be district pair-specific and directionally asymmetric. This reflects that physical, regulatory, and other costs vary not only across buyers and sellers but also by who they choose to transact with, and in which direction. Perfect competition leads to the result that prices equalize marginal valuations, up to transaction costs.

The primary purposes of the model are to (1) rationalize price dispersion in equilibrium, and (2) motivate the idea that water market participants buy at their marginal willingness to pay and sell at their marginal willingness to accept. Several alternative models can obtain these same basic results. Here I present a model with intermediaries and perfect competition because it is relatively simple and may be familiar from the international trade literature. In Appendix B, I describe a potentially more realistic bargaining model, in which water districts directly engage in simultaneous bilateral negotiations. As individual transaction volumes become small relative to total quantities consumed, prices given by the Nash bargaining solutions (within a Nash equilibrium across negotiations) converge to the result from the perfectly competitive model shown here.

Consider N water districts indexed by n (or by o for origin and d for destination). Each district has initial endowment E_n of a single homogeneous good (water) that is allocated efficiently among a continuum of consumers. Consumers' preferences can be aggregated such that each district has an inverse demand function $V_n(Q_n)$, which gives marginal valuations as a function of quantity consumed Q_n. Inverse demand is decreasing and twice differentiable.

Trade across districts is conducted by two layers of intermediaries. In each district, selling intermediaries ("sellers") can buy units of water from consumers (at their marginal valuations $V_o(Q_o)$) and sell to buying intermediaries ("buyers") in another district. Buyers, in turn, buy water from sellers and can sell to consumers in their own district (at their marginal valuations $V_d(Q_d)$). Sellers and buyers meet at exchange points unique to each pair of districts, where prices P_{od} are determined. Each transaction i generates transaction costs for both sellers and buyers:

Assumption 1. Per-unit transaction costs. *To complete a transaction of q_{iod} units of water, sellers and buyers each incur constant marginal transaction costs. They are non-negative and may be specific to each origin-destination district pair:*

$$\begin{aligned} &(Sellers) \qquad C^s_{od}(q_{iod}) \equiv \tau^s_{od} q_{iod} \\ &(Buyers) \qquad C^b_{od}(q_{iod}) \equiv \tau^b_{od} q_{iod}. \end{aligned} \tag{1}$$

These transaction costs capture not only physical transportation costs and legal contracting costs but also factors such as hassle costs, regulatory costs, and political pressures – anything that introduces a wedge between the value of the good for a consumer's own use and its value in a transaction. Note that this assumption rules out fixed costs of trading, which simplifies the model. This approximation still reflects many aspects of California's water market: some percentage of water is lost in conveyance, larger transactions receive greater regulatory scrutiny, owners of canals and pipelines often charge per-unit fees ("wheeling" charges), and larger transactions may generate more political backlash.

I make one more key simplifying assumption:

Assumption 2. Perfect competition. *Each district has enough intermediaries that they behave as price takers. That is, the quantity sold or purchased by any one intermediary does not affect the equilibrium price for any district pair:* $dP_{od}/dq_{iod} = 0$ *for all* o *and* d.

In Appendix A, I relax this assumption and conduct my empirical analysis allowing for markups or markdowns by either buyers or sellers. There, I find little evidence that market power explains the observable transaction costs that I estimate below.

Sellers and buyers choose non-negative quantities for each origin and destination pair od to maximize net profits:

$$\begin{aligned} &(Sellers) \quad \max_{q^s_{iod}} P_{od} q^s_{iod} - V_o(Q_o) q^s_{iod} - \tau^s_{od} q^s_{iod} \quad \text{s.t. } q^s_{iod} \geq 0 \\ &(Buyers) \quad \max_{q^b_{iod}} V_d(Q_d) q^b_{iod} - P_{od} q^b_{iod} - \tau^b_{od} q^b_{iod} \quad \text{s.t. } q^b_{iod} \geq 0 \end{aligned} \tag{2}$$

Each problem has two candidate solutions. First, sellers and buyers may not trade at all ($q^s_{iod} = 0$, $q^b_{iod} = 0$). If sellers' transaction costs are too large, or marginal valuations in the origin district are too high ($V_o(Q_o) + \tau^s_{od} > P_{od}$), sellers will not trade. If buyers' transaction costs are too large, or marginal valuations in the destination district are too low ($P_{od} > V_d(Q_d) - \tau^b_{od}$), buyers will not trade. In order for trading quantities to be positive, there must be non-negative marginal surplus between the seller and buyer: $V_d(Q_d) - \tau^b_{od} \geq V_o(Q_o) + \tau^s_{od}$.

If buyers and sellers do trade, taking first-order conditions yields my main result:

$$\begin{aligned} &(Sellers) \qquad P_{od} = V_o(Q_o) + \tau^s_{od} \\ &(Buyers) \qquad P_{od} = V_d(Q_d) - \tau^b_{od}. \end{aligned} \tag{3}$$

That is, if two districts trade at all, the price between each pair of districts equalizes sellers' marginal willingness to accept with buyers' marginal willingness to pay. Relative to marginal valuations, the negotiated price gives a premium to the seller, and a discount to the buyer, that is exactly large enough to compensate for the transaction costs that each incurs. These are the key conditions I use in my empirical analysis. As I will show, I directly observe prices, but not marginal valuations or transaction costs. However, because marginal valuations vary only by district, while transaction costs vary by district pair, I can separately identify them empirically.

Trading quantities traded are defined implicitly by combining the inverse demand curves $V_o(Q_o)$

and $V_d(Q_d)$ with the market clearing condition: the sum of quantities sold by sellers must equal the sum of quantities bought by buyers. Intuitively, in partial equilibrium (holding constant transactions in all other pairs of districts), sellers and buyers keep increasing quantities until the marginal surplus is arbitraged away and the first-order conditions are met. In general equilibrium, quantities adjust so that each district's marginal valuation is equalized across all of the other districts it sells to or buys from.

Consumer surplus for each district is calculated by comparing demand curves with transaction expenditures or revenues. For a district that buys water, its consumers obtain consumer surplus equal to the difference between marginal willingness to pay and transaction expenditures, summed over all quantities purchased in excess of the endowment. For a district that sells water, its consumers obtain consumer surplus equal to the difference between transaction revenues and willingness to accept, summed over all quantities sold from the endowment.[2] Expressing these mathematically, consumer surplus at quantity Q_n is defined as:

$$\begin{aligned} &(Buying\ District) \qquad CS_d(Q_d) \equiv \int_{E_d}^{Q_d} V_d(\varphi)d\varphi - \sum_o P_{od}q_{od} - \sum_o \tau^b_{od}q_{od} \\ &(Selling\ District) \qquad CS_o(Q_o) \equiv \int_{E_o}^{Q_o} V_o(\varphi)d\varphi + \sum_d P_{od}q_{od} - \sum_d \tau^s_{od}q_{od} \end{aligned} \tag{4}$$

where $q_{od} \equiv \sum_i q_{iod}$ is the sum of quantities traded by all intermediaries for that district. Because at equilibrium a district's marginal values are equalized across transactions ($V_d(Q_d) = P_{od} + \tau^b_{od}$ and $V_o(Q_o) = P_{od} - \tau^s_{od}$, from Equation 3), consumer surplus can also be written as:

$$CS_n(Q_n) = \int_{E_n}^{Q_n} \left[V_n(\varphi) - V_n(Q_n)\right] d\varphi. \tag{5}$$

3.3 From theory to estimation

Using this theoretical framework, I now derive my empirical procedure. My goals are to (1) find the competitive equilibrium without transaction costs, and (2) calculate the resulting change in consumer surplus. With knowledge of the demand functions $V_n(Q_n)$ and initial endowments E_n, the competitive equilibrium (in a counterfactual efficient market) can be found by solving the social planner's problem. That is: choose Q_n for all districts n to maximize the sum of all districts' consumer surplus, subject to the resource constraint that the sum of quantities must be equal to the sum of initial endowments. From there it is straightforward to calculate consumer surplus, using Equation 5.

The main empirical task is to estimate demand curves, since initial endowments can be read directly from the data. Demand curves are comprised of marginal valuations at different values of quantity. Given equilibrium marginal valuations and quantities in multiple time periods, I can

[2]This is identical to buyer's consumer surplus, for negative quantities relative to the endowment. On a graph, the seller's consumer surplus appears similar to producer surplus, but it is determined from a demand curve, not a supply curve – sellers in an exchange economy do not produce anything.

estimate price elasticities and reconstruct these demand functions. To estimate marginal valuations in each time period, I rely on the result that prices equalize up to transaction costs, and add two more assumptions. At this point I introduce time subscripts t, allowing all variables and parameters to vary arbitrarily across time periods and making further restrictions when needed. (In my empirical application I define a time period as one year; this is a reasonable approximation since almost two-thirds of observed transactions are arranged in the same three-month period of May through July.)

3.3.1 Obtaining marginal valuations by estimating transaction costs from observable determinants

The first task is to estimate marginal valuations for all districts in each period, $V_{nt}(Q_{nt})$. Ideally, marginal valuations would be observed directly in the data – for example, from prices on intra-district water markets. Unfortunately, these internal prices are difficult to obtain, and most water districts appear to not keep systematic records. Most observations in my dataset are cross-district transactions.

Instead, I can estimate marginal valuations by comparing prices across a district's transactions in each period. Combining the result that prices equal marginal willingness to pay (Equation 3) with the fact that transaction costs are non-negative (Assumption 1), it follows that $P_{odt} \leq V_{dt}$ for all buying districts d: all prices paid by a buying district are less than its marginal valuation. Therefore, I can lower-bound a buying district's marginal valuation with the highest price it paid: $V_{dt} \geq \max_o\{P_{odt}\}$. By a parallel argument, all prices accepted by a selling district are higher than its marginal valuation ($P_{odt} \geq V_{ot}$ for all selling districts o) so I can upper-bound a selling district's marginal valuation with the lowest price it accepted: $V_{ot} \leq \min_d\{P_{odt}\}$.

However, I can improve on these bounds by leveraging information across transactions. Although the overall levels of transaction costs τ^s_{odt} and τ^b_{odt} are unobserved, portions of these costs may be due to observable cost determinants, such as conveyance distance or specific regulatory reviews. Many of these observable determinants vary across transactions, generating within-district variation that I can use to estimate determinant-specific transaction costs. Then, I can use these estimates to adjust raw prices and obtain tighter bounds on marginal valuations.

For example, suppose I observe Redding selling both to Sacramento at \$50 and to Los Angeles at \$70, where the sale to Los Angeles undergoes a regulatory review that the sale to Sacramento does not. I bound Redding's marginal valuation as no more than \$50 and interpret the price difference – \$20 – as the cost of this review nominally incident on Redding. Then, if I also observe Fairfield selling to Los Angeles for \$80, and I know this transaction undergoes the same regulatory review as the Redding–Los Angeles sale, I can leverage this information. I know that this regulatory review costs sellers \$20, so I can improve my upper bound of Fairfield's marginal valuation from \$80 to \$60.

To formalize this intuition, I first make a further assumption on the functional form of transaction costs.

Assumption 3. Determinant-specific transaction costs are additively linear and constant across

transactions. *Transaction costs can be decomposed as* $\tau^s_{odt} = \tau^s \mathbf{B}_{od} + \tilde{\tau}^s_{odt}$ *(for sellers) and* $\tau^b_{odt} = \tau^b \mathbf{B}_{od} + \tilde{\tau}^b_{odt}$ *(for buyers), where* $\mathbf{B}_{od}$ *is a vector of observable transaction cost determinants and* τ^s *and* τ^b *are vectors of coefficients, and all parameter and cost determinant values are non-negative.*

This carries three substantive restrictions: first, transaction costs arising from specific cost determinants $\mathbf{B}_{od}$ are constant across consumers with the same values of $\mathbf{B}_{od}$; second, they are constant across time; and third, cost determinants do not interact with each other – when a transaction is subject to multiple cost determinants, the total costs equal the sum of their parts. In principle I could relax either the first or the second restriction, but not both; for τ^s and τ^b to be identified, they must be constant in at least one dimension. The third restriction in principle could be relaxed nonparametrically.

Next, recall that Equation 3 relates observed prices to transaction costs and marginal valuations. Both of the latter are unobserved in levels, but some components of transaction costs vary observably across transactions while marginal valuations do not. Combining that result with Assumption 3 yields

$$\begin{aligned} (Sellers) \qquad P_{odt} &= \alpha_{ot} + \tau^s \mathbf{B}_{od} + \varepsilon^s_{odt} \\ (Buyers) \qquad P_{odt} &= \alpha_{dt} + \tau^b \mathbf{B}_{od} + \varepsilon^b_{odt}, \end{aligned} \tag{6}$$

where the α's collect terms fixed within consumer and year ($\alpha_{ot} \equiv V_{ot} + \mathbb{E}_d[\tilde{\tau}^s_{odt}]$; $\alpha_{dt} \equiv V_{dt} + \mathbb{E}_o[\tilde{\tau}^b_{odt}]$) and the ε's collect miscellaneous terms ($\varepsilon^s_{odt} \equiv \tilde{\tau}^s_{odt} - \mathbb{E}_d[\tilde{\tau}^s_{odt}]$; $\varepsilon^b_{odt} \equiv \tilde{\tau}^b_{odt} - \mathbb{E}_o[\tilde{\tau}^b_{odt}]$) that I will treat as econometric errors.

Consistent estimators of τ^s and τ^b, which I describe in Section 5, then will allow me to more strongly bound marginal valuations. Before considering transaction costs, my best upper bound for sellers was the lowest price: $V_{ot} \leq P_{odt}\ \forall d$. Now, I can improve on this by adjusting these raw prices with estimated transaction costs:

$$\begin{aligned} (Sellers) \quad V_{ot} &\leq P_{odt} - \tau^s \mathbf{B}_{od} \quad \forall d \\ (Buyers) \quad V_{dt} &\geq P_{odt} + \tau^b \mathbf{B}_{od} \quad \forall o, \end{aligned} \tag{7}$$

which I obtain from Equation 3 and the assumption that all components of transaction costs are weakly positive (Assumption 3). Note that τ^s and τ^b are constant effects under Assumption 3, so these bounds hold exactly, not in expectation.

Finally, I assume that these bounds are "good enough," which allows me to interpret both marginal valuations and gains from trade as point estimates rather than bounds:

Assumption 4. Unobserved transaction costs are zero for at least one transaction. *Unobserved transaction costs* $\tilde{\tau}^s_{odt}$ *and* $\tilde{\tau}^b_{odt}$ *are zero for the transaction in each district and period with the lowest (for selling districts) or highest (for buying districts) adjusted price.*

This assumption implies that Equation 7 holds with equality for at least one transaction per district and period:

$$\begin{aligned} (Sellers) \qquad V_{ot} &= \min_d \{P_{odt} - \tau^s \mathbf{B}_{od}\} \\ (Buyers) \qquad V_{dt} &= \max_o \{P_{odt} + \tau^b \mathbf{B}_{od}\}. \end{aligned} \tag{8}$$

If this assumption is not true, then I will understate the dispersion in marginal valuations. Because more dispersion means more mutually beneficial transactions, my estimates of the potential gains from trade would likely be a lower bound. However, this assumption may not be unreasonable, since my empirical estimates of τ^s will be large enough to put some sellers' marginal valuations V_{ot} near zero.

Note that this procedure accounts for both observed and unobserved transaction costs. Observed transaction costs are reflected by adjusting prices for estimated costs of *observable* cost determinants. Even after adjusting prices, though, prices may vary across transactions due to differing *unobservable* cost determinants. Taking the minimum (or maximum) across these adjusted prices accounts for (at least some) unobserved transaction costs.

3.3.2 Estimating demand elasticities

My next task is to construct demand curves from these marginal valuations; specifically, inverse demand functions $V_n(Q_n)$ that describe how marginal valuations vary with quantities. The previous procedure yields marginal valuations in each period V_{nt}, which I can use along with temporal variation in quantities Q_{nt} to estimate the function $V_{nt} = f_n(Q_{nt})$. This can be done using any consistent estimator.

In practice, the classic endogeneity problem posed by the joint determination of prices and quantities makes it difficult to find such a consistent estimator. To overcome this, I instrument for quantities using yearly water entitlements. Entitlements are driven by the interaction of weather fluctuations with historically-determined allocation rules, making their year-to-year variation likely exogenous. The disadvantage of this quasi-experimental approach is that it cannot approximate $f_n(Q_{nt})$ very flexibly. I will assume that demand is isoelastic in the range between observed and optimal quantities, with constant elasticities per sector.

3.3.3 Finding the efficient allocation and calculating consumer surplus

Finally, upon obtaining demand curves $V_n(Q_n)$, I can find the efficient allocation by solving the social planner's problem. This is a constrained optimization exercise in which I choose Q_n for all districts n to maximize total consumer surplus, subject to the resource constraint that the sum of quantities must be equal to the sum of initial endowments. The total gain in consumer surplus is equal to the potential gains from an efficient water market without transaction costs (neither observed nor unobserved).

If Assumption 4 does not hold and instead there are unobserved transaction costs that affect all transactions, my estimate of this solution will be a lower bound of the true potential gains. This is because my estimates of V_n will be biased, understating the true dispersion in marginal valuations. To see this graphically in Figure 3, consider the efficient equilibrium between buying district d and selling district o calculated using their true demand curves, labeled with price P'. If I omit some unobserved transaction costs, I may underestimate marginal valuations for buyers and overestimate them for sellers, leading me to obtain the demand curves shown in dashed lines. Using these demand curves to calculate the efficient allocation would result in the equilibrium labeled with price P_{od}. The resulting consumer surplus is the triangle to the left of the dashed demand curves. Relative to this, the true equilibrium at P' reallocates a greater quantity of water, creating a larger area of consumer surplus.

I also note two implications of using consumer surplus as my welfare measure. First, it relies on the Pareto welfare criterion, which treats any voluntary transaction as inherently good and rules out any notion of merit goods – in which society might value certain parties holding possession of water more than the parties themselves do. Second, it does not account for any externalities associated with reallocation. Quantifying the externalities of water markets is beyond the scope of this paper, but a complete policy analysis should account for them.

3.3.4 Summary of empirical procedure

Summarizing the discussion in this section, I propose an empirical procedure to estimate the potential gains from trade in a thin exchange economy with transaction costs. This procedure has four steps:

- **Step 1: Estimate transaction costs from observable cost determinants.** Find cost determinants that are both observable and heterogeneous across transactions. Estimate the transaction costs associated with these cost determinants from variation in prices across transactions within consumer and year.

- **Step 2: Recover equilibrium marginal valuations.** First adjust prices by these observable transaction costs, applying the knowledge that consumers must be compensated for transaction costs in order to agree to the transaction. Then adjust for unobservable transaction costs by taking each buyer's highest adjusted price paid, and each seller's lowest adjusted price accepted, as final estimates of marginal valuations.

- **Step 3: Estimate demand curves.** Combine these equilibrium marginal valuations with data on quantities to measure the relationship between them. Use instrumental variables to overcome the problem of joint determination and obtain consistent estimates of demand parameters.

- **Step 4: Simulate counterfactual allocations.** Combine demand curves to find the quantity vector that maximizes total consumer surplus, and calculate the resulting consumer surplus.

A remaining concern is that some of the estimated price gaps may in fact represent markups, rather than true transaction costs. To see this, consider a decomposition $P_{odt} = V_{ot}(Q_{ot}) + \tau^s_{odt} + \mu^s_{odt}$, where $\mu^s_{odt} > 0$ is an ologopolistic markup; here, markups are empirically isomorphic to transaction costs. The assumptions of my model rule out the successful exercise of market power by either side of the market, but if these assumptions are false, then my estimation procedure for marginal transaction costs would also pick up differences in markups that are correlated with the transaction cost determinants. Such markups could arise if, for example, most sellers in a low-valuation region find it prohibitively expensive to sell to a high-valuation region, leaving buyers in the high-valuation region with access to only a small number of effective potential sellers.

I explore this possibility in Appendix A. There, I present an alternative model that allows both buyers and sellers to exploit market power, and I derive and perform an empirical procedure based on Atkin and Donaldson (2015) that adjusts raw prices for market power, using estimated pass-through rates as sufficient statistics. Then I repeat Step 1, again estimating marginal transaction costs, but using these adjusted prices. I find that estimates are noisier, but if anything, marginal transaction costs are larger. I conclude that the issue of market power is small relative to transaction costs, and so I proceed with the assumption of perfect competition.

4 New Data on California's Water Economy

To conduct my empirical analysis, I compile new data on California's water economy. I assemble for the first time the universe of yearly surface water allocations in California. I also use a proprietary dataset on open-market water transactions in California that is significantly more complete than publicly available datasets. To link across datasets and complete the analysis, I build a large crosswalk file, a geospatial dataset on user locations and boundaries, and a directed-graph hydrological network model of California's water infrastructure. For further analysis of agricultural outcomes (Appendix **??**), I further incorporate two uniquely high-resolution datasets on agriculture in California: (1) a satellite-derived remote sensing product for land use, and (2) microdata from the Census of Agriculture for agricultural finances.

4.1 Water transactions

No government agency or other institution maintains a centralized listing of all water transactions in California. Instead, I use a proprietary dataset compiled by WestWater Research, LLC. To my knowledge, this is the most complete dataset of water transactions in California, and it has been used in federal regulatory impact analyses. One water transactions dataset has been assembled and made publicly available by Gary Libecap at the University of California, Santa Barbara (e.g., Brewer et al. 2008); however, it has only a fraction of the transactions in California as the WestWater dataset. Another dataset has been assembled by Ellen Hanak at the Public Policy Institute of California, but it is not public. This dataset also appears to focus on transaction volume rather than prices (Hanak and Stryjewski 2012), whereas prices are mostly complete in the WestWater dataset.

I focus on (1) surface water transactions as opposed to groundwater, (2) the spot market (within-year leases) as opposed to permanent transfer of rights, (3) freely negotiated transactions as opposed to those where prices are set administratively, and (4) transactions involving at least one party that diverts water for consumption, as opposed to instream uses or storage. The WestWater dataset includes 6,264 transactions in total, but most are groundwater leases within adjudicated basins. Prices are available for 671 transactions that remain after applying these inclusion criteria. Table 1 shows summary statistics. Panel A shows that the distribution of volumes is highly dispersed and the mean price is $221 per acre-foot, in 2010 dollars. Panel B shows that the Sacramento River hydrologic region is the greatest net exporter and the Tulare Lake and South Coast regions are the greatest net importers. Figure 4 shows the distribution of prices on a base-10 logarithmic scale; transactions are centered around $100-300 but have substantial mass in the tails.

I identify the geographic location and sector (urban/municipal, agricultural, or environmental) of most buyers and sellers in the data, using several methods described in Appendix F.

4.2 Surface water deliveries and endowments

I assemble the universe of surface water endowments and deliveries in California, by user, sector, and year, from 1980 through 2016 (for endowments, and 1993-2016 for deliveries). By endowments I mean water that institutions (or individuals) are legally entitled to obtain on the basis of property rights or fixed long-term contracts. There are four sources of endowments: Central Valley Project (CVP) allocations, State Water Project (SWP) allocations, Lower Colorado Project entitlements, and surface water rights. Deliveries are the quantities of water actually taken or received from each of these four sources; they represent the best available wholesale-level estimate of surface water consumed. Deliveries can differ from endowments because of surface water transactions (either spot market or permanent), users voluntarily taking less than their endowments, or a variety of other programs that allow users to receive more or less water under various circumstances.

Endowments are calculated by multiplying a baseline maximum quantity by a year-varying allocation percentage. These allocation percentages, set according to weather and hydrologic conditions, are determined yearly for each of 14 separate contract types in the CVP and SWP. For the CVP, allocation percentages vary across both years and contract types; for the SWP, allocation percentages are constant across users, varying only across year. For water rights and the Lower Colorado Project, allocation percentages are always 100 percent.

Deliveries, maximum contract amounts, and yearly percentage allocations come from archives of the California Department of Water Resources (DWR) and U.S. Bureau of Reclamation (USBR). For surface water rights, I use self-reported diversions collected by the State Water Resources Control Board (SWRCB) rather than the face value of rights, which are often outdated for post-1914 appropriative rights and are not recorded for pre-1914 or riparian rights. This compilation uses recently available data made possible by a law that required all surface water rights holders to report their water use starting in 2010. It is reasonable to treat these reported diversions as the full, legally defensible value of present water rights, since they are public information and thus could potentially be used in future legal disputes.

Details of sources, cleaning, and processing of these variables are described in Appendix F.

4.3 Hydrologic network

To calculate characteristics of transaction conveyance paths, and to allow for physical transaction costs in counterfactual simulations, I construct a model of California's hydrological network. This is a set of nodes and edges corresponding to all major water conveyance channels in California: rivers, canals, aqueducts, and pipelines. Routes come from the National Hydrography Dataset of the U.S. Geological Survey. Each node and edge is parameterized with physical transportation costs drawn from the literature: pumping costs (for the energy required to lift water), and conveyance losses (to percolation, evaporation, and required outflows from the Sacramento–San Joaquin Delta). I run a graph-theory algorithm to obtain least-cost delivery pathways for a matrix of units defined by geography and sector (see Section 8 for details of this definition). As compared with more detailed hydrologic models such as CALVIN (e.g., Howitt et al. 1999), this model lacks comprehensive information on capacity constraints.

4.4 Crosswalk file and user locations file

To link users across datasets, I build a crosswalk file that accounts for variations and errors in names as well as mergers and name changes across time. This file has 28,764 entries (input names) pointing to 14,830 targets (output names). To identify the locations, boundaries, and areas of water users, I combine several publicly available shapefiles into a single geospatial dataset. Details of the sources and construction of these files are described in Appendix F.

5 Step 1: Estimating Transaction Costs from Observable Determinants

The first step in my empirical approach is to estimate marginal transaction costs from observable determinants. To do this, I find cost determinants that are both observable and heterogeneous across transactions, and then I estimate the transaction costs associated with these cost determinants from variation in prices. The intuition is that in equilibrium, a consumer selling to two buyers is indifferent between them at the margin, so any difference between them can be interpreted as marginal transaction costs.

5.1 Selecting cost determinants

Transaction costs may arise from a litany of cost determinants, which may fall into physical, administrative, policy-induced, or political-economy categories. I start with the typology of transaction cost determinants in Section 2.2, compiled from institutional knowledge. For some of these cost determinants, there are populations of both sellers and buyers for whom the determinant is incident on some trading partners but not others. Comparing prices across these trading partners, within these populations of sellers and buyers, can identify the marginal transaction costs.

From this list, I select the cost determinants that are both (a) observable, and (b) heterogeneous across trading partners. These are the cost determinants whose marginal transaction costs can be econometrically identified. Cost determinants that are common to all of a consumer's transactions (such as the costs of writing a contract, or disutility from market participation), are excluded because they are indistinguishable in price data from a level shift in marginal valuations.

In Table 2, I list and define the cost determinants that meet these criteria. They are: elevation lift, distance conveyed (in rivers, canals, and virtually), crossing the Sacramento–San Joaquin Delta, importing into or exporting from a federal or state waterr project, regulatory review by the State Water Boards, or whether the counterparty is a predominantly agricultural user.

5.2 Econometric specification

I stack the selected cost determinants into a vector $\mathbf{B}_{od}$, depending on seller (origin) o and buyer (destination) d. Most of these cost determinants are discrete so I use binary indicator variables; distances and elevation enter linearly.

To identify marginal transaction costs nominally incident on sellers, I regress price P_{jodt} in transaction j in year t on this vector $\mathbf{B}_{od}$ and seller-by-year fixed effects. Similarly, to identify costs incident on buyers, I use buyer-by-year fixed effects. The resulting regressions precisely follow the identification result derived earlier (Equation 6 in Section 3.3.1):

$$
\begin{aligned}
(Sellers)\ P_{jodt} &= \alpha_{ot} + \tau^s \mathbf{B}_{od} + \varepsilon^s_{jodt} \\
(Buyers)\ P_{jodt} &= \alpha_{dt} + \tau^b \mathbf{B}_{od} + \varepsilon^b_{jodt},
\end{aligned}
\tag{9}
$$

The seller regression measures price gaps across transactions within seller and year. That is, it isolates cases in which a seller completed transactions with two different buyers in the same year, of which one was subject to a particular cost determinant and the other was not. Each coefficient τ^s_h (on the hth cost determinant indexed in $\mathbf{B}_{od}$) then is a regression-weighted average of the price differences in all such cases. The intuition is that if the seller has to pay greater transaction costs to complete a transaction, the price received must be higher to compensate.

For this regression to produce an unbiased estimate of τ^s_h, I also need to assume selection on observables: that unobserved determinants of prices are uncorrelated with the cost determinants $\mathbf{B}_{od}$. In the special case where all other components of transaction costs are equal except those associated with the cost determinants $\mathbf{B}_{od}$, this condition follows immediately from my model. This is because in equilibrium, prices equalize marginal willingness to accept across trading partners, so prices are affected only by marginal valuations and transaction costs. Marginal valuations are absorbed by the seller-by-year fixed effects, as are transaction costs common to all transactions, leaving only variation in transaction costs that differ across transactions.

In the more general case where transaction costs vary arbitrarily across transactions, the assumption of mean independence is stronger. It requires that, conditional on seller and year, the selected cost determinants are uncorrelated with other unobserved cost determinants (or, without the assumptions of the model, other determinants of price). In robustness checks, I partially test this assumption by exploring whether the coefficient estimates change when I include other covariates.

The buyer regressions have exactly mirrored conditions and interpretations: τ_h^b measures differences in prices paid by a buyer to different sellers. I cluster standard errors in all regressions at the level of subregion-by-year[3] to allow for local spatial correlation. In some specifications, I use coarser fixed effects to explore robustness to specification.

Finally, I perform variable selection using the least absolute shrinkage selection operator (LASSO) to choose the best subset of cost determinants. Many factors are plausible determinants of transaction costs, yet not all may necessarily be empirically important. Using LASSO can reduce concerns about low statistical power, overfitting noisy data, and researcher degrees of freedom. If none of the proposed cost determinants have much predictive power for prices, they will all be zeroed out.

5.3 Results

LASSO-selected cost determinants First, Table 3, Panel B shows the marginal transaction costs associated with factors selected by LASSO. Columns (4) and (8) present my preferred specifications, with seller-by-year and buyer-by-year fixed effects. They show that several of these factors are indeed costly: for example, prices received by sellers are $76 per acre-foot higher for transactions that must cross the Sacramento–San Joaquin Delta, relative to other transactions by the same seller in the same year. Similarly, prices paid by buyers for transactions that export water from a federal or state water project are $194 per acre-foot lower than other transactions. These estimates are quite large, considering the mean price in the sample is $221.

Other factors that appear to be costly for sellers are distance conveyed in rivers and having an agricultural buyer. For buyers, distance of virtual conveyance and review by the State Water Boards appear to be costly. Not all coefficients are statistically distinguishable from zero, but the fact they were selected by LASSO suggests they matter for the overall model. The importance of some of these factors is also supported by the specifications with coarser fixed effects (columns 2-3 and 6-7), in which many of the same factors are selected, and the Delta-crossing coefficient is larger and more precise. Across these specifications, all coefficients selected by LASSO are positive for sellers and negative for buyers, offering internally consistent evidence that these regressions are measuring the expected marginal transaction costs.

All observable cost determinants In Panel A of Table 3, I show the results for all observable cost determinants, for completeness. Columns 1 and 5 show coefficients from separate regressions in which the vector $\mathbf{B}_{od}$ includes only one known cost determinant at a time. Columns 2-4 show the results from regressions in which all known cost determinants are included together. Specifications are otherwise identical to the corresponding columns of Panel B.

[3] I use "subregion" to refer to planning subarea (PSA) as defined by the California Department of Water Resources. California has 46 PSAs.

When including all proposed cost determinants, many coefficients are not statistically distinguishable from zero, and some have the opposite sign as expected. However, the overall patterns hold up and remain broadly consistent across specifications. Buyers require price discounts, and sellers require price premiums, to choose transactions subject to these cost determinants.

5.4 Discussion

This exercise estimates the price effects of several cost determinants that are both observable and heterogeneous across trading partners. This is a limited analysis of transaction costs for several reasons. One, this exercise misses cost determinants that are not econometrically identified in price data; many other types of transaction costs may be important as well, but my approach cannot measure them. Two, my specification only measures marginal transaction costs, capturing transaction costs that scale with the size of the transaction. If there are also fixed costs, either they are not picked up at all, or they may partially load onto the estimate of marginal transaction costs (i.e., as average transaction costs).

I emphasize that the estimated transaction costs are not necessarily limited to direct costs of regulatory reviews themselves. This analysis may capture other kinds of transaction costs – including physical, administrative, or political economy costs – that are collinear with the regulatory reviews. These estimates represent the marginal cost of selling or buying across a given institutional or geographic boundary, inclusive of all causes. In addition, despite my focus on costs, these regulatory reviews may bring important social benefits by preventing environmental externalities.

I next use these estimated transaction costs to obtain districts' marginal valuations. This relies on the notion that a seller will not accept a price unless it covers both the transaction costs and the seller's marginal valuation. For this exercise, I need to interpret my estimates as constant effects – costs incurred equally by everyone, as in Assumption 3 – not as average treatment effects. If the marginal transaction costs associated with my selected cost determinants instead differ across transactions, my results may be biased in ambiguous ways.

6 Step 2: Recovering Equilibrium Marginal Valuations of Water

Next, I combine these estimated transaction costs with observed prices to estimate marginal valuations of water in each district and year. I use a simple revealed-preference condition: a buying district's marginal valuation must be as high as its highest price paid, while a selling district's marginal valuation must be as low as its lowest price paid, after adjusting for the transaction costs estimated in Step 1. This is motivated by the structural knowledge that transaction costs insert a wedge between observed prices and marginal valuations: to compensate for higher transaction costs, a buyer requires a discount, while a seller requires a premium. Adjusting prices for estimated transaction costs corrects for *observed* transaction costs, while taking the minimum (for sellers) or maximum (for buyers) across transactions within a year corrects for some *unobserved* transaction costs.

To do so, I adjust observed prices for estimated transaction costs and then take the minimum (for sellers) or maximum (for buyers) of these adjusted prices within each district-year cell, as in equation 8:

$$\begin{aligned} (Sellers) \qquad \hat{V}_{ot} &= \min_j \left\{ P_{jodt} - \hat{\tau}^s \mathbf{B}_{od} \right\} \\ (Buyers) \qquad \hat{V}_{dt} &= \max_j \left\{ P_{jodt} + \hat{\tau}^b \mathbf{B}_{od} \right\} \end{aligned} \tag{10}$$

More specifically, I first classify districts as either net buyers or net sellers in each year on the basis of transaction volume (the vast majority only buy or sell, not both). Next, to calculate transaction costs $\tau^s \mathbf{B}_{od}$ and $\tau^b \mathbf{B}_{od}$, I use the disaggregated regressions from Step 1. From these regressions, I obtain fitted values from the cost determinants $\mathbf{B}_{od}$ only, not the fixed effects or controls. Then, I obtain adjusted prices by adding these transaction costs to the raw prices, imposing free disposal by censoring at zero when occasionally necessary. Finally, I obtain marginal valuations for each seller by taking the minimum across these adjusted prices within each year, and marginal valuations for each buyer by taking the maximum.

Figure 5 plots the kernel density of these marginal valuations, with the density of raw prices for comparison, for the district-years in which at least one price is observed. This exercise reveals that marginal valuations have even greater dispersion than observed prices, even though prices are quite dispersed themselves. Note that these marginal valuations are not yet directly comparable to each other: the graph pools multiple years, districts do not appear in the data every year, and districts select into the observed market according to their marginal valuations.

Still, this dispersion suggests that gains are available from reducing transaction costs, both observable and unobservable. That is, in a competitive equilibrium with smaller transaction costs, some consumers would purchase more water from other consumers, and both sides would enjoy welfare gains. To quantify these gains, I need to know how marginal valuations vary with quantity consumed, in order to extrapolate marginal valuations out of sample. In other words, I need a model of demand.

7 Step 3: Estimating Demand Elasticities

The third step in my empirical approach is to estimate price elasticities of demand for surface water in the wholesale market. Marginal valuations give me points on demand curves, but now I need demand elasticities in order to describe how steeply marginal valuations evolve with quantity.

To estimate demand elasticities, I exploit California's historical system of water allocation, which translates a given precipitation shock into vastly different supply shocks for different water users. Specifically, I use determinants of yearly surface water endowments as instruments for price. These initial endowments are relevant instruments because of transaction costs, which create inertia. In a frictionless market, the Coase Theorem suggests that the final allocation of quantities would depend only on demand, not on initial endowments. In a market with transaction costs, Coasian independence does not hold, and endowments persist.

I construct instruments by allowing each determinant of endowments to separately affect each agent's prices and quantities. Because this yields a large number of potential instruments, I then select optimal instruments using the IV Lasso algorithm of Chernozhukov et al. (2015). I omit year effects, which introduce bias by comparing agents within the same market, since trading creates mechanical spillovers. Although I can no longer control for unobserved statewide shocks to demand, the most important time-varying factors are likely related to water availability, which is flexibly captured by the instruments.

I estimate demand at the level of units, categories defined by geography and sector.[4] This aggregation serves two purposes. First, it ensures accurate matches between water deliveries and transactions data.[5] Second, shifting to unit-level analysis allows me to sidestep the issue of selection into the market. Individual district-level demand is difficult to estimate because I only observe prices for years in which a district chooses to trade. Instead of modeling both the intensive and extensive margin of trading, I estimate aggregate demand curves that subsume both margins. Unit-level marginal valuations can be viewed as the marginal valuation of the marginal district.

The approach I propose may be useful in estimating demand elasticities in settings beyond water markets. Endowments are often a convenient instrument in cap-and-trade and other exchange economies, and they may be assigned partly based on arbitrary factors such as historical conditions or fixed rules. Even in settings with fewer transaction costs than water markets, endowments may still persist because of behavioral factors or market failures. At the same time, spillovers mediated through market interactions are widespread but rarely addressed.

7.1 Empirical strategy

Estimating demand presents a classic endogeneity problem: quantities and prices are equilibrium outcomes jointly determined by the market interaction of all agents' demand curves. I need a set of supply shifters: instruments that create exogenous variation in the effective supply curves faced by each unit.

For instruments I use the yearly allocation percentages within the Central Valley Project and State Water Project. Each year, water users holding delivery contracts with these projects receive a specified percentage of their maximum contract amounts. Government agencies determine these allocation percentages each year, separately for each of 13 contract types, on the basis of reservoir levels and other environmental conditions – which in turn are primarily determined by precipitation during the previous winter.[6] These rules allow me to isolate variation driven by natural phenomena,

[4] That is, interactions of planning areas (hydro-geographical areas defined by the California Department of Water Resources), and sector (urban or agricultural).

[5] Often a wholesale district receives water deliveries but does not appear in transactions data, while a geographically-overlapping retail district has no water deliveries yet is recorded as selling water. This is because wholesale and retail water districts often have long-term purchasing agreements and other complex linkages, which are difficult to track. Aggregating to unit simplifies such situations, since it treats districts that share jurisdictions together as one agent.

[6] Water users who do not hold delivery contracts with either of these projects (i.e., contractors of the Lower Colorado Project or holders of appropriative or riparian rights) have never been required to significantly reduce their water use, so their allocation percentages are 100%.

not demand-side factors. Figure 7 plots these allocation percentages, aggregated to several regional categories.

One tempting approach might be to use each unit's own allocation percentage as the instrument, while controlling for unit and year fixed effects. This works in many settings, when each unit is an independent market, so units can serve as plausible counterfactuals for each other. However, in this setting, all units are connected by a single market, so changes to any one unit's prices and quantities alter the equilibrium and affect the prices and quantities of others. Here, year fixed effects violate the stable unit treatment value assumption (SUTVA) and introduce bias, which I prove in Appendix E.1. The intuition is that year effects difference out average outcomes, but average outcomes are themselves also affected by the instrument, so the estimated coefficient fails to isolate the individual treatment effect.

My solution is to use time-series variation and avoid within-year comparisons altogether. Instead of year fixed effects, I interact each of the 13 allocation-percentage series with unit indicator variables, such that each unit's endowments are allowed to have separate effects on each other unit's outcomes. For example, one unit's endowments may have a strong effect on its own outcomes, a moderate effect on the outcomes of a unit with which it frequently trades, and no effect on other units. This approach avoids bias from SUTVA violations because it analyzes each unit in isolation, simply pooling the final parameter.

This approach also handles the most obvious threats of omitted variables bias raised by the lack of year fixed effects. Allocation percentages are correlated, so if I used each unit's own allocation percentage as a single instrument, I would capture effects actually caused by other units' endowments. Using the full set of interactions allows me to flexibly model the effects of year-to-year changes in water availability in different parts of the state.

In the simplest specification, I regress the natural log of surface water quantity delivered for unit k in year t on log prices of transactions (made within that unit and year) and unit fixed effects:

$$\ln Q_{kt} = \eta \ln P_{jnklt} + \psi_k + \nu_{jnklt}. \tag{11}$$

Transactions j are made by agent n with a counterparty in unit l. The preferred specification also includes agent fixed effects and unit pair-specific fixed effects (to absorb differences in transaction costs across transactions within a unit), as well as unit-specific time trends (to control for changes in demand over time):

$$\ln Q_{kt} = \eta \ln P_{jnklt} + \psi_{nk} + \zeta_{kl} + \theta_k t + \nu_{jnklt}. \tag{12}$$

To avoid bias from spillovers, all coefficients are unit-specific; only the final elasticity η is pooled.

I use prices instead of marginal valuations in order to use all information in the data at this stage, and to avoid introducing estimation error from previous stages of analysis. Because observed prices differ from marginal valuations only by the amount of transaction costs, it suffices to control for the determinants of transaction costs across transactions within a unit. (In robustness checks, I use marginal valuations instead of prices and find that the results are not significantly different.)

I instrument log prices with log allocation percentages $\mathbf{z}_{kt}$, each fully interacted with unit and region indicator variables. Precisely stated, the instruments are $\mathbf{Z}_{kt} \equiv \mathbf{z}_{kt} \otimes \Gamma_k = (z_{1t}\Gamma_1 + z_{2t}\Gamma_1 + \ldots) + (z_{1t}\Gamma_2 + z_{2t}\Gamma_2 + \ldots) + \ldots$, where Γ_k are the indicators. These interactions allow the determinants of each unit's endowments to have separate effects on each other unit's prices and quantities.

These interactions yield a large number of instruments relative to the number of observations.[7] To avoid model overfitting and weak instruments, I estimate the model via post-Lasso two-stage least-squares, following the IV Lasso algorithm of Chernozhukov et al. (2015) as implemented in Stata by Ahrens et al. (2018). The goal is to choose enough instruments to avoid omitted variables bias, but not so many that the first stage is overfitted. IV Lasso uses data-driven penalization to choose an optimal subset of instruments that matter empirically.

Although the level of observation is transaction-agent, the regression estimates unit-level elasticities, as quantities are measured at the unit level. Accordingly, I cluster standard errors by unit-year, the level of variation in both my instruments and outcome variable. Because only the endogenous variable varies within unit-year, omitting fixed effects would make this regression equivalent to one that uses unit-year observations and mean prices.

Identification assumption The identification assumption has two key pieces. The first piece is conditional independence: conditional on covariates, changes in allocation percentages are not correlated with any other time-varying factors that independently affect prices or quantities. The second piece is the exclusion restriction: changes in allocation percentages affect quantities only through movements along demand curves, not through shifts in demand curves. In other words, conditional on observed prices and covariates, allocation percentages have no additional effects on quantities. Conditional independence ensures that the first stage and reduced form relationships are free from omitted variables bias, and the exclusion restriction ensures that the IV estimate can be interpreted as a causal relationship.

Conditional independence is a plausible assumption. Unit fixed effects absorb the influence of typical water availability, so the elasticity is estimated using only year-to-year variation in allocation percentages within each unit. These fluctuations are determined by mountain precipitation and reservoir conditions and cannot be manipulated by water users. In a different setting, one omitted variable might be local weather patterns, but in California, local rainfall meets a vanishingly small proportion of water demand. If a unit's own allocation percentage were the only instrument, another omitted variable might be water supplies in other parts of the state, since they are correlated and can all affect equilibrium outcomes, but I avoid this problem by using the full set of allocation percentages as instruments for prices faced by each unit. Demand may change over time in unobserved ways, but I control for linear time trends that vary by region. I cannot capture idiosyncratic shocks to demand, but these will not bias the results as long as they are not correlated with the instruments.

[7]There are 13 series of allocation percentages (each corresponding to a different project contract type), 9 regions, and 62 units, giving a total of 923 potential instruments. The contract types are: SWP Agricultural, SWP Municipal, CVP North of Delta Agricultural, CVP North of Delta Urban, CVP North of Delta Settlement Contractors, CVP American River Urban, CVP In Delta (Contra Costa), CVP South of Delta Agricultural, CVP South of Delta Urban, CVP South of Delta Settlement Contractors, CVP Eastside Division, CVP Friant Class 1, and CVP Friant Class 2.

The exclusion restriction is also plausible in this setting. Intuitively, percentage allocations are pure supply shocks. Increasing one unit's surface water endowment will increase quantities and decrease equilibrium prices, moving along demand curves without changing underlying preferences. Increasing other units' endowments will lower their marginal valuations, raising quantities traded, decreasing equilibrium prices and again increasing the first unit's quantity.

A potential threat to the exclusion restriction is substitution to groundwater or storage. A decrease in surface water endowments may lead a unit to extract more groundwater, reducing surface water quantity less than would occur otherwise. However, this need not violate the exclusion restriction, which simply requires that any endowment-driven changes in quantities also be reflected in prices. In a simple model, agents extract groundwater until the marginal cost equals the marginal valuation of water. Suppose the cost of groundwater extraction does not depend on surface water endowments, and groundwater is perfect substitute for surface water in the short run. Under these plausible conditions, year-to-year changes in groundwater quantities are fully determined by changes in the marginal valuations of water, and they need not enter demand as a separate term. Estimated elasticities measure the response of surface water quantities to prices, regardless of any shifts in groundwater extraction.

7.2 Results

Table 4 reports the results of these regressions, estimating demand for surface water in the wholesale market. As a first check that allocation percentages are indeed strong instruments, I estimate first stage and reduced form regressions by ordinary least squares using a single instrument: own-unit allocation percentage. The results are shown in Panel A for illustrative purposes; they are not used for further analysis because they do not account for the possible effects of other units' endowments. Columns 1 and 4 use the simplest specification (Equation 11), columns 2 and 5 include additional terms, and columns 3 and 6 use the preferred specification (Equation 12).

Across all specifications, the instrument appears to be relevant and strong. Effects are estimated fairly precisely and point in the expected directions: higher allocations decrease prices and increase quantity consumed. F-statistics exceed the standard rule-of-thumb values, although not by a large amount in the first stage; including additional Lasso-selected instruments will help. The much larger F-statistics for the reduced form are also reassuring, since in just-identified two-stage least squares the roles of the first stage and reduced form can be swapped: the elasticity is estimated by one ratio, while the inverse elasticity can be estimated by its reciprocal.

Panel B shows results of instrumental variable regressions, in which the IV Lasso algorithm chooses instruments from the full set of potential instruments. In column 1, the simplest specification with only unit fixed effects, the Lasso selects 10 instruments, and the instrumented effect of log prices on log quantities is -0.14. This is directly interpretable as an elasticity, implying that a 10% change in price results in a 1.4% change in quantity. Columns 2-6 report results from specifications that include additional sets of fixed effects and covariates. In the preferred specification (column 6), the Lasso selects 19 instruments, and the estimated elasticity is -0.20. Point estimates appear to be relatively stable across specifications, despite the Lasso selecting as many as 34 or as few as 7 instruments.

Although all first-stage F-statistics appear to be large, I also report the results of the sup-score test for weak instruments recommended by Belloni et al. (2012). Two of the six specifications fail to reject the null hypothesis, indicating they do not pass the test. However, this test is known to be conservative, and the point estimate from these specifications are not significantly different from the others.

In Panel C, I explore heterogeneity in demand elasticities. Each column estimates the preferred specification for a different subset of the sample. I find that urban units have a smaller elasticity (-0.10) than agricultural units (-0.23). Breaking up the agricultural sector into three regions, I find substantial heterogeneity even within this sector, with elasticities increasing from north to south. Elasticities are very small in the Sacramento Valley (-0.07), still somewhat small in the San Joaquin Valley (-0.11), and fairly large in the Tulare Lake Basin (-0.81). These estimates are the elasticities I carry forward into the counterfactual simulations.

In Table D1, I show several alternative specifications. Precision is similar when clustering more conservatively by unit (Panel A, columns 1-3). Magnitudes are similar when using estimated marginal valuations as the endogenous variable instead of prices (Panel A, columns 4-6); although standard errors appear to be smaller, they do not account for additional estimation error from the prior stage of estimating the marginal valuations. Results are also similar when providing the Lasso algorithm different sets of potential instruments: allocation percentages interacted with only region indicators (Panel B, columns 1-3) or with only unit indicators (Panel B, columns 3-6). Finally, including year fixed effects is known to introduce bias, and Panel C confirms that approach produces results that are unstable across specifications, are imprecise, have the wrong sign, and/or have weak first stages, both when using IV Lasso (columns 1-3) or a only single instrument (columns 4-6).

These are the first estimates of price elasticities of demand in the wholesale surface water market in California. Most prior price elasticities for water in the literature are individual-level estimates. Rather than measuring how households or farm operations alter their water consumption in response to retail water prices, my elasticities measure how much water the retail agencies themselves choose to buy or sell in response to wholesale water prices, representing the revealed preferences of market participants. These estimates tend to be smaller than most prior individual-level estimates but still within their broad range. For example, my elasticity of -0.10 for urban water demand is similar to previous findings for all of California (Buck et al. 2016, -0.14) and the city of Santa Cruz (Nataraj and Hanemann 2011, -0.12), but it is considerably smaller than an estimate for Riverside County (Baerenklau et al. 2014, -0.76) and the mean in a meta-analysis of urban water elasticities across the US (Dalhuisen et al. 2003, -0.41). My elasticity of -0.23 for agricultural water demand is considerably smaller than the mean in a meta-analysis across the US (Scheierling et al. 2006, -0.48), but similar to more recent estimates (Bruno 2017, -0.19; Hendricks and Peterson 2012, -0.10), and my subgroup elasticity of -0.81 for the Tulare Lake Basin is very similar to a previous estimate from that region (Schoengold et al. 2006, -0.79).

7.3 Fitted values

Next, I combine these demand models with the marginal valuations from Step 2 to obtain fitted values that I carry forward into simulations. Although the IV estimates above can be interpreted as weighted averages of heterogeneous elasticities, to extrapolate out of sample I now impose the assumption of constant effects: demand is isoelastic and constant across units and time.

To compute fitted values, I first invert Equation 11 to give log marginal valuation in terms of log quantity. Second, I assign each unit its corresponding elasticity from Table 4, Panel C (for the small number of agricultural units that fall outside of the three named regions, I apply the overall agricultural estimate). Third, I calculate each unit's intercept from its assigned elasticity and its means of log marginal valuations and log quantities consumed across years. Finally, I use the resulting models to obtain fitted values of marginal valuations for quantity values in three scenarios: median, wet, and dry years. I obtain these quantities by ordering each unit's annual quantity consumed from 1993 through 2015 and calculating the mean within each quintile. The median-year scenario corresponds to the middle quintile, the wet-year scenario to the least water-scarce quintile, and the dry-year scenario to the most water-scarce quintile.

Figure 6 maps these estimated marginal valuations for the median-year scenario.[8] Despite wide dispersion in marginal valuations among individual water districts, there is less dispersion in marginal valuations across geographical regions in a typical year. There are few broad patterns: with the exception of lower marginal valuations in much of urban Southern California, marginal valuations fall within $100-300 across most of the state. More familiar patterns emerge in the dry-year scenario (shown along with the wet-year scenario in Appendix D, Figure E1). Here, marginal valuations are quite high in much in urban Southern California, moderately high in the San Joaquin Valley and Tulare Basin regions, and generally low in the mostly-agricultural Sacramento Valley. The sets of unit-level marginal valuations and quantities in these three weather scenarios are direct inputs to my counterfactual simulations.

8 Step 4: Simulating the Gains from Trade

With a demand model, I can now perform welfare analysis. First, I calculate the gains in economic surplus achieved by observed transactions across units. Then, I simulate trading outcomes in counterfactual scenarios representing an efficient market. In these scenarios, unit-level marginal valuations are equalized up to purely physical transaction costs that cannot be avoided. All other transaction costs are eliminated. The results give the economic benefits available from reforming California's wholesale surface water market to reduce transaction costs for cross-region and cross-sector transfers.

[8]Each polygon in the map represents a "planning area," which contains two units, urban and agricultural. To generate the map, I average marginal valuations over the two units and weight by quantity consumed. (I also drop three units with very small quantities that may be mismeasured, by the criterion that their transaction volume is more than twice their quantity consumed. Including these units would not substantively change the results, since their quantities sum to 0.02 percent of total statewide quantity.)

Three points are worth keeping in mind. First, because I estimate demand at the level of units, all results capture only the gains from trade across sub-regions and across sectors. Although the results exclude potential gains from trading among local districts serving the same sector, such districts often have pre-existing arrangements with each other and may not be fully independent agents. My results instead capture the typical sense of a unified wholesale water market involving arms-length transactions throughout the state.

Second, gains are calculated and valued according to the implied objective functions of existing market participants, which are mainly retail water districts. Districts' own revealed preferences may differ from the aggregated preferences of their customers for a range of reasons, including political economy factors, regulatory distortions, or information frictions. Therefore, this approach does not capture gains from allowing individual households and farms to directly access the wholesale market, or from reallocation among individuals within districts. My simulations are best viewed as the result of reducing transaction costs in the existing wholesale market, without altering the local institutions that govern who is able to access the market.

Third, my approach lacks detailed information on capacity constraints. Most water conveyance routes in California have slack capacity in most years, but with additional trading it is possible that these physical constraints could begin to bind. Therefore, the simulation results should be interpreted as the benefits that would result from not only eliminating non-physical transaction costs but also building any extra conveyance capacity that becomes necessary. Building additional capacity would involve additional costs that would need to be weighed against the potential benefits.

8.1 Surplus from the observed market

I first use the demand model to calculate the economic surplus achieved to date by observed transactions across units. This surplus for each unit k is calculated as the difference between the demand curve and the equilibrium marginal valuation (i.e., price net of transaction costs), integrated between initial endowment E_k and baseline quantity Q_k^0:

$$Surplus_k(E_k, Q_k^0) \equiv \int_{E_k}^{Q_k^0} \Big[V_k(\varphi) - V_k(Q_k^0)\Big] d\varphi. \tag{13}$$

Baseline quantity is the observed post-trading quantity from the relevant weather scenario. Initial endowment is calculated by subtracting mean observed net purchase volume (again from the relevant weather scenario) from baseline quantity. The integral of the second term easily evaluates to $(Q_k^0 - E_k)V_k^0$, where $V_k^0 \equiv V_k(Q_k^0)$ is the fitted marginal valuation from Section 7.3. For the first term, the integral of inverse demand can be evaluated analytically given the isoelastic functional form:

$$\int_A^B V_k(\varphi) d\varphi = \exp\Big(-\frac{\psi_k}{\eta_k}\Big)\Big(\frac{1}{\eta_k}+1\Big)^{-1}\Big(B^{\frac{1}{\eta_k}+1} - A^{\frac{1}{\eta_k}+1}\Big) \tag{14}$$

where η_k is the unit's price elasticity of demand and ψ_k is the intercept of the log-log demand model computed in Section 7.3. Transaction costs are not explicitly included in the surplus calculation because marginal valuations already take them into account. Surplus is analogous to consumer surplus for units that buy water and producer surplus for units that sell water; it is positive for all units by construction.

8.2 Counterfactual simulations

Next, to simulate an efficient market, I solve the social planner's problem in a constrained optimization problem. Because an ideal market could implement the efficient allocation, the increase in surplus relative to baseline represents the potential gains from trade. To capture inescapable physical transaction costs, I calculate pair- and direction-specific conveyance costs in my hydrological network model and include them as costs in the objective function.

The social planner's objective is to choose the vector of bilateral transaction quantities q_{kl} (net volumes delivered from unit k to unit l, for all pairs of units) that maximizes aggregate valuation of water net of physical transaction costs, subject to the resource constraint. (Prices are irrelevant because they are merely transfers between units, canceling out in the sum.) Aggregate value is the area under each unit's demand curve up to the final quantity Q_k^f, summed over units.[9] Each unit's final quantity is its baseline quantity minus transaction quantities lost, plus transaction quantities gained. Physical transaction costs are directionally asymmetric: $c_{kl}q_{kl}$ for a positive transaction quantity (net delivery from k to l) and $c_{lk}q_{kl}$ for a negative transaction quantity (net delivery from l to k), where c_{kl} is the physical marginal cost of delivering water from k to l. The resource constraint requires that final quantities be positive, meaning that no unit may create new water in order to sell more than its baseline quantity. Together, the full optimization problem is:

$$\max_{\{q_{kl}\}_{k,l>k}} \sum_k \int_{Q_k^0}^{Q_k^f} V_k(\varphi)d\varphi - \sum_k \sum_{l>k} \Big[1(q_{kl}>0)c_{kl} - 1(q_{kl}<0)c_{lk}\Big] q_{kl} \tag{15}$$

subject to

$$\text{(definition of final quantities)} \qquad Q_k^f = Q_k^0 - \sum_{l>k} q_{kl} + \sum_{l<k} q_{lk} \;\forall k$$

$$\text{(resource constraint)} \qquad Q_k^f \geq 0 \;\forall k.$$

I solve this problem using the `patternsearch` solver in Matlab, using baseline quantities as the initial conditions and evaluating the integral with Equation 14. The solution to this planner's problem also satisfies the conditions of an efficient market (proven in Appendix E.2). Intuitively, marginal valuations must be equalized up to physical costs in both the planner's solution and market equilibrium; if they are not, then buying from one unit and selling to another would be profitable for an arbitrageur, while the same reallocation would increase welfare for the social planner.

[9] Because the value of isoelastic inverse demand is undefined at a quantity of zero, I instead maximize the difference in aggregate value relative to baseline quantities Q_k^0. These expressions are equivalent in a maximand since they differ only by a constant, $\int_0^{Q_k^0} V_k(\varphi)d\varphi$.

Unit-specific gains relative to baseline can be calculated in the same way as Equation 13, the area between a unit's demand curve and its post-trading marginal valuation:

$$H_k \equiv \int_{Q_k^0}^{Q_k^f} \left[V_k(\varphi) - V_k(Q_k^f)\right] d\varphi. \tag{16}$$

This expression uses the post-trading marginal valuation instead of marginal transaction costs, which avoids the need to directly calculate the incidence of these costs for each pair of buyers and sellers. Since marginal valuations equalize up to transaction costs, the sum of H_k over all units is equivalent to the maximand in Equation 15 (proven in Appendix E.3). These unit-specific gains, illustrated as the Harberger triangle in Figure 3, represent the gains from additional transactions that become newly profitable.

Besides these gains from new transactions, an efficient market also reduces the costs of transactions that were already taking place. These gains, illustrated in Figure 3 as the shaded rectangle, can be calculated as the difference between endowments and baseline quantities, multiplied by the difference between baseline marginal valuation and post-trading marginal valuation:[10]

$$R_k \equiv \left(Q_k^0 - E_k\right)\left(V_k^0 - V_k(Q_k^f)\right). \tag{17}$$

8.3 Main results

Scenario 1: Gains from observed spot-market transactions.

I first calculate the gains achieved to date by spot-market transactions, the same observations I use to estimate demand. Table 5, Panel A reports that the total volume of water traded across units in short-term transactions is 156,000 acre-feet for a typical year in my dataset (corresponding to the median quintile of water quantities consumed). Figure 8 maps the geographic patterns of net water sales among observed transactions. Sellers tend to be in the Sacramento Valley, northern San Joaquin Valley, and along the Colorado River; buyers tend to be in urban Southern California.

Applying my demand model to Equation 13, I calculate that these observed transactions result in economic surplus of \$13.4 million per year in the median-year scenario – an extremely small figure relative to annual water-related expenditures throughout California. In wet years (highest quintile), transaction volume and total gains are similar. In dry years (lowest quintile), transaction volumes and marginal valuations are more than twice as large. Total gains are substantially larger at \$87.7 million per year, although this is still small.

Scenario 2: Gains achieved by all observed transactions (including permanent transfers).

Surface water transactions in California are not limited to the spot market; many are permanent transfers of water rights or entitlements. Next I consider the gains achieved by all observed trans-

[10] I restrict R_k to be non-negative; negative values can sometimes arise if the simulation predicts net trading in the opposite direction of observed trading.

actions across units, including both spot-market and permanent transactions. Figure 8 shows that the geographic pattern of trade is similar when including permanent transactions, with additional buyers in the urban areas of both Northern and Southern California.

When including permanent transfers in the median-year scenario, gains increase in approximate proportion to the transaction volume – but they are still relatively small, at $35.9 million per year. However, in the dry-year scenario, gains are much larger: $822 million per year. This pattern of results suggests that permanent transfers have allowed many units to avoid massive welfare losses from reduced surface water availability during droughts.

Scenario 3: Gains from an efficient market, eliminating all non-physical transaction costs.

For my central result, I run the optimization problem including purely physical transaction costs as calibrated by my hydrological network model. This scenario eliminates all differences in marginal valuations up to physical transaction costs, increasing trading until no further transactions increase welfare. All quantities and gains are relative to the baseline scenarios, so they are additional to the gains achieved by the existing market (Scenarios 1 and 2).

Results are shown in Table 5, Panel B. In this simulation, total quantity traded increases dramatically, to over one million acre-feet per year. Total gains range from $85.8 million (in a median year) to $278 million (in a dry year). Although eliminating transaction costs does result in lower costs for transactions that already take place ($11.3 million in a median year), most of the gains arise from new transactions that would not have otherwise occurred ($74.5 million in a median year). These results suggest that reducing transaction costs could increase the gains achieved by the existing short-term market. But these gains are still quite small relative to the $30 billion per year that California is estimated to spend in the water sector (Hanak et al. 2014).

Geographic patterns of trading for the median-year and dry-year scenarios are mapped in Figure 8. In a median year, some water is sold within the Sacramento Valley, larger amounts are sold from the San Joaquin Valley to the lower Tulare Lake Basin, and some water moves from urban Southern California to other parts of the state. In a dry year, more water is sold from the Sacramento Valley to the San Joaquin Valley and points further south. Despite previous findings of costly water shortages in urban Southern California (Buck et al. 2016), these results suggest that districts in this region would not actually buy more water in drought years if transaction costs were lower, and in fact they might even sell water (or buy less) in non-drought years.

8.4 Extensions and sensitivity checks

Environmental constraints. Because my simulation does not take into account capacity constraints, a natural question is how important these might be. Even if additional pumping and conveyance infrastructure could be built, there may be points in the system at which additional flow would cause unacceptable environmental harm (for example, as defined by federal environmental regulations).

I explore the impact of such environmental constraints by examining the case of the Sacramento–San Joaquin Delta, which may be the single most ecologically sensitive juncture in California's hydrology. Specifically, I simulate an efficient market with the additional constraint that outflow from the Sacramento River may be no more than it is at present. This restriction addresses a common concern over the prospect of expanding water transactions in California: if additional water flows out of the Sacramento River and is pumped into canals, it could affect salinity levels and flow direction, harming water quality and the migration ability of endangered fish species.

Scenario 4 implements this simulation by separating the market into two segments, one north of the Sacramento–San Joaquin Delta (i.e., the mouth of the Sacramento River) and one south of it, and then simulating an efficient market in each segment as in Scenario 3. This allows water to be efficiently reallocated both within the Sacramento Valley and south of the Delta, but with no change in the outflow of the Sacramento River.

Table 5, Panel C shows that in this scenario, both trading volumes and economic gains would be virtually the same as in Scenario 3 – the environmental constraint barely seems to bind. This result is consistent with the geographic patterns of trading from Scenario 3, which show both purchases and sales within all regions of the state (the Sacramento Valley, the San Joaquin Valley, and Southern California). It suggests the gains from an efficient surface water market could be achieved almost entirely by trading within each side of the Sacramento–San Joaquin Delta, and they do not require relaxing current environmental regulations.

Functional form sensitivity. Next, I explore how sensitive my results are to the functional form for demand. In simulating counterfactual market equilibria and evaluating the resulting surplus, I rely on an assumption of isoelastic demand. This assumption may be innocuous if the change in quantity is small for all units, but otherwise it may be an important factor driving the results.

To test sensitivity to functional form, **Scenario 5** instead assumes that demand is linear. This scenario uses the same estimated elasticities and marginal valuations as in the previous scenarios but evaluates the integral of inverse demand using the corresponding expression for linear demand.[11] I find that the results are very similar using linear demand, with trading volumes and economic gains within 10 percent of the corresponding figures for Scenario 3. This suggests that no major aspect of my results is driven by the choice of functional form.

Physical transaction costs. I also investigate the extent to which my results are driven by purely physical transaction costs. Are the gains from an efficient market small because it is too expensive to arbitrage across regions, or simply because regions are not very heterogeneous? In **Scenario 6**, I simulate an efficient market with all transaction costs c_{kl} set to zero – implying, implausibly, that all water conveyance is costless.

The results are, once again, quite similar to the main results from Scenario 3. Gains in median and wet years are only slightly larger without physical transaction costs, and gains in a dry year are

[11] $\int_A^B V_k(\varphi)d\varphi = (1-1/\eta_k)(B-A)V_k^0 + (2\eta_k)^{-1}(B^2-A^2)(V_k^0/Q_k^0)$.

only 17 percent larger. This suggests that purely physical transaction costs are not a major obstacle to additional gains from trade.

Individual versus unit-level elasticities. My analysis uses the inferred preferences of retail water districts rather than the preferences of individual consumers and producers, who usually cannot directly access California's wholesale water market. However, there are many reasons to believe that individual-level preferences might differ from district-level behavior. In particular, the demand elasticities I estimated in Step 3 are smaller (less elastic) than the individual-level elasticities in much of the prior literature.

In **Scenario 7**, I simulate an efficient market using larger elasticities that are representative of the individual-level literature: -0.48 for agricultural units (from Scheierling et al. 2006),[12] and -0.143 for urban units (from Buck et al. 2016),[13] with all other parameters the same as in Scenario 3. I find that the gains from trade are much larger. Surplus nearly doubles in median and wet years, and in dry years it reaches \$711 million. These results suggest that the relatively small elasticities I estimate, particularly for agricultural water districts, are a major factor driving the surprisingly small results.

Dispersion of marginal valuations. Another way in which individual-level preferences may differ from district-level behavior is in the initial level of the marginal valuations. In particular, the district-level marginal valuations I estimate may be less dispersed than the prices (or shadow prices) faced by districts' retail customers, even after adjusting for the marginal costs of treatment and distribution.

To investigate the sensitivity of my results to the dispersion of marginal valuations, **Scenario 8** applies a transformation to estimated marginal valuations before calculating the fitted values in Step 3. Specifically, I double the variance of the logarithmic distribution of marginal valuations, by doubling the difference between each log marginal valuation and the grand mean of the log distribution and adding the grand mean to find each new marginal valuation. Doubling the log distribution is an arbitrary choice simply meant to illustrate the sensitivity.

I find that the gains are again larger when using these more-extreme marginal valuations, with surplus in median and wet years approximately triple the surplus achieved in Scenario 3. In **Scenario 9**, I apply the modifications of both Scenarios 7 and 8, using individual-level elasticities as well as more-extreme marginal valuations. Here, once again, the gains increase, ranging from \$392 to \$610 million per year depending on weather.

[12] This is the mean price elasticity in a meta-analysis of irrigation water demand in the United States. I apply it to all agricultural units except for those in the Tulare Lake Basin, which was already using a larger-magnitude elasticity.

[13] This is the most credible analysis of price elasticities that includes a large set of municipal districts in California; I use the central instrumental variables estimate.

8.5 Discussion

My main results suggest that reducing transaction costs in California's surface water market would achieve between \$86 and \$278 million dollars per year, depending on weather. These benefits are meaningful but significantly lower than often thought. Why are the potential benefits so small, given ample anecdotal evidence suggesting considerable misallocation of water throughout the state? Three explanations appear to be supported by evidence.

One explanation is that water allocation is simply not as inefficient as anecdotal evidence suggests. In most non-drought years, marginal valuations are not dramatically heterogeneous across regions and sectors, varying by factors of two or three, not orders of magnitude. In drought years, the benefits from permanent transfers (Scenario 2) in addition to the within-year spot market are quite large. This indicates that these long-term transactions over time have substantially reduced the cost of drought, and that marginal valuations otherwise would be much more dispersed.

Another explanation is that many of the potential benefits from water markets may be local rather than statewide. The gains I estimate are from simulating trading among units, i.e., across regions and sectors. However, as an intermediate step I also calculate marginal valuations for individual water districts within units. The variance of marginal valuations within units is similar to the variance across units,[14] suggesting that many of the potential gains are in fact local.

The final explanation is that water districts do not accurately represent the preferences of their retail customers when participating in this wholesale market. My approach infers and extrapolates from the objective functions of water districts, which may differ substantially from the aggregate preferences of retail customers for a range of reasons. My estimates imply that these districts behave conservatively in the wholesale surface water market, showing more inelastic demand than prior individual-level estimates in the literature. Simulations using these alternative elasticities find substantially larger gains, suggesting that the internal governance and pricing policies of these water districts are a major part of the explanation.

Overall, my analysis suggests that if there do exist large potential gains from reallocating water within California, they might not be achieved by reforms to reduce transaction costs in the wholesale water market without also addressing the internal workings of retail water districts. Policy reforms that could improve local allocation and make water districts' behavior more responsive to their customers' preferences include: (1) merging and consolidating water districts, (2) simplifying rate schedules, (3) raising per-unit rates to reflect wholesale opportunity costs in addition to other marginal costs of provision, (4) establishing buy-back programs at wholesale market prices, (5) adjusting rates dynamically to reflect changing conditions in the wholesale market, and (6) facilitating other ways for customers to participate in the wholesale market.

[14] The within-unit standard deviation (i.e., after partialing out both year and unit) is 194.3, while the between-unit standard deviation of marginal valuations (i.e., after taking the mean within unit-year and partialing out year) is 183.5.

9 Conclusion

In this paper, I use a revealed-preference approach to estimate the potential gains from increased trade in California's statewide water market. I develop a four-step empirical procedure to analyze welfare in thin markets with transaction costs, and I apply it to new, uniquely comprehensive data on California's water economy. First, I estimate marginal transaction costs associated with observable cost determinants by measuring price gaps across transactions within district and year. I find that several cost determinants give rise to large transaction costs, up to several hundred dollars per acre-foot. Second, I estimate districts' equilibrium marginal valuations by taking the maximum or minimum price after adjusting for estimated marginal transaction costs. I find that marginal valuations are substantially more dispersed than observed prices, implying that welfare gains are available from reallocation. Third, I estimate elasticities of inverse aggregate demand using weather-driven variation in yearly surface water entitlements. I find that both agricultural and urban sectors have fairly inelastic demand, although there is substantial heterogeneity within agriculture.

Having estimated equilibrium marginal valuations and demand elasticities, I combine them to construct demand curves and simulate counterfactual equilibria. I find that observed trading across regions and sectors achieves gains of just \$10 to \$88 million per year, which is perhaps unsurprising, given the low transaction volume. Reducing transaction costs in an efficient market would result in additional gains of \$86 to \$278 million per year – much larger than the existing market, but still small relative to total water sector expenditures in California.

The full optimization exercises are idealized but can serve as policy benchmarks. Essentially, I document empirical gaps in marginal valuations across water users, and the simulations quantify potential gains that could result from closing these gaps with an efficient statewide water market. Both gaps and gains, however, may be underestimated if there are additional unobserved transaction costs that do not vary across a user's transactions. Still, knowing a conservatively-estimated size of the potential benefits may help in making decisions about policy reforms.

My estimates also miss welfare gains from trading within water districts or allowing retail customers direct access to the statewide water market. This is intrinsic to my empirical approach, which infers the preferences of the water districts from their observed trading behavior. However, these gains may be more difficult to achieve than the ones I do measure, since they would require reforming the institutional structure of local governments and water utilities.

Finally, this paper estimates only the potential benefits of a more efficient water market, not the costs. A complete policy analysis should also consider the benefits of existing market-restricting regulations and the costs of creating stronger market-supporting institutions. Still, there may be many ways to develop more robust water markets without compromising environmental and other policy goals. Local hydrological externalities could be resolved via a streamlined and unified system for determining consumptive use, which would help reduce the large difference in regulatory treatment between on-site water use and off-site water transfers. Instream flow requirements could be met through direct environmental purchases or by capping and auctioning off instream capacity. Even with significant cross-region flow constraints (as might be preferred, for example, at California's Sacramento–San Joaquin Delta), economic gains are likely available from within-region water markets.

References

Adamopoulos, Tasso, Loren Brandt, Jessica Leight, and Diego Restuccia, "Misallocation, Selection, and Productivity: A Quantitative Analysis with Panel Data from China," 2017.

Ahrens, A., C.B. Hansen, and M.E. Schaffer, "pdslasso and ivlasso: Programs for post-selection and post-regularization OLS or IV estimation and inference," 2018.

Association of California Water Agencies, "Recommendations for Improving Water Transfers and Access to Water Markets in California," 2016.

Atkin, David and Dave Donaldson, "Who's Getting Globalized? The Size and Nature of Intranational Trade Costs," 2015.

Ayres, Andrew B., Eric C. Edwards, and Gary D. Libecap, "How Transaction Costs Obstruct Collective Action to Limit Common-Pool Losses: Evidence from California's Groundwater," 2017.

Baerenklau, Kenneth A, Kurt A Schwabe, and Ariel Dinar, "The Residential Water Demand Effect of Increasing Block Rate Water Budgets," *Land Economics*, 2014.

Belloni, Alexandre, Daniel Chen, Victor Chernozhukov, and Christian Hansen, "Sparse Models and Methods for Optimal Instruments with an Application to Eminent Domain," *Econometrica*, 2012, *80* (6), 2369–2429.

Bergquist, Lauren Falcao, "Pass-through, Competition, and Entry in Agricultural Markets: Experimental Evidence from Kenya," 2016.

Björnerstedt, Jonas and Johan Stennek, "Bilateral oligopoly - The efficiency of intermediate goods markets," *International Journal of Industrial Organization*, 2007, *25* (5), 884–907.

Brewer, Jedidiah, Robert Glennon, Alan Ker, and Gary Libecap, "2006 Presidential Address: Water Markets in the West: Prices, Trading, and Contractual forms," *Economic Inquiry*, 2008, *46* (2), 91–112.

Bruno, Ellen M, "Agricultural Groundwater Markets: Understanding the Gains from Trade and the Role of Market Power," 2017.

Bryan, Gharad and Melanie Morten, "Economic Development and the Spatial Allocation of Labor: Evidence From Indonesia," 2015.

Buck, Steven, Maximilian Auffhammer, and David Sunding, "Land markets and the value of water: Hedonic analysis using repeat sales of farmland," *American Journal of Agricultural Economics*, 2014, *96* (4), 953–969.

_ , _ , Stephen Hamilton, and David Sunding, "Measuring Welfare Losses from Urban Water Supply Disruptions," *Journal of the Association of Environmental and Resource Economists*, 2016, *3* (3), 743–778.

Burke, Marshall, Solomon M. Hsiang, and Edward Miguel, "Climate and Conflict," *Annual Reviews of Economics*, 2015, *7*, 577–617.

California Department of Water Resources, "State Water Project Delivery Capability Report 2015," 2015.

_ and U.S. Bureau of Reclamation, "Technical Information for Preparing Water Transfer Proposals (Water Transfer White Paper)," 2015.

California State Water Resources Control Board, "A Guide to Water Transfers," 1999.

Carey, Janis, David L. Sunding, and David Zilberman, "Transaction costs and trading behavior in an immature water market," *Environment and Development Economics*, 2002, *7* (04), 733–750.

Carey, Janis M. and David L. Sunding, "Emerging Markets in Water: A Comparative Institutional Analysis of the Central Valley and Colorado-Big Thompson Projects," *Natural Resources Journal*, 2001, (Spring).

Cason, Timothy N. and Lata Gangadharan, "Transaction Costs in Tradable Permit Market: An Experimental Study of Pollution Market Design," *Journal of Regulatory Economics*, 2003, *23* (2), 145–165.

Chernozhukov, Victor, Christian Hansen, and Martin Spindler, "Post-Selection and Post-Regularization Inference in Linear Models with Many Controls and Instruments," *American Economic Review: Papers & Proceedings*, 2015, *105* (5), 486–490.

Collard-Wexler, Allan, Gautam Gowrisankaran, and Robin S. Lee, ""Nash-in-Nash" Bargaining: A Microfoundation for Applied Work," 2017.

Culp, Peter W., Robert Glennon, Gary Libecap, George A. Akerlof, Roger C. Altman, Gordon S. Rentschler Memorial, and Laura D. Andrea Tyson, "Shopping for Water: How the Market Can Mitigate Water Shortages in the American West," 2014.

Dalhuisen, Jasper M., Raymond J. G. M. Florax, Henri L. F. de Groot, and Peter Nijkamp, "Price and Income Elasticities of Residential Water Demand: A Meta-Analysis," *Land Economics*, 2003, *79* (2), 292–308.

Davis, Lucas W. and Lutz Kilian, "The Allocative Cost of Price Ceilings in the U.S. Residential Market for Natural Gas," *Journal of Political Economy*, 2011, *119* (2), 212–241.

Donaldson, Dave, "Railroads of the Raj: Estimating the Impact of Transportation Infrastructure," 2012.

Gangadharan, Lata, "Transaction Costs in Pollution Markets: An Empirical Study," *Land Economics*, 2000, *76* (4), 601–614.

Glaeser, Edward L. and Erzo F. P. Luttmer, "The Misallocation of Housing Under Rent Control," *American Economic Review*, 2003, *93* (4), 1027–1046.

Gray, Brian, Ellen Hanak, Richard Frank, Richard Howitt, Jay Lund, Leon Szeptycki, and Barton "Buzz" Thompson, "Allocating California's Water: Directions for Reform," 2015.

Hahn, Robert W., "Market Power and Transferable Property Rights," *The Quarterly Journal of Economics*, 1984, *99* (4), 753–765.

Hanak, Ellen and Elizabeth Stryjewski, "California's Water Market, By the Numbers: Update 2012," 2012.

_, Brian Gray, Jay Lund, David Mitchell, Caitrin Chappelle, Andrew Fahlund, Katrina Jessoe, Josué Medellín-Azuara, Dean Misczynski, James Nachbaur, and Robyn Suddeth, "Paying for Water in California," *Public Policy Institute of California*, 2014.

Hendricks, Nathan P and Jeffrey M Peterson, "Fixed effects estimation of the intensive and extensive margins of irrigation water demand," *Journal of Agricultural and Resource Economics*, 2012, *37* (12), 1–19.

Ho, Kate and Robin S. Lee, "Insurer Competition in Health Care Markets," *Econometrica*, 2017, *85* (2), 379–417.

Howitt, Richard E., Jay R. Lund, K.W. Kirby, Marion W. Jenkins, and Andrew J. Draper, "Integrated Economic-Engineering Analysis of California's Future Water Supply: a report for the State of California Resources Agency," 1999.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *Quarterly Journal of Economics*, 2009, *124* (4).

Jenkins, Marion W., Jay R. Lund, and Richard E. Howitt, "Using Economic Loss Functions to Value Urban Water Scarcity in California," *American Water Works Association*, 2003, *95* (2), 395–429.

Kyle, Albert S., "Informed Speculation with Imperfect Competition," *The Review of Economic Studies*, 1989, *56* (3), 317–355.

Liski, Matti, "Thin versus Thick CO2 Market," *Journal of Environmental Economics and Management*, 2001, *41*, 295–311.

— **and Juan Pablo Montero**, "Market Power in an Exhaustible Resource Market: The Case of Storable Pollution Permits," *Economic Journal*, 2011, *121* (551), 116–144.

Malueg, David A. and Andrew J. Yates, "Bilateral oligopoly, private information, and pollution permit markets," *Environmental and Resource Economics*, 2009, *43* (4), 553–572.

Medellín-Azuara, Josué, Julien J. Harou, Marcelo A. Olivares, Kaveh Madani, Jay R. Lund, Richard E. Howitt, Stacy K. Tanaka, Marion W. Jenkins, and Tingju Zhu, "Adaptability and adaptations of California's water supply system to dry climate warming," *Climatic Change*, 2007, *87*.

Mérel, Pierre and Richard Howitt, "Theory and Application of Positive Mathematical Programming in Agriculture and the Environment," *Annual Review of Resource Economics*, 2014, *6* (1), 451–470.

Mukherjee, M. and K. Schwabe, "Irrigated Agricultural Adaptation to Water and Climate Variability: The Economic Value of a Water Portfolio," *American Journal of Agricultural Economics*, 2014, *97* (2009), 809–832.

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

Peterson, Deborah, Gavan Dwyer, David Appels, and Jane Fry, "Water trade in the southern Murray-Darling basin," *Economic Record*, 2005, *81* (255), S115–S127.

Qureshi, M. Ejaz, Tian Shi, Sumaira E. Qureshi, and Wendy Proctor, "Removing barriers to facilitate efficient water markets in the Murray-Darling Basin of Australia," *Agricultural Water Management*, 2009, *96* (11), 1641–1651.

Regnacq, Charles, Ariel Dinar, and Ellen Hanak, "The Gravity of Water: Water Trade Frictions in California," *American Journal of Agricultural Economics*, 2016, *98* (5), 1273–1294.

Restuccia, Diego and Raul Santaeulalia-Llopis, "Land Misallocation and Productivity," 2017.

Rosegrant, M. W., C. Ringler, D. C. McKinney, X. Cai, A. Keller, and G. Donoso, "Integrated economic-hydrologic water modeling at the basin scale: The Maipo river basin," *Agricultural Economics*, 2000, *24* (1), 33–46.

Rostek, Marzena and Marek Weretka, "Price Inference in Small Markets," *Econometrica*, 2012, *80* (2), 687–711.

— **and** — , "Dynamic Thin Markets," *Review of Financial Studies*, 2015, *28* (10), 2946–2992.

Scheer, Jennifer Beth, "Participation, Price, and Competition in the Sacramento Valley Water Market," 2016.

Scheierling, Susanne M., John B. Loomis, and Robert A. Young, "Irrigation water demand: A meta-analysis of price elasticities," *Water Resources Research*, 2006, *42* (1), 1–9.

Schlenker, Wolfram, W. Michael Hanemann, and Anthony C. Fisher, "Water availability, degree days, and the potential impact of climate change on irrigated agriculture in California," *Climatic Change*, 2007, *81* (1), 19–38.

Schoengold, Karina, David L. Sunding, and Georgina Moreno, "Price elasticity reconsidered: Panel estimation of an agricultural water demand function," *Water Resources Research*, 2006, *42* (9), 1–10.

Sekhri, Sheetal, "Wells, Water, and Welfare: The Impact of Access to Groundwater on Rural Poverty and Conflict," *American Economic Journal: Applied Economics*, 2014, *6* (3), 76–102.

Stavins, Robert N., "Transaction Costs and Tradeable Permits," *Journal of Environmental Economics and Management*, 1995, *29*, 133–148.

Stene, Eric A., "Central Valley Project Overview," 1995.

Sunding, David L., David Zilberman, Richard Howitt, Ariel Dinar, and Neal MacDougall, "Measuring the costs of reallocating water from agriculture: A multi-model approach," *Natural Resources Modeling*, 2002, *15* (2), 201–225.

UNDP, "Human Development Report 2006 - Beyond scarcity: Power, poverty and the global water crisis," 2006.

Western Water Company, "Statement for the State Water Resources Control Board Public Workshop on Water Transfers," 2000.

World Bank Group, "High and Dry: Climate Change, Water, and the Economy," 2016.

Figures

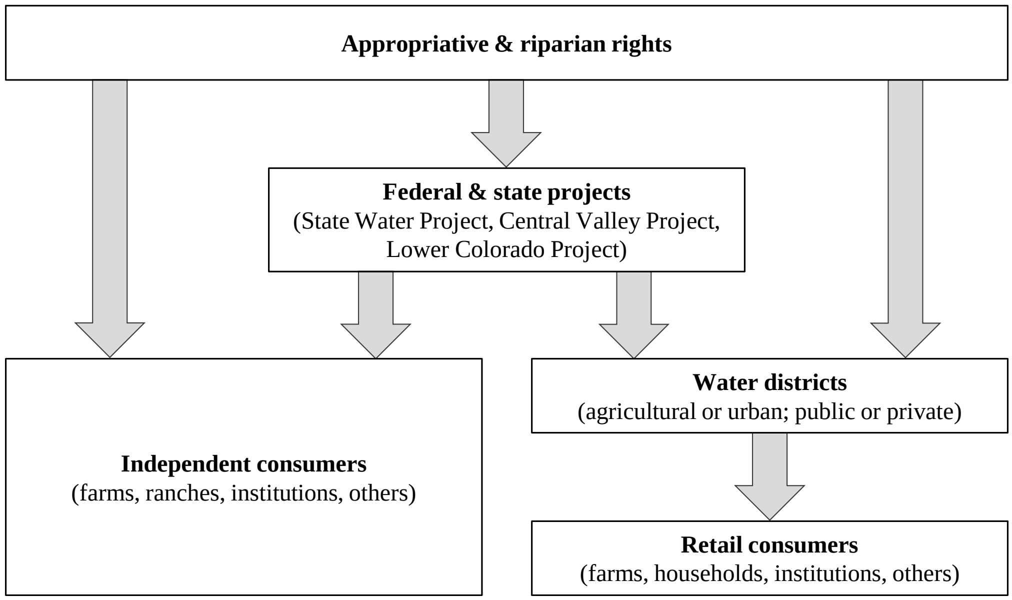

Figure 1: Structure of water entitlements in California. All water use is governed by a legal system of appropriative and riparian rights. Some water districts and independent consumers directly hold their own water rights; others hold long-term contracts with the federal and state water projects, which in turn hold water rights. Retail consumers purchase water from water districts. [Back]

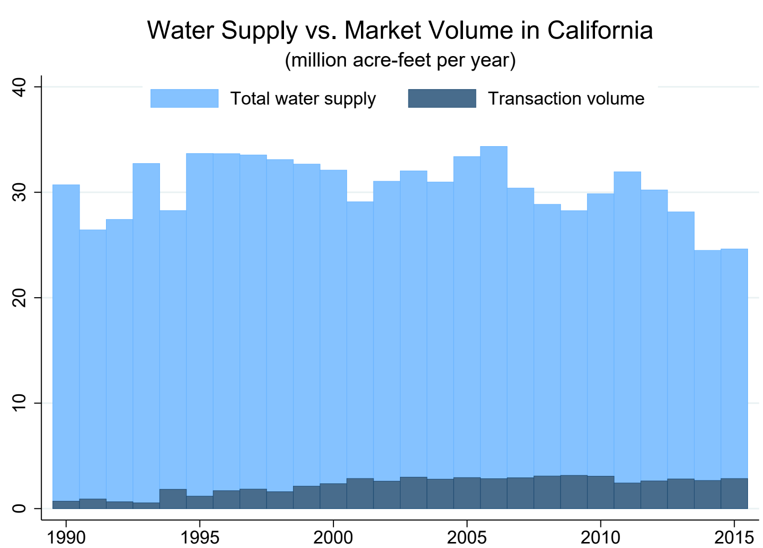

Figure 2: Observed volume of market transactions compared with total water supply. Market transactions include spot market transactions (within-year leases), longer-term leases, and permanent transfers of rights (which are counted in the transaction year and every year afterward). Water supply includes surface water rights, allocations from the federal and state water projects, and average annual groundwater supply. [Back]

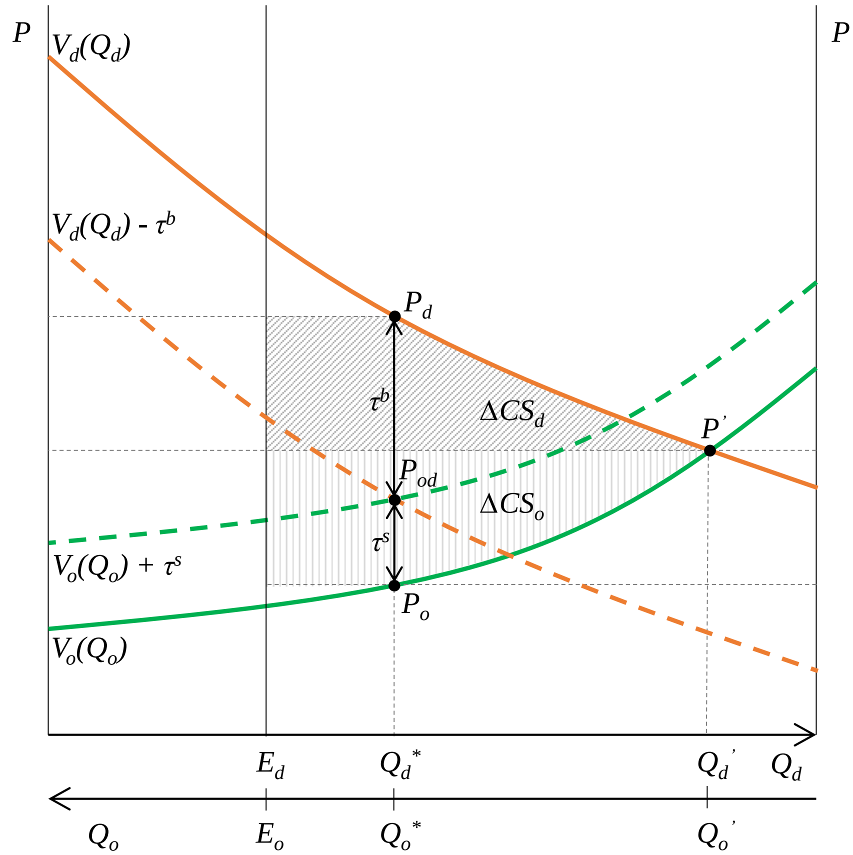

Figure 3: Trade between two districts, buying district d (for destination) and selling district o (for origin). Districts have endowments E_n and inverse demand $V_n(Q_n)$; demand for o is on a reversed axis so that the total width of the graph equals the sum of endowments. With costless trading, the competitive equilibrium among consumers in the two districts would result in price P'. If cross-market transactions incur per-unit transaction costs τ^s (for sellers) and τ^b (for buyers), buyers' marginal willingness to pay shifts down relative to demand, and sellers' marginal willingness to accept shifts up. Competitive equilibrium in this case would result in three distinct prices: a cross-market price P_{od} and within-market prices P_o and P_d, with $P_d - \tau^b = P_{od} = P_o + \tau^s$. Consumer surplus for buyers is shaded with diagonal lines, and consumer surplus for sellers (analogous to producer surplus) is shaded with vertical lines. [Back]

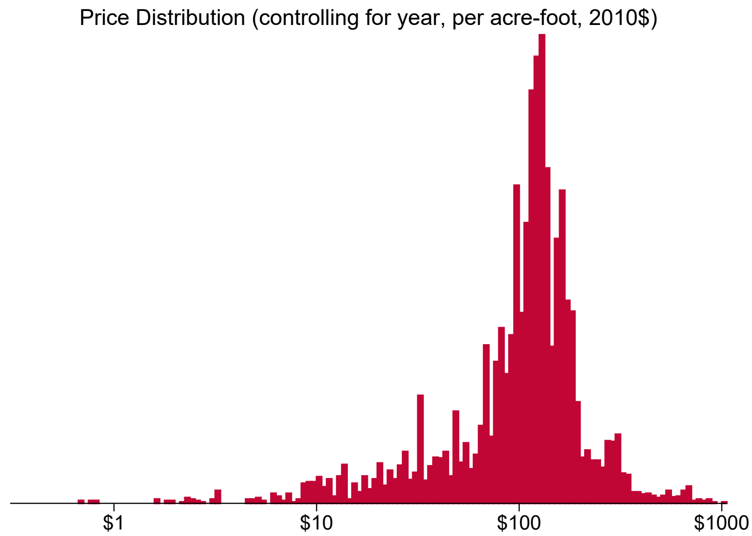

Figure 4: Distribution of prices, controlling for year, logarithmic scale. Graph plots a histogram of observed prices on California's wholesale surface water market, 1980-2015, converted to 2010 dollars using the CPI. I control for year by regressing log price on year fixed effects, taking the residual, and adding the grand mean. [Back]

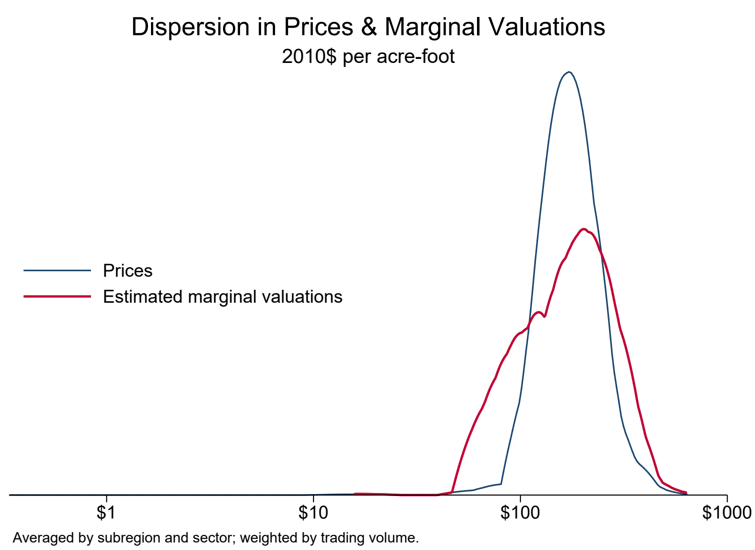

Figure 5: Distributions of prices and estimated equilibrium marginal valuations, logarithmic scale. Marginal valuations are constructed by taking the highest price paid (for buyers) or the lowest price accepted (for sellers) in each year, after adjusting prices for estimated transaction costs from known determinants. Distributions are estimated by kernel density, where an observation is one unit (defined as the intersection of subregion and sector), and units are weighted by trading volume. Within unit, prices and marginal valuations are averaged first across all users observed trading in a particular year and then across years. Unit-level valuations are direct inputs to the counterfactual simulations in Step 4. [Back]

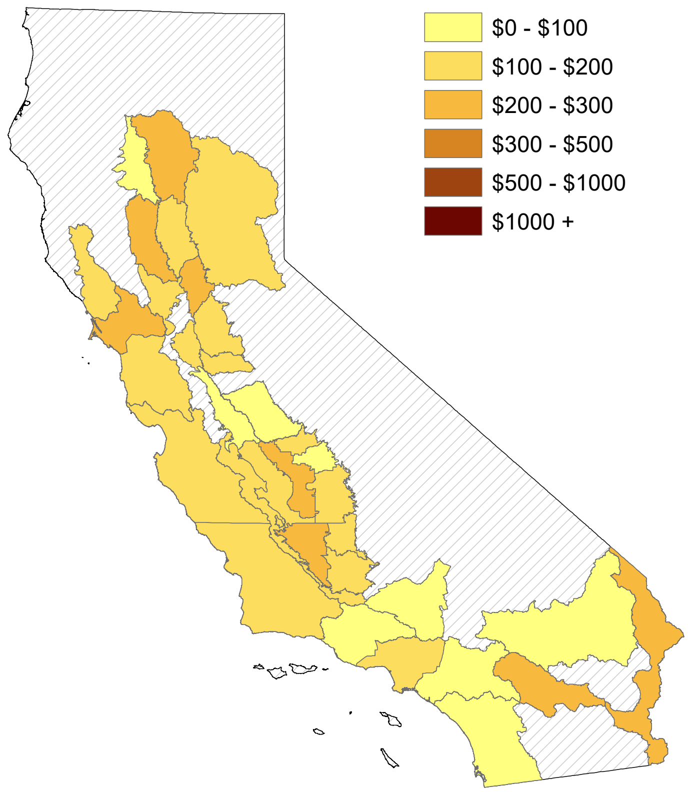

Figure 6: Estimated marginal valuations (per acre-foot) by geography for a median-year scenario. Geographic polygons correspond to subregions (i.e., planning areas as defined by the California Department of Water Resources); areas with diagonal shading have no observed transactions. Values shown are the fitted values from inverting the unit-specific demand models estimated in Step 3 and plugging in the median quantity consumed for each unit across years. (Unit is defined as the intersection of planning area with sector: urban or agricultural). The map shows the weighted average (by quantity) of unit-level marginal valuations across the two units within each planning area. Dry-year and wet-year scenarios are shown in Appendix D. Unit-level valuations in these three scenarios are direct inputs to the counterfactual simulations in Step 4. [Back]

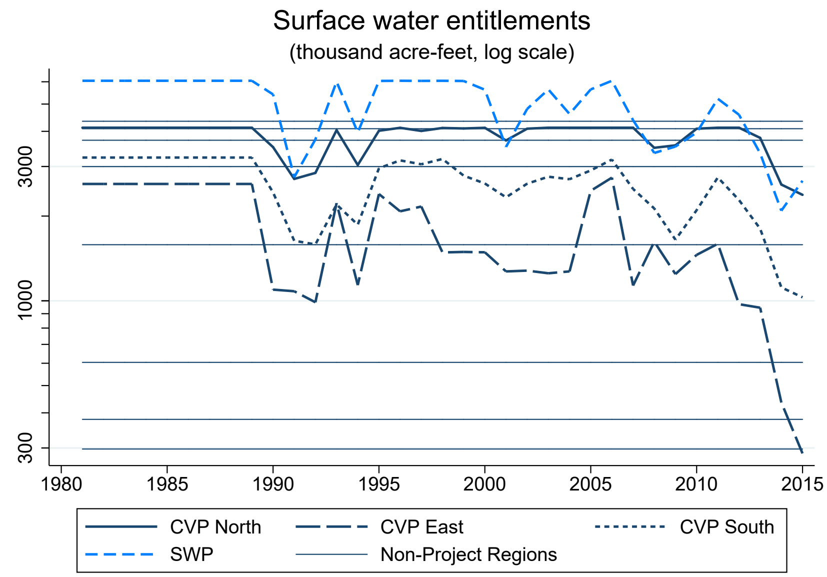

Figure 7: Entitlements over time for several categories of water endowments. This variation is used to estimate demand elasticities. Surface water entitlements are the sum of water rights and allocations from the federal and state water projects (CVP = Central Valley Project; SWP = State Water Project). Water rights are time-invariant, while project allocations vary year to year on the basis of weather conditions. This variability is set separately for each of 14 contract types. For clarity of illustration, each time series on the graph represents users aggregated by project and region. [Back]

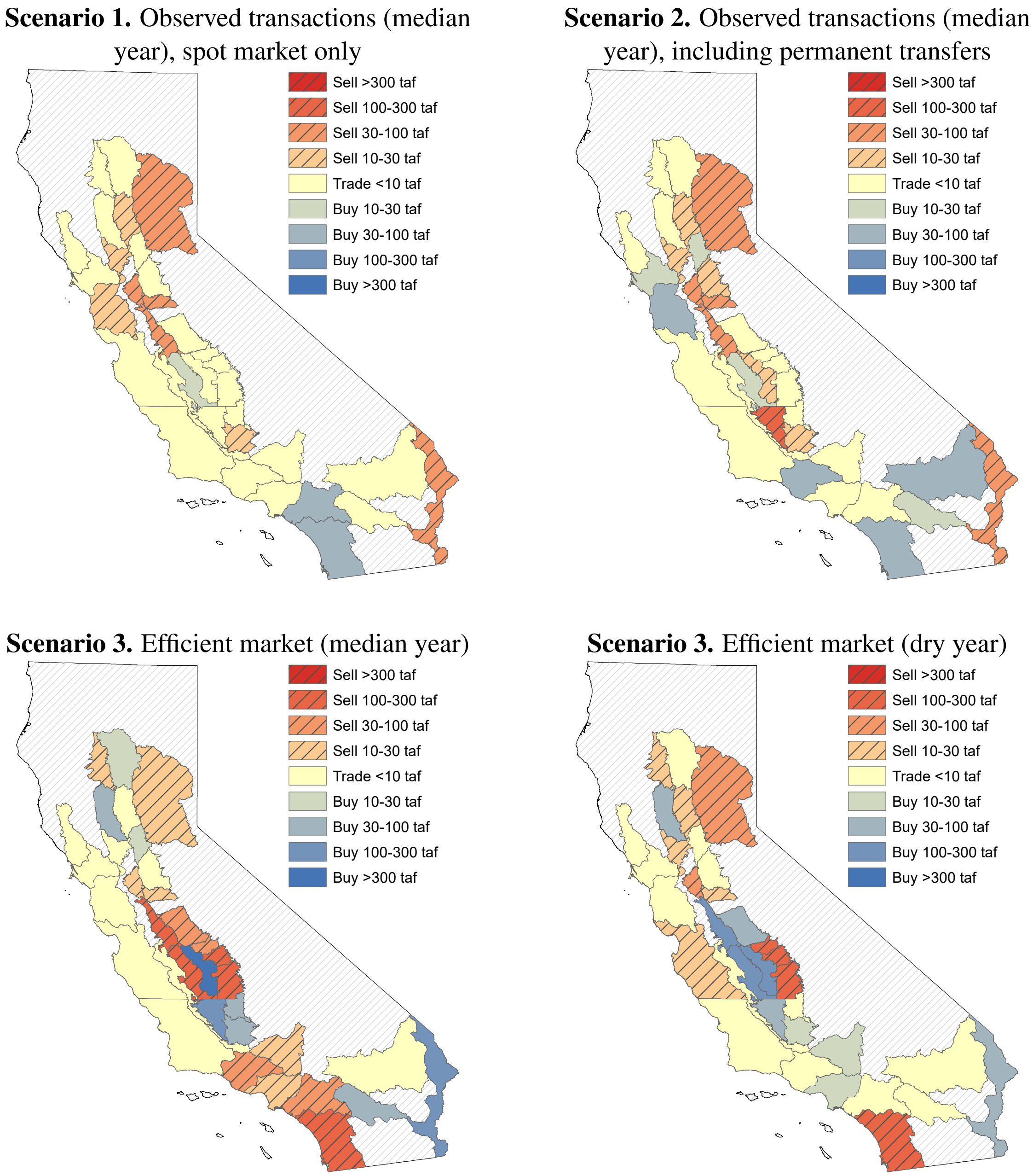

Figure 8: Net annual surface water quantities traded ("taf" = thousand acre-feet) in the existing market (Scenarios 1-2) and additional quantities traded in counterfactual simulations (Scenario 3), by geographic subregion. Median- and dry-year scenarios show trading relative to median and lowest-quintile water quantities consumed, respectively. Trading shown is limited to wholesale market transactions across units (defined by subregion and sector). Subregions are planning areas defined by the California Department of Water Resources; areas shaded with thin diagonal lines have no observed transactions. Quantities shown on map sum across sectors (i.e., units) sharing a subregion. For Scenario 3, transactions incur physical conveyance costs calibrated from a hydrological network model, but all other transaction costs are eliminated. [Back]

Tables

Table 1: Summary statistics of transactions data

Panel A: Summary statistics of transactions data

	Mean	Standard deviation	Observations
Volume (acre-feet)	13,629	31,213	671
Price (2010$/acre-foot)	221.3	297.4	647
Distance, Euclidean (km)	111.8	168.5	654
Distance conveyed (km)	149.3	234.9	665
Distance conveyed in rivers (km)	61.2	99.9	665
Distance conveyed in canals (km)	88.2	172.5	665
Distance of virtual conveyance (km)	72.4	117.0	665
Elevation lift (feet)	274.9	795.9	665
Crosses the Sacramento-San Joaquin Delta	18.3%	0.385	665
Reviewed by the State Water Boards	37.8%	0.483	671
Within project	40.8%	0.491	671

Panel B: Transactions by sector and region

	As Origin (Seller)		As Destination (Buyer)		
	Count	Volume (TAF)	Count	Volume (TAF)	Net sales (TAF)
By Sector					
Agricultural	597	7,606.3	359	2,195.4	5,410.9
Urban	101	833.0	217	2,661.0	-1,828.0
By Hydrologic Region					
North Coast	2	3.7	2	0.7	2.9
North Lahontan	0	0.0	0	0.0	0.0
Sacramento River	273	4,231.5	40	100.4	4,131.1
San Francisco Bay	17	189.7	47	346.1	-156.3
Central Coast	28	23.5	35	22.9	0.5
San Joaquin River	246	2,462.8	139	1,467.5	995.3
Tulare Lake	99	534.4	201	1,215.0	-680.6
South Coast	11	154.3	72	1,568.8	-1,414.4
South Lahontan	2	14.9	31	34.6	-19.7
Colorado River	20	845.2	9	100.5	744.8

TAF = thousand acre-feet. Panel A reports statistics for observed spot market (within-year) transactions of surface water in California that are freely negotiated and involve at least one party that diverts water for agricultural or urban consumption (i.e., not a predominantly environmental user nor a federal or state government agency). Variables are defined in Table 2. Panel B reports the count and total volume of transactions which begin (left) and end (right) in each sector or hydrologic region. Column sums do not exactly agree because a small number of transactions that involve more than two parties count in more than one hydrologic region. Net sales within category do not sum to zero because inclusion criteria are applied separately to each side of a transaction (i.e., many agricultural sales are purchased by environmental or government entities, which do not appear in the table as buyers). [Back]

Table 2: Cost determinants for which marginal transaction costs are econometrically identified in transactions data

	Cost determinant	Definition	Potentially costly because:
1	**Elevation gain (pump lift)**	Total vertical distance lifted.*	Energy is required to pump water uphill.
2	**Distance conveyed in rivers**	Total distance conveyed along a river.*	Some water is lost to evaporation and percolation.
3	**Distance conveyed in canals**	Total distance conveyed along a canal, aqueduct, or pipeline.*	Some water is lost to evaporation. Regulators and canal operators charge fees.
4	**Virtual distance**	Total distance of virtual movement, in which water is transferred against the direction of flow in a river or canal. Transaction quantity is diverted earlier, reducing channel flow.	No conveyance losses but may capture other non-physical costs that increase with distance between buyer and seller.
5	**Crosses the Sacramento–San Joaquin Delta**	Water must pass through the Delta and be pumped into a southward canal.	Fraction of each transaction must flow to the ocean to maintain salinity levels. Ecological requirements lead to delivery risks and heightened regulatory scrutiny.
6	**Import into federal or state water project**	Water is transferred into a project, instead of between contractors within the same project.**	Transactions within a project are subject to less regulatory scrutiny and may be easier to set up.
7	**Export from federal or state water project**	Water is transferred out of a project, instead of between contractors within the same project.**	Transactions within a project are subject to less regulatory scrutiny and may be easier to set up.
8	**State Water Boards (SWRCB) review**	Transaction involves a change in the place of use of a post-1914 appropriative water right.	Requires an extensive regulatory review.
9	**Agricultural counterparty**	Other party (buyer, for sellers; seller, for buyers) primarily uses water in the agricultural sector, rather than municipal/industrial.	Transactions involving agricultural users may incur a greater burden of ongoing monitoring and/or political challenges, or may be more difficult due to factors such as less market experience.

* Along the least-cost distance, as calculated by the hydrological network model. ** "Project" refers to either the federal Central Valley Project or the State Water Project. [Back]

Table 3: Marginal transaction costs from specific cost determinants

Dependent variable: **Price** (2010$/acre-foot)	**SELLERS** (positive coefficients are costly)				**BUYERS** (negative coefficients are costly)			
	Panel A: All known cost factors							
	Each alone	*All together*			*Each alone*	*All together*		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
River distance (km)	0.25 *	-0.34 *	-0.17	0.13	0.13 *	0.20	0.23 **	0.19 **
	(0.14)	(0.19)	(0.20)	(0.18)	(0.07)	(0.13)	(0.10)	(0.10)
Canal distance (km)	0.03	-0.30 *	-0.09	0.08	-0.10	-0.26	-0.21	-0.24
	(0.05)	(0.16)	(0.14)	(0.13)	(0.25)	(0.26)	(0.30)	(0.37)
Virtual distance (km)	-0.03	-0.11	0.02	-0.15	-0.32 **	0.00	-0.19 *	-0.13
	(0.09)	(0.13)	(0.12)	(0.10)	(0.16)	(0.16)	(0.11)	(0.11)
Pumping lift (ft)	0.00	0.04	0.02	-0.02	0.03	0.09 **	0.07	0.08
	(0.01)	(0.03)	(0.03)	(0.02)	(0.02)	(0.04)	(0.05)	(0.06)
Delta crossing	109.6 **	243.0 ***	139.9 **	74.3	27.8	-42.3	-47.3	-30.8
	(49.2)	(59.5)	(59.3)	(61.1)	(25.3)	(37.2)	(37.3)	(40.2)
State Water Boards review	8.9	253.8 ***	170.1	-29.8	-47.0 **	-82.2 **	-67.8 *	-68.1 *
	(16.1)	(95.5)	(119.3)	(43.1)	(23.1)	(38.8)	(34.5)	(36.6)
Import into project	10.9	23.3	69.1	0.0	-40.3	74.8	44.6	21.5
	(20.4)	(46.9)	(73.4)	(27.9)	(47.9)	(52.4)	(55.2)	(59.2)
Export from project	14.0	51.9	7.9	44.9	-207.7 **	-247.3 ***	-213.1 ***	-202.7 ***
	(16.5)	(50.4)	(63.9)	(43.9)	(99.9)	(74.5)	(65.1)	(67.5)
Agricultural counterparty	46.3 **	1.1	26.4	31.5 *	5.1	55.3	24.6	22.3
	(20.5)	(28.2)	(35.6)	(17.7)	(26.5)	(48.2)	(40.3)	(42.8)
	Panel B: Post-LASSO (Cost factors selected by LASSO)							
		(2)	(3)	(4)		(6)	(7)	(8)
River distance (km)				0.08				
				(0.15)				
Virtual distance (km)							-0.18	-0.16
							(0.11)	(0.11)
Delta crossing		106.6 ***	103.0 ***	75.7				
		(19.16)	(29.89)	(50.67)				
State Water Boards review		290.5 ***	162.7 *				-44.3 **	-50.3 **
		(90.85)	(83.21)				(20.05)	(21.90)
Import into project			68.2					
			(62.01)					
Export from project						-240.2 **	-211.7 ***	-193.7 **
						(97.9)	(79.5)	(78.0)
Agricultural counterparty				30.1 *				
				(17.0)				
Seller FE		✓	✓					
Seller's region by year FE		✓						
Seller's subregion by year FE			✓					
Seller by year FE	✓			✓				
Buyer FE						✓	✓	
Buyer's region by year FE						✓		
Buyer's subregion by year FE							✓	
Buyer by year FE					✓			✓
Observations	532	532	532	532	431	431	431	431
Clusters	507	507	507	507	304	304	304	304

Regressions of transaction price on cost determinants from Table 2; unitless variables are binary indicators. The left side (columns 1-4) include seller-side fixed effects, comparing prices across buyers. Positive coefficients here reflect a price premium, indicating that the cost determinant is costly to the seller. The right side (columns 5-8) include buyer-side fixed effects, comparing prices across sellers. Negative coefficients here reflect a price discount, indicating that the cost determinant is costly to the buyer. Fixed effects and counts at the bottom describe both Panels A and B. In Panel A, columns 1 and 5 report the results from separate regressions in which the row variable is the only regressor; columns 2-4 and 6-8 each show results from one regression containing all variables as regressors. Variables omitted from Panel B are not selected by LASSO in any of the specifications. Mean price in data is $221. Standard errors are shown in parentheses and clustered by transaction. Region and subregion are hydrological region and planning subarea as defined by the California Department of Water Resources. * p<.1, ** p<.05, *** p<.01. [Back]

Table 4: Demand elasticities

Panel A: First Stage & Reduced Form with One Instrument (OLS)

	First stage: Ln(Price)			Reduced form: Ln(Quantity consumed)		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Own allocation percentage)	-0.81 ***	-0.33 ***	-0.37 ***	0.36 ***	0.34 ***	0.34 ***
	(0.18)	(0.10)	(0.10)	(0.03)	(0.04)	(0.05)
Unit FE	✓	✓	✓	✓	✓	✓
Region-specific time trends		✓	✓		✓	✓
Buyer/seller unit pair FE		✓	✓		✓	✓
Cost factors X Direction			✓			✓
Agent FE			✓			✓
Observations	1031	1031	1031	1031	1031	1031
Clusters	331	331	331	331	331	331
F-statistic	19.8	11.2	14.4	109.9	59.4	51.7

Panel B: Price Elasticities of Demand (IV Lasso)

	Ln(Quantity consumed)					
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Price)	-0.14 ***	-0.14 ***	-0.21 ***	-0.16 ***	-0.16 ***	-0.20 ***
	(0.02)	(0.03)	(0.04)	(0.03)	(0.04)	(0.05)
Lasso-selected instruments	10	7	34	10	10	19
Unit FE	✓	✓	✓	✓	✓	✓
Region-specific time trends		✓	✓	✓	✓	✓
Buyer/seller unit pair FE			✓			✓
Cost factors X Direction				✓		✓
Agent FE					✓	✓
Observations	1053	1053	1053	1053	1053	1053
Clusters	345	345	345	345	345	345
First stage F-statistic	167.1	101.5	70.0	89.6	110.1	65.8
Sup-score weak-ID test	reject	reject	fail to reject	reject	reject	fail to reject

Panel C: Heterogeneous Price Elasticities of Demand (IV Lasso)

	Ln(Quantity consumed)				
			Agricultural regions		
	Urban	Agricultural	Sacramento Valley	San Joaquin Valley	Tulare Lake Basin
	(1)	(2)	(3)	(4)	(5)
Ln(Price)	-0.10 *	-0.23 ***	-0.07 ***	-0.11 ***	-0.80 ***
	(0.06)	(0.05)	(0.02)	(0.04)	(0.12)
Lasso-selected instruments	2	17	9	4	3
Unit FE	✓	✓	✓	✓	✓
Region-specific time trends	✓	✓	✓	✓	✓
Buyer/seller unit pair FE	✓	✓	✓	✓	✓
Cost factors X Direction	✓	✓	✓	✓	✓
Agent FE	✓	✓	✓	✓	✓
Observations	259	767	235	287	235
Clusters	88	243	79	70	84

Regressions estimating unit-level price elasticities of surface water demand by instrumental variables. Units aggregate all observed water users and market participants falling within the same geographic subregion (i.e., planning subarea as defined by the California Department of Water Resources) and sector (agricultural or urban). Observations are unit-transaction since multiple transaction prices are observed in the same unit-year, but quantities are aggregated within unit-year, and standard errors (in parentheses) are clustered by unit-year. Instruments are allocation percentages (of maximum contract amounts) determined yearly for each of 14 contract types within the Central Valley Project and State Water Project (set on the basis of precipitation and snowmelt), interacted with region and unit indicators. Panel A shows the effects on prices and quantities of a single instrument, own allocation percentage. These results are illustrative and not directly used in further analysis because they do not account for the possible effects of others' endowments. Panels B and C reports instrumental variables regressions in which the instruments are selected by Lasso from the full set of potential instruments: each contract type's allocation percentage interacted with unit indicators. Cost factors are the cost determinants listed in Table 2; they are interacted with transaction direction (i.e., whether buying or selling). Agent is the particular identity of the transactor. Region and subregion are hydrological region and planning subarea as defined by the California Department of Water Resources. * p<.1, ** p<.05, *** p<.01. [Back]

Table 5: Annual economic benefits from wholesale surface water markets in several scenarios

Panel A: Observed transactions in existing market

Scenario		Volume traded (acre-feet)	Average marginal valuation	Total gains (millions)
1 **Observed market**	Med	156,000	$ 151	$ 13.4
(within-year spot market	Wet	141,000	$ 121	$ 9.9
only)	Dry	364,000	$ 595	$ 87.7
2 **Observed market**	Med	293,000	$ 151	$ 35.9
(including permanent	Wet	251,000	$ 121	$ 23.5
transfers)	Dry	569,000	$ 595	$ 822.2

Panel B: Simulated efficient market

				Gains (millions)		
Scenario		Additional volume traded (acre-feet)	Average equilibrium price	From lower costs of observed transactions	From new transactions	Total
3 **Efficient market** (only	Med	1,085,000	$ 141	$ 11.3	$ 74.5	$ 85.8
physical transaction costs;	Wet	1,370,000	$ 111	$ 4.7	$ 96.6	$ 101.3
no capacity constraints)	Dry	1,009,000	$ 254	$ 52.9	$ 225.5	$ 278.3

Panel C: Extensions and sensitivity checks

Scenario		Additional volume traded (acre-feet)	Average equilibrium price	From lower costs of observed transactions	From new transactions	Total
4 **Key environmental**	Med	1,079,000	$ 145	$ 11.4	$ 73.9	$ 85.2
constraint (Sacramento	Wet	1,374,000	$ 102	$ 4.8	$ 90.3	$ 95.1
River outflow held fixed)	Dry	1,009,000	$ 254	$ 53.3	$ 225.5	$ 278.8
5 **Linear demand**	Med	1,026,000	$ 121	$ 6.4	$ 78.1	$ 84.5
(instead of isoelastic)	Wet	1,250,000	$ 86	$ 3.4	$ 98.3	$ 101.7
	Dry	1,096,000	$ 197	$ 55.7	$ 195.9	$ 251.7
6 **Zero transaction costs**	Med	1,120,000	$ 139	$ 12.3	$ 76.7	$ 89.0
(as if conveyance were	Wet	1,427,000	$ 111	$ 5.6	$ 102.6	$ 108.1
costless)	Dry	1,130,000	$ 237	$ 74.5	$ 250.2	$ 324.8
7 **Larger elasticities** (using	Med	1,869,000	$ 141	$ 10.3	$ 132.9	$ 143.2
individual-level estimates	Wet	2,537,000	$ 103	$ 3.8	$ 180.1	$ 183.8
from the literature)	Dry	2,530,000	$ 277	$ 56.9	$ 653.9	$ 710.7
8 **More-extreme marginal**	Med	1,900,000	$ 170	$ 19.9	$ 259.4	$ 279.3
values (using a twice-as-	Wet	2,156,000	$ 151	$ 13.2	$ 260.2	$ 273.4
dispersed log distribution)	Dry	1,671,000	$ 246	$ 81.8	$ 325.8	$ 407.6
9 **Both larger elasticities**	Med	3,139,000	$ 161	$ 14.5	$ 394.9	$ 409.4
and more-extreme	Wet	3,397,000	$ 146	$ 9.9	$ 382.5	$ 392.3
marginal values	Dry	3,130,000	$ 202	$ 75.0	$ 536.0	$ 610.9

Per-year welfare analysis of the existing market (Panel A) and counterfactual simulations (Panels B and C). Simulation trading volumes and prices are found by solving constrained optimization problems as described in Step 4, and trading volumes are relative to observed post-trading allocations. Each scenario is run in three versions: median ("med"), wet, and dry; these use baseline quantities corresponding to the middle, first, and last quintiles of quantities observed over time. For Scenarios 3-5 and 7-9, transactions incur physical conveyance costs calibrated from the hydrological network model. Scenarios 4-9 introduce extensions to the main simulation (Scenario 3) that are fully explained in the text. All dollar figures are in 2010 USD; gains are per year. [Back]

A Market Power as an Alternative Explanation for Price Gaps

In the main sections of this paper, I document price gaps associated with regulatory barriers and other cost determinants, and I interpret these price gaps as marginal transaction costs. In this appendix, I investigate whether these price gaps may instead be explained by market power.

Because my model of bilateral negotiations rules out the successful exercise of market power, I first develop an alternative model of spatial trade that allows both buyers and sellers to exploit market power. Then, I derive and perform an empirical procedure that adjusts raw prices for market power, using estimated pass-through rates as sufficient statistics. Both the model and empirical procedure are based in, and extend the approach of, Atkin and Donaldson (2015). Finally, using these adjusted prices, I re-estimate marginal transaction costs as in Section 5. I find little change in the overall pattern of marginal transaction costs, and so I conclude that the issue of market power is small relative to transaction costs.

Even if markups are non-negligible, however, the approach in the main sections of this paper are still interpretable. The price gaps I estimate in 5 would instead measure both marginal transaction costs and differences in markups (and markdowns) that are correlated with the selected cost determinants. Because both transaction costs and markups drive wedges between observed prices and marginal valuations, the bounds I estimate for marginal valuations using these estimated price gaps are still valid.[15] Then, as long as these differential markups are endogenous to the transaction costs (as opposed to arising from exogenous aspects of the market structure), eliminating transaction costs would also eliminate markups, and counterfactual price and quantity analysis proceeds through the same way. What markups do change is the interpretation of consumer surplus in welfare analysis. The deadweight loss triangles still represent true potential economic gains from removing the cost determinants, but now some of the price gap rectangle may simply represent transfers between parties without lost efficiency.

A.1 Model of spatial trade with market power

As in the main model, N water districts indexed by n (or by o for origin and d for destination) are given initial endowment E_n that is allocated efficiently among a continuum of consumers. Consumers' preferences can be aggregated such that each district has an inverse demand function $V_n(Q_n)$, which gives marginal valuations as a function of quantity consumed Q_n. Trade across districts is conducted by two layers of intermediaries. In each district, selling intermediaries ("sellers") can buy units of water from consumers (at their marginal valuations $V_o(Q_o)$) and sell to buying intermediaries ("buyers") in another district. Buyers, in turn, buy water from sellers and can sell to consumers in their own district (at their marginal valuations $V_d(Q_d)$). Sellers and buyers meet at exchange points unique to each pair of districts, where prices P_{od} are determined.

Each transaction i generates constant marginal transaction costs for both sellers and buyers, as

[15] Under an analogous assumption for markups as Assumption 3 for transaction costs: that differences in markups associated with specific cost determinants are constant and additively linear.

in Assumption 1. However, sellers and buyers may also incur fixed costs: $C^s_{od}(q_{iod}) = \tau^s_{od} q_{iod} + F^s_{od}$ and $C^b_{od}(q_{iod}) = \tau^b_{od} q_{iod} + F^b_{od}$.

Also unlike in the main model, strategic interactions are possible among sellers at specific destinations, and among buyers at specific origins. One seller's quantity decisions may affect the profits of other sellers to the same district through the aggregate quantity sold, and similarly for buyers. I summarize this strategic interaction using a competitiveness index, following Atkin and Donaldson (2015).

Assumption 5. Market structure. *Buying and selling intermediaries choose quantities q_{iod} to maximize profits subject to the expected response of other intermediaries. The competitiveness index $\phi^s_{od} \equiv \left(\frac{dQ_d}{dq_{iod}}\right)^{-1}\left(\frac{Q_d}{q_{iod}}\right)$ for sellers and $\phi^b_{od} \equiv \left(\frac{dQ_o}{dq_{iod}}\right)^{-1}\left(\frac{Q_o}{q_{iod}}\right)$ for buyers is fixed for each origin-destination district pair.*

This approach nests many specific models of market structure. For Cournot oligopoly with identical intermediaries, the competitiveness index ϕ^s_{od} equals Q_d/q_{iod}, or the number of intermediaries. For perfect collusion or a pure monopoly, $\phi^s_{od} = 1$, and for perfect competition, $\phi^s_{od} \longrightarrow \infty$. Note that both buyers and sellers cannot both successfully exercise market power in the same origin/destination pair. Market power comes from keeping quantities lower or higher than efficient; these are mutually exclusive. However, by permitting both in the same model, I allow the market power to appear on either side (which may even change from place to place).

Next, I define markups (also referred to as markdowns, for buyers) as any difference between prices and willingness to pay (or to accept):

$$\begin{aligned} (Sellers) \qquad & \mu^s_{od}(Q_o) \equiv P_{od}(Q_o) - V_o(Q_o) - \tau^s_{od} \\ (Buyers) \qquad & \mu^b_{od}(Q_d) \equiv V_d(Q_d) - \tau^b_{od} - P_{od}(Q_d). \end{aligned} \tag{18}$$

I then assume there are no strategic interactions *across* sellers and buyers.

Assumption 6. No interactions across sides of the market. *Markups set by sellers and buyers are not affected by the overall quantities consumed in their own districts: $\frac{\partial \mu^s_{od}}{\partial Q_o} = 0$ and $\frac{\partial \mu^b_{od}}{\partial Q_d} = 0$.*

This says that in choosing quantities, intermediaries consider only the other side's demand response, not any expected strategic change in markups. A sufficient condition for this assumption is that market power exists on only one side of the market: for any origin-destination pair, one side may exert market power, while the other consists of price-takers. (Appendix B describes a model of bargaining under bilateral oligopoly that allows for more symmetric analysis.)

With this setup in place, the first-order conditions for buyers and sellers yield expressions for markups:

$$\begin{aligned} (Sellers) \qquad & \mu^s_{od} = -\left(\frac{\partial V_d}{\partial Q_d}\right)\frac{Q_d}{\phi^s_{od}} \\ (Buyers) \qquad & \mu^b_{od} = +\left(\frac{\partial V_o}{\partial Q_o}\right)\frac{Q_o}{\phi^b_{od}}. \end{aligned} \tag{19}$$

Note that if markups are zero (such as under perfect competition), Equation 18 becomes identical to the result for price determination in the main model.

A.2 Deriving the estimation procedure

Following Atkin and Donaldson (2015), I use pass-through rates as sufficient statistics for the competitive structure of the market. I first define pass-through rates as the absolute rate at which costs (for sellers) or revenues (for buyers) are passed through to market equilibrium prices: $\rho^s_{od} \equiv \partial P_{od}/\partial MWTP_{od}$ and $\rho^b_{od} \equiv \partial P_{od}/\partial MWTA_{od}$, where $MWTP_{od} = V_o(Q_o) + \tau^s_{od}$ and $MWTA_{od} = V_d(Q_d) + \tau^b_{od}$. Then, I assume demand takes a particularly convenient functional form.

Assumption 7. Bulow-Pfeiderer demand. *Consumer preferences are time-invariant and take the form*

$$V_n(Q_n) = \begin{cases} a_n - b_n(Q_n)^{\delta_n} & \text{if } \delta_n > 0, a_n > 0, b_n > 0, 0 < Q_n < (a_n/b_n)^{1/\delta_n} \\ a_n - b_n \ln(Q_n) & \text{if } \delta_n = 0, a_n > 0, b_n > 0, 0 < Q_n < \mathrm{e}^{(a_n/b_n)} \\ a_n - b_n(Q_n)^{\delta_n} & \text{if } \delta_n < 0, a_n \geq 0, b_n < 0, 0 < Q_n < \infty. \end{cases}$$

Under this assumption, pass-through rates can be expressed as $\rho^s_{od} = \left(1 + \frac{\delta_d}{\phi^s_{od}}\right)^{-1}$ and $\rho^b_{od} = \left(1 + \frac{\delta_o}{\phi^b_{od}}\right)^{-1}$, and each is fixed within seller-buyer pair. Inserting this functional form for demand into buyers' and sellers' first-order conditions, and adding time subscripts, yields:

$$\begin{aligned} &(Sellers) \quad P_{odt} = \gamma^s_{od} V_{ot} + \rho^s_{od}\tau^s_{odt} - (1-\rho^s_{od})(\tau^b_{odt} - a_d) + (1-\rho^s_{od})(\frac{1}{\rho^b_{od}} - 1)a_o \\ &(Buyers) \quad P_{odt} = \gamma^b_{od} V_{dt} - \rho^b_{od}\tau^b_{odt} + (1-\rho^b_{od})(\tau^s_{odt} + a_o) + (1-\rho^b_{od})(\frac{1}{\rho^s_{od}} - 1)a_d, \end{aligned} \tag{20}$$

where $\gamma^s_{od} \equiv (\rho^s_{od} + \rho^b_{od} - 1)/\rho^b_{od}$ and $\gamma^b_{od} \equiv (\rho^s_{od} + \rho^b_{od} - 1)/\rho^s_{od}$.

From here, I can obtain estimates of transaction costs – accounting for markups – via a two-step procedure.

Step 1: Estimate pair-specific pass-through rates for buyers and sellers. First, to transform 20 into a regression model, I again adopt Assumption 3 (additively linear determinant-specific transaction costs). I absorb demand levels (terms depending on a_o and a_d) and time-invariant portions of transaction costs with buyer-seller pair fixed effects. Then, I regress prices on previously-estimated marginal valuations (also restoring Assumption 4 to ensure marginal valuations are point estimates rather than bounds):

$$\begin{aligned} &(Sellers) \quad P_{iodt} = \gamma^s_{od}\hat{V}_{ot} + \lambda^s_{od} + \varepsilon^s_{iodt} \\ &(Buyers) \quad P_{iodt} = \gamma^b_{od}\hat{V}_{dt} + \lambda^b_{od} + \varepsilon^b_{iodt}, \end{aligned} \tag{21}$$

where I assume movement in marginal valuations within buyer-seller pair is uncorrelated with other, unobserved determinants of prices. (The error terms arise from idiosyncratic time variation in transaction costs, $\tilde{\tau}^s_{odt}$ and $\tilde{\tau}^b_{odt}$.) Upon obtaining estimates $\hat{\gamma}^s_{od}$ and $\hat{\gamma}^b_{od}$, I can back out $\hat{\rho}^s_{od}$ and $\hat{\rho}^b_{od}$ by solving the system of two equations defining γ^s_{od} and γ^b_{od}: $\rho^s_{od} = \gamma^b_{od}/(\gamma^s_{od} + \gamma^b_{od} - \gamma^s_{od}\gamma^b_{od})$ and $\rho^s_{od} = \gamma^s_{od}/(\gamma^s_{od} + \gamma^b_{od} - \gamma^s_{od}\gamma^b_{od})$.

Step 2: Adjust prices for pass-through and re-estimate transaction costs from gaps in adjusted prices. Second, by inserting Assumption 3 into Equation 20 and rearranging, I obtain

$$\begin{aligned} &(Sellers) \qquad \frac{1}{\hat{\rho}^s_{od}}\left(\hat{\rho}^s_{od}P_{iodt} + (1-\hat{\rho}^s_{od})\hat{V}_{dt}\right) = \omega_{ot} + \tau^s\mathbf{B}_{od} + v^s_{iodt} \\ &(Buyers) \qquad \frac{1}{\hat{\rho}^b_{od}}\left(\hat{\rho}^b_{od}P_{iodt} + (1-\hat{\rho}^b_{od})\hat{V}_{ot}\right) = \omega_{dt} + \tau^b\mathbf{B}_{od} + v^b_{iodt}. \end{aligned} \tag{22}$$

These can be estimated as regressions, yielding new unbiased estimates of determinant-specific transaction costs $\hat{\tau}^s$ and $\hat{\tau}^b$. Here I have to assume that inverse demand levels a_n (absorbed by the error terms v_{iodt}) are uncorrelated with the cost determinants $\mathbf{B}_{od}$, which may appear to be a strong assumption but is also implicitly assumed in the main sections of this paper.

Finally, note that the full price gap as estimated in Section 5 (i.e., without adjusting prices for imperfect pass-through) identifies the sum of determinant-specific transaction costs and determinant-specific markups. Therefore, the difference between these two estimates identifies the determinant-specific markups.

A.3 Empirical implementation

To bring the regressions from Step 1 and Step 2 to the data, I make several more choices.

First, I coarsen the level of origin and destination *od* in order to conserve statistical power. Instead of estimating parameters at the consumer level, I use the level of groups defined by the interaction of hydrologic region and major cost determinants (the ones for which I measure transaction costs). Regression 21 is more demanding of the data than other regressions in this paper, as it requires estimating both an intercept and a slope (i.e., a fixed effect and a coefficient) for each origin-destination pair. If I continued to use consumer-level pairs, very few parameters would be identified, since most consumer pairs do not have repeated transactions over time. This coarser grouping of origins and destinations still allows markups to vary discontinuously across the cost determinants as well as across geography. Of 1,190 possible group-pair cells, 48 have at least 3 observations (the minimum necessary to identify both point estimates and standard errors).

Second, I instrument for marginal valuations using the leave-out mean of marginal valuations within each of these cells, estimating Equation 21 via two-stage least squares. This addresses the potential concern of mechanical correlation between prices and marginal valuations, since marginal valuations are constructed using within-consumer price data. It does not necessarily solve any other potential endogeneity concerns. Unfortunately, I cannot use the same instrument as in estimating demand elasticities, namely fluctuations in water entitlements. Here, I need to estimate cell-specific parameters, but water entitlements have variation within only a subset of the cells.

Third, I impute 1 for any parameters $\hat{\gamma}^s_{od}$ or $\hat{\gamma}^b_{od}$ that are unidentified, in order to still calculate values for $\hat{\rho}^s_{od}$ and $\hat{\rho}^b_{od}$. That is, I assume complete pass-through, as in the case of perfect competition. This affects origin-destination cells that have only 1 or 2 observations, which comprise 15% of my dataset.

Finally, in the second step, I drop observations for which the pass-through rate ($\hat{\rho}^s_{od}$ or $\hat{\rho}^b_{od}$) is below 0.2 or greater than 2, following Atkin and Donaldson (2015). This ensures that the estimates are not driven by noisy outliers in estimated pass-through rates, which may be amplified because they appear in the denominator on the left-hand side. This sample restriction affects 19% of my dataset.

A.4 Results

In Step 1, I estimate Equation 21 by two-stage least squares regression, obtain the estimates $\hat{\gamma}^s_{od}$ and $\hat{\gamma}^b_{od}$, and algebraically solve their definitions to obtain $\hat{\rho}^s_{od}$ and $\hat{\rho}^b_{od}$. These are pass-through rates for sellers and buyers, estimated using time-series variation in regional marginal valuations.

Figure B1 plots histograms of these estimated pass-through rates. Sellers' pass-through rates cluster tightly around 1, which is consistent with models in which sellers are not able to successfully exercise market power, including the model of perfect competition in the main sections of this paper. Buyers' pass-through rates are more dispersed, with a mode below 1, suggesting that buyers may have some ability to exercise market power. (Note that depending on market structure, market power may be consistent with pass-through rates either below or above 1).

In Step 2, I adjust raw prices for possible markups and markdowns and then re-estimate marginal transaction costs associated with the cost determinants selected in Section 5. Under the assumptions of the model in this appendix, the adjusted prices reflect what prices would be in a partial-equilibrium counterfactual without markups or markdowns. The prior estimates of transaction costs (in Section 5) then may be biased, since they capture both marginal transaction costs as well as any differential markups or markdowns that vary with the cost determinants (and presumably arise from the cost determinants, since transaction costs reduce competition). In contrast, the estimates here account for markups and markdowns, allowing me to better isolate the marginal transaction costs.

Table B1 shows the results of these marginal transaction cost regressions, from Equation 22. This table is identical to Table 3 from Section 5, except that the dependent variable is adjusted prices rather than raw prices. For sellers, adjusting for markups leaves the estimates essentially unchanged. For example, the coefficient on the aggregated cost determinant variable (Panel A, column 4) decreases from 88.5 to 72.4, but this difference is well within the range of statistical uncertainty. Coefficients on the disaggregated cost determinants are also individually similar to those in Table 3. For buyers, the overall pattern is similar as in Table 3, but the coefficients are considerably noisier. The point estimates are larger, suggesting that market power in this setting does not exacerbate marginal transaction costs, but rather acts to *reduce* the transaction costs associated with these cost determinants. However, once again, these new estimates (using adjusted prices) are not statistically distinguishable from the prior estimates (which used raw prices).

Overall, the results of this exercise suggest that market power does not explain the large marginal transaction costs I estimate in Section 5. After adjusting for possible markups, marginal transaction costs for sellers remain large and statistically distinguishable from zero. After adjusting for possible markdowns, marginal transaction costs for buyers are no smaller; if anything they may even be larger than when estimated with raw prices. I conclude that the assumption of no exercise of market power, maintained throughout the main sections of this paper, is reasonable, and that I can interpret marginal transaction costs as true deadweight loss rather than transfers between consumers.

Appendix A Figures and Tables

Figure B1: Estimated pass-through rates

Histograms of estimated pass-through rates for sellers and buyers. Each pass-through rate corresponds to a different origin-destination region pair, where regions are defined by the interaction of hydrologic region and major cost determinants. Pass-through rates are derived from transaction-level regressions of prices on marginal valuations (as estimated in Step 2 of the main part of this paper) in which both coefficients and fixed effects are estimated separately for each origin-destination pair. Each marginal valuation is instrumented with the leave-out regional mean of marginal valuations, meaning the coefficients are identified using time-series variation in regional marginal valuations.

Table B1: Marginal transaction costs, adjusting for incomplete pass-through

Dependent variable: **Price** ($/acre-foot)	**SELLERS** (positive coefficients are costly)				**BUYERS** (negative coefficients are costly)				**Total Cost**
Panel A: Aggregated cost factors									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Additional regulatory review	30.8	43.9 *	42.5	72.4 ***	-273	-594	-1103 *	-1132	1204 *
	(43.6)	(22.6)	(30.8)	(25.7)	(258)	(512)	(576)	(717)	(717)
Distance (km)	-5.21	-3.44 *	-0.97	-5.00	-3.51	-17.39 **	-33.41 ***	-29.21 **	24.21 **
	(3.94)	(1.98)	(2.20)	(3.13)	(12.21)	(8.46)	(9.27)	(11.87)	(12.28)
Cross-sector transfer (ag-to-urban)	-9.4 *	-9.0 ***	6.5	-6.3	17.9	-2.3	18.3 *	18.8	-25.1 *
	(4.8)	(3.0)	(4.4)	(4.0)	(16.2)	(11.6)	(10.6)	(13.4)	(14.0)
Seller's region by year FE	✓	✓	✓	✓					
Seller's subregion by year FE		✓		✓					
Seller FE			✓	✓					
Seller by year FE				✓					
Buyer's region by year FE					✓	✓	✓	✓	
Buyer's subregion by year FE						✓		✓	
Buyer FE							✓	✓	
Buyer by year FE								✓	
Observations	1,971	1,971	1,971	1,971	1,823	1,823	1,823	1,823	
Clusters	125	125	125	125	127	127	127	127	
Panel B: Disaggregated cost factors									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Delta crossing only	135 ***	130 ***	108 **	195 ***	-716 *	-914 **	-1681 ***	-1740 ***	1934 ***
	(36)	(48)	(47)	(49)	(406)	(460)	(559)	(636)	(638)
SWRCB review only	-26	35 ***	14	46 ***	18	144	-478	-262	308
	(26)	(12)	(26)	(13)	(386)	(658)	(739)	(1173)	(1173)
DWR review only	-37	-56 *	212	10	120	-153	-498	-68	79
	(39)	(34)	(195)	(33)	(396)	(211)	(350)	(111)	(116)
USBR review only	-91	-56	-541	-3	91	126	-454	-237	233
	(105)	(36)	(500)	(11)	(329)	(602)	(707)	(1109)	(1109)
Delta crossing + SWRCB review	-	-	-	-	-227	-986	-1265 **	-1383 *	-
	-	-	-	-	(496)	(595)	(625)	(719)	-
Delta crossing + DWR review	165 ***	105 **	129 ***	151 ***	-	-	-	-	-
	(37)	(43)	(46)	(50)	-	-	-	-	-
Delta crossing + USBR review	1	51 *	94 *	195 ***	-335	-875 *	-1244 **	-1532 **	1727 **
	(32)	(30)	(48)	(72)	(338)	(487)	(565)	(679)	(683)
County review	83.2 *	-24.5 *	-26.8	-30.4 **	34.1	-8.8	-25.3	-27.0	-3.4
	(43.2)	(14.5)	(19.1)	(14.1)	(80.4)	(64.0)	(73.8)	(64.5)	(66.0)
Distance (km)	-0.079	-0.006	0.047	0.002	-0.103	-0.524	-0.677	-0.326	0.328
	(0.074)	(0.037)	(0.040)	(0.056)	(0.404)	(0.355)	(0.579)	(0.591)	(0.593)
Seller's region by year FE	✓	✓	✓	✓					
Seller's subregion by year FE		✓		✓					
Seller FE			✓	✓					
Seller by year FE				✓					
Buyer's region by year FE					✓	✓	✓	✓	
Buyer's subregion by year FE						✓		✓	
Buyer FE							✓	✓	
Buyer by year FE								✓	
Observations	2,037	2,037	2,037	2,037	1,834	1,834	1,834	1,834	
Clusters	181	181	181	181	130	130	130	130	

Regressions of transaction price on cost determinants from Table 2, adjusting prices for incomplete pass-through (and therefore for some possible forms of market power). The left side (columns 1-4) include seller-side fixed effects, comparing prices across buyers. Positive coefficients reflect a price premium, indicating that the cost determinant is costly to the seller. The right side (columns 5-8) include buyer-side fixed effects, comparing prices across sellers. Negative coefficients reflect a price discount, indicating that the cost determinant is costly to the buyer. Column (9) calculates the linear combination of the absolute value of (4) and (8), the preferred specifications. Standard errors are shown in parentheses and clustered by region X year. Region and subregion are hydrological region and planning subarea as defined by the California Department of Water Resources. All regressions also include comparison-sample indicator variables. * p<.1, ** p<.05, *** p<.01.

B Nash-in-Nash Bargaining as an Alternative Model

The model in the main section of this paper relies on an assumption of perfect competition among intermediaries that trade across water districts. In this section, I propose a model with alternative foundations that may be more realistic. I show that, under certain conditions, this model yields approximately the same results on equilibrium prices as I use in my empirical approach. In this model, a finite number of districts engage in simultaneous bilateral negotiations to trade endowments of a homogeneous good. Both buyers and sellers incur nominal transaction costs that may be pair-specific and asymmetric. The solution concept, as in other bargaining models in the industrial organization literature (e.g., Ho and Lee 2017), is a Nash equilibrium in Nash bargaining solutions.

B.1 Basic bilateral negotiation

I begin with a general bargaining model. Here I make no assumptions on how negotiations proceed; the goal is simply to characterize the set of transaction prices that are individually rational.

An economy has N districts indexed by n who each have initial endowment E_n of water, a single homogeneous good. Each district has a total valuation function $\mathbb{V}_n(Q_n)$ that is positive, increasing, concave, and differentiable. The derivative of total valuation, $V_n(Q_n) \equiv d\mathbb{V}_n(Q_n)/dQ_n$ gives inverse demand, or marginal valuation at a given quantity demanded Q_n. Any pair of districts can trade water with each other; a transaction consists of quantity q_{od} (sold from origin o to destination d) and per-unit price P_{od}. In trading, districts incur pair-specific, asymmetric transaction costs: $C^s_{od}(q_{od})$ in selling and $C^b_{od}(q_{od})$ in buying. These cost functions are weakly positive, weakly increasing, and differentiable.

Consider a negotiation between a single pair of districts, holding fixed Q_n (the outcomes of all other negotiations). Net profits for a seller o and a buyer d are the additional profits obtained by a given transaction:

$$\begin{aligned} \pi^s_{od} &\equiv P_{od}q_{od} + \mathbb{V}_o(Q_o - q_{od}) - \mathbb{V}_o(Q_o) - C^s_{od}(q_{od}) \\ \pi^b_{od} &\equiv -P_{od}q_{od} + \mathbb{V}_d(Q_d + q_{od}) - \mathbb{V}_d(Q_d) - C^b_{od}(q_{od}). \end{aligned} \tag{23}$$

A transaction is individually rational if and only if net profits are positive for both seller and buyer: $\pi^s_{od} \geq 0$, $\pi^b_{od} \geq 0$. Rearranged, these conditions are:

$$\begin{aligned} P_{od}q_{od} &\geq WTA^s_{od} \equiv \mathbb{V}_o(Q_o) - \left[\mathbb{V}_o(Q_o - q_{od}) - C^s_{od}(q_{od})\right] \\ P_{od}q_{od} &\leq WTP^b_{od} \equiv \left[\mathbb{V}_d(Q_d + q_{od}) - C^b_{od}(q_{od})\right] - \mathbb{V}_d(Q_d). \end{aligned} \tag{24}$$

Combining these, the amount transferred must fall between the seller's willingness to accept and the buyer's willingness to pay: $WTA^s_{od} \leq P_{od}q_{od} \leq WTP^b_{od}$. Without loss of generality, I parameterize the proportion of bilateral surplus that is obtained by the seller as $\sigma_{od} \in [0,1]$. Then, all individually rational prices meet the condition $P_{od}q_{od} = \sigma_{od}WTP^s_{od} + (1-\sigma_{od})WTA^s_{od}$, or

$$P_{od} = (1-\sigma_{od})\left(\frac{\mathbb{V}_o(Q_o) - \mathbb{V}_o(Q_o - q_{od})}{q_{od}}\right) + \sigma_{od}\left(\frac{\mathbb{V}_d(Q_d + q_{od}) - \mathbb{V}_d(Q_d)}{q_{od}}\right)$$
$$+(1-\sigma_{od})\left(\frac{C^s_{od}(q_{od})}{q_{od}}\right) - \sigma_{od}\left(\frac{C^b_{od}(q_{od})}{q_{od}}\right). \quad (25)$$

In other words, individually rational prices are a linear combination of average incremental valuations and average transaction costs.

B.2 Nash equilibrium in simultaneous bilateral negotiations

Next, I specify conditions of these bilateral negotiations that will permit an equilibrium in prices and quantities. Throughout, I rely on results from Björnerstedt and Stennek (2007).

Each possible pair of districts in the economy engages in simultaneous Rubinstein-Stahl negotiations. Districts alternate bids in each round until a bid is accepted or negotation fails. Each bilateral negotiation is conducted by a separate delegated agent, who observes the entire history of bids and decisions up to that point but not bids in the same round.

While this setup does not literally describe the negotiation patterns in the California water market, its assumptions are weaker than the model in the main section of the paper. Instead of perfect competition among hypothetical intermediaries, this model allows bilateral oligopoly among districts, who conduct transactions directly. Collard-Wexler et al. (2017) relax the delegated-agent assumption, allowing consumers to use information across simultaneous negotiations, but their framework is not a close fit for my context, since it does not allow for endogenous trading-network formation nor for pair-specific transaction costs.

I again adopt Assumption 1, constant marginal transaction costs: $C^s_{od} = \tau^s_{od}q_{od}$ and $C^b_{od} = \tau^b_{od}q_{od}$. Björnerstedt and Stennek (2007) generalize this slightly, but fixed costs still need to be small. Agents maximize the sum of their districts' per-period profits, discounted by common determinant δ, where per-period profits are:

$$\Pi^s_{od} \equiv \mathbb{V}_o(E_o - \sum_d q_{od}) + \sum_d P_{od}q_{od} - \sum_d \tau^s_{od}q_{od} \quad (26)$$
$$\Pi^b_{od} \equiv \mathbb{V}_d(E_d + \sum_o q_{od}) - \sum_o P_{od}q_{od} - \sum_o \tau^b_{od}q_{od}.$$

Given this setup, Proposition 1 of Björnerstedt and Stennek (2007) ensures two things. First, there exists a sequential equilibrium in which agents immediately agree on the vector of transaction quantities q_{od} that maximizes bilateral surplus – i.e., the additional combined profit taking others as given, $\pi^s_{od} + \pi^b_{od}$. Taking the first-order condition shows that each pair of districts either does not trade ($q_{od} = 0$) or trades a positive amount implicitly defined by the condition

$$V_d(Q_d + q_{od}) - \tau^b_{od} = V_o(Q_o - q_{od}) + \tau^s_{od}. \quad (27)$$

This result is familiar from this paper's main model; it is identical to Equation 3. It says that the quantity traded equates the marginal valuations of buyer and seller, up to transaction costs.

Second, as the time between negotiation periods becomes small ($\delta \to 1$), prices give an equal split of the bilateral surplus. That is, $\sigma_{od} = \frac{1}{2}$ for all districts o and d. Finally, Proposition 2 of Björnerstedt and Stennek (2007) ensures that quantities in this equilibrium coincide with the Walrasian quantities; i.e., the quantities that would result if all districts were price-takers. These results also constitute a Nash equilibrium in Nash bargaining solutions; the outcome of each bilateral Rubinstein-Stahl negotiations (as the discount determinant approaches 1) maximizes that pair's bilateral Nash product, taking as given the outcomes of all other negotiations.

B.3 Small transaction quantities

With an equal split of surplus, prices in this equilibrium are given as

$$P_{od} = \frac{1}{2}\left(\frac{\mathbb{V}_o(Q_o) - \mathbb{V}_o(Q_o - q_{od})}{q_{od}}\right) + \frac{1}{2}\left(\frac{\mathbb{V}_d(Q_d + q_{od}) - \mathbb{V}_d(Q_d)}{q_{od}}\right) + \frac{1}{2}(\tau^s_{od} - \tau^b_{od}). \tag{28}$$

Note that the first two expressions in parentheses resemble the definition of a derivative. As individual transaction quantities become small relative to total quantity consumed ($q_{od}/Q_n \to 0$), prices converge to[16]

$$P_{od} = V_o(Q_o) + \tau^s_{od} = V_d(Q_d) - \tau^b_{od}. \tag{29}$$

Once again, prices equalize the seller's marginal willingness to accept with the buyer's marginal willingness to pay. This is exactly the same result as Equation 3 in the perfectly competitive model – the key identifying condition I use to estimate transaction costs and marginal valuations.

With Equation 29, this appendix suggests that my overall empirical approach does not hinge on the strong assumptions of the perfectly competitive model from Section 3. Instead, my key revealed-preference condition can be viewed as the approximate outcome of a significantly more realistic model of bilateral bargaining. This holds as long as transaction quantities are small relative to total quantity consumed – which is true for the vast majority of water districts in California.

B.4 Fixed costs

With arbitrary fixed costs, the equilibrium results of Björnerstedt and Stennek (2007) do not necessarily hold. (See Appendix A for a model that admits arbitrary fixed costs.) However, the set of individually rational prices can still be characterized.

Instead of Assumption 1, I allow both constant marginal transaction costs and fixed costs: $C^s_{od} =$

[16] This result is obtained from applying the definition of a derivative to Equation 28 and inserting Equation 27. It corresponds to Proposition 6 of Björnerstedt and Stennek (2007), which allows slightly more general transaction costs.

$\tau^s_{od} q_{od} + F^s_{od}$ and $C^b_{od} = \tau^b_{od} q_{od} + F^b_{od}$. Here, prices in observed transactions must meet the condition

$$
\begin{aligned}
P_{od} &= (1-\sigma_{od})\left(\frac{\mathbb{V}_o(Q_o) - \mathbb{V}_o(Q_o - q_{od})}{q_{od}}\right) + \sigma_{od}\left(\frac{\mathbb{V}_d(Q_d + q_{od}) - \mathbb{V}_d(Q_d)}{q_{od}}\right) \\
&\quad + (1-\sigma_{od})\tau^s_{od} - \sigma_{od}\tau^b_{od} + \frac{1}{q_{od}}\left[(1-\sigma_{od})F^s_{od} - \sigma_{od}F^b_{od}\right]. \qquad (30)
\end{aligned}
$$

for some value of $\sigma_{od} \in [0,1]$.

As transaction quantities become small relative to total quantity consumed ($q_{od}/Q_n \to 0$), assuming an equal split of surplus ($\sigma_{od} = \frac{1}{2}$), prices converge to:

$$
\begin{aligned}
P_{od} &= V_o(Q_o) + \tau^s_{od} + \frac{1}{2q_{od}}\left(F^s_{od} - F^b_{od}\right) \qquad (31) \\
&= V_d(Q_d) - \tau^b_{od} + \frac{1}{2q_{od}}\left(F^s_{od} - F^b_{od}\right).
\end{aligned}
$$

Quantities cannot become small relative to fixed costs, or the transaction would become too expensive. However, it is entirely possible that fixed costs are non-negligible yet individual transaction quantities are still small relative to overall quantities.

This result modifies Equation 29 under some conditions to account for fixed costs. What affects prices is not the absolute size of fixed costs, but the difference between the fixed costs incurred by buyer and seller. If the buyer and seller incur dramatically different fixed costs, prices look quite different from Equation 29. However, if both bargaining power and the size of fixed costs are similar between buyer and seller, prices are approximately the same as without fixed costs.

C Appendix Tables

Table D1: Demand elasticities: alternative specifications

Panel A: Price Elasticities of Demand -- Robustness Checks (I)

	Ln(Quantity consumed)					
	Errors clustered by unit			Using marginal valuations		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Price)	-0.14 ***	-0.15 ***	-0.20 **			
	(0.05)	(0.05)	(0.10)			
Ln(Marginal valuation)				-0.13 ***	-0.09 ***	-0.15 ***
				(0.03)	(0.02)	(0.03)
Lasso-selected instruments	10	9	18	12	9	17
Unit FE	✓	✓	✓	✓	✓	✓
Region-specific time trends		✓	✓		✓	✓
Cost factors X Direction		✓	✓		✓	✓
Buyer/seller unit pair FE			✓			✓
Agent FE			✓			✓
Observations	1031	1031	1031	574	574	574
Clusters	36	36	36	318	318	318
First stage F-statistic	165.8	76.7	65.4	71.3	39.2	29.7
Sup-score weak-ID test	fail to reject	reject	fail to reject	fail to reject	fail to reject	fail to reject

Panel B: Price Elasticities of Demand -- Robustness Checks (II)

	Ln(Quantity consumed)					
	Region instruments only			Unit instruments only		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Price)	-0.14 ***	-0.19 ***	-0.21 ***	-0.21 ***	-0.07 **	-0.31 ***
	(0.03)	(0.04)	(0.05)	(0.05)	(0.03)	(0.09)
Lasso-selected instruments	6	5	8	12	6	13
Unit FE	✓	✓	✓	✓	✓	✓
Region-specific time trends		✓	✓		✓	✓
Cost factors X Direction		✓	✓		✓	✓
Buyer/seller unit pair FE			✓			✓
Agent FE			✓			✓
Observations	1031	1031	1031	1031	1031	1031
Clusters	331	331	331	331	331	331
First stage F-statistic	27.6	19.9	14.1	19.5	4.0	13.7
Sup-score weak-ID test	reject	reject	fail to reject	fail to reject	fail to reject	fail to reject

Panel C: Price Elasticities of Demand -- With Year Fixed Effects (Biased)

	Ln(Quantity consumed)					
	IV Lasso			Single instrument (own allocation)		
	(1)	(2)	(3)	(4)	(5)	(6)
Ln(Price)	0.23 **	-0.08	-1.41 ***	5.17	11.95	-7.47
	(0.11)	(0.10)	(0.10)	(7.41)	(36.07)	(8.06)
Number of instruments	2	2	3	1	1	1
Unit FE	✓	✓	✓	✓	✓	✓
Year FE	✓	✓	✓	✓	✓	✓
Cost factors X Direction		✓	✓		✓	✓
Buyer/seller unit pair FE			✓			✓
Agent FE			✓			✓
Observations	1031	1031	1031	1031	1031	1031
Clusters	331	331	331	331	331	331
First stage F-statistic	8.6	0.7	191.7	0.5	0.1	0.7
Sup-score weak-ID test	reject	reject	fail to reject	-	-	-

Regressions estimating unit-level price elasticities of surface water demand by instrumental variables. Units aggregate all observed water users and market participants falling within the same geographic subregion and sector. Observations are unit-transaction but quantities are aggregated within unit-year. Regressions are similar to those in Table 4 but with modifications. Panel A shows results from clustering standard errors by unit instead of unit-year, and from using marginal valuations as the endogenous variable instead of prices. Panel B shows results from using subsets of candidate instruments: allocation percentages interacted with only region indicators or only unit indicators. Panel C shows results from biased regressions that include year fixed effects. * p<.1, ** p<.05, *** p<.01. [Back]

D Appendix Figures

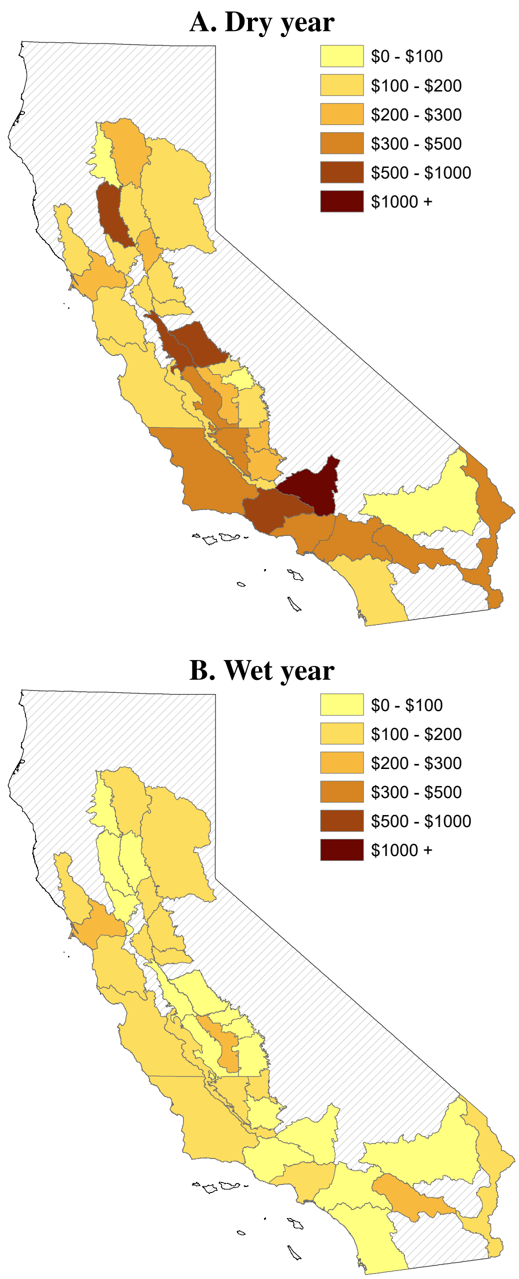

Figure E1: Estimated marginal valuations (per acre-foot) by geography for dry-year (A) and wet-year (B) scenarios. Geographic polygons correspond to subregions (i.e., planning areas as defined by the California Department of Water Resources); areas with diagonal shading have no observed transactions. Values shown are the fitted values from inverting the unit-specific demand models estimated in Step 3 and plugging in the median quantity consumed for each unit across years. (Unit is defined as the intersection of planning area with sector: urban or agricultural). The map shows the weighted average (by quantity) of unit-level marginal valuations across the two units within each planning area. The median-year scenario is shown in Table 6. Unit-level valuations in these three scenarios are direct inputs to the counterfactual simulations in Step 4. [Back]

E Proofs

E.1 Year effects bias elasticity estimates when comparing within a common market

Here I show that including year effects can bias elasticity estimates when prices and quantities are equilibrium outcomes of a common market. To keep the proof as simple as possible, I focus on the reduced form, showing that the effect of the instrument on quantities is biased. Consider a regression of quantities on prices estimated by two-stage least squares with a single instrument. Because the two-stage least squares (2SLS) estimate is equal to the ratio of the first-stage and reduced form coefficients, if the reduced form is biased, the 2SLS estimate is also biased.

Consider a simple data generating process involving two agents. In each year, each agent receives a fixed quantity α_i plus an observed time-varying entitlement z_{it}, of which they keep a fraction β and trade away the remainder to the other agent. Each agent also consumes an idiosyncratic shock ε_{it} that is uncorrelated with entitlements. Total quantities are:

$$q_{1t} = \alpha_1 + \beta z_{1t} + (1-\beta) z_{2t} + \varepsilon_{1t} \qquad \mathbb{E}[\varepsilon_{1t} | z_{1t}, z_{2t}] = 0$$
$$q_{2t} = \alpha_2 + \beta z_{2t} + (1-\beta) z_{1t} + \varepsilon_{2t} \qquad \mathbb{E}[\varepsilon_{2t} | z_{1t}, z_{2t}] = 0$$

This model captures a market with inertia; $\beta = 1$ would correspond to autarky while $\beta = 0.5$ would suggest no inertia, since entitlements given to either agent would be allocated evenly. For simplicity, the model is linear and the coefficient β is constant across the two agents.

First, under this data generating process, a simple fixed effects regression that includes both entitlements (each agent's own entitlement and the other agent's entitlement) would recover the correct, unbiased parameter β, because the econometric model would be identical to the data generating process.

Second, in general, a regression measuring the effect of agents' own entitlement must also control for the other agent's entitlement. An estimate of β from a regression containing only each agent's own entitlement would suffer from omitted variables bias unless the other agent's entitlement z_{-it} is uncorrelated with own entitlement z_{it}.

Third, using year effects will produce a biased estimate of β. Consider the regression

$$q_{it} = \alpha_i + \beta z_{it} + \theta_t + v_{it}.$$

Year effects are incidental parameters, so they can be eliminated by differencing the two agents:

$$\begin{aligned} (q_{1t} - q_{2t}) &= (\alpha_1 - \alpha_2) + \beta(z_{1t} - z_{2t}) + (\theta_t - \theta_t) + (v_{1t} - v_{2t}). \\ \Delta q_t &= \Delta\alpha + \beta \Delta z_t + \Delta v_t. \end{aligned}$$

This is now a simple one-variable ordinary least squares model, so the coefficient estimate $\hat{\beta}$ can be expressed as a ratio of covariances:

$$
\begin{aligned}
\hat{\beta} &= \frac{\text{cov}(\Delta z_t, \Delta q_t)}{\text{var}(\Delta z_t)} = \frac{\text{cov}(\Delta z_t, q_{1t} - q_{2t})}{\text{var}(\Delta z_t)} \\
&= \frac{\text{cov}\big(\Delta z_t, (\alpha_1 + \beta z_{1t} + (1-\beta) z_{2t} + \varepsilon_{1t}) - (\alpha_2 + \beta z_{2t} + (1-\beta) z_{1t} + \varepsilon_{2t})\big)}{\text{var}(\Delta z_t)} \\
&= \frac{\text{cov}\big(\Delta z_t, (2\beta - 1) z_{1t} - (2\beta + 1) z_{2t}\big)}{\text{var}(\Delta z_t)} = (2\beta - 1)\frac{\text{cov}(\Delta z_t, \Delta z_t)}{\text{var}(\Delta z_t)} = 2\beta - 1 \\
&= \beta - (1-\beta)
\end{aligned}
$$

which is not equal to β. Thus, year effects introduce a mechanical relationship such that the estimate $\hat{\beta}$ captures not only the correct effect of the entitlement on the agent's own quantities (β), but also the effect of the entitlement on the other agent's quantities ($1-\beta$). In autarky ($\beta = 1$) there would be no market spillovers and $\hat{\beta}$ would be unbiased. In the no-inertia case of $\beta = 0.5$, the estimated effect would be zero – falsely suggesting that raising entitlements does not increase quantities.

E.2 Solution to the planner's problem has the same necessary conditions as the market equilibrium

First, expand the first term of the maximand and rearrange it:

$$
\begin{aligned}
\sum_k \int_{Q_k^0}^{Q_k^f} V_k(\varphi) d\varphi &= \sum_k \exp\Big(-\frac{\psi_k}{\eta_k}\Big)\Big(\frac{1}{\eta_k}+1\Big)^{-1}\big((Q_k^f)^{\frac{1}{\eta_k}+1} - (Q_k^0)^{\frac{1}{\eta_k}+1}\big) \\
&= \sum_k \exp\Big(-\frac{\psi_k}{\eta_k}\Big)\Big(\frac{1}{\eta_k}+1\Big)^{-1}(Q_k^0)^{\frac{1}{\eta_k}+1}\Big[\Big(\frac{Q_k^f}{Q_k^0}\Big)^{\frac{1}{\eta_k}+1} - 1\Big] \\
&= \sum_k \exp\Big(-\frac{\psi_k}{\eta_k}\Big)\Big(\frac{1}{\eta_k}+1\Big)^{-1}(Q_k^0)^{\frac{1}{\eta_k}+1}\Big[\Big(1 + \frac{-\sum_{l>k} q_{kl} + \sum_{l<k} q_{lk}}{Q_k^0}\Big)^{\frac{1}{\eta_k}+1} - 1\Big].
\end{aligned}
$$

Then, take a first-order condition with respect to q_{od} by setting the derivative of the entire maximand equal to zero (assume unit o sells to unit d, without loss of generality):

$$
\begin{aligned}
0 &= \frac{d}{dq_{od}}\left\{\exp\left(-\frac{\psi_o}{\eta_o}\right)\left(\frac{1}{\eta_o}+1\right)^{-1}(Q_o^0)^{\frac{1}{\eta_o}+1}\left[\left(1+\frac{-\sum_{l>o}q_{ol}+\sum_{l<o}q_{lo}}{Q_o^0}\right)^{\frac{1}{\eta_o}+1}-1\right]\right\} \\
&\quad +\frac{d}{dq_{od}}\left\{\exp\left(-\frac{\psi_d}{\eta_d}\right)\left(\frac{1}{\eta_d}+1\right)^{-1}(Q_d^0)^{\frac{1}{\eta_d}+1}\left[\left(1+\frac{-\sum_{l>d}q_{dl}+\sum_{l<d}q_{ld}}{Q_d^0}\right)^{\frac{1}{\eta_d}+1}-1\right]\right\} \\
&\quad -\frac{d}{dq_{od}}\left\{\sum_{l>o}c_{ol}q_{ol}\right\} \\
&= \exp\left(-\frac{\psi_o}{\eta_o}\right)\left(\frac{1}{\eta_o}+1\right)^{-1}(Q_o^0)^{\frac{1}{\eta_o}+1}\frac{d}{dq_{od}}\left(1+\frac{-\sum_{l>o,l\neq d}q_{ol}+\sum_{l<o}q_{lo}-q_{od}}{Q_o^0}\right)^{\frac{1}{\eta_o}+1} \\
&\quad +\exp\left(-\frac{\psi_d}{\eta_d}\right)\left(\frac{1}{\eta_d}+1\right)^{-1}(Q_d^0)^{\frac{1}{\eta_d}+1}\frac{d}{dq_{od}}\left(1+\frac{-\sum_{l>d}q_{dl}+\sum_{l<d,l\neq o}q_{ld}+q_{od}}{Q_d^0}\right)^{\frac{1}{\eta_d}+1} \\
&\quad -c_{od}\frac{d}{dq_{od}}q_{od} \\
&= \exp\left(-\frac{\psi_o}{\eta_o}\right)\left(\frac{1}{\eta_o}+1\right)^{-1}(Q_o^0)^{\frac{1}{\eta_o}+1}\left(\frac{1}{\eta_o}+1\right)\left(\frac{Q_o^f}{Q_o^0}\right)^{\frac{1}{\eta_o}}\frac{-1}{Q_o^0} \\
&\quad +\exp\left(-\frac{\psi_d}{\eta_d}\right)\left(\frac{1}{\eta_d}+1\right)^{-1}(Q_d^0)^{\frac{1}{\eta_d}+1}\left(\frac{1}{\eta_d}+1\right)\left(\frac{Q_d^f}{Q_d^0}\right)^{\frac{1}{\eta_d}}\frac{1}{Q_d^0}-c_{od} \\
&= -\exp\left(-\frac{\psi_o}{\eta_o}\right)\left(Q_o^f\right)^{\frac{1}{\eta_o}}+\exp\left(-\frac{\psi_d}{\eta_d}\right)\left(Q_d^f\right)^{\frac{1}{\eta_d}}-c_{od}.
\end{aligned}
$$

Next, rearrange the demand model $Q_k^f=\left(V_k^f\right)^{\eta_k}e^{\psi_k}$ and substitute it for the parameters ψ_o and ψ_d:

$$
\begin{aligned}
0 &= -V_o^f\left(Q_o^f\right)^{-\frac{1}{\eta_o}}\left(Q_o^f\right)^{\frac{1}{\eta_o}}+V_d^f\left(Q_d^f\right)^{-\frac{1}{\eta_d}}\left(Q_d^f\right)^{\frac{1}{\eta_d}}-c_{od} \\
&= -V_o^f+V_d^f-c_{od}.
\end{aligned}
$$

Rearranging, and splitting transaction costs into seller and buyer components ($c_{od}=\tau_{od}^s+\tau_{od}^b$), the first-order conditions are:

$$V_d^f-V_o^f=\tau_{od}^s+\tau_{od}^b \quad \forall o,d \text{ s.t. } \quad q_{od}>0.$$

These are identical to the first-order conditions for the market equilibrium in Equation 3.

E.3 Sum of unit-specific gains equals the maximand

I need to prove that the sum of unit-specific gains in Equation 16 equals the maximand in Equation 15. The first term is identical in each expression, so it suffices to prove that the second terms are equal. Beginning with the second term of Equation 16 summed over all units k, I rearrange, switch indices twice, and expand:

$$\begin{aligned}
-\sum_k \int_{Q_k^0}^{Q_k^f} V(Q_k^f) d\varphi &= -\sum_k (Q_k^f - Q_k^0) V(Q_k^f) \\
&= -\sum_k \big(-\sum_{l>k} q_{kl} + \sum_{l<k} q_{lk}\big) V_k^f \\
&= \sum_k \sum_{l>k} q_{kl} V_k^f - \sum_k \sum_{l<k} q_{lk} V_k^f \\
&= \sum_k \sum_{l>k} q_{kl} V_k^f - \sum_l \sum_{k<l} q_{kl} V_l^f \\
&= \sum_k \sum_{l>k} q_{kl} V_k^f - \sum_k \sum_{l>k} q_{kl} V_l^f \\
&= \sum_k \sum_{l>k} q_{kl} (V_k^f - V_l^f) \\
&= \sum_k \sum_{l>k} q_{kl} \big[1(q_{kl} > 0) + 1(q_{kl} < 0)\big] (V_k^f - V_l^f) \\
& \sum_k \sum_{l>k} q_{kl} \big[1(q_{kl} > 0)(V_k^f - V_l^f) + 1(q_{kl} < 0)(V_k^f - V_l^f)\big]
\end{aligned}$$

Inserting the first-order conditions from the previous proof (i.e., $V_d - V_o = c_{od}$ for all o and d such that $q_{od} > 0$):

$$\begin{aligned}
-\sum_k \int_{Q_k^0}^{Q_k^f} V(Q_k^f) d\varphi &= \sum_k \sum_{l>k} q_{kl} \big[1(q_{kl} > 0)(-c_{kl}) + 1(q_{kl} < 0) c_{lk}\big] \\
&= -\sum_k \sum_{l>k} q_{kl} \big[1(q_{kl} > 0) c_{kl} - 1(q_{kl} < 0) c_{lk}\big]
\end{aligned}$$

which is the second term of Equation 15.

F Data Appendix

F.1 Transactions

WestWater Research, LLC provided a dataset containing 6,263 water transactions in California between 1990 and 2015. Variables include transaction date, volume, price, and duration; and name, latitude and longitude, and water use category of both origin and destination parties. I focus on the spot market, which I define as transactions that are initiated, delivered, and completed within one year. I drop multi-year leases and permanent transfers, leaving 4,906 spot market transactions. Of these, prices are available for 4,415.

Cleaning. I calculate price per acre-foot and deflate to 2010 dollars using the CPI. I reshape the data so there is one observation per party per transaction, creating 13,328 observations. For transactions with multiple buyers or multiple sellers, if more specific information is not available, I assume transaction volume is divided equally across parties.

Location. I classify all observations into one of 10 hydrologic regions (defined by California's Department of Water Resources [DWR]). When possible, I also generate latitude and longitude coordinates. I first attempt to use centroids from the user location file (see Section F.5), matching parties to users via the crosswalk file (see Section F.4). This matches 6,221 transaction-by-party observations. Second, I manually geolocate 65 users not appearing in the user location file but which are common in either transactions or entitlements data. For these users, I generate coordinates based on addresses, towns, or maps found via user websites and other publicly available documents; they match 180 additional observations. For remaining unmatched observations, I use location information from the original WestWater dataset. This assigns hydrologic region for all remaining observations and coordinates for 3,721 additional observations. This process leaves 3,206 observations for which location coordinates are unavailable. Finally, I use a user-written Stata command, -geoinpoly-, to locate coordinates within 8-digit watershed (hydrological unit code, as defined by the U.S. Geological Survey), sub-sub-region (detailed analysis unit, as defined by DWR), and county. These shapefiles are available from DWR's California Water Plan: http://www.water.ca.gov/waterplan/gis/index.cfm.

Sector. I classify all parties into one of three sectors: agriculture, urban/municipal, or environmental. I use the first successful method in the following order of priority:

1. Total historical entitlements (21% of observations). If the party appears in the entitlements dataset (see Section F.2), I assign to agriculture or municipal depending on which sector receives a majority of total historical entitlements.

2. Project agencies (3%). I classify the DWR (which runs the State Water Project) and US. Bureau of Reclamation (USBR, which runs the Central Valley Project) as environmental,

because they either devote the water to environmental flows or act as intermediaries. This essentially excludes them from sector-wise analysis.

3. Keywords (41%). I classify users based on the following keywords in their name. Agriculture: almonds, citrus, contractors, dairies, dairy, famers, family, farm, farmers, farming, grower, irrigating, irrigation, nurseries, nursery, orchard, ranch, river interests, trust. Municipal: archbishop, automobile, cement, cemetery, chemical, Chevron, church, city of, cold storage, college, communities, community services, companies, company, container, corporations, country club, developer, development, electric, energy, estate, foods, gardens, golf, gravel, homeowners, homes, housing, inc., Indians, industries, investment, K.O.A., landscaper, leasing, LLC, LP, military, mobile home, monastery, motor, municipal, mutual water, non-ag, oil, owners, park, paving, power authority, properties, property, railway, real estate, realty, recycled, refining, retail, rock, sanitation, school, speedway, Texaco, town of, tribe, university, ventures. Environmental: conservancy, duck hunting, ducks, fish & wildlife, forest service, forestry & fire prevention, water bank.

4. Original use categories (23%). I apply WestWater's original water use categories, based on agriculture, irrigation, and environmental, counting all other categories as urban.

5. Individual names (10%). I assume names of individual people are farmers and therefore agricultural.

6. Remainder (2%). I assume all remaining observations are urban.

F.2 Entitlements and Deliveries

I compile the universe of surface water entitlements (often called allocations) in California by user and year. These entitlements are from a combination of four water sources: Central Valley Project (CVP) allocations, State Water Project (SWP) allocations, and Lower Colorado Project entitlements, and water rights. All portions of entitlements are constructed by multiplying a time-invariant baseline maximum entitlement by a year-varying allocation percentage. This ensures that when using fixed effects, all identifying variation comes from the allocation percentages rather than changing maximum entitlements. For water rights and the Lower Colorado Project, these allocation percentages are simply 1. For the CVP, allocation percentages vary across both years and contract types; for the SWP, allocation percentages vary only across year.

Central Valley Project (CVP)

CVP entitlements are constructed by multiplying present maximum contract volume by yearly percentage allocations for each user, sector, and contract category. Entitlements are summed across contract categories within user and sector.

Maximum contract volumes are downloaded from the U.S. Bureau of Reclamation (USBR) at https://www.usbr.gov/mp/cvp-water/water-contractors.html. I sum the volume of contracts by user, sector (municipal & industrial vs. agricultural), percentage-allocation category (see below), and project vs. base supply. Base supply is contracts for delivery of water based on water rights predating the CVP, while project supply is contracts for delivery of new water made available by the CVP.

Percentage allocations by year are downloaded from the USBR at https://www.usbr.gov/mp/cvo/vungvari/water_allocations_historical.pdf. They are available for each contract year (the 12 months from March of the named year through February of following year) from 1977 to the present. Percentage allocations are determined separately for each of 14 contract categories (North of Delta Agricultural Contractors, North of Delta Urban Contractors (M&I), North of Delta Wildlife Refuges, North of Delta Settlement Contractors/Water Rights, American River M&I Contractors, In Delta - Contra Costa, South of Delta Agricultural Contractors, South of Delta Urban Contractors (M&I), South of Delta Wildlife Refuges, South of Delta Settlement Contractors/Water Rights, Eastside Division Contractors, Friant - Class 1, Friant - Class 2, Friant - Hidden & Buchanan Units). Some categories are combined in earlier years; when "American River M&I Contractors" and "In Delta – Contra Costa" are not specified separately, I impute the value for "North of Delta Urban Contractors". Maximum contract volumes and percentage allocations are merged on contract names via the crosswalk file (see Section F.4).

Relative to other sources of entitlements data, CVP entitlements are offset by two months because they are based on the contract year (Mar-Feb) rather than calendar year. However, only 4.7% of water deliveries occur in January and February, so the difference is small.

State Water Project (SWP)

SWP entitlements are constructed by multiplying maximum contract amount in 1990 by yearly percentage allocations. I choose 1990 because that is the first year all now-existing sections of the SWP were completed, and maximum Table A amounts stabilized. Maximum contract (Table A) amounts are taken from Table B-4 of Bulletin 132-16, downloaded from the DWR at http://www.water.ca.gov/swpao/bulletin_appendix_b.cfm. They are available by user, sector, and year from 1962 through 2016. Percentage allocations by year and sector are available from 1970 through 2017. For 1996-2017 they are taken from published notices to SWP contractors, downloaded from the DWR at http://www.water.ca.gov/swpao/deliveries.cfm. For 1970-1995 they are taken from Table 2-3 of the Monterey Plus Draft Environmental Impact Report, downloaded from http://www.water.ca.gov/environmentalservices/monterey_plus.cfm.

Lower Colorado Project

Lower Colorado Project entitlements are constructed from lists of Lower Colorado River water entitlements in California, downloaded from https://www.usbr.gov/lc/region/g4000/contracts/entitlements.html, and Appendix E of the Final Environmental Impact Statement of 2007 for the Colorado River Interim Guidelines for Lower Basin Shortages and Coordinated Operations for Lakes Powell and Mead, downloaded from https://www.usbr.gov/lc/region/programs/strategies/FEIS/index.html.

Post-2007 entitlements are used as baseline. Percentage allocations do not exist because a shortage has never been declared on the Colorado River.

Surface Water Rights

Surface water rights are not directly measured, so I construct entitlements from Statements of Diversion and Use and annual reports on file with the State Water Resources Control Board (SWRCB). Water rights rarely change nor are curtailed, so I treat them as permanent, fixed entitlements.

Reporting. Users holding post-1914 appropriative rights are required to submit annual reports of use. Riparian and pre-1914 appropriative rights were not systematically tracked by any government agency prior to 2010. However, since 2010 these rights holders must submit Statements of Diversion & Use, with civil penalties for noncompliance. From 2010 to 2016 this reporting requirement was once every three years; since 2016 rights holders must report every year. This means from 2015 onward, the SWRCB had at least one report of quantity diverted of nearly every water right claimed in California.

Although these diversion statements are self-reported, it is reasonable to treat them as the full, legally defensible value of present water rights. This is because appropriative rights are based on documented continuous beneficial use, and these statements are public information, so they could be used in future legal disputes. Therefore, users have incentives to neither report less than they would like to use in the future nor more than other evidence would support.

Data. All of SWRCB's records – water right permits, licenses, and Statements of Diversion & Use – are publicly available in the SWRCB's Electronic Water Rights Information Management System (eWRIMS). The online eWRIMS interface makes it difficult to view or download details for many records at once. Instead, I use a full extraction of the eWRIMS database, as of February 26, 2015, posted online as an exhibit in a 2016 administrative civil liability hearing for the Byron-Bethany Irrigation District. This was downloaded from http://www.waterboards.ca.gov/waterrights/water_issues/programs/hearings/byron_bethany/docs/exhibits/pt/wr70.xlsx. Variables include: amount diverted for select years from 2010 through 2014, face value of rights (for post-1914 rights), types of beneficial uses, status of right, year of first diversion, county, HUC12, and latitude & longitude of POD.

Cleaning. I follow the data cleaning and quality control procedures described by SWRCB in another exhibit ("Exhibit WR-11: Testimony of Jeffrey Yeazell", http://www.waterboards.ca.gov/waterrights/water_issues/programs/hearings/byron_bethany/docs/exhibits/pt/wr11.pdf), adding a number of further checks and corrections. I drop rights that are canceled, inactive, removed, or revoked, and those not yet active, and minor types of water rights (such as stock ponds and livestock), leaving only appropriative rights and statements of diversion and use.

The dataset has 95,535 observations at the level of right by point of diversion (POD) by beneficial use type, with a few duplicates. I drop duplicate observations so that the combination of these three variables form a unique key, then I reshape to the level of right by point of diversion, resulting in a dataset of 56,508 observations. For rights with multiple PODs, I keep only one so that a right is a unique record. SWRCB chooses the POD by alphabetical order on watershed name; I instead choose the POD from the watershed, source within watershed, and 12-digit hydrologic unit within source with the most PODs for that right; if there duplicates within 12-digit hydrologic unit, I keep the POD with the earliest number.

To construct the year a right first began, I use the year of first use when available (nearly all pre-1914 and riparian rights holders, and some post-1914 rights holders), followed by original permit issue date when available, license original issue date when available, and record status year when available. To construct the year a right ended, I take the first year a right was canceled, closed, inactive, rejected, or revoked.

I remove non-consumptive diversions by power-only and aquaculture-only, following SWRCB procedure. For rights that report no diversion to storage, I set diversions to zero. For diversions that do report diversion to storage, I subtract the amount used from the amount diverted, censoring negative values at zero.

I correct for over-reporting, following SWRCB procedure. For post-1914 rights, most observations include the face value of the rights, so if reported diversions in a year exceed the face value, I scale down that year's monthly reports so their total equals the face value. For pre-1914 and riparian rights, face value is not available but some report irrigated acres, so if reported diversions exceed 8 acre-feet per acre, I scale down that year's monthly reports so their total equals this limit. I add one more correction not performed by SWRCB: For post-1914 rights for which the face value is unavailable but irrigated acres is available, I apply the same acres-based correction, but conservatively only for observations whose total diversions exceed 80 acre-feet per acre.

I make further corrections to high outliers in a process not separately conducted by SWRCB. Many of these are likely errors in unit selection; there may also be low outliers, but I cannot detect them effectively. I calculate the standard deviation of the natural log of all monthly diversion values. For observations for which this standard deviation is greater than 2 and the average annual diversion exceeds the face value of the rights by more than 100 acre-feet, for years in which the total diversion exceed the smallest annual total by more than 100 times, I scale down each monthly value proportionally so that that year's total equals the smallest annual total. Although this correction process affects only 82 observations, it changes the total statewide reported diversions by more than 12 orders of magnitude. I also drop one riparian right held by an individual that implausibly reports an annual diversion of more than 100,000 acre-feet.

Further sample restrictions. I drop water held by federal and state projects (which are added to the entitlements dataset separately). I drop non-consumptive rights: those whose beneficial use is aesthetic, aquaculture, fish & wildlife, incidental power, power, recreational, or snow-making;

several known environmental or recreational users (California Department of Fish & Wildlife; California Department of Forestry & Fire Prevention; California Department of Parks & Recreation; Nature Conservancy; Pine Mountain Lake Association; Tuscany Research; U.S. Bureau of Land Management; U.S. National Park Service; U.S. Forest Service; U.S. Department of Fish & Wildlife; White Mallard, Inc.; Woody's on the River, LLC); two known electricity-generating users (Pacific Gas & Electric Co., Southern California Edison Company); and those whose name includes one of several keywords (duck club, gun club, power, preservation, shooting club, waterfowl, wetlands). I drop a small number of rights (157) with no location information available.

Sector. I categorize each right as agricultural or municipal based on whether its record lists irrigation or stockwatering as a beneficial use. I then set to municipal all users whose names include "city of" or "golf", and several known municipal users in Orange County (Irvine Ranch W.D., Orange County W.D., Santa Margarita W.D., Serrano W.D.).

Final variables. For each right, I average across reported annual diversions from 2010 through 2014. I then sum across rights within user and sector, keeping location information for the point of delivery with the largest volume. For each year from 1980 through 2015, I calculate the sum for rights that were held and active in that year. I use this sum as of 1990 for the baseline value, since that is the start of the water transactions dataset. Finally, CVP settlement and exchange contractors may have the same rights counted in both CVP and rights dataset. So as not to double count these, I subtract the maximum contract volume for base supply from their rights volumes.

Entitlements datasets

I combine these four sources to create three datasets of entitlements: user-by-year, market-by-year, and polygon-by-year.

User-by-year. I merge all sources on name and year, matching users via the crosswalk file (see Section F.4), and sum entitlements across sources. I restrict to 1981-2015, the years for which entitlements are available from all four sources. This results in a nearly balanced panel of 7,168 users over 35 years, for a total of 250,129 observations. For locations, I use centroids from the user location file (see Section F.5). For users not available in the user location file, I use the point of diversion listed in the rights dataset, assuming the place of use is in the same region as the point of diversion. For most users this is reasonable, but for some appropriative rights holders with large internal distribution systems, this could introduce substantial error. Unfortunately there is no way to systematically identify and correct these. For users not available in either the user location file or rights dataset, I merge to a dataset of 65 manually geolocated users. For these users, I generate coordinates based on addresses, towns, or maps found via user websites and other publicly available documents.

Market-by-year. I define 14 "markets" by geography and project status: SWP, CVP North, CVP East, CVP South, and non-project users in each of the 10 hydrologic regions. These are supersets of contract types and coincide with several major transaction cost determinants. For the few users that hold both SWP and CVP contracts, I choose the project from which the user draws the larger maximum entitlement. CVP North is defined by users north of the Sacramento–San Joaquin Delta, CVP East is users south of the Delta and east of the San Joaquin River, and CVP South is users south of the Delta and west of the San Joaquin River. Starting with the user-by-year dataset, I drop 408 observations without location information, which represent less than two-thousandths of a percent of total entitlements. I then sum across users within market and year.

Polygon-by-year. For agricultural analysis, I calculate per-acre entitlements for each of many overlapping polygons, comprised of the user locations file (see Section F.5) and 8-digit watersheds ("HUC8", hydrological unit code as defined by the U.S. Geological Survey). Starting with the user-by-year dataset, I match user entitlements to their own polygons when possible, and calculate entitlements and deliveries per acre of cropland within shape. When user polygons are not available, I aggregate users within 8-digit watersheds. (The users are mostly rights-holders only.) This assumes the place of use is in the same region as the point of diversion – which is reasonable for most users but could introduce substantial error for some appropriative rights holders with large private distribution systems; unfortunately there is no way to systematically identify and correct these. It also introduces measurement error via averaging; while most irrigation districts do typically allocate deliveries across retail customers on a per-acre basis, individual rights-holders may not be spread out so evenly. However, this problem is minor for my empirical analyses because I use water rights only for the overall level of water availability, not for identifying variation.

F.3 Quantity consumed

In some analyses I use market-level quantity consumed as an endogenous variable. (For definition of market, see "Entitlements datasets" above.) I do not measure water consumption directly but rather construct it from other datasets. I sum, by market and sector: (1) entitlements, (2) quantity purchased minus quantity sold, from transactions data, and (3) mean groundwater supply.

Groundwater consumption is not directly measured or monitored at all in California except in certain local areas. However, the DWR estimates groundwater supplies in several publications. I use average annual groundwater supply for 2005-2010 by hydrologic region, subregion (planning area), and sector, from the Volume 2 Regional Reports of DWR's California Water Plan Update 2013. I sum across planning areas within hydrologic region and sector. This measure of groundwater consumption does not capture changes in consumption from year to year, but my primary purpose is to accurately set the overall log-level of water consumption.

F.4 Crosswalk file

I create a crosswalk dataset that links water users by name across all other datasets used in this paper. To create it, I export raw names from each dataset and append them together. I strip punctuation and correct misspellings and other typos. I standardize common terms into acronyms (e.g., I.D. for irrigation district; M.W.C. for mutual water company; F.C.W.C.D. for flood control and water conservation district). For names of individual people, I match full names to entries with the same last name but only first initial(s) available. For agencies, when names are closely but not precisely similar I use agency websites and other publicly available documents to determine whether (a) one agency has changed its name, (b) one name is erroneous, or (c) they are indeed distinct agencies. I use footnotes and notes in original data sources to link users with name changes over time, keeping the most recent name. When a merger has occurred, I roll users up into the most aggregate version to maintain consistent definitions. The exception is companies with service in noncontiguous locations, for which I treat each location as a separate user. The final crosswalk file has 28,765 entries (input names) pointing to 14,830 targets (output names).

F.5 User location file

By combining all the relevant and publicly available georeferenced digital maps I can find, I create a dataset of the most accurate locations, areas, and boundaries for as many water users as possible. I combine the following datasets and link them via the crosswalk file. For each user, I keep one shape (feature) according to the following priority order:

1. DWR's Water Districts Boundaries, downloaded via the Query link found at https://gis.water.ca.gov/arcgis/rest/services/Boundaries/WaterDistricts/FeatureServer

2. Federal, State, and Private Water Districts shapefiles maintained by USBR and DWR, downloaded from the California Atlas at http://www.atlas.ca.gov/download.html.

3. Mojave Water Agency Water Companies, downloaded at https://www.mojavewater.org/geospatial-library.html.

4. California Environmental Health Tracking Program's Water Boundary Tool, downloaded at http://www.cehtp.org/page/water/download.

Before I append (merge) sources, I combine noncontiguous shapes for the same user (dissolve to create multipart features). After selecting one shape per user, I calculate the user's centroid (restricted to within shape), area, and cropland area (via zonal statistics). The latter uses the 2015 cropland mask from the USDA's Cropland Data Layer. I also construct versions of the user location file in which user shapes are interacted with either 8-digit watershed (hydrological unit code, as defined by the U.S. Geological Survey) or the intersection of sub-sub-region (detailed analysis unit, as defined by DWR) and county. These shapefiles are available from DWR's California Water Plan: http://www.water.ca.gov/waterplan/gis/index.cfm.