Liquid Markets:
An Empirical Analysis of a Water Exchange

Yunmi Kong

Rice University

September 5, 2023
Click here for the latest version.

Abstract

This paper empirically analyzes the performance of one of the world's most developed water exchanges, which operates as a primitive limit order market. Upon modeling participants' choice of order price and order type, I identify their latent value distributions from observed orders and trades. The model flexibly allows for dynamics, risk aversion, and default behavior. Counterfactual simulations suggest the observed exchange attains substantially lower trade surplus than the benchmark of periodic uniform-price market clearing. Droughts exacerbate the gap in surplus per unit traded between the observed exchange and the benchmark. I assess the role of volume frictions, price shading, and temporal dispersion in explaining the gap.

I would like to thank Javier Donna, Laurent Lamy, Mar Reguant, and conference/seminar participants at the UBC Summer IO Conference, Queen's University, MIT, KAEA mentoring workshop, UC Berkeley, Harvard, Caltech, Northwestern, Seoul National University, Rice-Sciences Po IO conference, Cowles Conference on Models & Measurement, and NBER Summer Institute IO for their comments. All remaining errors are mine.

E-mail: yunmi.kong.01@gmail.com

1 Introduction

According to UN-Water, which coordinates the United Nations' work on water and sanitation, water is the primary medium through which we will feel the effects of climate change.[1] Many countries "face water scarcity as a fundamental challenge to their economic and social development; by 2030 over a third of the world population will be living in river basins that will have to cope with significant water stress."[2] As a result, there is growing interest in water reform, particularly in the potential of water markets to help overcome these problems.[3] Drawing on water market data from Australia, the driest inhabited continent and world leader in market-based water management,[4] this paper investigates one piece of the broader puzzle to be solved: how the design of the trading mechanism, or market microstructure, impacts water market outcomes such as the gains from trade and how these impacts interact with drought. My approach to this question involves a structural analysis of a limit order market for water which models traders' pricing incentives while flexibly allowing for dynamics, risk aversion, and default behavior.

The spot market for water in Australia, particularly in the southern part of its Murray-Darling river basin (southern MDB), is large in terms of trade volume, monetary value, and participation. In the southern MDB, "allocation trades," which refer to spot trades of water, totaled 5527 gigaliters in the 2019-2020 financial year.[5] This amounted to over 40% of the total nominal volume of water allocatable to all water license holders, known as entitlements on issue, which was about 13000 GL in the southern MDB.[6] The monetary value of allocation trades across Australia was 708 million AUD in 2019-2020, when prices were particularly high due to one of the worst droughts on record, with the southern MDB comprising 93% of all volume.[7] In terms of participation, around 78% of irrigation farmers, responsible for the majority of water consumption, had conducted at least one water allocation trade by 2015.[8] Indeed, allocation trades are an important source of water for irrigators; only a small

[1] https://www.unwater.org/water-facts/water-and-climate-change.
[2] 2030 Water Resources Group (2009)
[3] National Water Commission (2011), Australia.
[4] Ibid.
[5] Bureau of Meteorology (2021), Australia.
[6] Ibid. Depending on water availability, the actual volume allocated may be lower than the nominal entitlements on issue.
[7] Australian Competition and Consumer Commission (2021) and Bureau of Meteorology (2021).
[8] Australian Competition and Consumer Commission (2021).

proportion of irrigators across the Basin use more sophisticated derivative products such as leases or forward contracts.[9] In 2018-2019, the average dairy, horticulture, and rice farm relied on allocation trades for 41%, 39%, and 23% of their water use, respectively.[10] These statistics illustrate that Australian water markets are well developed compared to that of other countries. At the same time, they are still at a relatively formative stage compared to markets for other goods and services,[11] making the study of these markets particularly interesting.

One of the main media for allocation trades in Australia are online exchanges, which operate as simple versions of a limit order market. To trade, a buyer or seller (trader) places either a limit order or a market order. Taking the buyer's perspective for expositional clarity, placing a limit order means the buyer lists the price and volume at which they want to buy, in the uncertain hope that a willing seller arrives to fulfill those terms. Placing a market order means the buyer selects an existing seller's limit order to fulfill, for guaranteed trade. There is no market maker; each distinct trade involves one buyer and one seller and can happen at a different price from other trades occuring at the same time. Thus, the water exchange features continuous, bilateral exchange at discriminatory prices. State approval authorities collect a flat fee per trade, where sourcing a purchase from multiple sellers would constitute multiple trades. Moreover, it is strictly optional for traders to allow their order volume to be split into smaller trades. As a result, I find evidence of trade frictions related to bilateral volume mismatch, which I refer to as "bilateral frictions." I also observe that limit orders with not very competitive prices still have a decent chance of becoming a trade. This provides stronger incentives for traders to list prices that are distant from their true valuations, which I refer to as "price shading."

The goal of assessing counterfactual exchange design requires recovery of the latent distribution of buyers' willingness to pay ("values") and sellers' willingness to accept ("costs") from observed data. Meanwhile, general limit order markets are complex dynamic games of which the solution is analytically intractable, and much is still unknown despite a long history of research in the finance literature.[12] This paper

[9]Ibid.

[10]Ibid. The remainder of farms' water use would be supplied by the water licenses ("entitlements") they own. See Section 2.1 about entitlements.

[11]National Water Commission (2011), Australia.

[12]Per Parlour and Seppi (2008), "there is still much we do not know about limit order markets. [...] Only very stylized environments have been studied thus far. Joint decisions about order aggressiveness and quantity have not been fully modelled [...]." Roşu (2009) explains that "one of the

does not provide a full equilibrium theory of limit order markets but focuses on deriving key identifying equations that exploit the data. In this vein, I model a trader's *best-response* action on the exchange given her value for water. Specifically, I model the trader's choice between placing a limit order and a market order and, conditional on placing a limit order, the choice of a limit order price. Again taking the buyer's perspective, buy limit orders with higher prices are more likely to be executed because sellers prefer higher prices. Thus, a buyer's optimal limit order price balances the marginal benefit of this increased execution probability against the marginal cost of paying a higher price conditional on execution. A key input to this tradeoff is the probability of limit order execution as a function of price, which is observed empirically. The buyers' optimization in light of this tradeoff generally leads them to shade their prices, meaning their limit order prices are lower than their latent values. Also, buyers with higher values list higher limit order prices. When it comes to the choice between market and limit orders, traders follow a threshold strategy: buyers place a market order if their value exceeds an indifference threshold and place a limit order otherwise. This is because the guarantee of trade offered by market orders is more valuable for higher-value buyers.

I extend this baseline model to allow dynamics, risk aversion, taxes and fees, and default behavior to affect traders' actions. First, dynamics arise with the possibility that traders may revise their unexecuted limit orders in a repeat attempt to trade. This introduces a positive continuation value of failing to trade, causing the buyer to behave as if her value is lower. Second, more risk averse traders engage in less price shading because they have a lower tolerance for the risk of failing to trade. Third, taxes and fees cause traders to behave as if their costs are higher (or values lower) than in the absence of taxes and fees. Finally, I allow for the possibility that a portion of traders may skip the optimal limit versus market order decision and place a limit order by default. This is empirically motivated by a market norm in which sellers tend to place limit orders.

The key to identifying the buyers' latent value distribution and sellers' latent cost distribution is the traders' first-order condition for limit order pricing per the model discussed above. This first-order condition takes as an input the empirical probability of limit order execution as a function of limit order price, in the spirit of Guerre, Perrigne, and Vuong (2000). The additional complexities of identifying the full model,

reasons for the scarcity of models of order-driven markets is the sheer complexity of the problem."

such as those arising from traders' selection into market versus limit orders and the dynamics of order revision, are formally addressed to establish nonparametric identification of the value and cost distributions. When it comes to risk aversion, I delineate the range of risk aversion levels that cannot rationalize the data if the supports of the value and cost distributions are non-negative. I then conduct the remainder of my analysis over a range of risk aversion levels that can rationalize the data. Since risk aversion affects the amount of traders' price shading, this flexible approach has the advantage of yielding a robust analysis across varying levels of price shading. My estimation procedure closely aligns with the identication argument above but also accounts for heterogeneity in the market environment by conditioning on a vector of covariates parametrically. These covariates include volume, water fundamentals such as the amount of water stored in the major lakes as well as descriptors of exchange conditions at the time the trader's action took place.

Upon estimating the latent value and cost distributions, I simulate the counterfactual benchmark of periodically crossing latent supply and demand. This benchmark differs from the observed exchange in that it temporally aggregates traders, it clears the market at a single market-clearing price, volumes are pooled so that trade is multilateral, and there is no price shading by construction. Comparing the observed exchange to the benchmark at the trade level, I find that the sorting of which buyers and sellers get to trade is different. Relatively low-value buyers and high-cost sellers are able to trade on the observed exchange by exploiting the time and price dispersion available there, and these can block more efficient trades involving higher-value buyers or lower-cost sellers.

Next, the comparison of total trade surplus depends on traders' risk aversion level. If traders' risk aversion were sufficiently high such that there was almost no price shading in the data, then the benchmark would yield about 18% higher surplus over the sample period. If traders' risk aversion were lower such that they engaged in meaningful price shading, then the surplus gap would be larger, exceeding 50% at the lowest risk aversion level that can rationalize the data. I find that periods of drought are when the observed exchange falls especially short in terms of trade surplus per unit (megaliter). This is because value heterogeneity is especially high during droughts, so there are higher per-unit gains to be had from sorting the highest value buyers and lowest cost sellers into trade. To supplement findings from the benchmark, I also simulate a closely-related dominant strategy double auction adapted from McAfee

(1992), which yields qualitatively similar results.

To assess the role of bilateral frictions, price shading, and temporal dispersion in explaining the surplus gap between the observed exchange and the benchmark, I additionally simulate counterfactual exchanges that implement one-megaliter splits of the sellers' volume. This allows buyers to fill their order volume with the cheapest units available on the exchange, regardless of how many different sellers these units come from and without having to buy the entirety of any seller's volume. Even holding limit order prices fixed as observed, I find that eliminating bilateral frictions in this way closes about half of the gap between the observed exchange and the benchmark while preserving a continuous-time market. This offers a practical, incremental policy option that could meaningfully improve market efficiency.

Related literature

This paper relates to the economics literature studying formal water markets and especially its intersection with the empirical tools of industrial organization. In the southern MDB, Rafey (2023) uses a production function approach to value the water market, quantifying the increase in agricultural output caused by the existence of water trading. In mid-twentieth century Spain, Donna and Espín-Sánchez (2018, 2023) study the frictions and performance of English auctions for water among farmers. Meanwhile, where formal water markets are relatively weak or lacking, such as in California, researchers have estimated the prospective gains that could be achieved if trade barriers were eliminated (e.g., Bruno and Jessoe (2021); Hagerty (2022)). The large literature on the broader economic, legal, policy, and scientific dimensions of water markets around the globe are surveyed by Kaiser and McFarland (1997), Chong and Sunding (2006), and Debaere et al. (2014) among others. I contribute to this literature by empirically analyzing the impact of exchange design, or market microstructure, on market outcomes. I do this by exploiting not only data on realized water trades but also data on traders' market *actions*—their buy and sell orders and order prices—some of which result in trade and some of which do not.

In structurally analyzing a limit order market, this paper relates to the pioneering work of Hollifield, Miller, and Sandås (2004) and Hollifield, Miller, Sandås, and Slive (2006) which study the Stockholm Stock Exchange and Vancouver Stock Exchange, respectively. The latter evaluates that the Vancouver Stock Exchange achieves about 90% of the maximum possible gains from trade. My approach is also related to work on structural estimation of auction models. Identifying and estimating players' latent values or costs through empirical expressions of their first-order condition, rather than through numerical simulation of potentially intractable equilibria, has become a workhorse of the auction literature since Guerre, Perrigne, and Vuong (2000). A related approach has been fruitfully employed to study electricity markets (e.g., Wolak (2000); Hortaçsu and Puller (2008); Reguant (2014)).

When it comes to exchange design, the finance literature has debated the merits of continuous markets versus periodic (batch) uniform-price auctions in the context of securities exchange (e.g., Madhavan (1992)).[13] Some high level insights from this literature include that temporal consolidation may improve efficiency at the cost of market inaccessibility between batches and that there is no one-size-fits-all solution; e.g., temporal consolidation may be more helpful for thinly traded securities.

The paper is organized as follows. Section 2 describes the institutional background of water exchange in Australia, the data, and suggestive evidence of market inefficiencies. Section 3 presents the baseline model and establishes identification of the model primitives. Section 4 discusses model extensions and describes the estimation procedure. Section 5 discusses empirical results. Section 6 assesses the performance of the observed exchange relative to the counterfactual benchmark of periodic uniform-price market clearing and studies the factors contributing to the gap. Section 7 concludes. The appendix collects all proofs.

2 Water exchange

2.1 Institutional background

This section provides some institutional background for water exchange in Australia, summarizing information gathered from the Australian Productivity Commission (2003), National Water Commission (2011), Goesch, Donoghoe, and Hughes (2019), the Australian Competition and Consumer Commission (2021), and the Murray-Darling Basin Authority website. The Murray-Darling Basin is a large area of south-eastern Australia where water flows through a system of interconnected rivers and lakes. The southern region of this basin or southern MDB, which incorporates the River Murray and its various tributaries, is considered one of the most sophisticated

[13]Batch auctions have also been proposed more recently as a policy response to high-frequency trading. See Budish, Cramton, and Shim (2015), Wah and Wellman (2013), Farmer and Skouras (2012).

water markets in the world. Water is traded under a cap-and-trade system as follows. The Murray-Darling Basin Agreement limits total water extraction, setting aside water for conveyance—water needed to "run the river" or deliver water without it evaporating or seeping into the riverbed—critical human water needs, and a conveyance reserve set aside for the next year.[14] Only after setting aside these volumes is remaining water allocated to water license holders, who may trade the water.

A water license or "entitlement" is a perpetual right to access an annual volume of water, subject to water availability. These entitlements were initially allocated by previous decisions of government, and new entitlements are generally not available for issue.[15] Historically tied to land, water access rights in the Basin are now generally separated from land ownership, though there are some exceptions. "Entitlement trade" refers to the trade of these water licenses, which involves transferring a perpetual right. Meanwhile, "allocation trade" refers to the one-time trade of a specific volume of water, as opposed to trading a license. This paper focuses on *allocation trades*, the trade of water.

As reported by Goesch et al. (2019), agriculture comprises around 70% of extractions for consumptive water use in Australia. This is mainly used for irrigation. As such, irrigation farmers known as "irrigators" are the main participants in Australian water markets, constituting the majority of commercial allocation trades. Irrigators' values for water exhibit substantial heterogeneity, both across and within crop types (Rafey (2023)). Other market participants include environmental water managers, irrigation infrastructure operators, and investors among others. The southern MDB is divided into 15 trading zones, often defined as areas within which trade can freely occur because users in a zone are subject to the same jurisdiction and draw water from the same source point, such as a particular storage or water course. Trades within a zone, known as intra-zone trades, do not usually cause physical movement of water. Rather, the traded water exits the seller's "allocation bank account" and enters the buyer's "allocation bank account". River operators do not generally release water to meet individual orders but manage bulk releases to meet forecasted aggregate demand based on historical usage data, weather, and other factors. Trade between zones is possible but subject to inter-zone trading rules, such as trade limits based on hydrological constraints. An irrigator is generally free to use the water in

[14] www.mdba.gov.au/water-management/allocations-states-mdba/water-critical-human-needs.
[15] Productivity Commission (2003).

their account at any time, though there are state rules governing the carryover of water past the end of the water year, such as a volume limit and a 5% deduction to account for evaporation losses in the state of Victoria. The water year begins on July 1st and ends on June 30th. Negative account balances are illegal and met with enforcement action,[16] supported by a comprehensive network of water meters to monitor irrigators' water use.[17]

Buyers and sellers are free to find a trade partner and set their trade price and volume in any manner they choose. With increased internet access, online exchanges have become one of the main avenues for trade. The provider of this paper's data sample estimates that over 60% of trades are now formed online, though no official statistics are available. Participants also trade by calling brokers or may trade with a friend without any intermediary involvement. Once a trade is formed, however, details of the transaction must be lodged on trade forms with state-owned approval or registration authorities. Nonetheless, information on trading activity is highly fragmented and incomplete; there is "currently no single entity responsible for, or capable of, gathering the necessary data."[18] Moreover, my research question requires information on the *pre-trade* actions of buyers and sellers—such as information on buy and sell orders, not all of which become trades—which is not collected by state authorities. Given the absence of data necessary to cover the entire market, I take a focused approach, studying the largest water exchange in one of the most actively traded geographic zones.

The exchange operates as follows. Both buyers and sellers, which I refer to collectively as traders, may post listings or "orders" on the exchange which list the price (per megaliter) and volume of water (megaliters) they wish to trade. These listings are visible to all traders. The listings do not reveal identities; the exchange operates as a blind market. If a buyer comes along and chooses one of the existing sell listings to fulfill, or a seller comes along and chooses one of the existing buy listings to fulfill, this results in a trade at the price and of the volume listed in the listing. So traders face a choice between fulfilling an existing listing posted by another trader, for guaranteed trade, or posting their own new listing in the uncertain hope that a counterparty will come along that wants to fulfill it. I use the term "market order"

[16] www.g-mwater.com.au/customer-services-resources/water-use-compliance/unauthorised-water-take-exceeding-your-water-entitlement.

[17] National Water Commission (2011).

[18] Australian Competition and Consumer Commission (2021).

to refer to the former action and the term "limit order" to refer to the latter action. I will refer to this market design as a "limit order market." Conditional on trade, sellers pay income taxes[19] on their sales revenue as well as a 2.5% fee to the exchange. Buyers pay a flat trade approval fee per trade to state authorities which was 47.50 AUD during my sample period.

The most familiar example of a limit order market is a stock exchange, but the reader should not infer that this water exchange shares any of the features of familiar stock exchanges other than what is described in the preceding paragraph. The water exchange is a pure limit order market with no market maker. It features an open limit order book, meaning each trader can observe the set of all existing listings, and they can choose which listing to fulfill if any. As a result of this market structure, each trade can happen at a different price even at a given moment in time. Volumes are not pooled across traders; all the water being transferred in a single trade is sourced from one seller and goes to one buyer, so that each trade is truly bilateral. It is strictly optional for traders to allow their limit order volume to be split into smaller trades, and such trades do not appear prevalent in my data. This is partially a consequence of the flat trade approval fee mentioned above which may in turn reflect path dependence in market development given the historically bilateral nature of traditional water brokerage. To summarize, the water exchange features continuous, bilateral exchange at discriminatory prices.

2.2 Data

The water exchange I study describes itself as Australia's largest independent water exchange. I observe the price, volume, start time, and end time of all buy and sell limit orders on the exchange, as well as the price and volume of all trades and the time the trade was approved by the state. The dataset covers the time period of April 2020 through the end of September 2021. This timeframe covers periods of both drought and relative plenty; to elaborate, water storage levels in Lake Hume, a reservoir impounded by a major dam across the River Murray, was at 13% of total capacity at the beginning of April 2020 and was at 97% of total capacity at the end

[19] According to OECD.Stat, the average personal income tax rate in Australia in 2020 was 24.1% while the corporate income tax is 25% for small and medium businesses and 30% for other companies. I use 24% for the income tax rate in my empirical analysis.

Table 1: Summary statistics of buy and sell orders in the sample

	count	mean	sd	median	min	max
sell orders						
price (AUD/ML)	2,182	228	140	200	55	950
volume (ML)	2,182	61	101	30	1	1,000
duration (days)	2,182	8	12	3	0	130
buy orders						
price (AUD/ML)	631	143	65	120	40	418
volume (ML)	631	131	165	87	1	1,000
duration (days)	631	6	7	3	0	46

of September 2021.[20] This variation will allow me to study the interaction of drought and exchange design in Section 6. The trading zone I study is zone 7, located on the Victorian side of the River Murray and one of the most actively traded zones. Due to the frictions associated with inter-zone trade, 90% of all allocation trades of which the source is zone 7 are intra-zone. I limit my analysis to intra-zone trades only. Within a zone, water is a commodity, i.e. one seller's water is no different from another seller's, because all water is drawn from the same source point and trade does not cause physical movement of water.

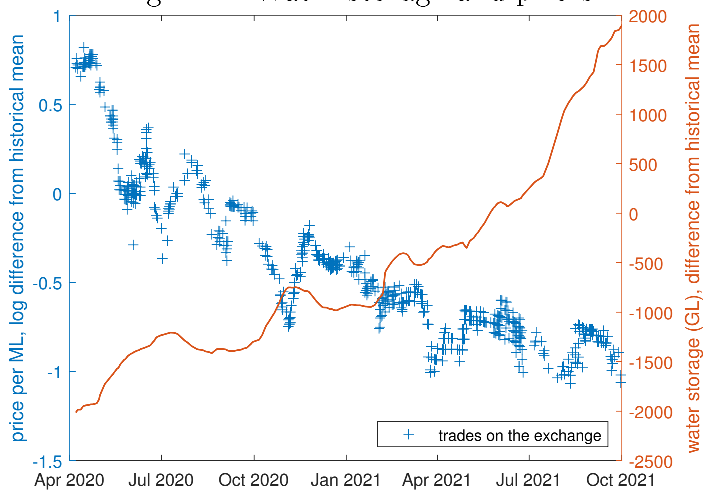
Though the water exchange is very developed by water market standards, it is thinner and slower than more familiar limit order markets for other goods. On the water exchange and for the trading zone I study, there are on average 6 limit orders per day, including both buy and sell limit orders. The mean duration between the start date and end date of a limit order is 8 days, though some limit orders have a duration as long as a few months. A large majority of limit orders, 78%, are sell limit orders, so the most representative path to trade on this exchange is for a seller to place a limit order and for a buyer to fulfill it through a market order. Table 1 provides additional descriptive statistics of the buy and sell orders.

When it comes to explaining water prices, the amount of water stored in major dams and lakes has a lot of explanatory power, as illustrated in Figure 1. In this figure, each cross marks a trade on the exchange. The x-axis shows the date of the trade and the left y-axis shows the log deviation of the trade's price from the historical calendar-day mean.[21] The solid line, read by the right y-axis, displays the

[20] Australia Bureau of Meteorology, https://mdbwip.bom.gov.au/175/#4.6/-31.5/147

[21] I compute the historical calendar-day mean price using trades during 2008–2019 from the relevant zone. Specifically, a kernel regression of price on the day number of the water year—where July 1st

Figure 1: Water storage and prices



deviation of total relevant water storage from its historical calendar-day mean.[22] The respective historical calendar-day means capture expected seasonal fluctuations, while deviations from these means capture drier-than-usual or wetter-than-usual conditions for the time of year. Figure 1 shows a strong negative correlation between deviations in water storage volumes and water prices during the sample period. Section 5.1 performs regression analyses of prices with a longer list of explanatory variables.

The data has the following limitations. First, trader identities are not provided, not even in anonymized form. Second, the data does not provide official linkages between buy and sell orders and final trades. The "lack of linkage between exchange [...] data and Basin State register data" on trades is not specific to my data sample but a general limitation that has been noted by the Australian government.[23] I will infer the linkages by matching the price, volume, and end dates of the buy and sell orders with the price, volume, and approval dates of the trades, allowing for a reasonable delay between the two dates as state approval of lodged trades is not instantaneous.

Naturally, not all limit orders are executed, so my analysis considers "revisions" of limit orders, in which a trader replaces an unexecuted limit order with a new limit order for that volume. Empirically, I search for instances in which a limit order ends unexecuted and subsequently a new limit order appears on the same day, for the same volume but with a different price. As I do not observe trader identities, there is no

is day 1—is used to predict the expected price for each month-day combination.

[22] I defer a detailed description of the storage variable to Section 5.1.

[23] Australian Competition and Consumer Commission (2021).

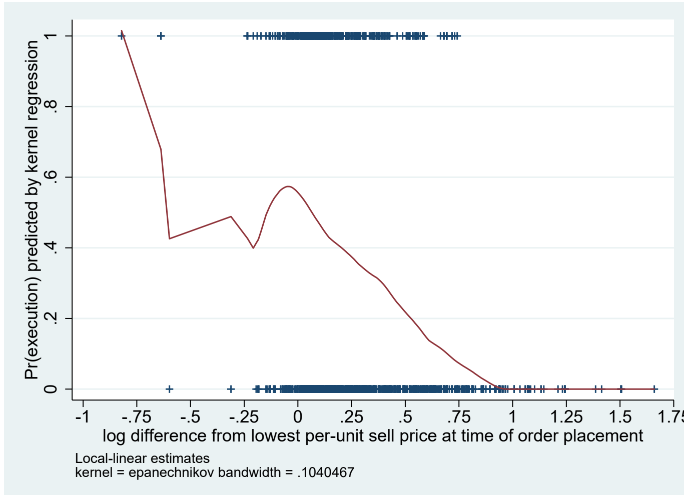
guarantee that all found instances are indeed revisions. Nonetheless, I will treat them as such in my analysis to account for the positive expected value of order revision in a manner that is as data-driven as possible.

2.3 Suggestive evidence of market inefficiencies

Bilateral frictions I examine every instance in which a buyer placed a market order, resulting in execution of a sell limit order. I find that in 73% of these instances, the buyer did not choose the sell listing with the lowest per-unit price available (the "low ask") but traded with another listing offering a different volume than the low-ask listing. This demonstrates that there are trade frictions associated with the bilateral, as opposed to multilateral, nature of water exchange. In a multilateral trade mechanism that freely splits and pools volumes, the buyer would fill her desired volume with the cheapest units available on the exchange, regardless of how many different sellers these units come from and without having to buy the entirety of any seller's volume. Then the low-ask units would always be first to be included in a market order purchase. On the observed exchange, however, the low-ask listing is often skipped. Frictions of this sort reduce effective market thickness. I refer to these as "bilateral frictions."

Incentives to shade limit order prices An important statistic that bears directly on traders' pricing incentives is the probability of limit order execution as a function of limit order price. In Figure 2, the x-axis displays the log difference between the price per-ML of each sell limit order and the lowest available sell limit order price at the time that limit order was placed. The latter price is meaningful as the lowest competing price the seller would have seen before choosing a price for her own limit order. I refer to this x-axis variable as relative price. On the y-axis, all the observed limit orders, marked by crosses, are at either 0 or 1, with 0 indicating that the limit order was not executed and 1 indicating that it was. Finally, the curve in the figure displays a kernel regression of limit order execution on relative price and can be interpreted as the predicted probability of order execution given the relative price of the order. Focusing on where data points are dense, roughly between the 1st and 99th percentiles of relative price or $x \in [-0.13, 1.07]$, the predicted Pr(execution) is decreasing in the seller's relative limit order price as expected; buyers prefer lower prices. What is surprising is that the slope of this decrease is rather flat; a sell limit

Figure 2: Sell limit order price and execution probability



order with a per-unit price 20% higher than the lowest available price still has a roughly 40% probability of eventual execution. The flatter is this slope, the greater is the incentive for sellers to choose a higher limit order price because the associated loss in Pr(execution) is smaller. I refer to traders listing prices different from their true valuation as "price shading." Even when there is room for a Pareto-improving trade between a buyer and seller, price shading can prevent it from happening. The observed exchange exhibits strong incentives for traders to shade their limit order prices.

3 Model and identification

In order to ultimately assess counterfactual exchange design, I set an intermediate goal of recovering the latent distribution of buyers' willingness to pay and sellers' willingness to accept from observed data. In service of this intermediate goal, I model a trader's pricing decision conditional on order volume as a best response to the market she observes. Let x be a finite-dimensional vector representing current market conditions, such as water storage levels, the season, the limit order book, and recent trade activity, as well as order volume. I refer to these collectively as "covariates" and describe them in detail in Section 5.1. Conditioning on relevant elements of x is implicit in this section unless explicit conditioning is needed for clarity. Section 4 discusses extensions of the baseline model to incorporate features such as risk aversion and taxes and fees.

3.1 Baseline model of limit order pricing

Buyers and sellers (traders) arrive exogenously to a market for water. I initially focus on buyers, returning to sellers in Section 3.4. Conditional on order volume m_i, buyer i obtains value v_i per unit upon receiving the water, with v_i distributed according to $F_b(\cdot|x_i)$ with a b subscript for the buyer.

Upon arrival, buyer i either places a buy listing at price p_i or fulfills an existing sell listing. Fulfilling a sell listing involves buying from the seller who placed the listing at the terms stated in the listing. In this paper, I refer to placing one's own listing as a "limit order" and fulfilling an existing listing as a "market order". While a market order always results in a trade, a limit order becomes a trade only if a counterparty decides to fulfill it. Thus, limit order execution is not guaranteed. I refer to the set of all live buy and sell limit orders not yet executed as the "limit order book".

Let $P_b(p_i|x_i)$ denote the probability that a buy limit order with price p_i is executed conditional on covariates x_i. This probability, an empirically observed object, is an important input to the buyer's best-response problem. I make the following assumptions regarding $P_b(p_i|x_i)$.

Assumption 1. *The execution probability $P_b(\cdot|x)$ for a buy limit order satisfies*

(i) $P_b(p|x)$ is strictly increasing in p.

(ii) $P_b(p|x)/P_b'(p|x)$ is strictly increasing in p.

(iii) $P_b(p|x) < 1$.

Assumption 1-(i) states that a buy limit order with a higher price is more likely to be executed. This is because the counterparties—sellers—prefer to receive higher prices. Assumption 1-(ii) is a technical condition involving the derivative $P_b'(p|x) \equiv \frac{\partial P_b(p|x)}{\partial p}$. For example, Assumption 1-(ii) would be satisfied if $P_b(\cdot|x)$ had a probit shape as $P_b(p|x) = \Phi(\beta_0 + \beta_1 p + \beta_2 x)$, where $\Phi(\cdot)$ is the standard normal cdf and $\beta_1 > 0$ per Assumption 1-(i). In that scenario, $\frac{P_b(p|x)}{P_b'(p|x)} = \frac{\Phi(\beta_0+\beta_1 p+\beta_2 x)}{\beta_1 \phi(\beta_0+\beta_1 p+\beta_2 x)}$ is increasing in p because $\frac{\Phi(\cdot)}{\phi(\cdot)}$ is an increasing function. The probability $P_b(p|x)$ is likely to have such an S-shape in p because the execution probability is insensitive to local price changes when p is extremely low or high (such extreme orders will be executed with probability near zero or one, respectively), but there is a range of p in between where the limit order book is denser and the execution probability responds more steeply to p. Assumption 1-(iii) states that execution of a buy limit order is not guaranteed.[24]

Choosing a limit-order price with no revision in the future

In Section 3.2, I allow for the possibility that a trader will revise her limit order after placing it. Here, consider first a baseline scenario: the buyer's choice of limit order price when she believes there will be no revision in the future. I adopt a best-response approach in the spirit of Guerre, Perrigne, and Vuong (2000).

For reasons that become clear shortly, I use an r subscript in notation for this scenario: Given her value $v_{r,i}$ and covariates $x_{r,i}$, buyer i chooses price $p_{r,i}$ to maximize her expected utility from the order in light of execution probability function $P_b(\cdot|x_r)$. Her maximization problem is

$$\max_{p_r} P_b(p_r|x_{r,i})(v_{r,i} - p_r), \tag{1}$$

where her objective function is the execution probability times her profit conditional on execution. Differentiating the objective function with respect to p_r, the first-order condition gives

$$p_r + \frac{P_b(p_r|x_{r,i})}{P_b'(p_r|x_{r,i})} = v_{r,i}. \tag{2}$$

Thus, rather than submitting a price equal to her value, she will "shade" her order price so that $p_r < v_{r,i}$. Let $p_r^*(v_{r,i}|x_{r,i})$ denote the buyer's optimal price that satisfies equation (2). By Assumption 1, the left-hand side of equation (2) is strictly increasing in p_r. It follows that $p_r^*(v_{r,i}|x_{r,i})$ is strictly increasing in $v_{r,i}$.

3.2 Limit order pricing with revision

I now allow for the possibility that a trader will revise her limit order after placing it. The purpose is to identify traders' latent values more accurately by adjusting for the (possible) continuation value of revision. After the passage of an exogenous, stochastic interval of time t_r since initial order placement, the buyer will revisit her limit order and, if it has not yet executed, exit the market with exogenous probability $1 - r$ or else revise the price with probability r. Motivated by the data, in which 1.5 percent of unexecuted limit orders seem to be revised more than once, I model traders

[24] Cohen et al. (1981) show that unless traders monitor a market continuously, the execution probability of a limit order is less than one no matter how closely it approaches the asking price.

revising their limit order up to one time. In this case, the buyer's problem of choosing a price in the revision stage is represented by (1). The model is generalizable to any fixed, finite maximum number of revisions by iterating the revision model up to that number. The buyer's value for one unit of water at the time of revision is denoted $v_{r,i}$. This $v_{r,i}$ has conditional distribution $F_b^r(\cdot|v_i)$ which is weakly stochastically ordered in v_i so that a buyer with higher v_i is weakly more likely to have a higher $v_{r,i}$. Persistent values, $v_{r,i} = v_i$, would be a special case of this more flexible setup.

Assumption 2. *Distribution $F_b^r(\cdot|v_i)$ of $v_{r,i}$ is weakly stochastically ordered in v_i.*

Let $C_b(v_i, p_i|x_i)$ denote the expected profit from revising a limit order in the future, as perceived by the buyer at the time she submits the initial limit order. The arguments v, p, x are those pertaining to the initial limit order. This $C_b(v_i, p_i|x_i)$ is analogous to a continuation value and will be an input to the buyer's optimization problem of choosing p. Mathematically, define $C_b(v_i, p_i|x_i)$ as

$$C_b(v_i, p_i|x_i) \equiv \mathrm{E}_{t_{r,i}, x_{r,i}, v_{r,i}} \Big[P_b\big(p_r^*(v_{r,i}|x_{r,i})\big|x_{r,i}\big)\big(v_{r,i} - p_r^*(v_{r,i}|x_{r,i})\big)\Big| v_i, p_i, x_i \Big], \quad (3)$$

where vector $x_{r,i}$ represents covariates at the time of revision. The buyer's optimal revision price as a function of her value $v_{r,i}$ is denoted $p_r^*(v_{r,i}|x_{r,i})$, and $P_b(p_r^*(v_{r,i}|x_{r,i})|x_{r,i})$ is the probability that a buy limit order with price $p_r^*(v_{r,i}|x_{r,i})$ executes.

Assumption 3. *Regarding the expected profit $C_b(v, p|x)$ of revising a limit order,*
(i) The buyer approximates that $C_b(v, p|x) = C_b(v|x)$.
(ii) $r|C_b(v'|x) - C_b(v|x)| < |v' - v|$ for all $v' \neq v$.

Assumption 3-(i) means the buyer approximates that her own choice of limit order price p will not influence the evolution of covariates x_r to such an extent that it would affect her revision payoff in the future. This is similar to a common assumption made when studying dynamic auctions on eBay, in which a buyer's continuation value is assumed not to depend on her previous actions; see, e.g., Backus and Lewis (2020), Bodoh-Creed et al. (2021), and Hendricks et al. (2021). One rationale in the case of eBay is that the influence of any one bid by one buyer dissipates quickly over time. Another rationale is that it represents realistic buyer behavior because forecasting the evolution of the entire market in response to one's individual action is complex while the gains of doing so for the buyer are small.

Assumption 3-(ii) means the expected profit of future revision, multiplied by the probability of revision r, is less sensitive to the argument v than is v itself. It is satisfied in the following two extreme models of the relationship between v_i and $v_{r,i}$. First, it is automatically satisfied if $v_{r,i}$ and v_i are conditionally independent with $F_b^r(v_r|v) = F_b^r(v_r)$. In this scenario, $C_b(v|x)$ does not depend on v so $C_b(v'|x) - C_b(v|x) = 0$. Second, the assumption is satisfied in the opposite scenario of perfectly persistent values, $v_{r,i} = v_i$.[25] In general, satisfaction of Assumption 3-(ii) depends on the function $P_b(\cdot|x_r)$ and the joint distribution of $(v_i, v_{r,i})$. The next corollary follows from the previous assumptions.

Corollary 1. $\partial C_b(v|x)/\partial v \geq 0$.

CHOOSING A LIMIT-ORDER PRICE ACCOUNTING FOR REVISION IN THE FUTURE
Now, supposing a buyer submits a limit order, consider her choice of price p for the order. Given her value v_i, buyer i chooses price p_i to maximize her expected utility from the order in light of the execution probability $P_b(p|x)$ and accounting for the possibility of future revision. Her maximization problem is

$$\max_p P_b(p|x_i)(v_i - p) + (1 - P_b(p|x_i))rC_b(v_i|x_i). \tag{4}$$

If the buyer's limit order is executed, she gets payoff $v_i - p$, and if it is not executed, she revises the order with probability r, yielding expected payoff $C_b(v_i|x_i)$ from the revision. Differentiating the objective function with respect to p, the first-order condition gives

$$p + \frac{P_b(p|x_i)}{P_b'(p|x_i)} = v_i - rC_b(v_i|x_i). \tag{5}$$

This first-order condition is similar to equation (2), but the buyer now chooses a price in light of a pseudo value that equals her value v_i less the continuation value $rC_b(v_i|x_i)$. In other words, since the buyer's expected payoff from an unexecuted order is now positive, she will behave as if her value is lower by that amount when choosing a limit order price.

By Assumptions 1 and 3, the left-hand side of equation (5) is strictly increasing in p and the right-hand side is strictly increasing in v. Therefore, a buyer's optimal

[25] In this scenario, the expression for $C_b(v|x)$ simplifies to $\mathrm{E}_{t_r,x_r}\left[P_b(p_r^*(v|x_r)|x_r)(v - p_r^*(v|x_r))\,|x\right]$, so by the envelope theorem, $\partial C_b(v|x)/\partial v = \mathrm{E}_{t_r,x_r}\left[P_b(p_r^*(v|x_r)|x_r)\,|x\right] < 1$ where the inequality follows from Assumption 1-(iii). Then $r|\int_{y=v}^{v'} \frac{\partial C_b(y|x)}{\partial y} dy| < |\int_{y=v}^{v'} 1 dy|$, satisfying Assumption 3-(ii).

choice of limit order price p is strictly increasing in her value v.

Corollary 2. *The buyers' best-response price function $p^*(v|x)$ for a limit order is strictly increasing in v.*

3.3 Choosing between a market order and a limit order

Previous sections considered the buyer's choice of price given that she places a limit order. Now consider the buyer's choice between placing a market order and a limit order. A rational buyer would compare the profit of the market order to that of a limit order and choose the one whose (expected) profit is higher.

Suppose a buyer chooses a market order with price ℓ and volume m. Conditional on m, let buyer i obtain value v_i per unit upon receiving the water. Execution of market orders is guaranteed by definition, so the buyer's per-unit profit from this market order is

$$M(v_i, \ell) \equiv v_i - \ell. \tag{6}$$

For the buyer to have placed this market order, it must be that $M(v_i, \ell)$ exceeds the expected profit from placing a limit order for m. To be precise, the buyer's expected per-unit profit from placing an optimally priced limit order (abstracting away from conditioning on x) is

$$L(v_i) \equiv P_b(p^*(v_i))\big(v_i - p^*(v_i)\big) + \big(1 - P_b(p^*(v_i))\big)rC_b(v_i). \tag{7}$$

The buyer choosing the market order implies that her v_i satisfies $M(v_i, \ell) \geq L(v_i)$. The next proposition states that the buyer's preference between $M(v_i, \ell)$ and $L(v_i)$ follows a threshold strategy such that buyers with values lower than a threshold denoted $\tilde{v}(\ell)$ would prefer the limit order and buyers with values higher than the threshold would prefer the market order.

Proposition 1. *Among $v > \ell$, there is at most one value such that $L(v) = M(v, \ell)$, i.e., such that a buyer with this value is indifferent between an optimally priced limit order and a market order of price ℓ. Denoting that value as $\tilde{v}(\ell)$, the buyer would prefer the limit order if $v_i < \tilde{v}(\ell)$ and prefer the market order if $v_i \geq \tilde{v}(\ell)$.*

Meanwhile, data suggest that some traders arriving in the market might default to placing a limit order, rather than placing a limit order only if it is more profitable than a market order. I allow for this default behavior in buyers and sellers with probability α_b and α_s, respectively. With the remaining probability, traders act in accordance with Proposition 1.

3.4 The seller

Now consider the seller. Conditional on order volume m_i, seller i's opportunity cost of relinquishing the water is c_i per unit, distributed according to $F_s(\cdot|x_i)$. The seller's problem is largely a mirror image of the buyer's. The seller analogs of Assumptions 1-3 are named Assumptions 1S-3S and stated in the Appendix. This section concisely summarizes seller analogs of the results we discussed in previous sections for buyers. Note that unlike buy limit orders, sell limit orders with a higher price are *less* likely to be executed because the counterparties—buyers—prefer to pay lower prices.

Choosing a limit-order price with no revision in the future

Consider a seller's choice of limit order price when she believes there will be no revision in the future. Given her opportunity cost $c_{r,i}$ and probability of order execution $P_s(p_r|x_r)$, her maximization problem is $\max_{p_r} P_s(p_r|x_{r,i})(p_r - c_{r,i})$. Differentiating this objective function with respect to p_r, the first-order condition gives

$$p_r + \frac{P_s(p_r|x_{r,i})}{P_s'(p_r|x_{r,i})} = c_{r,i}. \tag{8}$$

Thus, rather than submitting a price equal to her cost, she will choose a higher price $p_r > c_{r,i}$, since $P_s'(p_r|x_{r,i}) < 0$. Let $p_r^*(c_r|x_r)$ denote the seller's optimal price that satisfies equation (8). By Assumption 1S, the left-hand side of equation (8) is strictly increasing in p_r. It follows that $p_r^*(c_r|x_r)$ is strictly increasing in c_r.

Choosing a limit-order price accounting for revision in the future

Now consider the possibility that the seller will revise her limit order after placing it. Her per-unit opportunity cost of relinquishing water at the time of revision is denoted $c_{r,i}$. Let $C_s(c_i|x_i)$ denote the expected profit from revising a limit order in the future, as perceived by the seller at the time she submits the initial limit order: $C_s(c_i|x_i) \equiv \mathrm{E}_{t_{r,i},x_{r,i},c_{r,i}}\left[P_s(p_r^*(c_{r,i}|x_{r,i})|x_{r,i})(p_r^*(c_{r,i}|x_{r,i}) - c_{r,i})\,|c_i, x_i\right]$. When choosing a price p for such a limit order, her maximization problem is $\max_p P_s(p|x_i)(p - c_i) + (1 - P_s(p|x_i))rC_s(c_i|x_i)$. Differentiating this objective function with respect to p, the first-order condition gives

$$p + \frac{P_s(p|x_i)}{P_s'(p|x_i)} = c_i + rC_s(c_i|x_i). \tag{9}$$

Corrollaries 1S and 2S are seller analogs of Corrollaries 1 and 2.

Corrollary 1S. $\partial C_s(c|x)/\partial c \leq 0$.

Corrollary 2S. *The sellers' best-response price function $p^*(c|x)$ for a limit order is strictly increasing in c.*

Choosing between a market order and a limit order

Suppose a seller chooses a market order with price h and volume m. Conditional on m, let seller i's opportunity cost of relinquishing the water be c_i per unit. Then the seller's per-unit profit from this market order is $M(c_i, h) \equiv h - c_i$. For the seller to have placed this market order, it must be that $M(c_i, h)$ exceeds the expected profit from placing a limit order for m. The seller's expected per-unit profit from placing an optimally priced limit order (abstracting away from conditioning on x) is $L(c_i) \equiv P_s(p^*(c_i))(p^*(c_i) - c_i) + (1 - P_s(p^*(c_i)))rC_s(c_i)$.

Proposition 1S. *Among $c < h$, there is at most one value of c such that $L(c) = M(c, h)$, i.e., such that a seller with this cost is indifferent between an optimally priced limit order and a market order of price h. Denoting that value as $\tilde{c}(h)$, the seller would prefer the limit order if $c_i > \tilde{c}(h)$ and prefer the market order if $c_i \leq \tilde{c}(h)$.*

3.5 Identification

The broad intuition for identification of buyers' values and sellers' costs conditional on order volume is provided by the first-order conditions for limit order pricing derived above. However, there are additional complexities to address because i) traders may choose to place market orders rather than limit orders, and ii) traders may revise their limit order in the future.

I assume the following are observed for each limit order: the price p and covariates x; indicators for whether it is buy/sell, an initial order/revision, executed/not; and the associated revision order, if any. The following are observed for each market order: the trade price p, covariates x, and whether it is buy/sell. For the purpose of assessing counterfactual exchange design, the necessary model objects to identify are F_b, F_b^r and F_s, F_s^r. These are the distribution of buyer values, the conditional distribution of buyers' revision values, and the analogs for sellers. While I use this distributional notation in keeping with convention and for the sake of expositional brevity, identification is in fact obtained at a more granular level: order-level values and costs are identified for all limit orders, both initial orders and revisions. As an intermediate step, I also identify the execution probabilities, $P_b(\cdot|x)$ and $P_s(\cdot|x)$, and the probabilities of defaulting to limit orders, α_b and α_s. For the remainder of this section, I abstract away from conditioning on x and x_r.

Identification of $P_b(\cdot)$,$P_s(\cdot)$ and $\tilde{F}_b^r(\cdot|\cdot)$,$\tilde{F}_s^r(\cdot|\cdot)$

The probability of limit order execution $P_b(p)$ and $P_s(p)$ for the buyer and seller, respectively, are identified as $\mathrm{E}[\mathbb{1}_{\text{executed}}|p]$, where $\mathbb{1}_{\text{executed}} = 1$ if a limit order is executed and 0 otherwise. Next, consider identification of the buyer's value v_r or seller's cost c_r associated with each observation of a *revised* limit order with price p_r. Having identified $P_b(\cdot)$ and $P_s(\cdot)$, equations (2) and (8) identify the type as $v_r = p_r + \frac{P_b(p_r)}{P_b'(p_r)}$ and $c_r = p_r + \frac{P_s(p_r)}{P_s'(p_r)}$ for buyers and sellers, respectively.

By Corrollaries 2 and 2S, the limit order price function $p^*(\cdot)$ of a trader is monotonic in the trader's value or cost. To reduce clutter in notation, I use $p^*(\cdot)$ to denote both the buyers' and sellers' price functions though they are different functions for the two sides. Using the monotonicity of $p^*(\cdot)$, I define new notation $\tilde{F}_b^r(\cdot|\cdot)$ and $\tilde{F}_s^r(\cdot|\cdot)$ such that $\tilde{F}_b^r(v_r|p^*(v)) = F_b^r(v_r|v)$ and $\tilde{F}_s^r(c_r|p^*(c)) = F_s^r(c_r|c)$.[26] As the type underlying each revised limit order is identified, and the price of the initial order associated with each revised order is observed, the distribution $\tilde{F}_b^r(\cdot|p)$ conditional on any p is identified as the empirical distribution of v_r conditional on that initial order price p. Analogous logic applies to identification of the seller's $\tilde{F}_s^r(\cdot|p)$.

Identification of the v or c associated with each limit order

Using the monotonicity in Corrollaries 2 and 2S, I define new notation $\tilde{C}(\cdot)$ for the expected profit of revising a limit order such that $\tilde{C}_b(p^*(v)) = C_b(v)$ for a buyer and $\tilde{C}_s(p^*(c)) = C_s(c)$ for a seller. Specifically, this involves replacing $F_b^r(v_r|v)$ with $\tilde{F}_b^r(v_r|p^*(v))$ inside the definition of $C_b(\cdot)$. For a buyer, this gives $\tilde{C}_b(p) \equiv \int_{v_r} P_b(p_r^*(v_r))(v_r - p_r^*(v_r))d\tilde{F}_b^r(v_r|p)$. As $P_b(\cdot)$, $\tilde{F}_b^r(v_r|p)$, and the value v_r associated

[26] In an analysis of sequential auctions, Kong (2021) uses the monotonicity of first-auction bids to similarly define the distribution of second-auction values conditional on observed first-auction bids and establishes identification of that conditional distribution.

with each observed revision price $p_r^*(v_r)$ are identified as explained above, $\tilde{C}_b(p)$ is identified. Analogous logic applies to identification of the seller's $\tilde{C}_s(p)$.

Having identified $\tilde{C}_b(\cdot)$ and $\tilde{C}_s(\cdot)$, the buyer's v or seller's c associated with an initial limit order is identified by first-order conditions (5) and (9) as

$$v = p + \frac{P_b(p)}{P_b'(p)} + r\tilde{C}_b(p), \tag{10}$$

$$c = p + \frac{P_s(p)}{P_s'(p)} - r\tilde{C}_s(p). \tag{11}$$

Knowledge of the v or c associated with any order price p implies identification of the inverse limit-order price function, $p^{*-1}(\cdot)$. Having identified $\tilde{F}^r(\cdot|\cdot)$ above, knowledge of $p^{*-1}(\cdot)$ allows identification of $F^r(\cdot|\cdot)$ using $F_b^r(\cdot|p^{*-1}(p)) = \tilde{F}_b^r(\cdot|p)$ and $F_s^r(\cdot|p^{*-1}(p)) = \tilde{F}_s^r(\cdot|p)$.

Identification of α_b and α_s

Per Proposition 1, a rational buyer i prefers a market order of price ℓ_i over a limit order if $v_i \geq \tilde{v}(\ell_i)$, where the threshold $\tilde{v}(\ell_i)$ is known as the solution to the indifference condition $L(\tilde{v}(\ell_i)) = M(\tilde{v}(\ell_i), \ell_i)$. The probability α_b that buyers default to placing a limit order is then identified as the probability that they place a limit order despite $v_i \geq \tilde{v}(\ell_i)$:

$$\alpha_b = \frac{\#\{\text{limit orders}|v_i \geq \tilde{v}(\ell_i)\}}{\#\{\text{market orders}\} + \#\{\text{limit orders}|v_i \geq \tilde{v}(\ell_i)\}}. \tag{12}$$

Similarly, by Proposition 1S, the probability α_s that sellers default to placing a limit order is identified as the probability that they place a limit order despite $c_i \leq \tilde{c}(h_i)$:[27]

[27] In equations (12) and (13), we need the ℓ_i or h_i relevant to each observed limit order to identify α_b and α_s. To be clear, the researcher does *not* need to know ℓ_i or h_i—the best available market order price—when it comes to identifying the values and costs underlying each limit order, per equations (10) and (11). But ℓ_i or h_i *is* needed to define whether an observed limit order constitutes irrational behavior. When available market order opportunities have heterogeneous volumes, what constitutes this best available market order price is less obvious because it is not clear which sell listings were relevant for a buyer who ultimately ordered volume m_i in the data.

At a minimum, the set of market order opportunities relevant to this buyer i includes all live sell listings offering *at least* volume m_i, because the buyer could buy and leave unused any excess over m_i. To compare with i's observed limit order, define ℓ_i as the lowest of (total price)/m_i among this set. (Note ℓ_i is then likely to come from a listing with a volume relatively close to m_i, because larger volumes are less likely to have the lowest total price.) A rational buyer choosing to place the observed limit order despite the availability of this ℓ_i implies that her value cannot exceed $\tilde{v}(\ell_i)$. Therefore, if her limit order actually has $v_i \geq \tilde{v}(\ell_i)$, it is classified as "default" behavior.

Imposing arbitrary assumptions that enlarge the set of sell listings relevant to buyer i can only

$$\alpha_s = \frac{\#\{\text{limit orders} | c_i \leq \tilde{c}(h_i)\}}{\#\{\text{market orders}\} + \#\{\text{limit orders} | c_i \leq \tilde{c}(h_i)\}}. \tag{13}$$

Identification of $F_b(\cdot)$,$F_s(\cdot)$

While the buyer's v associated with each initial limit order is point identified as explained above, all we can know about the v associated with each observed market order of price ℓ_i is that $v_i \geq \tilde{v}(\ell_i)$. Thus, for a fixed ℓ, we would not be able to identify $F_b(\cdot)$ above $\tilde{v}(\ell)$ if $\alpha_b = 0$; we would only be able to identify the shape of $F_b(\cdot)$ for values up to $\tilde{v}(\ell)$ and identify the total mass but not the shape of values above $\tilde{v}(\ell)$. On the other hand, if $\alpha_b > 0$, we would also see limit orders associated with values above $\tilde{v}(\ell)$. Since we can identify the v_i associated with each limit order, this allows for the nonparametric identification of $F_b(\cdot)$ on its entire support. Analogously, if $\alpha_s > 0$, $F_s(\cdot)$ is nonparametrically identified on its entire support. The proof of Proposition 2 provides the details of this step.

Proposition 2. *Under Assumptions 1–3 and 1S–3S, the buyer value v or seller cost c associated with each limit order is identified in addition to the conditional distributions of revision values and costs, $F_b^r(\cdot|\cdot)$ and $F_s^r(\cdot|\cdot)$, and the probabilities of defaulting to limit orders, α_b and α_s. If $\alpha_b, \alpha_s > 0$, the value and cost distributions $F_b(\cdot)$ and $F_s(\cdot)$ are also identified nonparametrically.*

Discussion

This section established identification of buyers' values and sellers' costs *conditional* on the volumes they ordered in the data. To the extent that bilateral frictions on the exchange distort a trader's observed order volume away from their ideal, it is not possible to identify from my data what volume the trader would have ordered in the absence of said frictions. Stated in broader terms, the value-volume function (i.e., demand or supply curve) of individual trader i cannot be identified from their scalar order price and volume. There is also no clear relationship between order volumes and prices in the aggregate (see Section 5.1, Table 2). In light of the non-identification of latent target volumes, the counterfactual simulations of Section 6 will have traders retain the same order volumes as seen in the data. This conservative

lower ℓ_i. Lower ℓ_i implies lower $\tilde{v}(\ell_i)$, leading to a higher probability that an observed limit order constitutes default behavior. I use the definition above which minimizes the implied rate of defaulting. For a seller i who placed a sell limit order of volume m_i in the data, the analogous definition of h_i is the highest of (total price)/m_i among all live buy listings that seek *up to* volume m_i.

approach understates the efficiency gains of the counterfactual mechanisms I study, which eliminate bilateral frictions. If traders were to order different, undistorted volumes in response to mechanisms that eliminate volume frictions, the full efficiency gains would be greater than the ones I present in Section 6.

4 Extensions and estimation

4.1 Extensions

Risk aversion The main participants in Australian water markets are irrigators (farmers). In light of evidence from the literature, e.g., Bajari and Hortaçsu (2005), that real behavior is often better represented by risk-averse models than risk-neutral ones, I allow for the possibility that traders are risk averse. A risk-averse buyer with no future revision chooses a limit order price to maximize $P_b(p_r)U((v_r - p_r)m)$, where $U(\cdot)$ is the utility function. Under the constant relative risk aversion (CRRA) specification of $U(\cdot) = \frac{(\cdot)^{1-\eta}}{1-\eta}$, the first-order condition for this maximization problem gives

$$\frac{P_b(p_r)}{P_b'(p_r)} = \frac{U(v_r - p_r)}{U'(v_r - p_r)} = \frac{v_r - p_r}{1 - \eta}. \tag{14}$$

A CRRA parameter of $\eta = 0$ represents risk neutrality while $\eta > 0$ represents risk aversion. For $\eta \in [0, 1)$ the right-hand side of (14) is increasing in η. Given Assumption 1, a more risk-averse buyer will thus choose a higher limit order price to satisfy (14); i.e., she will shade her price less. The intuition is that a risk-averse trader is less tolerant of the risk of not trading and is therefore willing to accept lower profit to reduce this risk. For any given η, the v_r associated with an observed limit order price p_r is identified by rearranging equation (14) to obtain

$$v_r = p_r + (1 - \eta)\frac{P_b(p_r)}{P_b'(p_r)}. \tag{15}$$

Meanwhile, when a buyer anticipates she will have an opportunity to revise the limit order with probability r and the expected utility of revision is $C_b(v)$, her optimal limit order price p satisfies

$$\frac{P_b(p)}{P_b'(p)} = \frac{U(v - p) - rC_b(v)}{U'(v - p)} = \frac{v - p}{1 - \eta} - rC_b(v)(v - p)^{\eta}. \tag{16}$$

As explained in Section 3.5, we can replace $C_b(v)$ with $\tilde{C}_b(p)$ in equation (16) if $p = p^*(v)$. This yields $\frac{P_b(p)}{P_b'(p)} = \frac{U(v-p)-r\tilde{C}_b(p)}{U'(v-p)}$. Given that $U'(\cdot) > 0$ and $U''(\cdot) \leq 0$, the right-hand side of this equation is increasing in v. Thus, for any given η, the v associated with any observed p is identified as the unique value that satisfies this equation.

It remains to identify η. The empirical literature shows that risk aversion can be identified, for example, through various exclusion restrictions.[28] Here, I will take a more flexible approach: rather than pinning down a single value of η through an exclusion restriction, the empirical analysis will be conducted over a range of η values that can rationalize the observed data under Assumption 4 below, a natural assumption regarding the value and cost distributions. Per equation (15), an important role of η is to affect the amount of price shading, or the difference between the trader's chosen limit order price and her underlying value. Conducting the analysis and comparing empirical findings across a range of risk aversion levels will yield additional insights about the impact of price shading and conclusions that are robust to varying levels of price shading.

Assumption 4. *The distributions of buyer values F_b, F_b^r and of seller costs F_s, F_s^r each have non-negative support.*

Some values of η may not be able to rationalize the observed data while satisfying Assumption 4. For example, consider the seller's analog of equation (15), $c_r = p_r + (1-\eta)\frac{P_s(p_r)}{P_s'(p_r)}$. Because $P_s'(p_r) < 0$ for the seller, the right-hand side may be negative at some values of η, which violates Assumption 4.

Taxes and fees In the empirical application, income taxes and exchange fees are levied on the seller as a fraction τ of sales revenue if a trade occurs. Then a seller with CRRA utility and no future revision would choose a limit order price to maximize $P_s(p_r)U([(1-\tau)p_r - c_r]m)$. The first-order condition for this maximization problem gives

$$p_r + (1-\eta)\frac{P_s(p_r)}{P_s'(p_r)} = \frac{c_r}{1-\tau}.$$

[28] Example exclusion restrictions used in the auction literature include conditional independence of value distributions and the number of bidders (Guerre, Perrigne, and Vuong (2009)) or auction formats (Lu and Perrigne (2008), Kong (2020)). Perrigne and Vuong (2021) and Vasserman and Watt (2021) survey additional strategies for identifying risk aversion.

When a seller anticipates she will have an opportunity to revise the limit order with probability r and the expected utility of revision is $C_s(c)$, her optimal limit order price p satisfies

$$p + (1-\eta)\frac{P_s(p)}{P_s'(p)} = \frac{c + rC_s(c)(1-\eta)[(1-\tau)p - c]^{\eta}}{(1-\tau)}. \tag{17}$$

Meanwhile, if a trade occurs, buyers pay a flat trade approval fee of 47.50 AUD per trade to the state water authority during the sample period. Denoting this flat fee as f, the buyers' first order conditions for limit order pricing follow (14) and (16) with $v - f/m$ (unit value net of fees) replacing v in those equations.

4.2 Estimation

While the estimation procedure closely aligns with the identification argument, a key task of estimation is to condition estimated functions on the vector of covariates x in the context of a limited sample. Therefore, some functions will be estimated with parametric specifications though they are identified nonparametrically. To allow a range of different risk aversion levels, I conduct the estimation procedure once for each value of the CRRA parameter η over a grid in [0,1) at intervals of 0.01.

Estimation of $P_b(p|x)$ I specify the execution probability function $P_b(p|x)$ with a probit model of the form $P_b(p|x) = \Phi(f(p;\beta_p) + x\beta_x)$, where $\Phi(\cdot)$ is the standard normal CDF, $f(p;\beta_p)$ is a fourth order polynomial of p to allow flexible shapes in p, and β_p and β_x are probit coefficient vectors to be estimated. To satisfy the monotonicity property $P_b'(p|x) > 0$, the probit coefficients are estimated via maximum likelihood subject to the constraint that $f'(p;\beta_p) > 0$ for buyers.

Estimation of $\tilde{C}_b(p|x)$ Having estimated $\hat{P}_b(p|x)$, the buyer's value v_r associated with each *revision* order price p_r is estimated in accordance with first-order condition (15) as

$$\hat{v}_r - f/m = p_r + (1-\eta)\frac{\hat{P}_b(p_r|x_r)}{\hat{P}_b'(p_r|x_r)}.$$

The buyer's expected utility from such a revision order is $\hat{P}_b(p_r|x_r)([\hat{v}_r - p_r]m - f)^{1-\eta}/(1-\eta) \equiv m^{1-\eta}\hat{\pi}_r$. Then, as defined in Section 3.5, the expected utility of revision $\tilde{C}_b(p|x)$ is the expectation of $\hat{\pi}_r$ given p and x of the initial limit order. I estimate $\hat{\hat{C}}_b(p|x)$ via random forest regression of each revision order's $\hat{\pi}_r$ on the p and x of the associated initial order.

Estimation of the $\hat{v}$ associated with each limit order Having estimated $\hat{\hat{C}}_b(p|x)$, the buyer's value v associated with each initial limit order price p is estimated by solving for the $\hat{v}$ that satisfies first-order condition (16), of which the empirical analog is

$$\frac{\hat{P}_b(p|x)}{\hat{P}_b'(p|x)} = \frac{\hat{v} - f/m - p}{1-\eta} - r\hat{\hat{C}}_b(p|x)(\hat{v} - f/m - p)^{\eta}.$$

Estimation of α_b Having estimated the $\hat{v}$ of each limit order, the probability α_b that buyers default to placing a limit order is estimated using the empirical analog of equation (12), namely

$$\hat{\alpha}_b = \frac{\#\{\text{limit orders}|\hat{v}_i \geq \hat{\tilde{v}}(\ell_i)\}}{\#\{\text{market orders}\} + \#\{\text{limit orders}|\hat{v}_i \geq \hat{\tilde{v}}(\ell_i)\}}.$$

The market/limit threshold $\hat{\tilde{v}}(\ell_i)$ applicable to each limit order i is found by solving the indifference condition that a buyer with value $\hat{\tilde{v}}(\ell_i)$ is indifferent between a limit order and a market order, given the risk aversion level η.

Estimation of $F_b(v|x)$ Note that $\hat{v}_i$ is not available for the observed market orders, about which we know only that $v_i \geq \hat{\tilde{v}}(\ell_i)$. Therefore, even though we have order-level $\hat{v}_i$ for all limit orders, we will also want to estimate the value distribution $F_b(v|x)$ to enable simulations in the post-estimation analysis. I specify a lognormal distribution for $F_b(v|x)$ because the empirical distribution of $\hat{v}$ resembles a lognormal distribution. Specifically, I let v be distributed according to Lognormal$(x\theta_\mu, \exp(x\theta_\sigma))$, where $x\theta_\mu$ and $\exp(x\theta_\sigma)$ are the mean and standard deviation of $\ln(v)$, respectively. The parameter vectors θ_μ and θ_σ are estimated by maximizing the likelihood of observed limit and market orders. The likelihood of a market order is $(1 - F_b(\hat{\tilde{v}}(\ell_i)|x_i))(1 - \alpha_b)$ because a market order implies $v_i \geq \hat{\tilde{v}}(\ell_i)$ and that the buyer did not default to a limit order. Letting $f_b(v|x)$ denote the pdf of $F_b(v|x)$, the likelihood of a limit order with $\hat{v}_i \geq \hat{\tilde{v}}(\ell_i)$ is $f_b(\hat{v}_i|x_i)\alpha_b$, and the likelihood of a limit order with $\hat{v}_i < \hat{\tilde{v}}(\ell_i)$ is $f_b(\hat{v}_i|x_i)$.

The seller analogs of each of the objects above are estimated in the same manner using the seller's version of the model equations and definitions.

5 Estimation results

5.1 Covariates

Elements of the vector of covariates x broadly fall into two categories. The first category, which I collect in the vector x_w, concerns the fundamentals of water. I account for water availability using the daily total amount of water in the major storages of the Victorian Murray River, as provided by the Australia Bureau of Meteorology.[29] To capture true deviations of storage levels from long-term norms as opposed to expected seasonal fluctuations, I regress the log of the daily total storage amounts during February 2012 to February 2022 on month-day fixed effects and refer to the regression residuals as "seasonally adjusted" log storage. Seasonality is accounted for separately by *waterday*, which is a day number: the first day of the water year, July 1st, is day 1, and the last day of the water year, June 30th, is day 365 or 366 depending on the year. To be clear, *waterday* enters linearly (for ease of interpretation in Tables 2 and 3) or in a polynomial (in the structural estimation) and contains only month-day (seasonal) information; it is not a year-month-day fixed effect. The order volume is also included in x_w.

The second category of covariates, which I collect in the vector x_e, concerns conditions of the exchange at the time the trader's action being analyzed took place. While the exchange is a high dimensional object, I include a feasibly limited set of variables: the total volume in megaliters of live sell limit orders (ask depth) and buy limit orders (bid depth); the lowest available per-unit sell limit order price (low ask) or the highest available per-unit buy limit order price (high bid) and their difference (bid-ask spread); total trade volume and the mean and standard deviation of trade prices during the preceding week. If there were no trades during the preceding week, I substitute with the mean and standard deviation of trade prices from the closest preceding week in which trades occurred. If the ask (bid) depth is zero, I use the most recent non-empty low ask (high bid) as a substitute for the low ask (high bid).

Table 2 shows that these covariates are able to explain a large fraction of the

[29]These storages are Lake Dartmouth, Lake Hume, the Menindee Lakes, and Lake Victoria.

sample variance of limit order prices. The regressions in columns (1) and (2) pool buy and sell limit orders, while columns (3) and (4) use only sell limit orders and buy limit orders, respectively. The $R^2 = 0.78$ of column (1) shows that elements of x_w, namely the amount of water storage and seasonality, have high explanatory power when it comes to limit order prices. In particular, limit order prices decrease in the amount of water stored. Column (2) shows that limit order prices are highly correlated with the mean trade price of the preceding week, the inclusion of which increases the R^2 from 0.78 to 0.88. Columns (3) and (4) include the remaining x_e elements as regressors. In general, these regression coefficients need not have causal interpretations. Appendix Table A1 explores the explanatory power of additional candidates including rainfall, evapotranspiration, and announcements of additional water being allocated. In light of coefficient signs on these variables that go against economic logic and in the absence of improvement in the R^2, I opt for the set in Table 2 to avoid increasing the number of parameters to estimate in the structural model.

Table 3 investigates the relationship between a trader's choice of limit order price and probability of order execution, controlling for the covariates in x. To capture, in a manner that is interpretable across periods of heterogeneous price distributions, how competitive the chosen price is relative to the best per-unit price available on the exchange, I normalize the limit order price as (price $-$ best available price at time of order placement)/(standard deviation of traded prices during the preceding week) and use this normalized price in the probit regression. In line with economic intuition, column (1) shows that execution probability is decreasing in order price for sell orders. Meanwhile, consistent with the raw patterns seen in Section 2.3, the response of the predicted execution probability to the order price is not very steep. To elaborate, the standard deviation of traded prices is on average about 7% of the low ask, though this varies across time. Interpreting Table 3, column (1) at sample mean values of x, a sell limit order with a 7% higher per-unit price than the low ask would have an approximately 3.2% lower execution probability than an order placed at the low ask, ceteris paribus.

Over 95% of realized trades in the sample come from execution of sell limit orders as opposed to buy limit orders, so the slope of buyers' execution probability is more difficult to estimate. To exploit a larger number of "ones" in estimating this slope, I first define $\tilde{P}_b(p|x)$ as the probability that either a buy limit order executes or a trade of very similar price and volume occurs within its duration; very similar price

Table 2: Regression of log limit order price on covariates

	(1)	(2)	(3)	(4)
daily storage (log), seasonally adjusted	-2.461	-0.472	-0.551	-0.089
	(0.027)	(0.048)	(0.061)	(0.071)
waterday/366	-0.824	-0.204	-0.141	-0.061
	(0.017)	(0.019)	(0.022)	(0.027)
volume (log)	-0.004	0.003	0.003	0.004
	(0.004)	(0.003)	(0.003)	(0.005)
mean trade price, prev week (log)		0.800	0.580	0.331
		(0.017)	(0.029)	(0.047)
st dev of trade price, prev week			0.004	0.001
			(0.001)	(0.001)
total trade volume, prev week (1000 ML)			-0.009	0.001
			(0.002)	(0.002)
ask depth (1000 ML)			-0.008	0.011
			(0.005)	(0.008)
bid depth (1000 ML)			0.049	0.008
			(0.008)	(0.009)
bid-ask spread (100 AUD)			0.011	0.004
			(0.003)	(0.004)
low ask (log)			0.268	
			(0.024)	
high bid (log)				0.528
				(0.045)
sell dummy	0.130	0.130		
	(0.013)	(0.010)		
constant	5.299	0.986	0.825	0.608
	(0.023)	(0.095)	(0.114)	(0.155)
Data sample: sell or buy limit orders	both	both	sell	buy
Observations	2813	2813	2182	631
R^2	0.782	0.876	0.891	0.879
Adjusted R^2	0.782	0.876	0.890	0.877

Standard errors in parentheses

Table 3: Probit regression of whether a limit order executes

	(1) sell	(2) buy*
limit order price, normalized	-0.086	0.106
	(0.008)	(0.034)
daily storage (log), seasonally adjusted	-2.221	-3.221
	(0.462)	(0.890)
waterday/366	-0.631	-0.683
	(0.156)	(0.293)
volume (log)	-0.140	0.035
	(0.025)	(0.058)
mean trade price, prev week	-0.005	-0.006
	(0.001)	(0.003)
st dev of trade price, prev week	-0.005	-0.020
	(0.004)	(0.013)
total trade volume, prev week (1000 ML)	-0.068	0.079
	(0.018)	(0.025)
ask depth (1000 ML)	0.107	0.136
	(0.032)	(0.089)
bid depth (1000 ML)	0.163	0.089
	(0.055)	(0.097)
bid-ask spread (100 AUD)	-0.050	-0.016
	(0.021)	(0.046)
low ask	0.001	
	(0.001)	
high bid		0.002
		(0.003)
constant	1.441	-0.613
	(0.230)	(0.493)
Observations	2182	631
Pseudo R^2	0.106	0.089

Notes: *The probit regression in (2) estimates the probability that either a buy limit order executes or a trade of very similar price and volume occurs within its duration; very similar price and volume is defined as $(p' - p)^2 + (m' - m)^2 \leq 1$. See the end of Section 5.1 for details. Standard errors in parentheses.

and volume is defined as $(p' - p)^2 + (m' - m)^2 \leq 1$. I then use an approximation that $P_b(p|x) = \omega \tilde{P}_b(p|x)$, where $\omega \equiv \frac{\int\int P_b(p|x)dpdx}{\int\int \tilde{P}_b(p|x)dpdx}$. Table 3, column (2) reports the relationship between $\tilde{P}_b(p|x)$ and its arguments via a probit regression. This $\tilde{P}_b(p|x)$ is increasing in order price for buy orders.

5.2 Estimation results

This section discusses the results of the estimation procedure detailed in Section 4.2. As described there, the execution probabilities $P_b(p|x)$ and $P_s(p|x)$ are estimated as an intermediate step. The results of this estimation are similar to that shown in Table 3, with polynomials of limit order price and day number added for more flexibility. The polynomial of price is constrained to be monotonic. The appendix reports the estimated coefficients in Table A2.

Recall that I conduct the estimation procedure once for each value of the CRRA parameter η over a grid in [0,1) at intervals of 0.01. To respect Assumption 4 while allowing for econometric error, the post-estimation analysis of Section 6 focuses on the range of η values given which at most 1% of all sell limit orders have $\hat{c} < 0$. This narrows η to $[0.61, 1)$. The remainder of the paper will report results at the two extremes of this range, $\eta = 0.61$ and $\eta = 0.99$.

The exposition, for the sake of brevity, often refers to value or cost distributions such as $F_s(\cdot)$ and $F_b(\cdot)$. Estimates are in fact obtained at a much more granular level: order-level $\hat{v}$ or $\hat{c}$ are estimated for all limit orders. To describe these order-level value estimates, it is helpful to relate them to the respective observed order prices. As discussed previously, different values of η lead to different amounts of price shading. Given $\eta = 0.99$, the median amount of price shading $|[p_i - (\hat{v}_i - f/m_i)]/(\hat{v}_i - f/m_i)|$ among all buy limit orders is 0.4%, and the median amount of price shading $((1-\tau)p_i - \hat{c}_i)/\hat{c}_i$ among all sell limit orders is 0.7%. In other words, there is very little price shading when traders are this risk averse, resulting in $p \approx v$. Meanwhile, given lower risk aversion at the level of $\eta = 0.61$, the analogous medians are 13% for buy limit orders and 36% for sell limit orders. When it comes to the lognormal specifications of $F_s(\cdot|x_w)$ and $F_b(\cdot|x_w)$—the end of Section 4.2 explains why these are estimated—the appendix reports the estimated parameters in Table A3. Unlike the probability of limit order execution, the distribution of latent values (or costs) is specified as a policy-invariant primitive that depends only on water fundamentals x_w

and not on exchange conditions x_e. Similarly to what we saw in Table 2, values for water are decreasing in the amount of water available in the major storages of the Victorian Murray River. The appendix also provides an illustration of how well the shape of the lognormal distribution approximates that of the estimated order-level valuations (Figure A1).

Finally, the estimated probability α of placing a limit order by default varies widely with η. This is because buyers' $\hat{v}$ and sellers' $\hat{c}$ also vary with η, and α is estimated by assessing the rationality of choosing a limit order versus a market order in light of these $\hat{v}$ or $\hat{c}$. Nonetheless, estimates given both high and low η indicate that sellers are much more likely to default to limit orders than are buyers, in harmony with the empirical observation that 78% of limit orders on the exchange are sell orders. Specifically, $\hat{\alpha}_s = 0.93$ and $\hat{\alpha}_b = 0.18$ given $\eta = 0.61$, while $\hat{\alpha}_s = 0.62$ and $\hat{\alpha}_b = 0.04$ given $\eta = 0.99$.

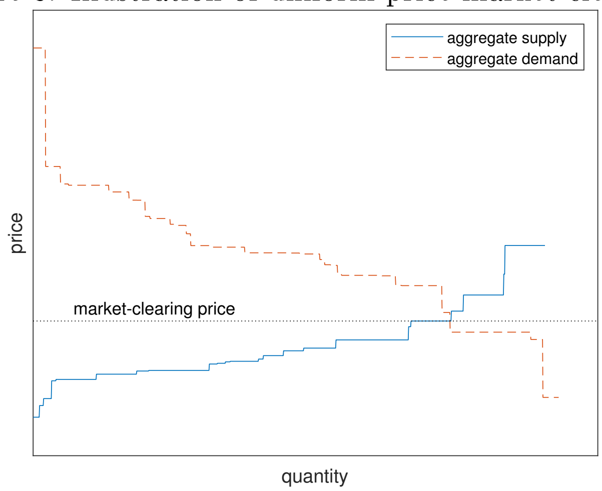
6 Counterfactual analysis

6.1 Benchmark: batch uniform-price market clearing

Consider the benchmark of periodically crossing latent supply and demand, as illustrated in Figure 3. A market clearing price and quantity are determined by where the period's aggregate supply and aggregage demand curves cross. Within each period, this would give the efficient allocation. In particular, I will consider market clearing on a weekly basis for comparability with the observed exchange, where the mean duration of a limit order is about 8 days. In finance, call or batch auctions at discrete times have long been considered an alternative to continuous trading; Economides and Schwartz (1995), Handa and Schwartz (1996), and others have argued the merits of concentrating orders in a call auction. Also, Rustichini, Satterthwaite, and Williams (1994) and Satterthwaite and Williams (2002) show that trade at a single market-clearing price through a double auction makes the worst-case inefficiency converge to zero at the fastest possible rate as the number of traders increases.

My benchmark differs in important ways from the continuous, bilateral exchange observed in the data. The benchmark has temporal aggregation of traders as the market is cleared at discrete time intervals. Trade is multilateral rather than bilateral as the volumes are pooled; i.e., the volume that the benchmark allocates to a given

Figure 3: Illustration of uniform-price market clearing



trader may come from/go to multiple counterparties. This avoids bilateral frictions. Unlike the observed exchange where multiple trade prices occur even at a given moment, the benchmark generates a single market-clearing price that sorts which buyers and sellers get to trade and which do not. Finally, the benchmark crosses latent supply and demand, so there is no price shading by construction. To supplement the benchmark, which quantifies an ideal, I also consider a closely related incentive compatible mechanism adapting McAfee (1992) in Section 6.2.

I use the model primitives estimated per Sections 4 and 5 to counterfactually simulate the market-clearing benchmark discussed above. Aside from the change from limit order market to batch market clearing, I endeavor to keep all else the same as on the observed exchange. First, traders arrive at the same times, with the same volumes, and with the same values $\hat{v}$ or costs $\hat{c}$ as on the observed exchange. There are two exceptions that arise from tracking revisions. One exception occurs if an order is not executed in the data but the corresponding order does become a trade in the counterfactual world: then any revisions of such orders are deleted in the counterfactual world because it would be double counting to resubmit a volume that has already traded. The other exception occurs if an initial order and associated revision order happen in the same week in the data: then the revision order is deleted from the counterfactual to avoid double counting that volume in a single week. By reducing volume in the counterfactual benchmark, these exceptions work against the benchmark in a comparison versus the observed exchange. Second, each trader still submits a single price for their order. Third, traders still account for dynamics, meaning there is a non-zero continuation value of failing to trade. Since continuation values depend on exchange design, the continuation values traders account for in the counterfactual are recomputed by simulation per the details below. Finally, sellers still pay the same rate of taxes and fees. Buyers still pay the same flat fee per trade, but now a purchase resulting from a single buy order counts as a single trade for purposes of the fee, regardless of the number of sellers from which that purchase is sourced. If the market clearing counterfactual leads to only part of their order executing, I have each buyer pay 47.50 AUD $\times \frac{\text{executed volume}}{\text{order volume}}$ to preserve the same fee per traded unit they would have paid on the observed exchange. Implementation details follow; readers who wish to skip the technical details may proceed to the counterfactual simulation results.

Implementation details

In any given week, I observe buyers and sellers arrive and make a limit or market order on the observed exchange. For each observed limit order, I have an estimated $\hat{v}_i$ or $\hat{c}_i$ to use in simulating the counterfactual. For each observed market order, I know only that $v_i \geq \hat{\tilde{v}}_i$ for buy orders and $c_i \leq \hat{\tilde{c}}_i$ for sell orders, where $\hat{\tilde{v}}_i$ and $\hat{\tilde{c}}_i$ are the relevant thresholds for choosing a market order. Therefore, for each buy market order I randomly draw a v_i from the left-truncated lognormal distribution $[\hat{F}_b(\cdot|x_{w,i}) - \hat{F}_b(\hat{\tilde{v}}_i|x_{w,i})]/[1 - \hat{F}_b(\hat{\tilde{v}}_i|x_{w,i})]$, and for each sell market order I randomly draw a c_i from the right-truncated lognormal distribution $\hat{F}_s(\cdot|x_{w,i})/\hat{F}_s(\hat{\tilde{c}}_i|x_{w,i})$. I now have a value or cost associated with every trader arrival. For ease of notation, I will refer to both estimated and simulated valuations collectively using the hat notation as $\hat{v}_i$ and $\hat{c}_i$. Given no future revision, a buyer's market-facing value equals her $\hat{v}_i - f/m_i$ with flat fee f, while a seller's market-facing cost is $\hat{c}_i/(1-\tau)$ due to taxes and fees of rate τ. With the possibility of future revision, a buyer's market-facing value is $\hat{v}_i - f/m_i - U^{-1}(r\hat{C}_{CF,b}(\hat{v}_i|x_i))$, and a seller's market-facing cost is $[\hat{c}_i + U^{-1}(r\hat{C}_{CF,s}(\hat{c}_i|x_i))]/(1-\tau)$, where the CF subscripts in $\hat{C}_{CF,b}(\hat{v}_i|x_i)$ and $\hat{C}_{CF,s}(\hat{c}_i|x_i)$ denote that they are continuation values in the *counterfactual*, as opposed to the observed, design.

The counterfactual continuation values are simulated by iteration as follows. Using estimated continuation values $\hat{C}_b(\hat{v}_i|x_i)$ and $\hat{C}_s(\hat{c}_i|x_i)$ from the observed exchange as initial guesses, I simulate the counterfactual benchmark week by week in calendar sequence. This allows me to compute the realized utility $\hat{\pi}_r$ of each revision in the counterfactual. Next, I estimate the expected utility of revision given $\hat{v}_i$ and x_i

of the initial order via random forest regression of $\hat{\pi}_r$ on $\hat{v}_i$ and x_i. This estimation gives updated counterfactual continuation values $\hat{C}_{CF,b}(\hat{v}_i|x_i)$ and $\hat{C}_{CF,s}(\hat{c}_i|x_i)$. Then the counterfactual benchmark is re-simulated using the updated continuation values, and this procedure is iterated until the difference in $U^{-1}(r\hat{C}_{CF,b}(\hat{v}_i|x_i))$ and $U^{-1}(r\hat{C}_{CF,s}(\hat{c}_i|x_i))$ between iterations is less than 0.10 Australian dollars.

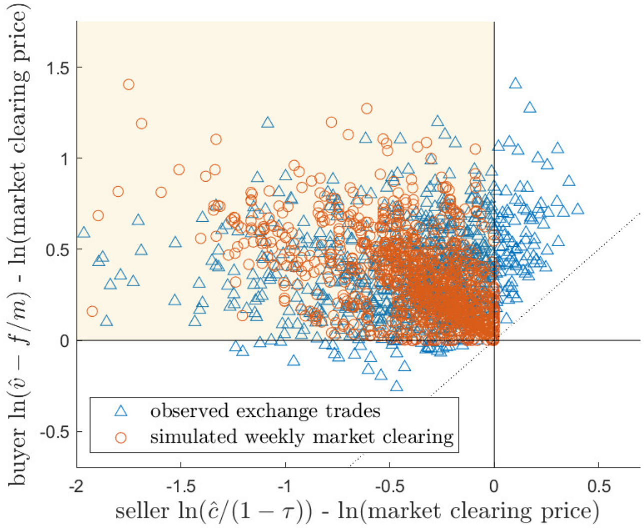
Armed with market-facing valuations for all trader arrivals, I construct the aggregate supply and demand curves for each calendar week. As each trader submits a single price for their volume, the aggregate supply and demand curves are not smooth but involve steps. If the curves cross at a single point, the market-clearing price and quantity are immediately determined by that point. It is also possible that the curves cross along a vertical or horizontal line due to their steps. In this case, the market-clearing quantity (or total trade volume) is defined such that every unit for which the market-facing value of the buyer is greater than or equal to the market-facing cost of the seller is traded. If the supply and demand curves meet along a vertical line or do not cross at all, then the market-clearing price is defined to be the midpoint of the buyer's and seller's market-facing valuations for the marginal traded unit. Finally, I note that clearing the market in this way may result in partial execution of the volume of one marginal trader.

Counterfactual simulation results

Before discussing counterfactual simulation outcomes in the aggregate, it is informative to examine the simulated benchmark at the trade level, comparing counterfactual trades to observed trades. Figure 4 provides a visual comparison at the trade level. The triangles mark trades in the observed limit order market while the circles mark trades in the counterfactual uniform-price market clearing benchmark. The x-axis represents the log difference between the seller's tax-adjusted cost $\hat{c}/(1-\tau)$ and the price that would have cleared the market the week the trade occurred. The y-axis represents the log difference between the buyer's value $\hat{v} - f/m$ and the price that would have cleared the market the week the trade occurred. The dotted line is a 45-degree line; all trades should lie above this line because trade is possible only if $\hat{v} - f/m \geq \hat{c}/(1-\tau)$.

In the counterfactual benchmark, there is a single market-clearing price p each week, and only those buyers with $v - f/m \geq p$ and sellers with $c/(1-\tau) \leq p$ can qualify for trade. Thus, all the circles in Figure 4 lie in the shaded upper-left quadrant. Meanwhile, the triangles that fall outside the shaded quadrant illustrate

Figure 4: Comparison of counterfactual versus observed trades, $\eta = 0.61$



that in the observed limit order market, relatively low-value buyers and high-cost sellers are able to trade by exploiting the time and price dispersion available there. These trades are still individually Pareto-improving, as they are above the 45-degree line, but they may happen at the cost of more efficient potential trades involving higher-value buyers or lower-cost sellers. The figure illustrates that the *sorting* of which buyers and sellers get to trade is different in the market-clearing benchmark versus in the observed exchange.

Table 4 reports how aggregate outcomes of the counterfactual benchmark compare to those of the observed limit order market. The table reports outcomes given two different risk aversion levels, $\eta = 0.99$ and $\eta = 0.61$. To compare counterfactual versus observed outcomes during the same time period and under the same risk aversion level, it displays each respective $\frac{\text{counterfactual outcome}-\text{observed outcome}}{\text{observed outcome}} \times 100\%$. The primary outcome of interest is total trade surplus, defined as the sum of buyer's profit $\hat{v} - f/m - p$ and seller's profit $(1-\tau)p - \hat{c}$ for all traded units of water. In all reported categories, trade surplus is higher in the market-clearing benchmark than in the observed limit order market. How much higher it would be varies with the extent of risk aversion and the time period. Consider first the outcomes given $\eta = 0.99$. Recall that there would be very little price shading on the observed exchange given this level of risk aversion, so price shading is effectively excluded as a possible source of discrepancy with the benchmark. Setting aside price shading, Table 4 reports that

Table 4: Percent change of counterfactual outcomes from observed exchange

time period	storage residual (ML, mean)	total surplus	surplus per unit traded	total trade volume
			benchmark $\eta = 0.99$	
full sample	−591,242	+18%	+13%	+4%
Apr. 2020–Sep. 2020	−1,483,614	+26%	+38%	−9%
Oct. 2020–Mar. 2021	−799,056	+9%	+8%	+1%
Apr. 2021–Sep. 2021	507,807	+23%	+1%	+22%
			$\eta = 0.61$	
full sample	−591,242	+58%	−20%	+97%
Apr. 2020–Sep. 2020	−1,483,614	+40%	+9%	+28%
Oct. 2020–Mar. 2021	−799,056	+45%	−12%	+65%
Apr. 2021–Sep. 2021	507,807	+140%	−23%	+212%
			dominant strategy double auction $\eta = 0.61$	
full sample	−591,242	+52%	−16%	+80%
Apr. 2020–Sep. 2020	−1,483,614	+37%	+13%	+21%
Oct. 2020–Mar. 2021	−799,056	+39%	−9%	+53%
Apr. 2021–Sep. 2021	507,807	+129%	−18%	+179%

Notes: Table reports the percent change in simulated counterfactual outcomes from corresponding outcomes of the observed limit order market for the same time period and the same risk aversion level.

trade surplus would be 18% higher in the benchmark over the full sample period. Meanwhile, if price shading does occur on the observed limit order market, there would be additional efficiency losses coming from that. Given the incentives for price shading produced by the derivative of the execution probability with respect to price, Section 5.2 suggests a considerable amount of price shading would take place at lower risk aversion levels. Indeed, Table 4 reports that the surplus gains of the benchmark would be much larger given $\eta = 0.61$, exceeding 50% over the full sample period.

6.2 A dominant strategy double auction

To supplement the benchmark, I also simulate a closely-related incentive compatible mechanism by adapting the dominant strategy double auction proposed by McAfee (1992). In this double auction, prices are determined by the bids of the buyer and seller who marginally disqualify for trade. Thus, a trader's bid does not affect her price conditional on trade but only affects whether she trades. By the usual Vickrey argument (as applied in, e.g., a second-price auction), bidding one's true valuation is a dominant strategy in this auction. However, relying on excluded bids to determine prices leads to, in some situations, marginal trades being dropped, so incentive compatibility comes at the cost of a surplus loss relative to the benchmark. Also, the mechanism sometimes needs to pay sellers a lower price than it collects from buyers and makes money as a result.

The McAfee (1992) auction is for traders demanding or supplying one unit each. In my multi-unit setting, there is an additional incentive to bid untruthfully arising from the possibility of partial execution. If the market-clearing quantity results in partial execution of the marginal qualifying bid, then there are potential gains to be had from lying to move up in rank among trade-qualified bids. To restore incentive compatibility, I remove the link between bid rank and portion executed by executing an equal fraction of all trade-qualifying bids on the same side of trade. For interested readers, the next paragraph specifies further details of the multi-unit dominant strategy double auction that I implement. Aside from the mechanism, all other implementation details remain the same as in Section 6.1.

In each calendar week, buyers report their per-unit value b, and sellers report their per-unit cost s. The mechanism will charge buyers a price p_b and pay sellers p_s per traded unit, to be determined below. Let demand units be ranked, highest per-unit value first, and let supply units also be ranked, lowest per-unit cost first. In auction version (A), let buyer k be the worst-ranked buyer for whom all demanded units have a value greater than or equal to the costs of correspondingly ranked supply units. Let seller $j(k)$ be the counterpart seller for k's last unit. If next-ranked buyer $k+1$ and seller $j(k)+1$ exist, define $p_0 = \frac{1}{2}(b_{k+1} + s_{j(k)+1})$. Then,

$$\begin{cases} \text{if } b_{k+1} \leq s_{j(k)+1} \text{ and } p_0 \in [s_{j(k)}, b_k], & p_b = p_s = p_0; \\ \text{otherwise, if } b_{k+1} \geq s_{j(k)+1}, & p_b = b_{k+1},\ p_s = s_{j(k)+1}. \end{cases}$$

In either of the above scenarios, buyers 1 through k will buy their demanded units in full, and sellers 1 through $j(k)$ will sell, with any partial execution applied as an equal fraction across them. If neither of the above scenarios applies, I move up the ordered buyer list in sequence until a buyer $k-i$ is found such that $b_{k-i+1} \geq s_{j(k-i)+1}$. Then $p_b = b_{k-i+1}$, $p_s = s_{j(k-i)+1}$, and buyers 1 through $k-i$ will buy their demanded units in full. If this condition is not met by any i, no trade occurs. In auction version (B), I reverse the role of the buyer and seller above. Namely, let seller k be the worst-ranked seller for whom all supplied units have a cost lower than or equal to the values of correspondingly ranked demand units, let buyer $j(k)$ be the counterpart buyer for k's last unit, and implement rules analogous to those of auction version (A). For each calendar week, I then adopt the version that yields higher trade volume.

The bottom of Table 4 displays the results for $\eta = 0.61$. For $\eta = 0.99$, traders already choose $p \approx v$ due to their risk aversion; a dominant strategy auction is not separately needed to induce truthtelling. Comparing with corresponding results from the benchmark, we see that the percent changes of the dominant strategy auction are of similar but reduced magnitude. For example, given $\eta = 0.61$, the total trade surplus increase over the full sample period would be 58% in the benchmark versus 52% in the dominant strategy auction.

One caveat in interpreting these results is that this comparison of counterfactuals versus the observed exchange holds trader entry constant at the observed level. If a new mechanism increases trade surplus, it may in practice attract more participants, increasing market thickness and generating even larger surplus gains. Here I provide computations for a conservative baseline maintaining the observed level of market thickness.

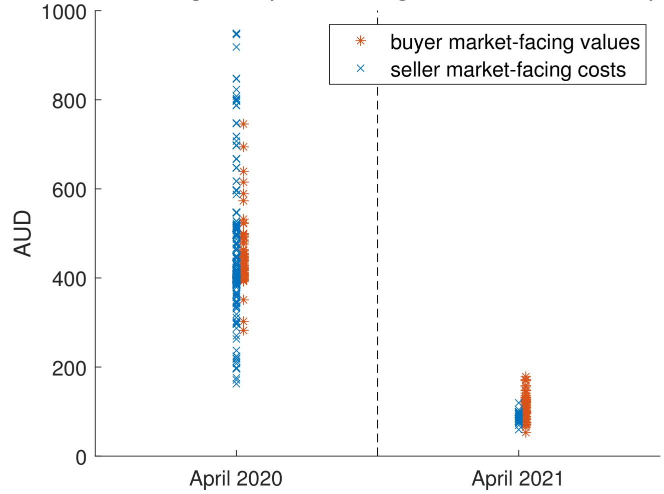
6.3 Interaction with drought

Table 4 also divides the data sample into three six-month periods and reports period-specific outcomes. As water availability steadily improved during my data sample, this allows me to examine how exchange design might interact with drought levels. As a measure of water availability, the table displays the mean of seasonally adjusted water storage for each period, which is the residual or deviation of water stored in the major storages of the Victorian Murray River from predicted levels given the day of the year; see Section 5.1. A negative storage residual indicates there is less water than the historical average; a positive storage residual indicates there is more.

How does exchange design interact with drought? Decomposing total surplus into "surplus per united traded" and "total trade volume" helps us understand this better. First, relative surplus per unit traded decreases with water availability at both risk aversion levels reported in Table 4. In other words, periods of drought are when the counterfactual design would be especially effective at increasing trade surplus per unit. This is because value heterogeneity is especially high during droughts, so there are higher per-unit gains to be had from sorting the highest value buyers and lowest cost sellers into trade. To illustrate this point, Figure 5 visualizes the impact of drought on value heterogeneity by plotting buyers' market-facing values and sellers' market-facing costs at the order level given $\eta = 0.99$, including for orders that did not execute, at two points in time. The first is April 2020, the month in my data when water availability was lowest, and the second is one year later in April 2021, when water was plentiful in comparison. Indeed, the support of traders' values is much wider in April 2020. With these same traders in April 2020, the benchmark would have arranged counterfactual trades with an average seller opportunity cost $\hat{c}$ of about 218 AUD/ML versus 291 AUD/ML for trades seen on the observed exchange, substantially reducing the average opportunity cost of traded water during drought.

Meanwhile, per the "total trade volume" column in Table 4, the benchmark's ability to expand trade volume increases with water availability; it seems that when water is abundant, a more efficient mechanism can expand trade volume proportionally more than it can during drought. To summarize, Table 4 suggests that a more efficient trade mechanism would lead to better trades during droughts and more trades during times of abundance.

Figure 5: Value heterogeneity in drought versus in plenty, $\eta = 0.99$



6.4 A decomposition

How much do bilateral frictions, price shading, and temporal dispersion each contribute to the gap between the observed exchange and the counterfactual benchmark? I conduct two additional simulations to isolate the effect of each feature. First, I simulate a counterfactual scenario that maintains the same limit order prices (price shading) and continuous-time nature (temporal dispersion) as the observed exchange, but eliminates frictions arising from bilateral volume mismatch. Comparing this new counterfactual to the observed exchange isolates the contribution of bilateral frictions. The way I eliminate bilateral frictions is by implementing 1 ML unit splits of the sellers' volume. This allows buyers to fill their order volume with the cheapest units available on the exchange, regardless of how many different sellers these units come from and without having to buy the entirety of any seller's volume. To respect potential lumpiness in buyers' demand and to understate rather than overstate the counterfactual gains, I let buyers buy only if they can fill their entire order volume with units that are cheaper than their value. If the supply of such units is less than a buyer's order volume, she does not buy at all. Buyers pay a flat fee f per trade as in the benchmark; a purchase resulting from a single buy order counts as a single trade for purposes of the fee, regardless of the number of sellers from which the purchase is sourced. Sellers pay taxes and fees of rate τ on sales revenue as before. Noting that over 95% of observed trades are generated by buyers fulfilling seller limit orders, I simplify the simulation by having sellers submit limit orders only and buyers submit

Table 5: Percent change of counterfactual total surplus from observed exchange

		$\eta = 0.99$	$\eta = 0.61$
(a)	observed limit order market	0%	0%
Counterfactual simulations:			
(b)	limit order market, 1ML splits, price as observed	+11%	+27%
(c)	limit order market, 1ML splits, no price shading	–	+52%
(d)	benchmark, batch uniform-price market clearing	+18%	+58%

Notes: Table reports the percent change of the total surplus of the simulated counterfactual scenarios from that of the observed exchange.

market orders only.

Next, I simulate a counterfactual exchange that additionally eliminates price shading. Here, sellers' limit order prices are set equal to their market-facing costs, while other details of the simulation are the same as in the previous paragraph. Differences between this simulation and that of the previous paragraph isolate the contribution of price shading. Finally, differences between this simulation and the benchmark isolate the contribution of temporal dispersion and the sorting of buyers and sellers.

Table 5 reports how the simulated total trade surplus generated by these counterfactual scenarios differs from that of the observed exchange over the full sample period. The difference between rows (a) and (b) isolates the contribution of bilateral frictions, i.e., the effect of moving from bilateral to multilateral exchange. Interestingly, the 1 ML splits in (b) close about half of the surplus gap between the observed exchange and the benchmark, (d)-(a). This suggests that ameliorating bilateral frictions could go a long way towards making trade more efficient while preserving a continuous-time market. Next, the difference between rows (b) and (c) isolates the contribution of price shading. If $\eta = 0.99$, there is effectively no price shading, so there is no difference between (b) and (c). But if $\eta = 0.61$, there is substantial price shading, and eliminating that price shading expands trade volume substantially, closing most of the surplus gap with the benchmark. This is consistent with Table 4, where volume expansion is a key driver of the benchmark's surplus gains for $\eta = 0.61$ but less so for $\eta = 0.99$. Lastly, the difference between rows (c) and (d) isolates the contribution of temporal consolidation and sorting the highest value buyers and lowest cost sellers into trade. Table 5 suggests this element is relatively more important if there is less price shading, such as when $\eta = 0.99$; for a given set of observed limit order prices, less price shading implies higher seller costs and lower buyer values, and efficient buyer-seller sorting matters more in such an environment. This also is consistent with Table 4, where the benchmark's gains in surplus-per-unit-traded are more prominent given $\eta = 0.99$ than they are given $\eta = 0.61$.

A practical takeaway from this section, robust across low and high η, is that encouraging multilateral trade through free volume splits would meaningfully improve the efficiency of this market. Compared to a complete overhaul of the market mechanism, this incremental reform could be more readily absorbed into the existing market system. As implemented in my counterfactual simulations, such a push would need to be accompanied by a change in state trade approval fee policy so that the fee paid no longer depends on the number of counterparties through which an order is fulfilled.

7 Conclusion

In this paper, I assess the performance of a limit order market for water through an empirical analysis of its buy and sell orders and trades. This analysis intersects the literatures on water markets, industrial organization, and finance. Combining order-level exchange data with a model of participants' best-response order type and order price, I identify the latent values underlying the orders. To provide market assessments that are robust to realistic behavior, I extend the model to account for order revision, a range of risk aversion levels, and default choices. I use the latent values thus identified to counterfactually simulate a performance benchmark and to decompose the gap between the observed market and the benchmark. In addition, wide variation in water availability during my sample period combined with the analysis above allows me to see how different aspects of market performance interact with drought. I find that moving from a system built around bilateral trade towards more multilateral exchange, where, e.g., a single buy order could be sourced from more than one seller and vice versa without undue expense, is a policy that could move the needle on market efficiency.

There is much more work to be done on water markets, and economists' contributions to this work will continue to grow. For example, this paper focuses on intra-zone trades, or trades within one trading zone, which constitute the vast majority of allocation trades in Australia. But the economics of *inter*-zone trade is also interesting and could be especially important during times of geographical water imbalance. The small share of inter-zone trades points to the greater trade frictions involved, and one avenue for future research is how to improve the design of inter-zone markets while respecting the hydrological constraints of crossing zones. Also, this paper studies water exchange in a region where the history, hydrology, and institutions were conducive to an early emergence of formal water markets. An important, ongoing question is how to design markets in regions like the American West where water rights laws and hydrological conditions differ and have so far been less conducive to flourishing trade.

References

2030 Water Resources Group (2009): "Charting Our Water Future: Economic Frameworks to Inform Decisionmaking," Tech. rep., 2030 Water Resources Group.

Australian Competition and Consumer Commission (2021): "Murray-Darling Basin Water Markets Inquiry-Final Report," Tech. rep., Australian Competition and Consumer Commission, Canberra, Australia.

Backus, M. and G. Lewis (2020): "Dynamic Demand Estimation in Auction Markets," https://mbackus.github.io/research.html.

Bajari, P. and A. Hortaçsu (2005): "Are Structural Estimates of Auction Models Reasonable? Evidence from Experimental Data," *Journal of Political Economy*, 113, 703–741.

Bodoh-Creed, A. L., J. Boehnke, and B. Hickman (2021): "How Efficient are Decentralized Auction Platforms?" *The Review of Economic Studies*, 88, 91–125.

Bruno, E. M. and K. Jessoe (2021): "Missing markets: Evidence on Agricultural Groundwater Demand from Volumetric Pricing," *Journal of Public Economics*, 196, 104374.

Budish, E., P. Cramton, and J. Shim (2015): "The High-Frequency Trading Arms Race: Frequent Batch Auctions as a Market Design Response," *The Quarterly Journal of Economics*, 130, 1547–1621.

Bureau of Meteorology (2021): "Australian Water Markets Report 2019–20," Tech. rep., Commonwealth of Australia, Melbourne, Australia.

Chong, H. and D. Sunding (2006): "Water Markets and Trading," *Annual Review of Environment and Resources*, 31, 239–264.

Cohen, K. J., S. F. Maier, R. A. Schwartz, and D. K. Whitcomb (1981): "Transaction Costs, Order Placement Strategy, and Existence of the Bid-Ask Spread," *Journal of Political Economy*, 89, 287–305.

Debaere, P., B. D. Richter, K. F. Davis, M. S. Duvall, J. A. Gephart, C. E. O'Bannon, C. Pelnik, E. M. Powell, and T. W. Smith (2014): "Water Markets as a Response to Scarcity," *Water Policy*, 16, 625–649.

Donna, J. D. and J.-A. Espín-Sánchez (2018): "Complements and Substitutes in Sequential Auctions: The Case of Water Auctions," *The RAND Journal of Economics*, 49, 87–127.

——— (2023): "The Illiquidity of Water Markets," Available at SSRN: https://ssrn.com/abstract=2667654.

Economides, N. and R. Schwartz (1995): "Electronic Call Market Trading: Let Competition Increase Efficiency," *Journal of Portfolio Management*, 21(3), 10–18.

Farmer, J. D. and S. Skouras (2012): "Review of the benefits of a continuous market vs. randomised stop auctions and of alternative Priority Rules (policy options 7 and 12)," Tech. rep., Foresight, UK Government Office for Science, London, UK.

Goesch, T., M. A. Donoghoe, and N. Hughes (2019): "Snapshot of Australian Water Markets," Tech. rep., Australian Bureau of Agricultural and Resource Economics and Sciences, Canberra, Australia.

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-price Auctions," *Econometrica*, 68, 525–574.

——— (2009): "Nonparametric Identification of Risk Aversion in First-Price Auctions under Exclusion Restrictions," *Econometrica*, 77, 1193–1227.

Hagerty, N. (2022): "Liquid Constrained in California: Estimating the Potential Gains from Water Markets," https://www.nickhagerty.com/research.

Handa, P. and R. A. Schwartz (1996): "How Best to Supply Liquidity to a Securities Market," *Journal of Portfolio Management*, 22(2), 44–51.

Hendricks, K., A. Sorensen, and T. Wiseman (2021): "Dynamics and Efficiency in Decentralized Online Auction Markets," https://www.ssc.wisc.edu/ sorensen/.

Hollifield, B., R. A. Miller, and P. Sandås (2004): "Empirical Analysis of Limit Order Markets," *The Review of Economic Studies*, 71, 1027–1063.

Hollifield, B., R. A. Miller, P. Sandås, and J. Slive (2006): "Estimating the Gains from Trade in Limit-Order Markets," *The Journal of Finance*, 61, 2753–2804.

Hortaçsu, A. and S. L. Puller (2008): "Understanding Strategic Bidding in Multi-Unit Auctions: A Case Study of the Texas Electricity Spot Market," *The RAND Journal of Economics*, 39, 86–114.

Kaiser, R. A. and M. McFarland (1997): "A Bibliographic Pathfinder on Water Marketing," *Natural Resources Journal*, 37, 881.

Kong, Y. (2020): "Not Knowing the Competition: Evidence and Implications for Auction Design," *The RAND Journal of Economics*, 51, 840–867.

——— (2021): "Sequential Auctions with Synergy and Affiliation across Auctions," *Journal of Political Economy*, 129, 148–181.

Lu, J. and I. Perrigne (2008): "Estimating Risk Aversion from Ascending and Sealed-Bid Auctions: The Case of Timber Auction Data," *Journal of Applied Econometrics*, 23, 871–896.

Madhavan, A. (1992): "Trading Mechanisms in Securities Markets," *the Journal of Finance*, 47, 607–641.

McAfee, R. P. (1992): "A Dominant Strategy Double Auction," *Journal of Economic Theory*, 56, 434–450.

National Water Commission (2011): "Water Markets in Australia: A Short History," Tech. rep., National Water Commission, Canberra, Australia.

Parlour, C. A. and D. J. Seppi (2008): "Limit Order Markets: A Survey," *Handbook of Financial Intermediation and Banking*, 5, 63–95.

Perrigne, I. and Q. Vuong (2021): "Econometric Methods for Auctions," vol. 7 of *Handbook of Econometrics*, forthcoming.

Productivity Commission (2003): "Water Rights Arrangements in Australia and Overseas: Annex C, Victoria," Tech. rep., Productivity Commission, Melbourne, Australia.

Rafey, W. (2023): "Droughts, Deluges, and (River) Diversions: Valuing Market-Based Water Reallocation," *American Economic Review*, 113, 430–471.

Reguant, M. (2014): "Complementary Bidding Mechanisms and Startup Costs in Electricity Markets," *The Review of Economic Studies*, 81, 1708–1742.

Roşu, I. (2009): "A Dynamic Model of the Limit Order Book," *The Review of Financial Studies*, 22, 4601–4641.

Rustichini, A., M. A. Satterthwaite, and S. R. Williams (1994): "Convergence to Efficiency in a Simple Market with Incomplete Information," *Econometrica*, 1041–1063.

Satterthwaite, M. A. and S. R. Williams (2002): "The Optimality of a Simple Market Mechanism," *Econometrica*, 70, 1841–1863.

Vasserman, S. and M. Watt (2021): "Risk Aversion and Auction Design: Theoretical and Empirical Evidence," *International Journal of Industrial Organization*, 79, 102758.

Wah, E. and M. P. Wellman (2013): "Latency Arbitrage, Market Fragmentation, and Efficiency: A Two-Market Model," in *Proceedings of the Fourteenth ACM Conference on Electronic Commerce*, 855–872.

Wolak, F. A. (2000): "An Empirical Analysis of the Impact of Hedge Contracts on Bidding Behavior in a Competitive Electricity Market," *International Economic Journal*, 14, 1–39.

A Appendix

A.1 Proofs

Assumption 1S. *The execution probability $P_s(\cdot|x)$ for a sell limit order satisfies*

(i) $P_s(p|x)$ is strictly decreasing in p.

(ii) $P_s(p|x)/P_s'(p|x)$ is strictly increasing in p.

(iii) $P_s(p|x) < 1$.

Assumption 2S. *Distribution $F_s^r(\cdot|c_i)$ of $c_{r,i}$ is weakly stochastically ordered in c_i.*

Assumption 3S. *Regarding the expected profit $C_s(c, p|x)$ of revising a limit order,*

(i) The seller approximates that $C_s(c, p|x) = C_s(c|x)$.

(ii) $r|C_s(c'|x) - C_s(c|x)| < |c' - c|$ for all $c' \neq c$.

Proof of Corrollary 1

Proof. By the envelope theorem, the derivative of $P_b(p_r^*(v_{r,i}|x_{r,i})|x_{r,i})(v_{r,i} - p_r^*(v_{r,i}|x_{r,i}))$ with respect to $v_{r,i}$ is equal to $P_b(p_r^*(v_{r,i}|x_{r,i})|x_{r,i}) \geq 0$, so this expression is weakly increasing in $v_{r,i}$. Then, by the weak stochastic ordering of $F_b^r(v_{r,i}|v_i)$ given by Assumption 2 and the definition of $C_b(v|x)$ in equation (3), $C_b(v'|x) \geq C_b(v|x)$ for any $v' > v$. □

Proof of Proposition 1

Proof. The proof proceeds by showing that $0 < \frac{\partial L(v)}{\partial v} < \frac{\partial M(v,\ell)}{\partial v}$ for all $v > \ell$. If this is true, there will be at most one value $\tilde{v}(\ell) > \ell$ at which $L(v) = M(v, \ell)$, with $L(v) > M(v, \ell)$ for $v \in (\ell, \tilde{v}(\ell))$ and $L(v) < M(v, \ell)$ for $v \in (\tilde{v}(\ell), \infty)$. In the remainder of this proof, we can focus our attention on $v \in (\ell, p^{*-1}(\ell))$. This is because, first, buyers with $v \leq \ell$ will place a limit order because $M(v, \ell) \leq 0 \leq L(v)$ for $v \leq \ell$. Second, for v such that $p^*(v) \geq \ell$, buyers will place a market order because it guarantees execution at a weakly better price than the optimal limit order.

By the envelope theorem, $\frac{\partial L(v)}{\partial v} = P_b(p^*(v)) + (1 - P_b(p^*(v)))rC_b'(v)$. Given Assumptions 1 and 3, $P_b(p^*(v)) < 1$ for $p^*(v) < \ell$ and $rC_b'(v) < 1$. Meanwhile, $P_b(p^*(v)) \geq 0$ and $r \geq 0$ by definition of a probability and $C_b'(v) \geq 0$ by Corrollary 1. Thus, $0 \leq \frac{\partial L(v)}{\partial v} < 1 = \frac{\partial M(v,\ell)}{\partial v}$. This completes the proof. □

Proof that Assumption 1S-(ii) is satisfied if $P_s(\cdot|x)$ has a probit shape with $P_s(p|x) = \Phi(\beta_0 + \beta_1 p + \beta_2 x)$, where $\Phi(\cdot)$ is the standard normal cdf and $\beta_1 < 0$ per Assumption 1S-(i).

Proof. If $P_s(p|x) = \Phi(\beta_0 + \beta_1 p + \beta_2 x)$, then $P_s(p|x)/P_s'(p|x) = \frac{\Phi(\beta_0+\beta_1 p+\beta_2 x)}{\phi(\beta_0+\beta_1 p+\beta_2 x)\beta_1} \equiv k(p)$. Defining $j(\cdot) \equiv \frac{\Phi(\cdot)}{\phi(\cdot)}$, $k(p) = \frac{1}{\beta_1} j(\beta_0 + \beta_1 p + \beta_2 x)$ and $k'(p) = j'(\beta_0 + \beta_1 p + \beta_2 x)$. Since it is known that $\frac{\Phi(\cdot)}{\phi(\cdot)}$ is a strictly increasing function, $j'(\cdot) > 0$. Thus, $k'(p) > 0$. $\square$

Proof of Corrollary 1S

Proof. By the envelope theorem, the derivative of $P_s(p_r^*(c_{r,i}|x_{r,i})|x_{r,i})(p_r^*(c_{r,i}|x_{r,i}) - c_{r,i})$ with respect to $c_{r,i}$ is equal to $-P_s(p_r^*(c_{r,i}|x_{r,i})|x_{r,i}) \leq 0$, so this expression is weakly decreasing in $c_{r,i}$. Then, by the weak stochastic ordering of $F_s^r(c_{r,i}|c_i)$ given by Assumption 2S and the definition of $C_s(c|x)$ in Section 3.4, $C_s(c'|x) \leq C_s(c|x)$ for any $c' > c$. $\square$

Proof of Proposition 2

Proof. The only part of the identification argument not fully explained in the main text is the nonparametric identification of $F_b(\cdot)$ and $F_s(\cdot)$ given $\alpha_b > 0$ and $\alpha_s > 0$. The proof of that step follows. First, consider the buyers. Let the value threshold $\tilde{v}$, determining whether a rational buyer chooses a limit order or a market order, be fixed at an arbitrary value t. Buyers with values $v \geq \tilde{v} = t$ will either rationally submit a market order or submit a limit order by default. Thus,

$$1 - F_b(t) = \frac{\#\{\text{buy market orders}|\tilde{v} = t\} + \#\{\text{buy limit orders with } v \geq t|\tilde{v} = t\}}{\#\{\text{all (market and limit) buy orders}|\tilde{v} = t\}}.$$

Having identified $F_b(t)$ using this ratio of observables, $F_b(\cdot|v \geq t) \equiv \frac{F_b(\cdot)-F_b(t)}{1-F_b(t)}$ is nonparametrically identified as the empirical cdf of all the values associated with $\{\text{buy limit orders with } v \geq t|\tilde{v} = t\}$—this is the set of limit orders submitted by buyers that (mistakenly) default to limit orders, allowing us to identify the shape of $F_b(\cdot)$ for values above t. Also, $F_b(\cdot|v < t) \equiv \frac{F_b(\cdot)}{F_b(t)}$ on $[0, t)$ is nonparametrically identified as the empirical cdf of all the values associated with $\{\text{buy limit orders with } v < t|\tilde{v} = t\}$. Since $F_b(t)$ is identified above, identification of $F_b(\cdot|v \geq t)$ and $F_b(\cdot|v < t)$ implies identification of $F_b(\cdot)$ over its support.

Next, consider the analogous argument for sellers. Let the cost threshold $\tilde{c}$, determining whether a rational seller chooses a limit order or a market order, be fixed at an arbitrary t. Sellers with costs $c \leq \tilde{c} = t$ will either rationally submit a market order or submit a limit order by default. Thus,

$$F_s(t) = \frac{\#\{\text{sell market orders}|\tilde{c} = t\} + \#\{\text{sell limit orders with } c \leq t|\tilde{c} = t\}}{\#\{\text{all (market and limit) sell orders}|\tilde{c} = t\}}.$$

Having identified $F_s(t)$ using this ratio of observables, $F_s(\cdot|c > t) \equiv \frac{F_s(\cdot)-F_s(t)}{1-F_s(t)}$ is nonparametrically identified as the empirical cdf of all the values associated with $\{\text{sell limit orders with } c > t|\tilde{c} = t\}$. Also, $F_s(\cdot|c \leq t) \equiv \frac{F_s(\cdot)}{F_s(t)}$ on $[0, t)$ is nonparametrically identified as the empirical cdf of all the values associated with $\{\text{sell limit orders with } c \leq t|\tilde{c} = t\}$—this is the set of limit orders submitted by sellers that (mistakenly) default to limit orders, allowing us to identify the shape of $F_s(\cdot)$ for costs below t. Since $F_s(t)$ is identified above, identification of $F_s(\cdot|c > t)$ and $F_s(\cdot|c \leq t)$ implies identification of $F_s(\cdot)$ over its support. $\square$

A.2 Supplementary tables, figures, and descriptions

Table A1: Regression of log sell limit order price on covariates

	(1)	(2)	(3)
daily rainfall (mm)	-0.001		
	(0.003)		
daily evapotranspiration (mm)	-0.006		
	(0.003)		
weekly mean rainfall (mm)		0.012	
		(0.005)	
weekly mean evapotranspiration (mm)		-0.007	
		(0.005)	
monthly mean rainfall (mm)			0.041
			(0.010)
monthly mean evapotranspiration (mm)			-0.015
			(0.006)
fraction entitlement volume allocated to date	-0.008	0.000	0.117
	(0.050)	(0.069)	(0.077)
daily storage (log), seasonally adjusted	-0.587	-0.617	-0.699
	(0.064)	(0.068)	(0.072)
waterday/366	-0.162	-0.175	-0.261
	(0.041)	(0.054)	(0.063)
volume (log)	0.003	0.003	0.004
	(0.003)	(0.003)	(0.003)
mean trade price, prev week (log)	0.569	0.529	0.489
	(0.029)	(0.032)	(0.033)
st dev of trade price, prev week	0.003	0.003	0.003
	(0.001)	(0.001)	(0.001)
total trade volume, prev week (1000 ML)	-0.009	-0.008	-0.007
	(0.002)	(0.003)	(0.003)
ask depth (1000 ML)	-0.007	-0.002	-0.001
	(0.005)	(0.005)	(0.005)
bid depth (1000 ML)	0.044	0.043	0.036
	(0.008)	(0.008)	(0.008)
low ask (log)	0.268	0.297	0.327
	(0.025)	(0.027)	(0.027)
bid-ask spread (100 AUD)	0.012	0.013	0.013
	(0.003)	(0.003)	(0.003)
constant	0.932	0.975	1.002
	(0.128)	(0.133)	(0.131)
Observations	2182	2182	2182
R^2	0.892	0.892	0.893
Adjusted R^2	0.891	0.891	0.892

Standard errors in parentheses

Table A2: Estimated parameters of $\hat{P}_s(p|x)$ and $\hat{\tilde{P}}_b(p|x)$

	sell	buy
limit order price, normalized		
polynomial term 1	-0.091	0.087
polynomial term 2	1.3E-04	-0.005
polynomial term 3	9.7E-05	0.001
polynomial term 4	-4.2E-06	-3.0E-05
daily storage (log), seasonally adjusted	-2.503	-0.902
volume (log)	-0.137	0.009
mean trade price, prev week	-0.006	-0.002
st dev of trade price, prev week	-3.6E-04	-0.016
total trade volume, prev week (1000 ML)	-0.060	0.031
ask depth (1000 ML)	0.104	0.039
bid depth (1000 ML)	0.195	0.037
bid-ask spread (100 AUD)	-0.041	-0.052
low ask	0.001	
high bid		0.002
waterday/366		
Bernstein polynomial term 1	1.117	-4.153
Bernstein polynomial term 2	-0.482	2.872
Bernstein polynomial term 3	-0.006	-2.152
Bernstein polynomial term 4	-0.297	-1.022
constant	1.067	-0.257

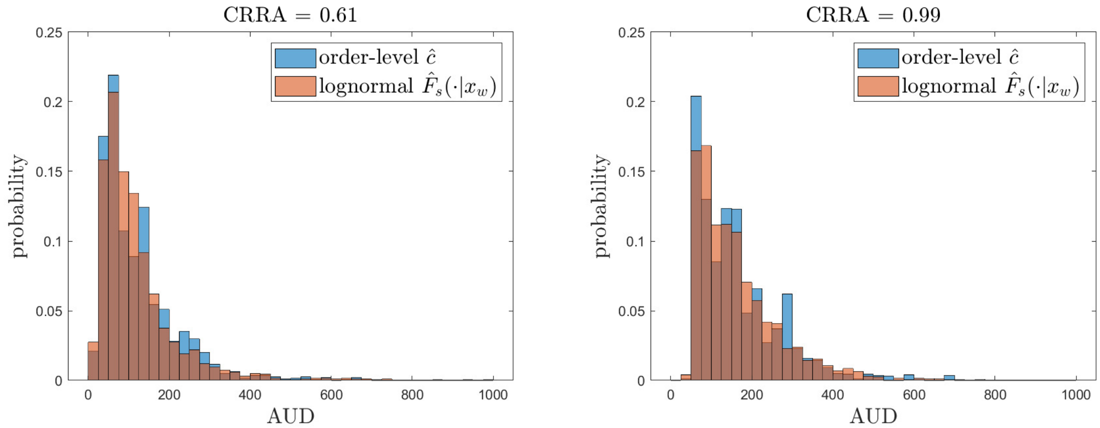
Notes: The execution probability for sell limit orders is estimated according to a probit model, subject to the constraint that the probability is decreasing in price. For buy limit orders, I estimate $\tilde{P}_b(p|x)$ as defined at the end of Section 5.1.

Table A3: Estimated parameters of lognormal $\hat{F}_s(c|x_w)$ and $\hat{F}_b(v|x_w)$

	$\eta = 0.61$				$\eta = 0.99$			
	seller		buyer		seller		buyer	
	$\hat{\theta}_\mu$	$\hat{\theta}_\sigma$	$\hat{\theta}_\mu$	$\hat{\theta}_\sigma$	$\hat{\theta}_\mu$	$\hat{\theta}_\sigma$	$\hat{\theta}_\mu$	$\hat{\theta}_\sigma$
daily storage (log), seasonally adjusted	-2.45	-0.84	-2.38	-0.02	-2.69	-1.20	-2.40	-0.29
waterday/366								
Bernstein polynomial term 1	-0.08	1.95	0.43	-0.03	0.95	1.30	0.43	-1.69
Bernstein polynomial term 2	-0.33	-1.65	-0.50	0.10	-1.25	-1.97	0.02	3.16
Bernstein polynomial term 3	-1.56	2.33	-0.84	0.03	-0.89	1.55	-1.15	-2.04
Bernstein polynomial term 4	-0.62	-0.36	-0.38	-0.22	-0.47	-0.79	-0.20	0.07
volume (log)	0.07	-0.18	-0.03	2.4E-03	0.02	-0.06	-0.05	3.4E-03
constant	4.55	-0.74	5.52	-1.50	4.93	-1.46	5.43	-1.26

Notes: θ_μ and θ_σ are parameters of the distribution Lognormal$(x_w\theta_\mu, x_w\theta_\sigma)$, where $x_w\theta_\mu$ and $x_w\theta_\sigma$ are the mean and standard deviation of $\ln(v)$ for buyers or $\ln(c)$ for sellers, and x_w is a vector of the covariates listed in the left column.

Figure A1: Histogram of order-level $\hat{c}$ versus lognormal $\hat{F}_s(\cdot|x_w)$



Notes: These figures compare the histogram of sellers' estimated order-level costs for initial limit orders, $\hat{c}$, to a histogram of random draws from the estimated lognormal distribution $\hat{F}_s(\cdot|x_w)$. The purpose is to check whether the lognormal specification is a good approximation. Specifically, for each $\hat{c}_i$, one draw is randomly drawn from the lognormal distribution conditional on the covariates $x_{w,i}$ corresponding to that order. The left figure is conditional on CRRA parameter 0.61, while the right figure is conditional on CRRA parameter 0.99. The figures are not expected to be identical because $\hat{F}_s(\cdot|x_w)$ also includes costs of sellers that placed market orders.