# The Illiquidity of Water Markets: Efficient Institutions for Water Allocation in Southeastern Spain

Donna, Javier and Espin-Sanchez, Jose

Ohio State University, Northwestern

16 August 2018

# The Illiquidity of Water Markets: Efficient Institutions for Water Allocation in Southeastern Spain*

Javier D. Donna and José-Antonio Espín-Sánchez

The Ohio State University    Yale University

First version: July, 2013

This version: August 16, 2018

*We are indebted to our advisors and members of our dissertation committees helpful discussions, guidance, and support. Donna: Rob Porter (committee chair), Meghan Busse, Aviv Nevo, and Florian Zettelmeyer. Espín-Sánchez: Joel Mokyr (committee chair), Joseph Ferrie, Regina Grafe, and Rob Porter. We especially thank Tiago Pires and Igal Hendel for their many and helpful suggestions. Discussions with Jason Blevins, Michael Grubb, and Rick Steckel have greatly benefited this work. We thank the participants of the seminars at Yale, NYU-Stern, Berkeley-Haas, IIOC (Chicago, Philadelphia), NBER-DAE Summer Institute, EARIE (Milan), Jornadas de Economia Industrial (Barcelona), Stanford, University of British Columbia, London School of Economics, Society for Economic Dynamics (Toronto), Kent State University, Boston College, the University of Texas at Austin, Dartmouth College, The Econometric Society's North American Summer Meetings (Minnesota), Arizona State University, Advances in Policy Evaluation Conference, NBER-IO Summer Institute, Barcelona GSE Summer Forum Applied Industrial Organization, Banff Empirical Microeconomics Workshop, and The Ohio State University. We also thank Jakub Kastl, Erin Mansur, and Paquale Schiraldi, as well as Fabrizio Zilibotti, Alberto Abadie, and five anonymous referees for valuable comments. We would also like to express our gratitude to Fernanda Donna and Antonio Espín for their help for collecting part of the information used in this project, to Juan Gutiérrez for his help with archival data, and to Kelly Goodman for editorial advice. We thank the AEMET for providing us with the meteorological data. Javier D. Donna acknowledges financial support from the CSIO at Northwestern University, the allocation of computing time from the Ohio Supercomputer Center, and financial support from the SBS at The Ohio State University. José-Antonio Espín-Sánchez also acknowledges financial support from the WCAS Robert Eisner Economics and the CSIO fund at Northwestern University, and Fundación Caja Madrid. All errors are our own.

**Abstract**

We investigate the efficiency of a market institution (an auction) relative to a non-market institution (a quota) as a water allocation mechanism in the presence of frictions, by exploring a particular historical institutional change in Mula, Spain. We estimate a structural dynamic model under the auction accounting for the three main features in the empirical setting: intertemporal substitution, liquidity constraints, and seasonality. We use the estimated model to compute the welfare under auctions, quotas, and the highest-valuation allocation. We find that the institutional change in Mula, from auctions to quotas, was welfare improving for the apricot farmers considered.

**JEL Codes**: D02, D53, L11, L13, G14, Q25.

**Keywords**: Institutions, Financial Markets, Demand, Dynamic Bidding, Market Efficiency.

Javier D. Donna
Department of Economics
The Ohio State University
1945 N High St, 425 Arps Hall
Columbus, OH 43210
Phone: 614-688-0364
Email: donna.1@osu.edu

José-Antonio Espín-Sánchez
Department of Economics
Yale University
27 Hillhouse Ave, Room 38
New Haven, CT 06511-3703
Phone: 203-432-0890
Email: jose-antonio.espin-sanchez@yale.edu

# 1 Introduction

Water allocation is a central concern of policy discussions around the world. Water scarcity is extremely acute in places such as India, Latin America, and the U.S. (Vörösmarty *et al.*, 2010). Seventy percent of fresh water usage worldwide is for irrigation. Water markets are emerging as the preferred institution to allocate irrigation water used by farmers in the developed world, particularly in dry regions of the U.S. and Australia (Grafton *et al.*, 2011). In the absence of *frictions*, water markets are efficient because they allocate water according to farmers' valuations. When frictions are present, however, markets may not be efficient. Consider, for example, the friction that arises when some farmers do not have enough cash to pay for water in the market (*i.e.* some farmers are liquidity constrained). A market allocates water to the farmer who has the highest valuation, and is not liquidity constrained. A market failure occurs if some of the farmers who are liquidity constrained have higher valuations than farmers who are not liquidity constrained. In this case, a non-market institution may allocate water more efficiently than a market.

In this paper, we investigate the efficiency of a market institution (an auction) relative to a non-market institution (a quota as described below) as a water allocation mechanism in the presence of frictions. To perform the empirical analysis, we use data from water markets in Mula, Spain. Frictions arose in Mula because some farmers did not have enough cash during the summer to purchase water in the market. The market price of water increased substantially during summer because: (i) the agricultural products cultivated in the region, such as apricots, needed more water during this season of rapid fruit growth, thus increasing demand for water; and (ii) weather seasonalities in southern Spain generated low rainfall during the summer. These conditions made summer the dry or "critical" season.

In the leading article of the first volume of the *American Economic Review,* Coman (1911) pointed out this issue about liquidity constraints during the critical season: "In southern Spain, where this system obtains and water is sold at auction, the water rates mount in a dry season to an all but prohibitive point." During the critical season, only wealthy farmers could afford to buy water.[1] But poor farmers with the same production technology (*i.e.* who grew the same agricultural products) would also benefit from buying water during the critical season. Indeed, we find that poor farmers bought less water during the critical season than wealthy farmers who had the same crop and number of trees.

To perform the efficiency analysis we exploit four unique features of our data. First, for over 700 years from 1244 until 1966, the citizens of Mula used an unregulated market (an auction as described in Donna and Espín-Sánchez, 2018) to allocate water from the river among farmers. This is an unusual scenario because when water markets are used, they are typically regulated markets (Grafton *et al.*, 2011; Libecap, 2011). Changes in regulatory frameworks over time or across geographical markets preclude to infer gains from trade using price differences. Recovering demand in such cases requires strong assumptions about market participants (Libecap, 2011). Second, water in this setting is an intermediate good used to produce crops, the final products. Water demand is determined by the technological constraint imposed by the crop's production function, which in turn determines the seasonal water need of the tree, as we explain below. Thus, the demand for water is independent of the wealth of the farmer, as long as the farmer has enough cash to pay for the water (*i.e.* there are no income effects). We focus on the set of farmers who only grow apricot trees and, thus, have the same production function. Third, some of the farmers in Mula were part of the wealthy elite. We identify the wealthy farmers by merging urban real estate tax records with auction data.[2] We use that the wealthy farmers were not liquidity constrained as argued in Section 2, and the previous feature—that water is an intermediate good—to estimate the transformation rate of the production function that characterizes the demand system for *all* apricot farmers. This approach allows us to use the empirical context of Mula to differentiate liquidity constraints from unobserved heterogeneity, as we discuss in

[1] In our analysis we define a farmer as "wealthy" if the value of urban real estate of the farmer obtained from the urban real estate tax data is above the median among the apricot farmers, and "poor" otherwise.

[2] In Donna and Espín-Sánchez (2017) we use a different criterion to identify wealthy farmers, whether a farmer uses the honorific title *Don*. We show that the behavior of poor and wealthy farmers thus defined is also consistent with the presence of liquidity constraints.

Section 7. Finally, in 1966 the market in Mula was replaced by a non-market institution, a quota.[3] Under the quota system, farmers who owned a plot of fertile land were entitled to a fixed amount of water—proportional to the size of their plot—for irrigation, and paid only a small annual fee for maintenance costs. A natural question arises: How did the institutional change from auctions to quotas affect welfare in the presence of liquidity constraints?

In this paper, we empirically investigate how this institutional change—from auctions to quotas—affected efficiency as a measure of welfare. With output data before and after the institutional change, computing welfare would be straightforward. However, output data is not available. We build a structural econometric model that allows us to compute the output under auctions and quotas. The econometric model uses detailed input data (units of water purchased, rainfall amount, number of apricot trees, *etc.*) along with the apricot's production function, that transforms these inputs into apricots, to compute the counterfactual output before and after the institutional change. In the model, water for irrigation has diminishing returns, and farmers are heterogeneous on two dimensions: their willingness to pay (productivity) and their ability to pay for the water (cash holdings). On the one hand, in the absence of liquidity constraints, markets are efficient. On the other hand, in the absence of heterogeneity in productivity, a fixed quota system is efficient, due to the decreasing marginal returns of water.[4] In our empirical setting, although farmers are *ex-ante* homogeneous in productivity, they are *ex-post* heterogeneous in productivity because they receive an *i.i.d.* productivity shock in each period. In addition, some farmers are liquidity constrained. In this general case, the efficiency of auctions relative to quotas is ambiguous as discussed in Section 6. It is then an empirical question to assess which institution is more efficient. To the best of our knowledge, no empirical study has investigated the efficiency of auctions relative to quotas in the presence of liquidity constrained bidders.

[3]In this paper we do not exploit the institutional change from 1966 as a source of variation due to lack of output data. See Espín-Sánchez (2017) for details about the historical institutional change.

[4]In a static setting, if farmers are sufficiently wealthy, markets are efficient. If farmers are homogeneous, quotas are efficient. If all farmers are homogeneous and sufficiently wealthy, then both markets and quotas are efficient. In a dynamic setting the characterization of efficiency is more complex, and it includes the probability distribution of the evolution of the supply of water and future irrigation.

We begin our empirical analysis by estimating demand for water under the auction system. To estimate demand, we account for three features of the empirical setting. First, irrigation increases the moisture level of the land, thus reducing future demand for water. Irrigation creates an intertemporal substitution effect, where water today is an imperfect substitute for water tomorrow because it evaporates over time. The resulting dynamics in the irrigation demand system are similar to those in the storable goods demand system. Soil's moisture level in the former plays an analogous role to inventory in the latter (*e.g.* Hendel and Nevo, 2006). Second, some farmers are liquidity constrained. Wealthy, unconstrained, farmers strategically delay their purchases, and buy water during the critical season, when agricultural products need water the most. Poor farmers, who may be liquidity constrained, buy water before the critical season in anticipation of an increase in price. Finally, weather seasonality increases water demand during the critical season, when fruit grows more rapidly. Seasonality shifts the whole demand system, conditional on intertemporal substitution and liquidity constraints.[5]

Ignoring the presence of liquidity constraints biases the estimated demand elasticity downwards. To see this, consider the decrease in demand due to an increase in price during the critical season. When farmers are liquidity constrained, the decrease in demand has two components: (1) the decrease in demand due to the price being greater than the valuation of certain farmers; and (2) the decrease in demand due to some farmers being liquidity constrained, even when their valuation is above the prevailing price. If one does not account for the second component, one would attribute this decrease in demand to greater price sensitivity. Thus, one would incorrectly interpret liquidity constraints as more elastic demand, biasing the estimated demand downwards.

In our econometric model the farmer's utility has three components. First, the apri-

[5]To account for the intertemporal substitution effect, we condition water demand on the moisture level of the soil. The moisture level is obtained by applying the findings from the literature in agricultural engineering described in Appendix A.2, and the observed rainfall and irrigation. This procedure allows to construct a moisture variable for each farmer, that is treated as an observable state variable for the estimation (Hendel and Nevo, 2006). The moisture variable measures the amount of water accumulated in each farmer's plot.

cot's production function that transforms water into apricots. Second, the cost of producing apricots, measured as the amount spent on water plus an irrigation cost. Finally, an idiosyncratic productivity shock that is farmer specific. Conditional on the soil's moisture level, the type of agricultural product (*i.e.* apricot), and the number of trees, farmers' productivity is assumed to be homogeneous up to the idiosyncratic shock. This gives us the exclusion restriction to identify the other source of heterogeneity, liquidity constraints. To estimate the econometric model we construct a conditional choice probability estimator (Hotz and Miller, 1993) using only data on the wealthy farmers, where liquidity constraints are not binding.

We use the estimated dynamic demand system to compute welfare under auctions and quotas. We show that: (i) the type of quota closest to the one implemented in Mula increased welfare relative to auctions; and (ii) the welfare under auctions is greater than under quotas with random assignment. These results show the importance of the choice of the type of quota system.

In summary, we make three main contributions: (1) we combine a unique data set, that includes detailed financial information and individual characteristics, with a novel econometric approach to estimate demand in the presence of storability, liquidity constraints, and seasonality; (2) we investigate the efficiency of auctions relative to quotas in the presence of liquidity-constrained bidders by exploring a particular historical institutional change; (3) from an historical perspective, we conclude that the institutional change in Mula was welfare improving for the apricot farmers analyzed in this paper because the quota system more often allocated water units following farmers' valuations than did the market.

## Related Literature

Scholars studying the efficiency of irrigation communities in Spain have proposed two competing hypotheses to explain the coexistence of auctions and quotas. On the one hand, Maass and Anderson (1978) claimed that, absent operational costs, auctions are more efficient than quotas. They argued that both systems nevertheless existed because the less efficient system (quotas) was simpler, and easier to maintain. Hence, once operational costs are take into account, quotas were more efficient than auctions in places with less water scarcity. This hypothesis is supported by observations of auctions in places where water was extremely scarce (Musso y Fontes, 1847; Pérez Picazo and Lemeunier, 1985). On the other hand, Garrido (2011) and González Castaño and Llamas Ruiz (1991) argued that owners of water rights had political power, and were concerned only with their revenues, regardless of the overall efficiency of the system.

The theoretical literature on auctions with liquidity constraints is recent (*e.g.* Che and Gale, 1998). Our model is closest to that of Che et al. (2013). The authors assume that agents can consume at most one unit of a good with linear utility in their type. They conclude that markets are always more efficient than quotas, although some non-market mechanisms can outperform markets when resale is allowed. In our model, we allow agents to consume multiple units with a concave utility function, and we incorporate dynamics (intertemporal substitution). In our setting, there is no strict ranking between markets and quotas, but non-market mechanisms with resale can still outperform both markets and quotas.

Auctions with liquidity constraints can be seen as a particular case of asymmetric auctions. Athey *et al.* (2011) and Krasnokutskaya and Seim (2013) conclude that preferential auctions decrease efficiency if they reallocate from high-bid bidders to low-bid bidders. If some bidders face liquidity constraints, however, giving preferential treatment to those bidders could increase efficiency, similarly to Marion (2007). If bidders with liquidity constraints have higher valuations than unconstrained bidders, this reallocation would increase efficiency. Identifying valuations from liquidity constraints is necessary to estimate efficiency gains in preferential auctions. Ignoring the presence of liquidity constraints in preferential auctions could bias the estimated distribution of valuations downwards. Moreover, if firms face capacity constraints, as in Jofre-Bonet and Pesendorfer (2003), then small firms would be more efficient than large firms when the latter have high capacity contracted. Because small firms are also more likely to face liquidity constraints, the presence of capacity constraints would further increase the bias against small firms. A normative implication is that the government or auctioneer increases efficiency by treating small firms' bids favorably. A positive implication is that treating bids of small firms as unconstrained bids underestimates the productivity of small firms.

Recent macroeconomic research points to the importance of financial constraints, and the dynamics of wealth accumulation in the real economy (Moll, 2014). Imperfect capital markets are important in developing countries (Banerjee and Moll, 2010). Rosenzweig and Wolpin (1993) estimate a structural model of agricultural investments in the presence of credit constraints. Udry (1994) studies how state-contingent loans are used in rural Nigeria to insure against some portion of output's variability. Laffont and Matoussi (1995) show how insufficient working capital affects contract arrangements in rural Tunisia. Jayachandran (2013) demonstrates that the presence of liquidity constraints among land owners in Uganda renders upfront payment in cash more effective than promised future payments. Bubb *et al.* (2016) study rural India, where liquidity constraints in water markets reduce efficiency, as in our case.

We are not aware of any empirical paper analyzing the effect of liquidity constraints in an auction setting. Pires and Salvo (2015) find that low income households buy smaller sized storable products (detergent, toilet paper, *etc.*) than do high income households, even though smaller sized products are more expensive per pound. They attribute this puzzling result to low income households' liquidity constraints.

We estimate a dynamic demand model with seasonality and storability. There is a vast empirical industrial organization literature on dynamic demand (*e.g.* Boizot *et al.,* 2001; Pesendorfer, 2002; Hendel and Nevo, 2006; Gowrisankaran and Rysman, 2012).[6] However, none of this work examines how liquidity constraints affects demand. To the best of our knowledge, this paper is the first to propose and estimate a demand model with storability,

[6] See Aguirregabiria and Nevo (2013) for a recent survey.

seasonality, and liquidity constraints. Timmins (2002) studies dynamic demand for water, and is closest to our paper, although he estimates demand for urban consumption rather than demand for irrigation. Moreover, while Timmins (2002) uses parameters from the engineering literature to estimate the supply of water, we use parameters from the literature in agricultural engineering to determine both the demand structure and soil's moisture levels (see Appendix A.2). To estimate the parameters that characterize demand we use only data from the wealthy farmers (excluding data from the poor farmers who may be liquidity constrained), and then project the inferred preferences from these "trusted choices" onto the welfare of the poor farmers in the counterfactual analysis. This approach of using trusted choices for the welfare analysis is similar to the ones by, *e.g.*, Handel and Kolstad (2015) and Ketcham *et al.* (2016), who use the choices or revealed preferences of informed consumers to identify risk preferences or to proxy for concealed preferences of misinformed consumers, respectively.[7]

# 2 Environment and Data

## 2.1 Environment

Southeastern Spain is the most arid region of Europe. The region is located to the east of a mountain chain, the Prebaetic System. Rivers flowing down the Prebaetic System provide the region with irrigation water. Most years are dryer than the average. There are only a few days of rain, but they are of high intensity.[8]

Weekly prices of water in the auction are volatile. These prices depend on the season of the year, and the amount of rainfall. Because rainfall is difficult to predict, it is also difficult to predict the need for cash to buy water in the auction. Water demand is seasonal, peaking

[7] A related approach is to investigate separately at the choices of trusted experts in the industry as in, *e.g.*, Bronnenberg *et al.* (2015) and Johnson and Rehavi (2016).

[8] For example, 681 millimeters of water fell in Mula on one day, $10^{th}$ October 1943, while the yearly average in Mula is 326 *mm*. Summers are dry, and rainfall occurs most often during fall and spring.

during the weeks when fruit grows most rapidly, before the harvest. Farmers sell their output after the harvest, once per year. Only then do farmers collect cash (revenue) from growing their agricultural products. Hence, the weeks when farmers need cash the most to pay for water in the auctions, the weeks before the harvest, are the weeks furthest away from the previous harvest, when they collected the revenue the last time. As a consequence, poor farmers who do not have other sources of revenue may be liquidity constrained.

Given that demand is seasonal, farmers take into account the joint dynamics of water demand and price of water, when making auction's purchasing decisions. Water today is an imperfect substitute for water tomorrow. Future water prices are difficult to predict. Farmers consider current prices of water, and form expectations about future prices of water. A farmer who expects to be liquidity constrained during the critical season—when the demand is highest—may decide to buy water several weeks before the critical season, when the price of water is lower.

Farmers are "hand-to-mouth" consumers in that they have only enough money for their basic necessities (González Castaño and Llamas Ruiz, 1991). A farmer who expects to be liquidity constrained in the future would attempt to borrow money. However, poor farmers in Mula did not have access to credit markets.[9] Even if a credit market is in place, lenders may not grant loans. In the presence of limited liability (*i.e.* the farmer is poor), and non-enforceable contracts (*i.e.* poor institutions), endogenous borrowing constraints emerge (see Albuquerque and Hopenhayn, 2004, for a model of endogenous liquidity constraints). Hence, even if a credit market exists, non-enforceable contracts would prevent farmers from having cash when they need it most.[10]

[9]Interviews with surviving farmers confirm that some farmers were liquidity constrained—they did not have enough cash to buy their desired amount of water—yet they did not borrow money from others. See Appendix D.2 for details.

[10]In contrast to the German credit cooperatives from Guinnane (2001), the farmers in southeastern Spain were not able to create an efficient credit market. Spanish farmers were poorer than German farmers, and the weather shocks were aggregate (not idiosyncratic) shocks, and greater in magnitude. Hence, to reduce the risk, Spanish farmers should have resorted to external financing. However, external financing had problems, such as monitoring costs and information acquisition, that credit cooperatives did not have.

## 2.2 Institutions

**Auctions.** Since the $13^{th}$ century, Spanish farmers used a sequential outcry ascending price (or English) auction to allocate water. The basic structure of the sequential English auction remained unchanged from the $13^{th}$ century until 1966, when the last auction was run. The auctioneer sold each of the units sequentially, and independently of each other. The auctioneer tracked the buyer's name, and price for each unit of water sold. Farmers had to pay in cash on the day of the auction.[11]

Water was sold by *cuarta* (quarter), a unit that denoted the right to use water flowing through the main channel during three hours at a specific date and time. Property rights to water, and land were independent. Some individuals, not necessarily farmers, were Waterlords. Waterlords owned the right to use the water flowing through the channel. The farmers who participated in auctions owned only land. Water was stored at the main dam, and a system of channels delivered it to the farmer's plot. Water flowed from the dam through the channels at approximately 40 liters per second, so each unit of water sold at auction (*i.e.* the right to use water from the canal for three hours) carried approximately 432,000 liters of water. During our period of analysis, auctions were held once a week, every Friday. During each session, 40 units were auctioned: four units for irrigation during the day (from 7:00 AM to 7:00 PM), and four units for irrigation during the night (from 7:00 PM to 7:00 AM) on each weekday (Monday to Friday). Our sample consists of all water auctions in Mula from January 1955 until July 1966, when the last auction was run.

**Quotas.** On August 1, 1966 the water allocation system switched from an auction to a fixed quota system. Under this system, water allocation for each farmer was tied to land ownership. Each plot of land was assigned a fixed amount of water every three weeks, called a *tanda*. The amount of time allocated to each farmer was proportional to the size of the

[11] Allowing the farmers to pay after the critical season would have helped to mitigate the problems created by the liquidity constraints, and would have increased the revenue in the auction. However, it was written in their bylaws that the payment had to be in cash. This suggests that the water owners were concerned about repayment after the critical season due to non enforceable contracts (poor Spanish institutions).

farmer's plot. Every December, a lottery assigned a farmer's order of irrigation within each round. The order did not change during the entire year. At the end of the year, farmers paid a fee to the *Sindicato,* proportional to the size of their plot. Farmers paid after the critical season, and were not liquidity constrained.[12] The fee covered the year's operational costs, which included guard salaries and maintenance of the dam. In the counterfactual analysis we use this quota system, a non-market institution, as point of reference against which we compare the welfare of markets.

## 2.3 Data

We examine a unique panel data set where each period represents one week, and each individual represents one farmer. The unit of observation is a farmer-week. We collected the data from four sources. The first source is the weekly auction. For the period from January 1955 until the last auction, in July 1966, we observe the price paid, the number of units bought, the date of the purchase, and the date of the irrigation. This data was obtained from the municipal archive of Mula. The second source is rainfall measurements, obtained from the Spanish National Meteorological Agency, *AEMET*. The third source is a cross sectional agricultural census from 1955. The census data contains information regarding the farmer's plots, including type of agricultural product, number of trees, production, and sale's price. The final source is urban real estate tax records from 1955. We use this information to identify liquidity constraints. See Appendix A.1 for details, and summary statistics.

**Auction Data.** Auction data encompass 602 weeks, and can be divided into three categories based on bidding behavior and water availability: (i) normal periods (300 weeks), when for each transaction the name of the winner, price paid, date and time of the irrigation for each auction were registered; (ii) no-supply periods (295 weeks), when due to water shortage in the river or damage to the dam or channel—usually because of intense rain—no

[12] The farmer was the owner of the water under the quota system, so the price that the farmer paid was the average cost of operation, which was smaller that the average price paid per unit of water under the auction system.

auction was carried out; and (iii) no-demand periods (7 weeks), when some units were not sold due to lack of demand because of recent rain, and the price dropped to zero. For the empirical analysis we use data for the period 1955-66.

**Rainfall Data.** We link auction data to daily rainfall data for Mula, which we obtain from the *Agencia Estatal de Metereología*, AEMET (the Spanish National Meteorological Agency). In regions with Mediterranean climate, rainfall occurs mainly during spring and fall. Peak water requirements for products cultivated in the region are reached in spring and summer, between April and August. The coefficient of variation of rainfall is 450 percent ($^{37.08}/_{8.29} \times 100$), indicating that rainfall varies substantially.

**Agricultural Census Data.** We also link auction data to the data that we collected from the 1954/55 agricultural census from Spain, which provides information on individual characteristics of farmers' land. The Spanish government conducted an agricultural census in 1954/55 to enumerate all cultivated soil, crop production, and agricultural assets available in the country. The census recorded the following individual characteristics about farmers' land: type of land and location, area, number of trees, production, and the price at which this production was sold in the year of the census. We match the name of the farmer on each census card with the name of the winner of each auction.

**Urban Real Estate Tax Data.** Finally, we link the previous data to the urban real estate tax registry in 1955. To identify the source of financial constraints, we need a variable related to farmers' wealth but unrelated to their demand for water. We use urban real estate taxes to identify the wealthy farmers, as explained next.

### 2.4 Preliminary Analysis

In this subsection, we provide descriptive patterns from the data. Four main fruit tress grow in the region: orange, lemon, peach, and apricot. Oranges are harvested in winter, when water prices are low; thus farmers are unlikely to face liquidity constraints. The other three types of fruit are harvested in the summer. We focus on apricots because they are the most common of these summer crops.

**Wealthy Farmers.** We define a farmer as "wealthy" if the value of urban real estate of the farmer obtained from the urban real estate tax data is above the median among the apricot farmers, and "poor" otherwise. Farmers grow their agricultural products in rural areas, thus, urban real estate constitutes non-agricultural wealth. In the empirical analysis, we use the set of wealthy farmers, and exploit that they were never liquidity constrained. We make two observations. First, the value of urban real estate owned by the farmer should not affect the farmer's production function (*i.e.* the farmer's willingness to pay for water), conditional on the type of agricultural product, the size of the plot, and the number of trees. Hence, after accounting for these variables, the value of the urban real estate should not be correlated with a farmer's demand for water, which is determined by the production function of the agricultural product, apricots in our case. We later use this as an "exclusion restriction" to identify liquidity constrained farmers. Second, we argue that wealthy farmers in Mula were never liquidity constrained, because of the value of their urban real estate properties. These farmers were very wealthy, owning several urban properties. For instance, on average wealthy farmers had an annual rental income of 5,702 *pesetas* derived from the urban real state. Whereas their average annual water expenditure was 500 *pesetas.* In 1963, the year with the highest water expenditures in the sample, their average annual water expenditure was 1,619 *pesetas.* None of the poor farmers owned any urban property.

**Water Demand and Apricot Trees.** Table 1 displays the seasonal stages of the typical apricot tree that is cultivated in Mula, the *búlida* apricot. These trees need water the most during the late fruit growth (stages II and III), and the early post harvest (EPH).[13] This defines the critical irrigation season for these apricot trees. Stage III corresponds to the period when the tree "transforms" water into fruit at the highest rate. The EPH period is important because of the hydric stress the tree suffers during the harvest (see Appendix A.2

[13] The beginning of the post-harvest period coincides with week 24. In the model in Section 3we assume that all the harvest takes place on week 24. In practice the harvest would take several weeks. The tree is vulnerable during the early post-harvest weeks, and the moisture of the tree during those weeks would affect the harvest of the current year.

Table 1: Seasonal Stages for "Búlida" Apricot Trees.

<table>
<tr><td>JAN</td><td>FEB</td><td>MAR</td><td>APR</td><td>MAY</td><td>JUN</td><td>JUL</td><td>AUG</td><td>SEP</td><td>OCT</td><td>NOV</td><td>DEC</td></tr>
<tr><td>DORM</td><td>FLOW</td><td colspan="3">FRUIT GROWTH</td><td colspan="6">POST-HARVEST</td><td>DORM</td></tr>
<tr><td></td><td></td><td>I</td><td>II</td><td>III</td><td colspan="2">EARLY</td><td colspan="4">LATE</td><td></td></tr>
</table>

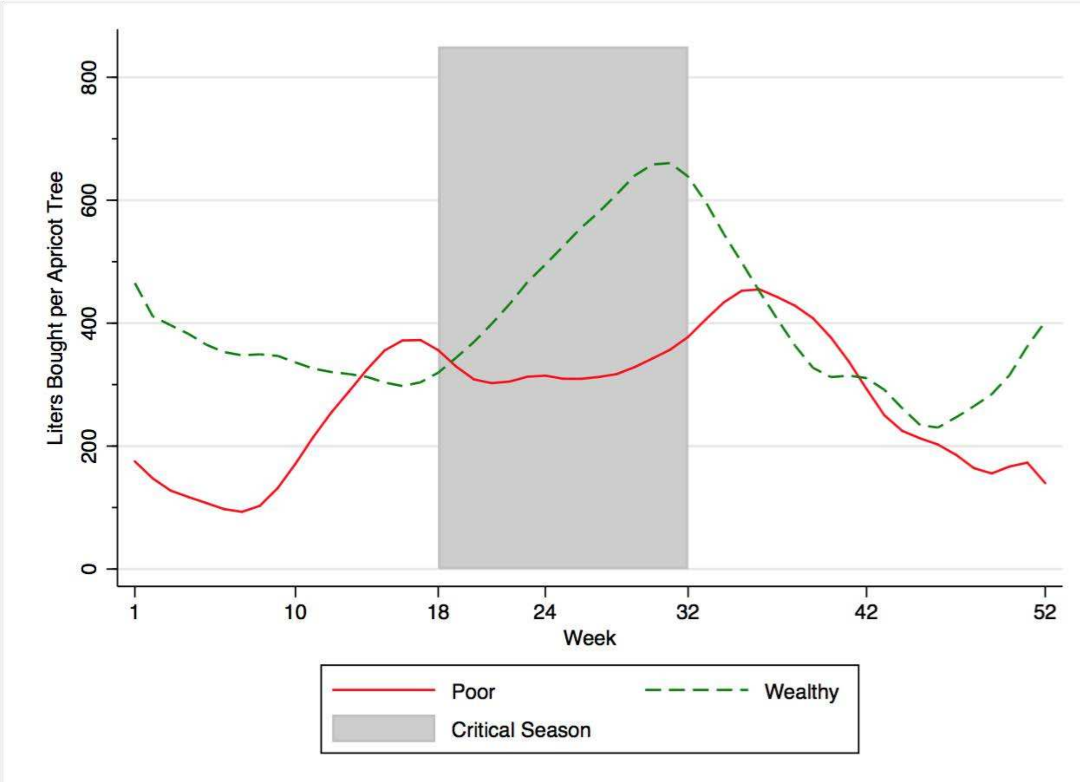
*Notes*: Obtained from Pérez-Pastor *et al.* (2009). DORM refers to Dormancy. FLOW refers to Flowering.

for details).

Figure 1 shows the purchasing patterns of wealthy, and poor apricot's farmers. The figure displays the average liters per tree that each type of farmer purchased in the auction. The shaded area corresponds to the critical season as defined above. The price of water increases substantially during the critical season because (i) apricots, along with other products cultivated in the region, require more irrigation during this season, increasing the demand for water in the auction; and (ii) weather seasonalities in southern Spain generate less rainfall during these months (see Appendix A.1 for details). Wealthy farmers—who are not liquidity constrained—demand water as predicted by Table 1. Wealthy farmers strategically delay their purchases, and buy water during the critical season, when the apricot trees need water the most. Poor farmers—who may be liquidity constrained—display a bimodal purchasing pattern for water inconsistent with Table 1. The first peak occurs before the critical season, when water prices are relatively low. Poor farmers buy water before the critical season because they anticipate that they may not be able to afford water during the critical season, when prices are high. A fraction of this water will evaporate, but the rest remains as soil's moisture. The second peak occurs after the critical season, when water prices are again relatively low. After the critical season, poor farmers' plot have a low moisture level if they were unable to buy sufficient water during the critical season. Thus, poor farmers buy water after the critical season to prevent their trees from withering. This purchasing pattern for the poor farmers—high purchases before and after the critical season, and low purchases during the critical season—is explained by the model that we present in Section 3, which includes seasonality, storability, and liquidity constrains.[14]

[14]Table 2, discussed next, shows that the differences in purchases between poor and wealthy farmers are only significant during the critical season. Although our model has clear predictions for the difference in

Figure 1: Seasonality and Purchasing Patterns of Wealthy and Poor Farmers.




*Notes*: The figure displays the average liters bought per farmer and per tree disaggregated by wealthy and poor farmers using a least squares smoother. A farmer is defined as wealthy if the value of urban real estate of the farmer is above the median. A farmer is defined as poor if the value of urban real estate of the farmer is below the median. The shaded area corresponds to the critical season as defined in Table 1.

Table 2: Demand for Water per tree and Urban Real Estate.

| # units bought per tree | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Wealthy | 0.0131*** | 0.0073 | 0.0066 | 0.0017 | 0.0058 | -0.0002 |
| | (0.0042) | (0.0044) | (0.0044) | (0.0047) | (00.49) | (0.0051) |
| (Wealthy) | | | 0.0374*** | 0.0315*** | 0.0383*** | 0.0326*** |
| × (Critical Season) | | | (0.0091) | (0.0094) | (0.0093) | (0.0094) |
| (Wealthy) | | | | | 0.0039 | 0.0104 |
| × (Weeks 1-10) | | | | | (0.0092) | (0.0101) |
| Covariates | No | Yes | No | Yes | No | Yes |
| Number of observations | 14,448 | 14,448 | 14,448 | 14,448 | 14,448 | 14,448 |

*Notes:* All regressions are OLS specifications. The sample is restricted to farmers who grow only apricots. The dependent variable is the number of units bought per tree by each individual farmer during a given week. "Wealthy" is a dummy variable that equals 1 if the value of urban real estate of the farmer is above the median, and 0 otherwise. "Critical season" is a dummy variable that equals 1 if the observation belongs to a week during the critical season, and 0 otherwise. "Weeks 1-10" is a dummy that equals 1 if the observation belongs to one of the first ten weeks of the year, and 0 otherwise. "Covariates" are the price paid by farmers in the auction, the amount of rainfall during the week of the irrigation, and the farmer's soil's moisture level. Standard errors in parentheses. * p<0.10; ** p<0.05; *** p<0.01.

Table 2 shows similar evidence to the one in Figure 1 using OLS regressions. As in Figure 1, in Table 2 we restrict attention to farmers who grow only apricot trees. We regress the number of units bought per tree by each farmer in a given week on several covariates. The variable "wealthy" is a variable that equals 1 if the value of the urban real estate owned by the farmer is greater than the sample median, and 0 otherwise.[15] This dummy variable identifies farmers who are wealthy, thus not liquidity constrained. For farmers who are not liquidity constrained, demand for water should not be correlated with wealth. Consider two farmers who grow apricots, have the same number of trees, and are not liquidity constrained. Water demand is determined by a tree's water need according to the apricot production function as predicted by Table 1. These two farmers should have the same demand for water up to an idiosyncratic shock. Therefore, there should be no relationship between the demand for water, and the monetary value of urban real estate for unconstrained farmers. Table 2 displays the regressions using the number of units bought per tree as the dependent variable to account for the size of the farmer's plot.[16] Column 1 shows that wealthy farmers buy more water overall. However, the coefficient is not statistically different from zero in column 2, when we include the covariates. This is consistent with wealthy and poor farmers buying the same amount of water throughout the year. In columns 3 and 4 we include an interaction between "wealthy" and "critical season." The variable "critical season" is a dummy variable that equals 1 if the observation belongs to a week during the critical season, and 0 otherwise, as defined in Table 1. The interaction term is positive, and statistically different from zero in all specifications. Wealthy farmers demand more water per tree during the critical season

purchasing during the critical season, the predictions outside the critical season are ambiguous and depend on the severity of liquidity constraints. In particular, poor farmers buy less water than wealthy farmers outside the critical season only when liquidity constraints are severe.

[15]This definition of "wealthy" farmers is conservative. Both in the historical narrative and in our estimates of the probability of being liquidity constraints in Appendix C.6, we find that some of the farmers who are potentially liquidity constrained, *i.e.*, non wealthy farmers, are not liquidity constrained. This results in a smaller sample of wealthy farmers to estimate the demand for water. However, because the our main goal is to study the behavior of liquidity constrained farmers, we chose to be conservative in this definition. The results are similar if we select the $40^{th}$ or the $60^{th}$ percentile, instead of the median.

[16]Wealthy farmers own larger plots, and farmers can only buy whole units of water. So there may be economies of scale in water purchases only available to wealthy farmers.

than poor farmers who have the same agricultural products (apricots). The effect of liquidity constraints on the demand for water is concentrated on the critical season, when the price for water in the auction is high. For robustness, in columns 5 and 6 we also include the interaction between "wealthy" and an indicator for purchases during the first 10 weeks of the year. The coefficient of this interaction is not statistically different from zero, as expected. Appendix C.6 presents additional evidence about the presence of liquidity constraints for the poor farmers, using the estimates of the model in the next section.

# 3 The Econometric Model

In this section, we present the econometric model that we use to compute welfare under auctions and quotas. Computing welfare would be straightforward with output data (*i.e.* production data) before and after the institutional change, but such output data is not available. Therefore, we use detailed input data (units of water purchased, amount of rainfall, number of apricot trees, *etc.*) along with the apricot production function to compute the output under auctions and quotas. We proceed in three steps. First, we present the econometric model, which uses the apricot production function and incorporates the three features of our setting: storability, liquidity constraints, and seasonality. Second, we estimate the model using the input data for the wealthy farmers. Finally, we use the estimated model to compute the counterfactual output of apricot, as a measure of welfare, under auctions and quotas for all farmers (*i.e.* before and after the institutional change).

Farmers used a sequential outcry ascending price (or English) auction to allocate water. Every week during each session, 40 units were auctioned: four units for irrigation during the day and four units for irrigation during the night on each weekday (Monday to Friday). In this paper we do not model the auction game and, thus, we abstract from the within-week variation in prices, which is very low (see Donna and Espín-Sánchez 2018). We translate the auction mechanism into a simpler dynamic demand system, whereby individual farmers take prices as exogenous. This allows us to focus on the dynamic behavior of farmers across weeks. We focus only on the demand system of the 24 farmers who only grow apricot trees. This is the largest group of farmers that grow one single crop. Note, however, that there are more than 500 farmers in the data set who can participate in the auction. Hence, we assume that the distribution of the highest valuation among the other 500 farmers is exogenous to the valuation of a given farmer, conditional on the week of the year, on the price, and on the rain in the previous week. This is a credible assumption in our setting because it is unlikely that any individual apricot's farmer could affect the equilibrium price in the auction.[17]

We now describe the econometric model. The economy consists of $N$ rational and forward-looking farmers, indexed by $i$, and one auctioneer. Water increases the farmer's soil's moisture level. So, from the point of view of the farmer, there are two goods in the economy: moisture, $M$, measured in liters per square meter and money, $\mu$, measured in *pesetas*. Time is denoted by $t$, the horizon is infinite, and the discount between periods is $\beta \in (0,1)$. Demand is seasonal, and we denote the season by $w_t \in \{1,2,...,52\}$, representing each of the 52 weeks in a given year. In each period, the supply of water in the economy is exogenous. Farmers receive only utility for water consumed during the critical season. Water is an intermediate good. Hence, the utility refers to the farmer's profit, and is measured in *pesetas*, *not* in utils. Water purchased in any period can be carried forward to the next period, but it "evaporates" as indicated by the evolution of soil's moisture in the equation in (1) that we describe below. That is, water "depreciates" at some rate $\delta \in (0,1)$. Farmers' preferences are represented by: $u\left(j_{it}, M_{it}, w_t, p_t, r_t, \varepsilon_{ijt}, \mu_{it}; \gamma, \sigma_\varepsilon, \zeta\right) = h\left(j_{it}, M_{it}, w_t; \gamma\right) + \varepsilon_{ijt} - \zeta_j - p_t j_{it}$, where $h\left(\cdot\right)$, is the apricot's production function, that is common to all farmers and is strictly increasing in the moisture level of the farmer, $M_{it}$; $j_{it} \in \{0,1,...,J\}$, is the number of units that farmer $i$ purchases in period $t$; $\varepsilon_{ijt}$ is an additive productivity shock to farmer $i$ in period $t$, given that the farmer bought $j_t$ units of water that we describe below; $\mu_{it}$ is the amount

[17] In Appendix C.10 we present dynamic auction model that accounts for the within-week variation in prices. In the structural estimation we abstract from it because there is little variation in average prices (among several units) within a week across farmers. See Donna and Espín-Sánchez (2018) for details.

of cash that farmer $i$ has in period $t$ that we describe below; $p_t$ is a scalar that represents the price of water in the auction in period $t$ that we describe below; $(\gamma, \sigma_\varepsilon, \zeta)$ is a vector of parameters to be estimated that we describe below. Note that $\zeta_j$ represents the cost that the farmer incurs when irrigating (this disutility could result, for example, if the farmer hires a laborer to help with irrigation). We restrict attention to the case where farmers do not incur irrigation costs if they do not irrigate ($\zeta_0 = 0$) and irrigation costs are constant across units ($\zeta_j = \zeta$ if $j > 0$). Farmers optimization problem is subject to the constraints that we describe when we explain the equation in (3). The function $u(\cdot)$ depends implicitly on $r_t$ and $\sigma_\varepsilon$, because the moisture and the additive productivity shocks depend on $r_t$ and $\sigma_\varepsilon$, respectively, as described below.

Motivated by the historical context of Mula, we assume that farmers are "hand-to-mouth" consumers, *i.e.*, we require that $(\mu_{it} - p_t j_{it}) \geq 0$, $\forall j_{it} > 0$ (limited liability). Further, we assume that wealthy farmers obtain cash flow from their non-agricultural wealth. So they always have enough cash, and the limited liability constraint is never binding. However, the constraint could be binding for poor farmers. Poor farmers anticipate that the constraint may be binding in the future (*e.g.* during the critical season) and, thus, they buy water before the critical season, when prices of water are low. Farmers in the economy differ from each other in two ways. First, they differ in their productivity shock, $\varepsilon_{ijt}$. Second, they differ in their wealth levels, $\mu_{it}$. Both, $\varepsilon_{ijt}$ and $\mu_{it}$, are private information. We describe the evolution of the wealth level below.

## State Variables and Value Function

There are six state variables in the model.

**Moisture.** The moisture, $M_{it}$, measures the amount of water accumulated in the farmer's plot. This is a deterministic, observable variable measured in liters per square meter. The trees on a farmer's plot die if the soil's moisture level falls below the permanent wilting point, $PW$, which is a scalar obtained from the literature in agricultural engineering. So each farmer $i$ must satisfy the constraint: $M_{it} \geq PW$, $\forall t$. The evolution of $M_{it}$ is given by:[18]

$$M_{it} = min\left\{M_{i,t-1} + r_{t-1} + \frac{j_{it-1} \cdot 432,000}{area_i} - ET\left(M_{it-1}, w_{t-1}\right), FC\right\}, \quad (1)$$

where $r_t$ is the amount of rainfall, measured in liters per square meter, in Mula during period $t$; $432,000$ is the number of liters in each unit of water; $area_i$ is the farmer's plot area, measured in square meters; $ET\left(M_{it}, w_t\right)$ is the adjusted evapotranspiration in period $t$ described in Appendix A.2; and $FC$ is the full capacity of the farmer's plot also described in Appendix A.2. Moisture and seasonality are the main determinants of water demand. The moisture level increases with rain and irrigation, and decreases over time as the accumulated water "depreciates" (evapotranspiration). We use the equation in (1) to compute the moisture level. Note that equation 1 accounts implicitly for decreasing marginal returns (concavity) of water in two ways. First, because there is a maximum capacity in the farmer's plot represented by $FC$, farmers "waste" water if the moisture level increases above $FC$. Second, evapotranspiration (*i.e.* depreciation of the water in the farmer's plot) is greater for higher levels of moisture. Thus, farmers with high levels of moisture in their plots "waste" more water *via* greater evapotranspiration. Note that there are declining returns of irrigation water (units bought by the farmer), even when the production function is linear in moisture.

**Weekly Seasonal Effect.** The week of the year, $w_t$, is the weekly seasonal effect. This is a deterministic variable with support on $\{1, 2, ..., 52\}$ that evolves a follows: $w_t = w_{t-1} + 1$ if $w_{t-1} < 52$, and $w_t = 1$ otherwise. Farming is a seasonal activity, and each crop has different water requirements depending on the season. The requirement of water for the apricot trees is captured by its production function, $h\left(j_{it}, M_{it}, w_t; \gamma\right)$. Because the market for water has a weekly frequency, we include a state variable with a different value for each week.

**Price of Water and Rainfall.** For each week $t$, the price of each unit of water in the auction, $p_t$, and the amount of rainfall in the town, $r_t$, are two random variables whose joint

[18] See Appendix A.2 for details. The variable moisture implicitly accounts for the decreasing marginal returns of water on area because larger plots receive smaller increase of moisture after purchasing a unit of water as can be seen in the equation in (1).

probability distribution is described below. The price of water is measured in *pesetas*, and the rainfall is measured in liters per square meter. We model the joint probability distribution of prices and rainfall to capture two main empirical regularities from our setting. First, the major determinant of the price of water in the auction is weather seasonality. Second, the variation of prices and rainfall across years is low, conditional on the week of the year (which captures seasonality).[19] Our data in this paper cover a sample of 12 years. We model the joint evolution of the price of water in the auction in period $t$ and rainfall in period $t-1$ assuming that, holding fixed the week of the year, farmers jointly draw a price-rain pair, $(p_t, r_{t-1})$, *i.i.d.* among the 12 pairs (*i.e.* the 12 years of the same week) available in the data with equal probability.[20] Note that water for each week is auctioned on the Friday of the previous week. So when a farmer jointly draws a pair price-rain, the rain corresponds to the rain during the week previous to the irrigation. Thus, prices for the week of the irrigation are drawn conditional on the week of the year, and the rainfall during the previous week. The rain during the previous week captures the dynamics of droughts, *i.e.,* prices are systematically higher when there is no rain. See Appendix C.4 for further discussion.

**Productivity Shock.** The productivity shock, $\epsilon_{it} \equiv (\epsilon_{i0t}, ..., \epsilon_{iJt})$, is a choice-specific component of the utility function.[21] We assume that the productivity shocks, $\varepsilon_{ijt}$, are drawn *i.i.d.* across individuals and over time from a Gumbel distribution with CDF $F(\varepsilon_{it}; \sigma_\varepsilon) = e^{-e^{-\varepsilon_{it}/\sigma_\varepsilon}}$, where $\sigma_\varepsilon$ is a parameter to be estimated. The variance of this distribution is given by $\sigma_\varepsilon^2\pi^2/6$. The higher the value of the parameter $\sigma_\varepsilon$, the more heterogeneous the distribution of productivity. In addition, productivity shocks are drawn *i.i.d.* across the choice of not buying, $j = 0$, and buying, $j > 0$. So every farmer receives one shock, but the shock is the same for all $j > 0$. Formally, let $\hat{j} \in \{0, 1\}$, where $\hat{j} = 0$ if $j = 0$ and $\hat{j} = 1$ if $j > 0$. Then the

[19] See Donna and Espín-Sánchez (2018) for details.

[20] We obtain similar results by estimating the joint distribution of prices and rain nonparametrically conditional on the week of the year, and then drawing price-rain pairs from this distribution, conditional on the week of the year.

[21] Alternatively, one could refer to these shocks as a component of the costs of irrigation. Note that these shocks have no impact on the marginal productivity of moisture. See Section 6 for a discussion of their impact on welfare.

productivity shocks $\varepsilon_{i\hat{j}t}$ are drawn *i.i.d.* across $\hat{j} \in \{0,1\}$ and the shock is the same for every unit, so $\varepsilon_{ijt} = \varepsilon_{i\hat{j}t}$ for $j = 0$ and $\varepsilon_{ijt} = \varepsilon_{i\hat{j}t}$ for $j > 0$. We present closed-form expressions for the conditional choice probabilities using this specification in Appendix C.3.[22]

**Cash Holdings.** The cash holdings, $\mu_{it}$, measure the amount of cash that farmer $i$ has in period $t$. The variable $\mu_{it}$ is measured in *pesetas* and evolves according to:

$$\mu_{it} = \mu_{i,t-1} - p_{t-1} j_{i,t-1} + \Phi_t \left( re_i; \phi \right) + \eta_{it} + \nu_{it}, \tag{2}$$

where $\Phi_t \left( re_i; \phi \right) = -\phi_{i0} + \phi_1 re_i$ captures the weekly cash flow function derived from the real estate value, $\phi re_i$, minus individual $i$'s average weekly consumption that is constant over time, $\phi_{i0}$; $\eta_{it}$ is the revenue that the farmer obtains from selling the harvest discussed in the equation in (5); and $\nu_{it}$ are idiosyncratic financial shock that are drawn *i.i.d.* across individuals and over time from a normal distribution. The revenue, $\eta_{it}$, is zero all weeks of the year, except the week after the harvest, when farmers sell their products and collect revenue for the whole year. Note that $\phi_{i0}$ represents the average weekly consumption. Therefore, weekly consumption is not necessarily constant over time.

The value function is given by:

$$\begin{gathered} V \left( M_{it}, w_t, p_t, r_t, \mu_{it}, \epsilon_{ijt} \right) \equiv \underset{j_{it} \in \{0,1,\ldots,J\}}{max} \left\{ h \left( j_{it}, M_{it}, w_t; \gamma \right) + \varepsilon_{ijt} - \zeta_j - p_t j_{it} + \right. \\ \left. + \beta \mathbb{E} \left[ V \left( M_{i,t+1}, w_{t+1}, p_{t+1}, r_{t+1}, \mu_{i,t+1}, \epsilon_{i,t+1} \right) | M_{it}, w_t, p_t, r_t, \mu_{it}, \epsilon_{i,t}, j_{it} \right] \right\}, \\ \text{s.t. } M_{it} \geq PW, \quad j_{it} p_t \leq \mu_{it}, \quad \forall j_{it} > 0, \end{gathered} \tag{3}$$

subject to the evolution of the state variables as described above. The expectation is taken over $r_t$, $p_t$, $\varepsilon_{ijt}$, and $\nu_{it}$. For wealthy farmers we assume that the constraint $j_{it} p_t \leq \mu_{it}$ is not binding.

[22] Note that the choice is not binary, $j_{it} \in \{0, 1, ..., J\}$. In Appendix C.3 we describe two specifications for the productivity shocks. First, for the case of *i.i.d.* shocks across choice alternatives, where each choice alternative involves the purchase of a different number of units. Second, the one presented above, which is our preferred specification and where the productivity shocks are drawn *i.i.d.* across the choice of not buying, and buying. For robustness, in Subsection C.5 in the appendix we present the demand estimates of the structural model using both specifications, which yield similar results. This is because farmers buy more than one unit very seldom (see Table A1 in Appendix A.1). That is, the extensive margin is what matters.

### The Apricot's Production Function

The production function of the apricot tree is given by Torrecillas *et al.* (2000):

$$h\left(j_{it}, M_t, w_t; \gamma\right) = \left[\gamma \cdot \left(M_t - PW\right) \cdot KS\left(M_t\right) \cdot Z\left(w_t\right)\right], \tag{4}$$

where $h\left(j_{it}, M_t, w_t; \gamma\right)$ is the harvest at period $t$; $\gamma$ is a parameter that measures the transformation rate of water into apricots during the fruit's growth season and the early post-harvest stress season, and it is measured in *pesetas* per liter; $KS\left(M_t\right)$ is the hydric stress coefficient, which is a weakly increasing function of moisture and is described in Appendix A.2; $Z\left(w_t\right)$ is a dummy variable that equals 1 during weeks 18 to 32 and 0 otherwise, and it captures the seasonal stages of the typical apricot tree cultivated in Mula, as emphasized when we discussed Table 1. The farmer's revenue in a given year is:[23]

$$Revenue = \sum_{w_t=18}^{32} \gamma \cdot \left(M_t - PW\right) \cdot KS\left(M_t\right). \tag{5}$$

# 4 Estimation

We estimate the parameters that characterize demand, $\Theta \equiv (\gamma, \sigma_\varepsilon, \zeta)$, using data from wealthy farmers, and excluding data from poor farmers who may be liquidity constrained. To perform the estimation we assume that there is no persistent unobserved heterogeneity that affects the production function of wealthy and poor farmers differently (*i.e.* no dynamic sample selection on unobservables). We also assume that wealthy farmers are never liquidity constrained. Although the latter assumption is not necessary to identify the model, it simplifies the estimation (see Appendices C.1 and C.8 for discussions), and is motivated by the empirical context as discussed in Section 2.

[23] The production function measures production in *pesetas*. The actual price at which the production is sold is determined in the output market. We do not have data on the price at which this production is sold. So we recover the revenue of the farmers up to this constant (the common price at which the production of all farmers is sold in apricot's market). This price only shifts the revenue function of all (wealthy and poor) farmers. So it does not affect our welfare analysis.

## 4.1 Demand Estimates

We construct a two-step conditional choice probability (CCP) estimator (Hotz and Miller, 1993) to estimate the parameters that characterize demand.

**Step 1.** We compute transition probability matrices for the following observable state variables: moisture, week, price, and rain. As described above, the productivity shocks, $\varepsilon_{ijt}$, are assumed to be *i.i.d.* Gumbel, so they can be integrated analytically. The evolution of moisture depends on both the farmers' decisions to buy water and rainfall. Therefore, certain values of moistness are never reached in the sample, even when their probability of occurrence is nonzero. To estimate demand, however, we need to integrate the value function for each possible combination of state variables in the state space. Thus, we first estimate the CCP using the values of the state space reached in the sample, using only data on wealthy farmers. Then we use the CCP estimator to predict the CCP on the values of the state space unreached in the sample.[24]

**Step 2.** We build an estimator similar to the one by proposed by Hotz *et al.* (1994). We use the transition matrices to forward simulate the value function from the equation in (3).[25] This gives us the predicted CCP by the model as a function of the parameter vector, $\Theta$. We estimate the parameter vector $\Theta$ using a GMM estimator based on the moment conditions proposed by Hotz *et al.* (1994).

**Identification.** The exclusion restriction is that wealthy farmers are not liquidity constrained. Under this exclusion restriction, the identification of $\Theta$ follows the standard arguments (*e.g.* see Hotz Hotz and Miller, 1993; Hotz et al., 1994; Rust, 1996; Magnac and Thesmar, 2002; and Aguirregabiria, 2005). In our case the transformation rate, $\gamma$, is identified from the variation in purchasing patterns across seasons, and the variation in moisture levels across farmers within the same season. The irrigation cost, $\zeta$, is constant across units,

[24]We estimate the CCP both non-parametrically, using kernel methods to smooth both discrete and continuous variables, and parametrically, using a logistic distribution, *i.e.*, a multinomial logit regression. See Appendix C.8 for details.

[25]For the initial condition of the moisture we follow Hendel and Nevo (2006, p. 1,647), and use the estimated distribution of moisture to generate its initial distribution, as described in Appendix C.1.

Table 3: Structural Estimates

| | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Transformation rate ($18 \leq week \leq 32$): | | | | |
| – Linear term: $\hat{\gamma}_L$ | 0.1584<br>(0.0067) | 0.1790<br>(0.0094) | 0.2124<br>(0.0026) | 0.0734<br>(0.0064) |
| – Quadratic term: $\hat{\gamma}_Q$ | – | 1.36e-04<br>(1.59e-05) | – | 6.19e-05<br>(8.82e-06) |
| Irrigating cost: $\hat{\zeta}$ | 24.3755<br>(6.9848) | 182.174<br>(13.2357) | 78.8924<br>(3.1994) | 34.3495<br>(6.76245) |
| Scale parameter of Gumbel distribution: $\hat{\sigma}_\varepsilon$ | 1.0100<br>(0.2568) | 1.0854<br>(0.1286) | 0.9361<br>(0.0393) | 1.0144<br>(0.1048) |
| Area heterogeneity | No | No | Yes | Yes |

*Notes*: Bootstrapped standard errors in parenthesis. See Subsection 6.1 for details about this table.

and independent of the moisture level. It is identified from the variation in the level of prices, and the farmer's decision of buying *vs.* not buying, holding constant the level of moisture. Finally, the parameter $\sigma_\varepsilon$ represents the inverse of the marginal utility of income in our model, and is identified because our specification for the utility function of the farmer on page 18 does not include a parameter that multiplies the price of water. Such parameter is typically called "$\alpha$" in the industrial organization literature (see Berry *et al.*, 1995; Hendel and Nevo, 2006), and measures the income sensitivity of a consumer. Because we are estimating a production function, the utility function on page 18 is in *pesetas*, not in utils.

# 5 Estimation Results

Table 3 displays the estimation results for the demand parameters, $\Theta \equiv (\gamma, \sigma_\varepsilon, \zeta)$, from the demand model in equation in 3 using the estimation procedure from Subsection 4.1. We present two sets of estimates with different different discrete types of farmers, who differ only in their plot's area as described below. In columns 1 and 2 we perform the estimation with only one type of farmer who has the median number of trees from the sample ("Area heterogeneity: No" ). This means that when we forward simulate the value function (as outlined in Subsection 4.1), $area_i$ from the equation in (4) is set to the median area for all individual farmers $i$. In column 1 we use the apricot production function as outlined in the equation in (4). The estimated transformation rate is $\hat{\gamma}_L = 0.16$. For robustness, in column 2 we add a quadratic term for moisture, $\gamma_Q$, to the specification in column 1 to incorporate potential increasing or decreasing marginal returns explicitly. However, the estimated coefficient for the quadratic term of the transformation rate is small in magnitude, $\hat{\gamma}_Q = 1.36e-04$. In columns 3 and 4 we repeat the estimation from the previous two columns using ten different discrete types of farmers, who differ only in their plot's area ("Area heterogeneity: Yes" ). The area of each type corresponds to the number of trees owned by the wealthy farmers in the data. There are twelve wealthy farmers in the data, but there are two pairs of farmers with the same area. Each discrete type has the same probability. This means that when we forward simulate the value function as outlined in Subsection 4.1, the value of $area_i$ from the equation in (4) is drawn uniformly at random from a distribution with discrete support at the points $\{area_1, area_2, \ldots, area_{10}\}$. In Table 3 we report the mean $\Theta \equiv (\gamma, \sigma_\varepsilon, \zeta)$ across types. The estimated scale parameter of the distribution of idiosyncratic productivity, $\hat{\sigma}_\varepsilon$, is similar in magnitude across the specifications. The higher the parameter $\sigma_\varepsilon$, the higher the variance of the distribution of idiosyncratic productivity. When $\sigma_\varepsilon = 1$, the distribution of idiosyncratic productivity is a standard Gumbel. Finally note that the estimated irrigation cost has the expected sign and is sensible in magnitude.[26]

[26]In Appendix C.5 we present additional estimates of the model using a different specification for the productivity shocks. The overall estimates are similar to the specifications in Table 3. In a previous draft we also obtained similar results to the ones in Table 3 using: (i) an specification that allows for different transformation rates for pre-season ($18 \leq week \leq 23$) and on-season ($24 \leq week \leq 32$), and (ii) an autoregressive specification for the productivity error term.

# 6 Welfare

In this section we use the estimated demand system to compare welfare under auctions, quotas, and the highest-valuation allocation.[27] In our model there are two potential sources of inefficiency in the allocation of water. First, the allocation could be inefficient because some farmers receive water at a time when they are relatively unproductive. This inefficiency arises because farmers are *ex post* heterogeneous in productivity, so we call it *inefficiency due to heterogeneity.* Second, the allocation could be inefficient because some farmers receive water when their soil moisture level is relatively high. This inefficiency arises because the production function is concave in water, so we call it *inefficiency due to concavity.* Quotas allocate water units uniformly, so they always create inefficiency due to heterogeneity, but never inefficiency due to concavity in units bought. Markets would correct both inefficiencies if there were no liquidity constraints, but would create both inefficiencies when liquidity constraints are present. If farmers are heterogeneous and the production function is linear in the number of units bought, markets are always more efficient than quotas. Quotas are more efficient than markets when there is large heterogeneity in wealth, and small heterogeneity in productivity. Markets are more efficient in the opposite case. In the general case where there is heterogeneity in both wealth and productivity, the efficiency of markets relative to quotas is ambiguous.

In our empirical setting, large heterogeneities in wealth create liquidity constraints. Due to the dynamics created by the moisture, liquidity constraints create inefficiency due to concavity by allocating water to wealthy farmers with relatively high moisture levels. The heterogeneity in productivity is captured by the productivity shocks, $\varepsilon_{ijt}$. Although these shocks are drawn *i.i.d.* across individuals and over time, the estimated value of $\sigma_\varepsilon$ measures the degree of heterogeneity. The higher the value of the parameter $\sigma_\varepsilon$, the more heterogeneous the distribution of productivity. Because ours is a discrete choice model and the

[27]The $HV$ corresponds to the static first-best allocation. However, due to dynamics and the possibility of strategic delaying in the decisions to purchase water it may not coincide with the dynamic first-best allocation, which is a complex problem that is outside the scope of this paper.

error term, $\epsilon_{ijt}$, is choice-specific, the relevant measure for efficiency are the differences in $\epsilon_{ijt}$ across choices conditional on the choice, not the $\epsilon_{ijt}$ by itself, nor the unconditional difference. For example, in the case in which $J = 1$, the farmer chooses whether to buy one unit or not to buy. The farmer balances the difference in utility between buying or not, considering both the observable and unobservable (for the econometrician) components. The probability of a farmer buying water increases with the expectation of the difference in $\epsilon_{ijt}$, *i.e.*, with $\mathbb{E}\left[\epsilon_{i1t} - \epsilon_{i0t}\right]$. By construction, the unconditional mean of the differences in the error term is zero. Hence, in the quotas system, because the farmers cannot choose when to irrigate, the conditional and unconditional expectations of the difference in the error terms are zero: $\mathbb{E}\left[\epsilon_{i1t} - \epsilon_{i0t}\right] = \mathbb{E}\left[\epsilon_{i1t} - \epsilon_{i0t}|j = 1\right] = \mathbb{E}\left[\epsilon_{i0t} - \epsilon_{i1t}|j = 0\right] = 0$. However, in the auction system, farmers *do* choose when to irrigate, and the conditional expectation is always positive. Farmers are more likely to irrigate when their unobserved utility of irrigation is positive, $\epsilon_{i1t} > \epsilon_{i0t}$, and they are more likely not to irrigate when their unobserved utility for no irrigation is positive, $\epsilon_{i0t} > \epsilon_{i1t}$. This implies that under the auction system: $\mathbb{E}\left[\epsilon_{i1t} - \epsilon_{i0t}|j = 1\right] > 0$ and $\mathbb{E}\left[\epsilon_{i0t} - \epsilon_{i1t}|j = 0\right] > 0$. In other words, gains from trade are realized in the auction system. The greater the parameter $\sigma_{\varepsilon}$, the greater these gains from trade. In our empirical setting, gains from trade are translated into the timing of irrigation. Farmers trade with each other to irrigate during their preferred weeks of the year.

## 6.1 Welfare Measures

We now describe how we construct the welfare measures. We compute two welfare measures, the yearly mean revenue and welfare, both per tree and per farmer. For both measures we do not take into account water expenses because they represent transfers, and we are interested in welfare. The only difference between revenue and welfare is due to the choice specific unobservable component, $\epsilon_{ijt}$, as explained above.[28] We compute the welfare measures for

[28]Welfare is always greater than revenue under the auction system because the former accounts for the differences in the choice specific unobservable component. Not accounting for these differences would underestimate welfare under auctions.

Table 4: Welfare Results

| | Auctions complete units | | Quotas complete units | | | | High Valuation complete units |
|---|---|---|---|---|---|---|---|
| | Ac Revenue | Ac Welfare | Qc | Qc75% | Qc50% | Qc25% | HVc |
| **Welfare measures:** (mean per farmer, per tree, per year) | | | | | | | |
| - All farmers pre-season (24 farmers) | 972.22 | 978.82 | 825.80 | 907.90 | 978.37 | 1042.04 | 1,055.03 |
| - All farmers on-season (24 farmers) | 392.05 | 396.85 | 321.83 | 367.73 | 405.66 | 438.43 | 445.28 |
| - Poor farmers whole season (12 farmers) | 1,178.04 | 1,189.44 | 1,152.79 | 1,295.77 | 1,395.70 | 1,444.91 | 1,448.40 |
| - Wealthy farmers whole season (12 farmers) | 1,550.50 | 1,561.91 | 1,142.48 | 1,255.47 | 1,372.75 | 1,516.03 | 1,552.22 |
| - All farmers whole season (24 farmers) | 1,364.27 | 1,375.67 | 1,147.64 | 1,275.62 | 1,384.02 | 1,480.47 | 1,500.31 |
| **Amount of water allocated:** (mean number of units per farmer) | | | | | | | |
| - Poor farmers whole season (12 farmers) | 19.42 | 19.42 | 25.62 | 27.09 | 27.44 | 26.52 | 26.39 |
| - Wealthy farmers whole season (12 farmers) | 33.33 | 33.33 | 27.13 | 25.66 | 25.31 | 26.23 | 26.36 |
| - Total units allocated whole season (24 farmers) | 633 | 633 | 633 | 633 | 633 | 633 | 633 |

Notes: See Appendix C.7 for a discussion about the computation of the welfare measures.

the following allocation mechanisms: (1) auctions using complete units, $Ac$, wherein complete water units are assigned to the farmer who bought them as observed in the data; (2) quotas with random assignment of complete units, $Qc$, wherein every time we observe a farmer purchasing a unit of water under the auction system, the complete unit of water is assigned uniformly at random, proportional to their amount of land, among all farmers; (3) quotas with sequential assignment of complete units, $QcX\%$, wherein every time we observe a farmer purchasing a unit of water under the auction system, the complete unit of water is assigned uniformly at random, proportional to their amount of land, among the X% of farmers who did not receive irrigation the longest; and (4) the highest-valuation allocation using complete units, $HVc$, wherein every time we observe a farmer purchasing a unit of water under the auction system, the complete unit of water is assigned to the farmer who values the water the most. In all cases, $Ac$, $Qc$, $QcX\%$, and $HVc$, we compute the welfare measures using the actual allocation of water from the data under the auctions system (*i.e.* the total amount of water allocated in all mechanisms is the same), and the estimates from column 3 from Table 3.[29] As explained in Subsection 2.2, the quota system in Mula allocates units in sequential rounds of three weeks. So $Qc25\%$ is closest to this system. Below we describe how we compute the welfare measures under each mechanism. See Appendix C.7 for details.

### Auctions using Complete Units ($Ac$)

We compute revenue and welfare, as explained above. For both poor and wealthy farmers, we use the moisture level resulting from their actual purchase decisions.

### Quotas ($Qc$)

Revenue and welfare coincide under the quota system because farmers do not choose when to irrigate. We only report one measure that we call "welfare." As explained in Section 2, in this paper we focus on the 24 farmers who only grow apricot trees. These farmers bought

[29] We obtain similar results simulating the purchase decisions under the auction system, and then using the resulting allocation to compute the welfare under quotas and $HVc$.

633 units of water under the auction system during the period under analysis. Under the quota system, we allocate the same number of units of water, 633 units, in the same week when these units were bought under the auction. We consider several quota scenarios that we call $QcX\%$, and are defined as follows. Every time we observe that a farmer bought a unit of water during the auction on a particular date, the complete unit of water is assigned uniformly at random, proportional to their amount of land, among the $X\%$ of farmers who did not receive irrigation the longest, on the same date. In other words, units are allocated sequentially. Note that $Qc100\%$ corresponds to the special case of a quota with random assignment of complete units, called $Qc$ above. In Table 4 we report the mean welfare measures across simulations.

**Highest Valuation using Complete Units ($HVc$)**

We compute the highest-valuation allocation using complete units ($HVc$) as follows. Every time we observe that a farmer bought a unit of water during the auction on a particular date, the complete unit of water is assigned to the farmer who has the lowest moisture level on that date. This corresponds to the farmer who has been the longest without irrigation.

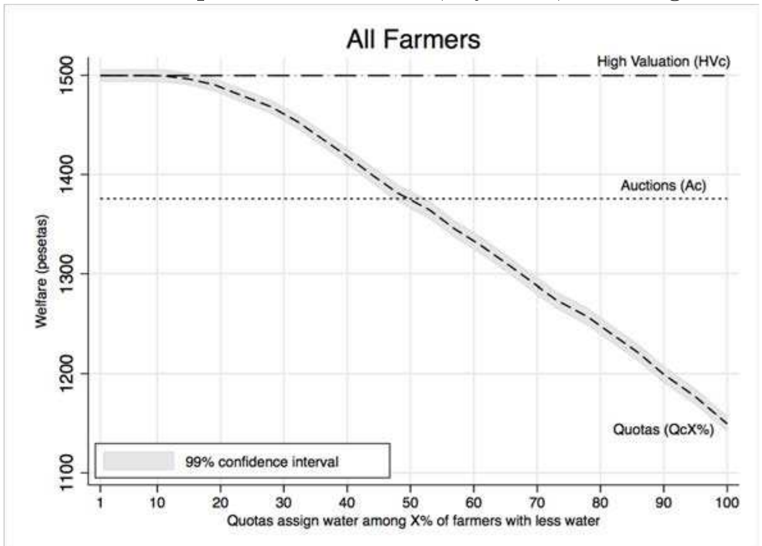
## 6.2 Welfare Results

Table 4 displays welfare results under auctions, quotas, and the $HVc$ allocation. For the welfare analysis in Table 4 we use the results from specification 3 in Table 3. We report the mean welfare per farmer, per tree, and per year. The bottom part of the table shows the mean number of units per farmer during the whole period under analysis under each mechanism. The total amount of water, 633 units, is the same across all mechanisms. The differences in welfare across columns are a consequence of differences in moisture levels across farmers.

As expected under the auction system, poor farmers have a lower welfare than wealthy farmers. The quota system increases poor farmers' welfare, and decreases wealthy farmers' welfare. Overall the quotas $Qc25\%$ produce 7.6% ($^{(1,480.47-1,375.67)}/_{1,375.67}$) more welfare per tree than auctions, $Ac$. Table 4 shows that the following ranking holds in terms of efficiency: $HVc > Qc25\% > Qc50\% \cong Ac$, where a "greater than" inequality indicates greater welfare, and where the symbol "$\cong$" indicates that the welfare is not statistically different. That is, randomly allocating the complete units of water, in proportion to amount of land, results in a decrease in efficiency relative to auctions. In $Qc50\%$, complete units of water are allocated among the 50 percent of farmers who have received less water in the past, in proportion to their amount of land. The welfare under $Qc50\%$ is not statistically different than the welfare under $Ac$. In $Qc25\%$, complete units of water are allocated among the 25 percent of farmers who have received less water in the past, in proportion to their amount of land. The welfare under $Qc25\%$ is greater than under $Ac$. In Mula, the quota allocation mechanism was similar to $Qc25\%$ because every farmer was assigned a certain amount of water every three weeks (the duration of a round, or *tanda*), proportional to their plot's size.

**Auctions, Quotas, and Highest Valuation.** Figure 2 shows the welfare comparison among auctions $Ac$, the $HVc$ allocation, and quotas $QcX\%$ for different values of $X$. Note that auctions $Ac$ and $HVc$ are constant in $X$. The figure shows the mean welfare per farmer, per tree, and per year. The welfare measures are the same as in Table 4. The main difference between $HVc$ and $Ac$ is that poor farmers do not buy much water during the critical season under $Ac$. Randomly allocating complete units of water results in a decrease in efficiency relative to auctions. Lower welfare results from the decreasing marginal returns of water of the apricot production function. Although all farmers receive the same amount of water per tree, the timing of the allocation is important. For example, consider the case of two identical farmers, A and B. Suppose that there are four units of water to be allocated in four consecutive weeks, 1, 2, 3, and 4. Allocating the first two units during weeks 1 and 2 to farmer A, and the second two units during weeks 3 and 4 to farmer B, results in a lower welfare than allocating the first unit to A, the second unit to B, the third unit to A, and the fourth to B. As $X$ decreases, the quota system $QcX\%$ allocates units among the farmers

Figure 2: Welfare Comparison: Auctions, Quotas, and Highest Valuation

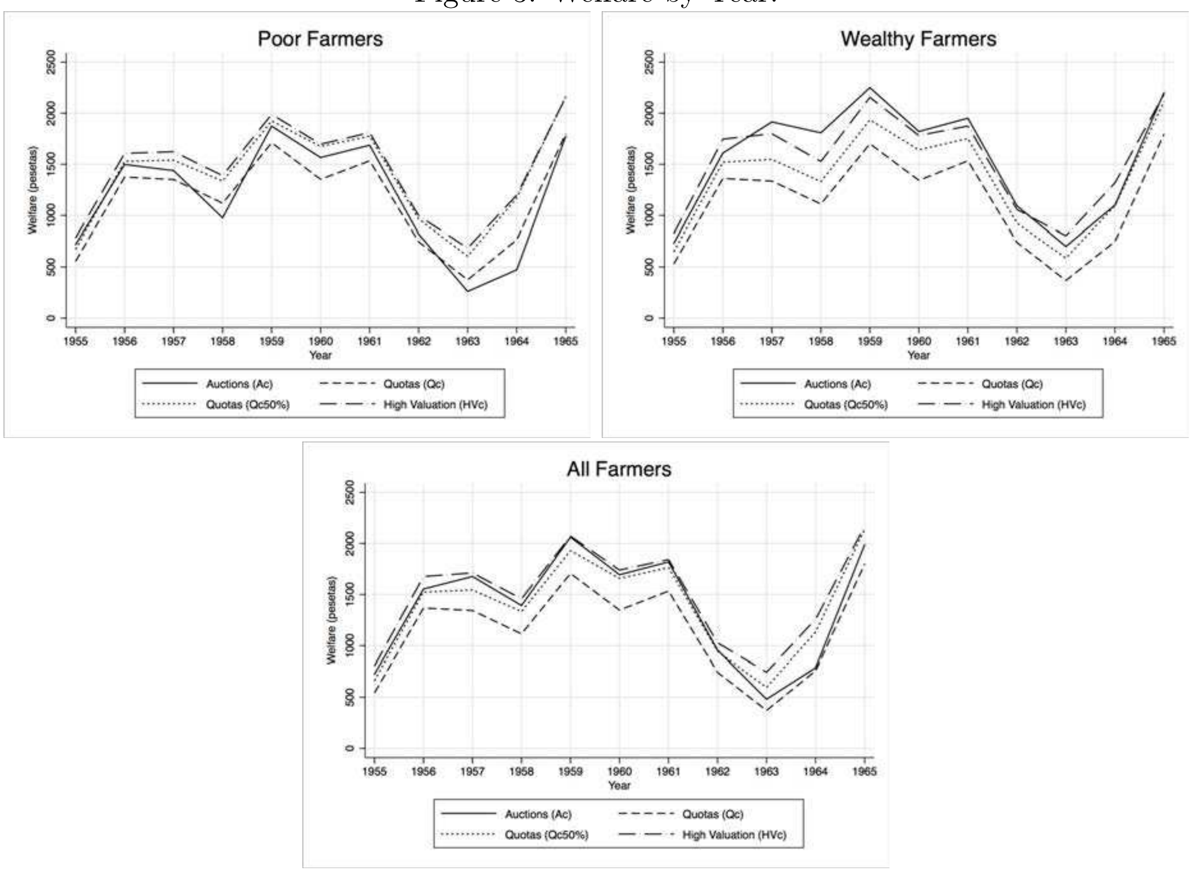



*Notes*: See Appendix C.7 for a discussion about the computation of the welfare measures. Confidence intervals account for uncertainty about the estimated parameters (by drawing from the asymptotic distribution) and across simulations.

who irrigated less in the past. This is similar to the $HVc$ allocation, where water is allocated to the farmer who values water the most. In the limit, as $X$ decreases enough, the welfare under $QcX\%$ is similar to the welfare under $HVc$. In our empirical setting, varying $X$ is equivalent to varying the duration of the round. Long rounds indicate that farmers do not irrigate often, while short rounds indicate that farmers incur irrigation costs often.

**Yearly Results.** Figure 3 shows the welfare results by year, from 1955 to 1965, and by allocation mechanism, $Ac$, $Qc$, $Qc50\%$, and $HVc$. There is substantial variation across years due to variation in rainfall. Revenue is the lowest for both poor and wealthy farmers during 1962-63, the driest years in our sample. The top two panels in Figure 3 display welfare disaggregated by poor and wealthy farmers under $Ac$, $Qc$, $Qc50\%$, and $HVc$. Although the overall performance of $Ac$ is similar to $Qc$, the distribution is different. Wealthy farmers perform better under $Ac$ than under $Qc50\%$, as expected. Poor farmers perform better under $Qc50\%$ than under $Ac$. During dry years, such as 1963 or 1964, poor farmers perform better under $Qc$ than under $Ac$. The difference between $Ac$ and $HVc$ is the highest in 1963, the year with the lowest rain in the sample. The drought increased the price of water relative to

Figure 3: Welfare by Year.




*Notes:* See Appendix C.7 for a discussion about the computation of the welfare measures.

the other years in the sample. The negative impact of this drought on poor farmers under $Ac$ was larger than its positive impact on wealthy farmers.

In this paper we analyzed the welfare implications of the institutional change from auctions to quotas for farmers who grew only apricots. Note that the welfare results do not apply necessarily to farmers who grew other crops, or to farmers who had a mix of several crops. In particular, farmers who had a summer (*e.g.* apricot) and a winter (*e.g.* oranges) crop can smooth spending throughout the year, and may not benefit much from the quotas.

# 7 Discussion

In this section we discuss the fit of our model to the empirical setting in Mula. See Appendix D for a thorough discussion.

**Unobserved Heterogeneity.** The production differences in Table 4 are attributable to differences in soil's moisture levels because our specification assumes that all farmers are equally productive, up to an idiosyncratic productivity shock. An alternative explanation would be that production differences are due to unobserved differences in productivity. For example, it could be that wealthy farmers used additional productive inputs, such as manure, in greater quantities than did poor farmers. Thus, poor farmers' production would be lower than wealthy farmers' production due to both differences in soil moisture levels, and greater use of these additional productive inputs.

Although we cannot rule out this argument explicitly, it does not affect our counterfactual results from Table 4. We cannot rule it out explicitly because we have no data about the relative use of these additional productive inputs, and our econometric specification does not allow for persistent differences in productivity among farmers. However, it does not affect our counterfactual results in the historical context of Mula. Artificial fertilizers were not introduced in Mula until the 1970s. Farmers did use manure and mules when farming the land. If poor farmers faced liquidity constraints when buying water, it is reasonable that they also faced liquidity constraints when buying these additional inputs. Therefore, if wealthy farmers used additional productive inputs in greater quantities than did poor farmers under the market system, then the transition from markets to quotas would increase the poor farmers' production more than we predicted in the counterfactual from Table 4. Under quotas, farmers do not have to make large payments for water, leaving them extra cash to buy additional productive inputs. In other words, poor farmers are less likely to be liquidity constrained to buy additional inputs under quotas. Thus, even if poor farmers were less productive than wealthy farmers under the market, they would be as productive as wealthy farmers under quotas. We further explore this issue in Appendix D.2, where we generalize the model and allow for correlation between farmers' wealth and productivity.[30]

**Liquidity Constraints *vs.* Risk Aversion or Impatience.** One concern when identifying liquidity constraints is the similar empirical implications when agents face liquidity constraints and when agents are risk averse. In particular, if poor farmers are more risk averse, their purchase of water before the critical season (*i.e. before* uncertainty about rain is realized) is consistent with both liquidity constraints and risk aversion. We could use the response of poor farmers to their purchase timing to investigate this concern. The main difference in farmers' behavior under liquidity constraints and risk aversion occurs during the summer, when prices are high. On the one hand, if poor farmers face liquidity constraints, they would not be able to buy water when the price is high, even if the moisture level in their plots is low. On the other hand, if farmers are unconstrained but risk averse, they would have the same demand for water as wealthy farmers during the summer (*i.e. after* uncertainty about rain is realized), conditional on soil's moisture levels. In Table 2, column 4 we show that holding the moisture level fixed, poor farmers buy less water than wealthy farmers. Following the results in this table, along with the opinions presented above, we conclude that poor farmers faced liquidity constraints. The same argument rules out the

[30] An alternative would be to allow for persistent differences in productivity among farmers, thus allowing for dynamic sample selection on unobservables. For some papers in this active area of research see, *e.g.*, Aguirregabiria and Mira (2007), Norets (2009), Arcidiacono and Miller (2011), Hu and Shum (2012), Blevins (2016), Connault (2016), and the references there.

possibility that the results are driven by poor farmers being more impatient (lower discount factor) than wealthy farmers. If poor farmers were more impatient, their moisture level would be always lower than that of wealthy farmers because an extra *peseta* spent on water has an immediate cost, and a future reward. However, poor farmers have higher moisture levels than wealthy farmers before the critical season, lower moisture levels during the critical season, and again higher moisture levels right after the critical season (Figure 1). This behavior rules out differences in discount factors between wealthy and poor farmers.

# 8 Concluding Remarks

We empirically investigate the welfare effect of a historical institutional change from markets to quotas in the presence of frictions. Both systems allocated water to farmers for irrigation. A market institution was active for more than 700 years in the southern Spanish town of Mula. In 1966, a fixed quota system replaced the auction. Under the quota system, farmers who owned a plot of land in the irrigated area were entitled to a fixed amount of water, proportional to the size of their plot. In the absence of frictions, a market is efficient because it allocates water according to the valuation of farmers. When frictions are present, however, markets may not be efficient. Frictions arose in Mula because farmers had to pay in cash for the water that they purchased, but did not always have enough cash during the critical season, when they needed the water the most. When farmers are liquidity constrained, the efficiency of auctions relative to quotas is theoretically ambiguous. It is then an empirical question to assess which institution is more efficient. We show that some farmers were liquidity constrained in Mula, as some historians have suggested. Poor farmers bought less water than wealthy farmers during the critical season, and obtained lower revenue per tree as a result. To compute welfare under auctions and quotas, we estimated the dynamic demand system under the auction accounting for three main features of the empirical setting: intertemporal substitution, liquidity constraints, and seasonality.

We used the estimated demand system to compare welfare under markets, quotas, and the highest-valuation allocation.

The contributions of this paper are twofold. First, from a historical perspective, we provide empirical evidence of a source of inefficiency in water markets. Second, from an industrial organization perspective, we propose a dynamic demand model that includes storability, seasonality, and liquidity constraints. Ignoring the presence of liquidity constraints biases the estimated inverse demand and demand elasticity downwards. To perform the estimation we used only the choices of farmers who were not liquidity constrained; then, we used the model to infer the conduct of all farmers in a counterfactual setting in which no one was liquidity constrained. Our approach may also be applied to other settings by identifying, for example, rational agents, informed consumers, buyers affected by inertia, or other frictions, and simulating counterfactuals where those features are not present.

Our analysis is subject to several limitations. First, the empirical results in this paper apply only to the empirical setting in Mula. One should not conclude that all water markets are inefficient. We have presented an empirical framework incorporating the main components found in other water markets: seasonal demand, storability, and liquidity constraints. Our framework can be adapted to assess the efficiency of water markets in other empirical settings. Second, the results from our welfare analysis in Mula only apply to the 24 apricot farmers used in our sample. Finally, our model does not allow for systematic (or permanent) differences in productivity across farmers. We believe this is sensible in the empirical context of Mula for the reasons discussed in the paper. Our goal was to investigate how the institutional change, from markets to quotas, affected efficiency for apricot farmers in Mula. However, allowing for systematic or serially correlated differences in productivity may generate dynamic sample selection on unobservables, and may be a central explanation in favor of the efficiency of markets in general. Enriching the model in such dimension is an avenue for future research.

# References


[1] Aguirregabiria, V., 2005, "Nonparametric Identification of Behavioral Responses to Counterfactual Policy Interventions in Dynamic Discrete Decision Processes," *Economics Letters*, Vol. 87, 393–398.

[2] Aguirregabiria, V., and Mira P., 2007, "Sequential estimation of dynamic discrete games," *Econometrica*, Vol. 75, No. 1, 1–53.

[3] Aguirregabiria, V., and Nevo, A., 2013, "Recent Developments in Empirical IO: Dynamic Demand and Dynamic Games," *Advances in Economics and Econometrics: Theory and Applications: Tenth World Congress*.

[4] Albuquerque, R., and Hopenhayn, H. A., 2004, "Optimal Lending Contracts and Firm Dynamics," *Review of Economic Studies*, Vol. 71, No. 2, 285-315.

[5] Arcidiacono, P., and Miller, R. A., 2011, "Conditional choice probability estimation of dynamic discrete choice models with unobserved heterogeneity," *Econometrica*, Vol. 79, No. 6, 1823-1867.

[6] Athey, A., Coey, D. and Levin, J., 2013, "Set-Asides and Subsidies in Auctions," *American Economic Journal: Microeconomics*, Vol. 5, No. 1, 1-27.

[7] Banerjee, A. V., and Moll, B., 2010, "Why Does Misallocation Persist?," *American Economic Journal: Macroeconomics*, Vol. 2, No. 1, 189-206.

[8] Berry, S., Levinsohn, J., and Pakes, A., 1995, "Automobile Prices in Market Equilibrium," *Econometrica*, Vol. 63, No. 4, 841-90.

[9] Blevins, J. R., 2016, "Sequential Monte Carlo Methods for Estimating Dynamic Microeconomic Models," *Journal of Applied Econometrics*, Vol. 31, No. 5, 773-804.

[10] Boizot, C., Robin, J. M., and Visser, M., 2001, "The demand for food products: an analysis of interpurchase times and purchased quantities," *Economic Journal*, Vol. 111, No. 470, 391-419.

[11] Bronnenberg, B. J., Dubé, J. P., Gentzkow, M., and Shapiro, J. M., 2015, "Do pharmacists buy Bayer? Informed shoppers and the brand premium," *Quarterly Journal of Economics*, Vol. 130, No. 4, 1669-1726.

[12] Bubb, R., Kaur, S., and Mullainathan, S., 2016, "Contracting Failures in the Village Economy," working paper.

[13] Che Y.-K., and Gale, I., 1998, "Standard Auctions with Financially-Constrained Buyers," *Review of Economic Studies*, Vol. 65, No. 1, 1-21.

[14] Che Y.-K., Gale, I., and Kim J., 2013, "Assigning Resources to Budget-Constrained Agents," *Review of Economic Studies*, Vol. 80, No. 1, 73-107.

[15] Coman, K., 1911, "Some Unsettled Problems of Irrigation," *American Economic Review*, Vol. 1, No. 1, 1-19.

[16] Connault, B., 2016, "Hidden rust models," working paper.

[17] Donna, J. D., and Espín-Sánchez, J. A., 2017, "Let the Punishment Fit The Criminal," working paper.

[18] Donna, J. D., and Espín-Sánchez, J. A., 2018, "Complements and Substitutes in Sequential Auctions: The Case of Water Auctions," *RAND Journal of Economics,* Vol. 49, No. 1, 87–127.

[19] Espín-Sánchez, J. A., 2017, "Institutional Institutional Inertia: Persistent Inefficient Institutions in Spain," *Journal of Economic History*, Vol. 77, No. 3, 692-723.

[20] Garrido, S., 2011, "Governing scarcity. Water Markets, Equity and Efficiency in pre-1950s Eastern Spain," *International Journal of the Commons*, Vol. 5, No. 2., 513-534.

[21] González Castaño, J., and Llamas Ruiz, P., 1991, *El agua en la Ciudad de Mula, S. XVI-XX.,* Comunidad de Regantes Pantano de la Cierva*,* Mula.

[22] Gowrisankaran, G., and Rysman, M., 2012, "Dynamics of Consumer Demand for New Durable Goods," *Journal of Political Economy*, Vol. 120, No. 6, 1173-1219.

[23] Guinnane, T. W., 2001, "Cooperatives As Information Machines: German Rural Credit Cooperatives, 1883 1914," *Journal of Economic History*, Vol. 61, No. 2, 366-389.

[24] Grafton, R. Q., Libecap, G., McGlennon, S., Landry, C., and O'Brien, B., 2011, "An Integrated Assessment of Water Markets: A Cross-Country Comparison," *Review of Environmental Economics and Policy*, Vol. 5, No. 2, 219-239.

[25] Handel, B. R., and Kolstad, J. T., 2015, "Health Insurance for "Humans": Information Frictions, Plan Choice, and Consumer Welfare," *American Economic Review*, Vol. 105, No. 8, 2449-2500.

[26] Hendel, I., and Nevo, A., 2006, "Measuring the implications of sales and consumer inventory behavior," *Econometrica*, Vol. 74, No. 6, 1637-1673.

[27] Hotz, J., and Miller, R., 1993, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, Vol. 60, No. 3, 497-529.

[28] Hotz, J., Miller, R., Sanders, S., and Smith, J., 1994, "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies*, Vol. 61, No. 2., 265-289.

[29] Hu Y., and Shum M., 2012, "Nonparametric identification of dynamic models with unobserved state variables," *Journal of Econometrics*, Vol. 171, 32–44.

[30] Jayachandran, S., 2013, "Liquidity Constraints and Deforestation: The Limitations of Payments for Ecosystem Services," *American Economic Review Papers and Proceedings*, Vol. 103, No. 2, 309-313.

[31] Jofre-Bonet, M., and Pesendorfer, M., 2003, "Estimation of a Dynamic Auction Game," *Econometrica*, Vol. 71, No. 5, 1443-1489.

[32] Johnson, E. M., and Rehavi, M. M., 2016, "Physicians treating physicians: Information and incentives in childbirth," *American Economic Journal: Economic Policy*, Vol. 8, No. 1, 115-141.

[33] Ketcham, J. D., Kuminoff, N. V., and Powers, C. A., 2016, "Estimating the heterogeneous welfare effects of choice architecture: An application to the medicare prescription drug insurance market," NBER working paper No. w22732.

[34] Laffont, J.-J., and Matoussi, M. S., 1995, "Moral Hazard, Financial Constraints and Sharecropping in El Oulja," *Review of Economic Studies*, Vol. 62, No. 3, 381-399.

[35] Libecap, G. D., 2011, "Institutional Path Dependence in Adaptation to Climate: Coman's "Some Unsettled Problems of Irrigation," *American Economic Review*, Vol. 101, No. 1, 64-80.

[36] Maass, A., and Anderson, R. L., 1978, *... and the Desert Shall Rejoice: Conflict, Growth, and Justice in Arid Environments*, The MIT Press.

[37] Magnac, T., and D. Thesmar, 2002, "Identifying Dynamic Discrete Decision Processes," *Econometrica*, Vol. 70, No. 2, 801–816.

[38] Marion, J., 2007, "Are bid preferences benign? The effect of small business subsidies in highway procurement auctions," *Journal of Public Economics*, Vol. 91, 1591-1624.

[39] Musso y Fontes, J., 1847, *Historia de los riegos de Lorca*, Imprenta José Carle-Palacios.

[40] Moll, B., 2014, "Productivity Losses from Financial Frictions: Can Self-Financing Undo Capital Misallocation?," *American Economic Review*, Vol. 104, No. 10, 3186-3221.

[41] Norets A., 2009, "Inference in dynamic discrete choice models with serially correlated unobserved state variables," *Econometrica*, Vol. 77, 1665–1682.

[42] Ohio Supercomputer Center. 1987, "Columbus OH: Ohio Supercomputer Center," http://osc.edu/ark:/19495/f5s1ph73.

[43] Pérez-Pastor, A., Domingo, R., and Torrecillas, A., 2009, "Response of apricot trees to deficit irrigation strategies," *Irrigation Science*, Vol. 27, 231-242.

[44] Pérez Picazo, Ma. T., and Lemeunier, G., 1985, "Agua y coyuntura económica. Las transformaciones de los regadíos murcianos (1450-1926)," *Cuadernos de Geografía Humana*.

[45] Pires, T., and Salvo, A., 2015, "Cash-Constrained Households and Product Size," working paper.

[46] Rosenzweig, M. R., and Wolpin, K. I., 1993, "Constraints, Consumption Smoothing, and the Accumulation of Durable Production Assets in Low-Income Countries: Investments in Bullocks in India," *Journal of Political Economy*, Vol. 101, No. 2, 223-244.

[47] Rust, J., 1996, "Numerical Dynamic Programming in Economics," in *Handbook of Computational Economics*, Vol. 1, ed by. H. Amman, D. Kendrick, and J. Rust. Amsterdam: North- Holland, 619–729.

[48] Timmins, C., 2002, "Measuring the Dynamic Efficiency Costs of Regulators' Preferences: Municipal Water Utilities in the Arid West," *Econometrica*, Vol. 70, No. 2, 603-629.

[49] Torrecillas, A., Domingo, R., Galego, R., and Ruiz-Sánchez, M.C., 2000, "Apricot tree response to withholding irrigation at different phenological periods," *Scientia Horticulturae*, Vol. 85, 201-205.

[50] Udry, C., 1994, "Risk and Insurance in a Rural Credit Market: An Empirical Investigation in Northern Nigeria," *Review of Economic Studies*, Vol. 61, No. 3, 495-526.

[51] Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., Glidden, S., Bunn, S. E., Sullivan, C. A., Reidy Liermann, C., and Davies, P. M., 2010, "Global threats to human water security and river biodiversity," *Nature*, Vol. 467, 555-561.

# Appendix to "The Illiquidity of Water Markets: Efficient Institutions for Water Allocation in Southeastern Spain" (For Online Publication)

Javier D. Donna and José-Antonio Espín-Sánchez

The Ohio State University | Yale University

Javier D. Donna
Department of Economics
The Ohio State University
1945 N High St, 425 Arps Hall
Columbus, OH 43210
Phone: 614-688-0364
Email: donna.1@osu.edu

José-Antonio Espín-Sánchez
Department of Economics
Yale University
27 Hillhouse Ave, Room 38
New Haven, CT 06511-3703
Phone: 203-432-0890
Email: jose-antonio.espin-sanchez@yale.edu

# Contents

**A Additional Description of the Data, Moisture, and Preliminary Analysis A-4**

A.1 Additional Description of the Data. . . . . . A-4

A.1.1 Auction Data. . . . . . A-4

A.1.2 Rainfall Data. . . . . . A-7

A.1.3 Agricultural Census Data. . . . . . A-7

A.1.4 Real Estate Tax Data. . . . . . A-8

A.1.5 Summary Statistics. . . . . . A-9

A.2 Additional Description of the Moisture and the Production Function. . . . . A-10

A.2.1 Evapotranspiration Under Hydric Stress. . . . . . A-12

A.2.2 Details about the Apricot Production Function. . . . . . A-15

A.3 Additional Preliminary Analysis. . . . . . A-16

**B Welfare Comparison of Markets *vs.* Quotas A-18**

**C Details about the Estimation Procedure and the Structural Model A-21**

C.1 Demand Estimates. . . . . . A-21

C.2 Properties of the Demand Estimator. . . . . . A-24

C.3 Specification of the Productivity Shock. . . . . . A-27

C.3.1 Specification with *i.i.d.* shocks across choice alternatives. . . . . . A-27

C.3.2 Specification with shocks non *i.i.d.* across choice alternatives. . . . . A-28

C.4 Serial Correlation Over Time in Rainfall and Auction Water Price. . . . . . A-29

C.5 Demand Estimates: Estimates of Additional Specifications of the Model. . . A-34

C.6 Probability of Being Liquidity Constrained. . . . . . A-36

C.7 Welfare Measures. . . . . . A-42

C.8 Alternative Estimation Methods and Fit to the Data. . . . . . A-46

C.9 Dynamic Demand System. . . . . . A-50

C.10 Dynamic Auction Model with Within-Week Price Variation. . . . . . A-51

**D Extended Discussion about the Fit of the Model to the Data A-53**

D.1 Unobserved Heterogeneity. . . . . . A-53

D.2 Correlation Between Wealth and Productivity. . . . . . . . . . . . . . . . . . . A-54
D.3 Strategic Supply. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A-60
D.4 Strategic Size and Sunk Cost. . . . . . . . . . . . . . . . . . . . . . . . . . A-60
D.5 Optimal Mix of Crops and Size. . . . . . . . . . . . . . . . . . . . . . . . . A-61
D.6 Trees. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A-62
D.7 Collusion. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A-62
D.8 Liquidity Constraints *vs.* Risk Aversion or Impatience. . . . . . . . . . . . . A-63
D.9 Attrition. . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . . A-64

# Appendix

This is the web appendix for "The Illiquidity of Water Markets: Efficient Institutions for Water Allocation in Southeastern Spain," by Javier D. Donna and José-Antonio Espín-Sánchez.

## A Additional Description of the Data, Moisture, and Preliminary Analysis

In this section we present detailed information regarding the data collection, and the moisture computation.

### A.1 Additional Description of the Data.

#### A.1.1 Auction Data.

The mechanism to allocate water to the farmers was a sequential outcry ascending price (or English) auction. The auctioneer sold each of the units sequentially and independently of each other. The auctioneer recorded the name of the buyer of every unit and the price paid by the winner. The farmers could not store water in their plots. Reselling water was forbidden.

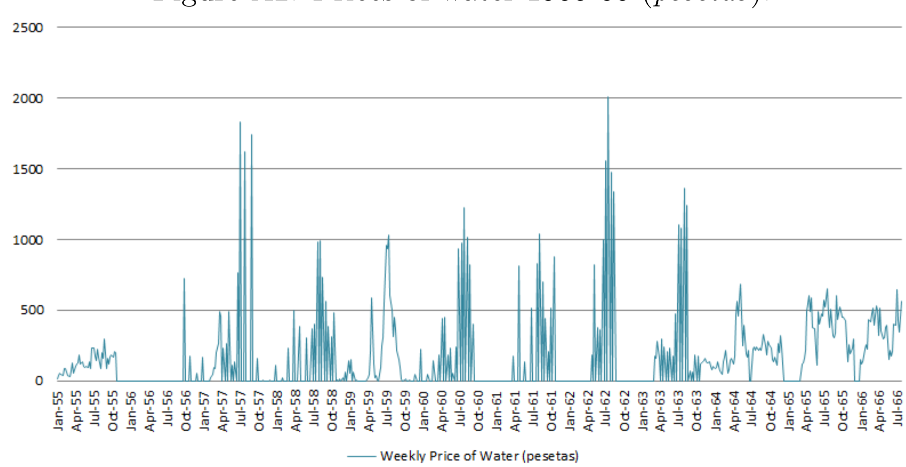
The basic selling unit was a *cuarta* (quarter), the right to use water that flowed through the main channel for three hours. Water was stored at the De La Cierva dam. Water flowed from the dam through the channels at approximately 40 liters per second. As a result, one *cuarta* carried approximately 432,000 liters of water. During our sample period, auctions were carried out every Friday. Every week, 40 *cuartas* were auctioned: four *cuartas* for irrigation during the day (from 7:00 AM to 7:00 PM) and four *cuartas* for irrigation during the night (from 7:00 PM to 7:00 AM), for each weekday (Monday to Friday). The auctioneer first sold the 20 *cuartas* corresponding to the night-time, and then the 20 *cuartas* corresponding to the day-time. Within each day and night group, units were sold beginning with Monday's four *cuartas*, and finishing with Friday's.

Auction data encompasses 602 weeks, and can be divided into three categories based on bidding behavior and water availability: (i) Normal periods (300 weeks), when for each transaction the name of the winner, price paid, date and time of the irrigation was registered for each auction; (ii) No-supply periods (295 weeks), when due to water shortage in the river or damage to the dam or channel—usually because of intense rain—no auction was carried out; and (iii) No-demand periods (7 weeks), when some units were not sold due to lack of demand due to recent rain, and the price dropped to zero. The sample for the empirical analysis focuses on the period from 1955 until 1966.

Figure A1 shows a sample from original data for May 17, 1963, obtained from the historical archive. Units 1 to 4 are the units bought on Monday during day (unit 1 corresponds to the right to irrigate from 7AM to 10AM, unit 2 from 10AM to 1PM, unit 3 from 1PM to 4PM, and unit 4 from 4PM to 7PM). Similarly, units 5 to 8 are the units bought on Tuesday during day; units 9 to 12 are the units on Wednesday during day; units 13 to 16 are the units on Thursday during day; and units 17 to 20 are the units on Friday during day. From the data we observe the name of the farmer who won each of the auctions, and the price paid by each farmer.

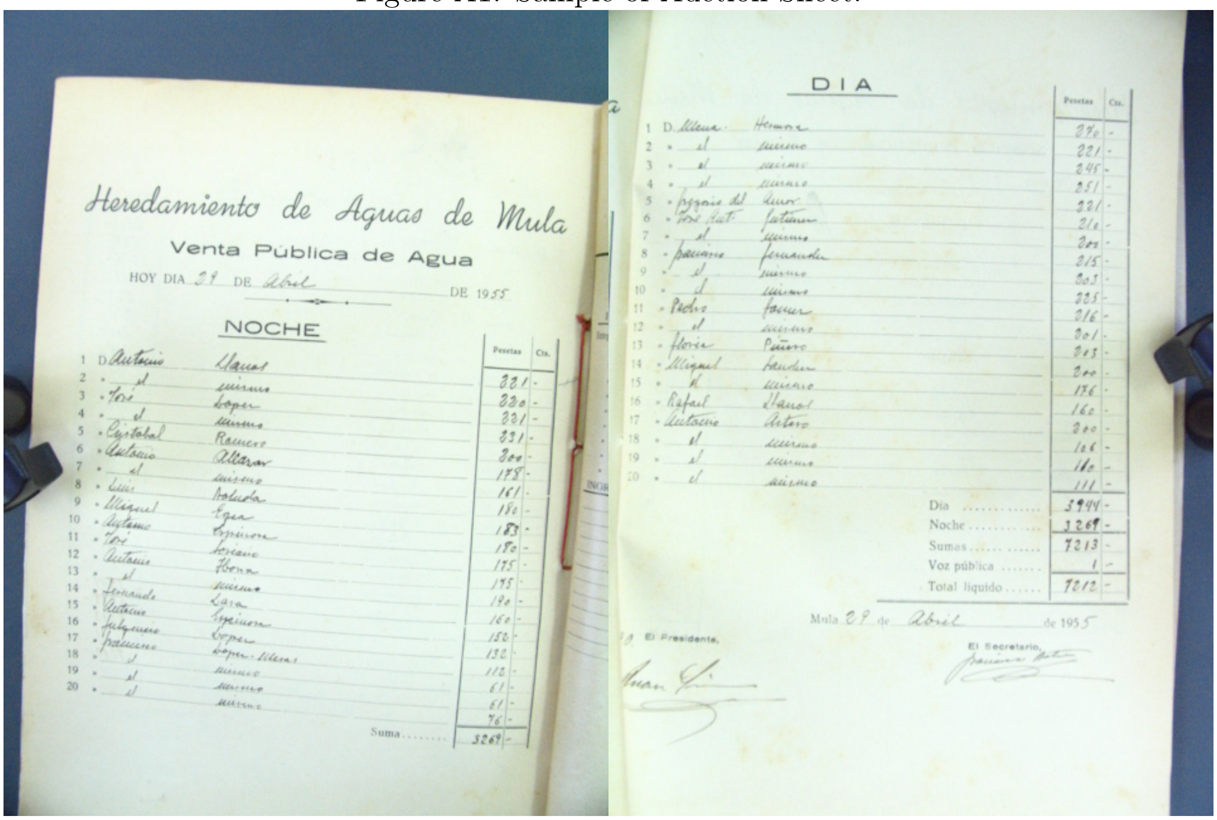
Figure A2 shows the weekly average price paid by the farmers during our sample period. There is substantial variation in prices, that range from 0.005 to 2007 *pesetas*. In the fall of 1955, a large flood damaged the dam for several months. Thus, auctions were not run until the next fall. In some dry years, like 1961-63, auctions were not run in winter, causing the prices to soar in spring and summer.

Figure A2: Prices of water 1955-66 (*pesetas*).



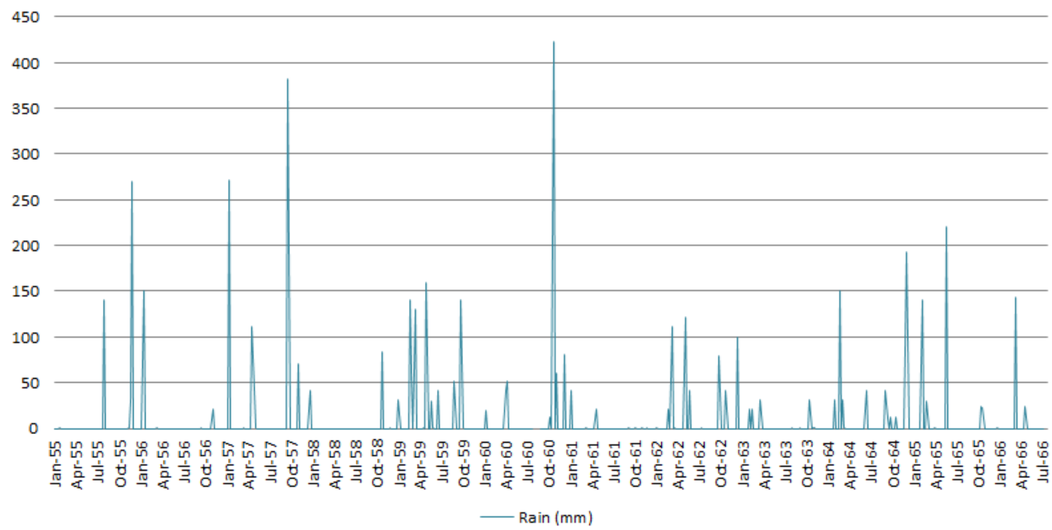

*Notes:* Weekly average price of the water sold at auction in Mula, from January, 1955 until July 1966, when the last auction was run.

Figure A1: Sample of Auction Sheet.




*Notes:* Sample pictures of the data from the Municipal Archive in Mula, Section *Heredamiento de Aguas* (HA). This pictures correspond to the same sheet of paper, containing the information of the winners and price paid for the 40 units sold on April 29, 1955.

Figure A3: Weekly Rainfall in Mula ($mm$).




*Notes:* Weekly rainfall in Mula from date from the AEMET.

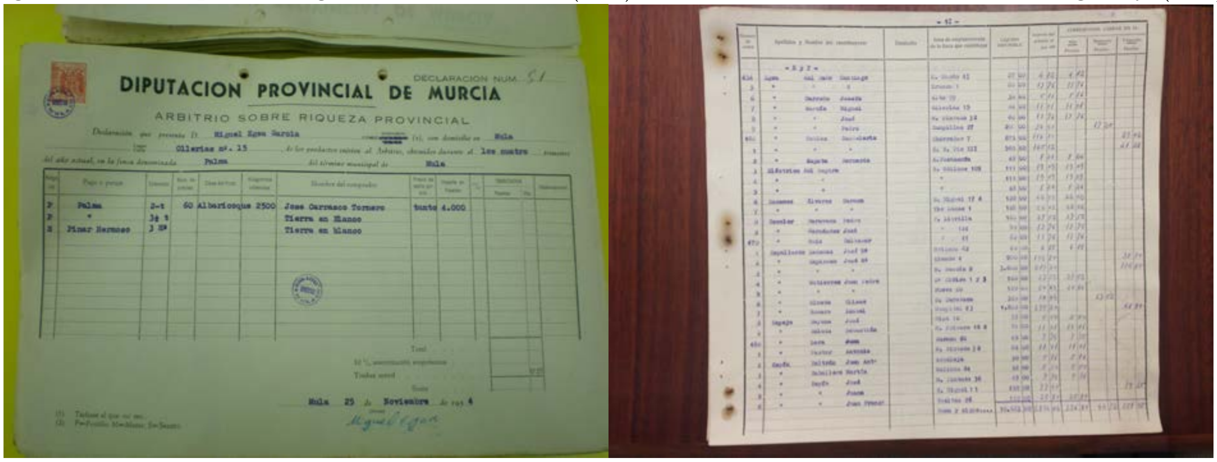
### A.1.2 Rainfall Data.

We also link auction data to daily rainfall data for Mula, which we obtain from the *Agencia Estatal de Metereología*, AEMET (the National Meteorological Agency). Mediterranean climate rainfall occurs mainly in spring and fall while peak water requirements for products cultivated in the region are reached in spring and summer. During these months, from April to August, more frequent irrigation is recommended because the tree's production quality is more sensitive to water deficits. Figure A3 shows that there are only a few weeks with positive rainfall. In our sample, the weekly rainfall exceeded the yearly average on two occasions, in September 1957 and in October 1960.

### A.1.3 Agricultural Census Data.

We also link auction data to the data that we collected from the 1954/55 agricultural census from Spain, which provides information on individual characteristics of farmers' land. The census was conducted by the Spanish government to enumerate all cultivated soil, production crops, and agricultural assets available in the country. Individual characteristics of farm land owned by potential bidders, who we link with the names in the auctions data, include the type of land and location, area, number of trees, production, and the price at which this production was sold in the census year. There are approximately 500 different bidders in our

Figure A4: Samples of Agricultural Census (left) and Urban Real Estate Taxes registry (left).




*Notes:* Sample pictures of the data from the Municipal Archive in Mula. Left: card from the Agricultural Census in 1955. The farmer (Miguel Egea Garcia) lived in Mula (15 Ollerias). He was an owner of three plots, one one of them uncultivated, with an extension of two Tahúllas, with 60 apricot trees. In 1954 he obtained 2,500 *Kg* of apricots which he sold for 4,000 *pesetas* in bulk. Right: sheet from the Urban Real Estate Taxes registry, corresponding to 1954. The citizen in registry 458 (Miguel Egea Garcia) owned a house in 15 Ollerias for which he paid 64 *pesetas* in taxes.

sample.

Figure A4 (left) shows a sample card of a farmer taken from the agricultural census data. It can be seen in Table 1 in the paper, that *Area* and *the number of trees* vary considerably across farmers. For the case of apricot-only farmers, on average each farmer had 86 trees and bought 31.5 units of water during the period 1955-66.

As regards the composition of the farmers' plots in Mula, the most common agricultural trees were oranges (33 percent) and apricots (29 percent), followed by lemons (12 percent) and peaches (5 percent). These farmers grew a wide variety of vegetables, including tomatoes, red peppers, cucumbers, and also potatoes. Vegetables and potatoes were complementary to the trees. Fruit trees produce greater returns than vegetables, but require irrigation at specific times of the year, and up to five years to reach maturity. By contrast, vegetables can be harvested a few months after being sowed, but they have lower returns. Hence, they can produce high output during a rainy year, and their cost of drying up during drought is low because they can be sown again the year after.

#### A.1.4 Real Estate Tax Data.

Figure A4 (right) shows a sample sheet taken from the urban real estate tax registry. We can see how the name and addresses match uniquely, so we know the person in the auction data, the agricultural census and the tax registry is indeed the same farmer. The data is held at the Municipal Archive in Mula, in the General Section. It contains the public records of the real estate income taxes paid annually for each individual who owned an urban property in the town of Mula. We first link the names in the auction data to those in the agricultural census data. Then we link those names to the urban real estate data. The value in the real estate data records corresponds to the taxable income for urban real estate only. Farmers had to pay an annual tax equal to 17% of the taxable income. That is, 17% of the rental value of the properties, not the stock value of the properties. The rural real estate holdings were subject to different taxes and are kept in a different directory.

Among the poor farmers, 10 out of 12 owned no urban real estate, the other two that we consider "poor" owned a small house, with a taxable value that is approximately one hundred times lower than that of the wealthiest of the farmers. In our estimates we found that the two "poor" farmers who owned a house were also never liquidity constrained. The average taxable base for urban real estate is 969 pesetas for the wealthy farmers and 6 *pesetas* for the poor farmers.

The taxable base is useful for relative comparisons, because it uses the same formula for all urban properties. This is amplified in the case of apricot-only farmers because the comparison is between farmers (ten out of twelve were poor) who mostly owned no urban real estate at all, making the taxable base particularly indicative of the wealth difference between the farmers. Note, however, that the conversion from taxable base to actual value is not straightforward. The taxable base is equivalent to the estimated annual rent (net from maintenance costs) that the owner could get from their property. The average value of the tax base of a house in the town was about 40 *pesetas*. The values that we obtain from wealthy farmers are much higher because they include multiple properties, and for the case of the very wealthy, mansions.

#### A.1.5 Summary Statistics.

Table A1 shows the summary statistics of selected variables used in the empirical analysis.

Table A1: Summary Statistics of Selected Variables.

| Variable | Mean | St. Dev. | Min | Med | Max | No. obs. |
|---|---|---|---|---|---|---|
| Weekly rain (*mm*) | 8.29 | 37.08 | 0 | 0 | 423.00 | 602 |
| Water price (*pesetas*)[a] | 326.16 | 328.45 | 0.005 | 217.9 | 2,007 | 602 |
| Real estate tax (*pesetas*) | 482.10 | 1,053.6 | 0 | 48 | 8,715 | 496 |
| Area (*ha*) | 2.52 | 5.89 | 0.024 | 1.22 | 100.1 | 496 |
| Number of trees[b] | 311.3 | 726.72 | 3 | 150 | 12,360 | 496 |
| Units bought | 0.0295 | 0.3020 | 0 | 0 | 4 | 145,684 |

*Notes*: The sample refers to all farmers. There are 496 census cards in the archive. We matched 242 individuals to the auction data. The agricultural census include farmers who have only *secano,* or dry, lands and thus, are not in our sample. The sample after the matching process consists of 602 weeks, and 242 individuals for a total of 145,684 observations.
[a] Water price is the weekly average price in the auction. [b] Number of trees includes vines.

## A.2 Additional Description of the Moisture and the Production Function.

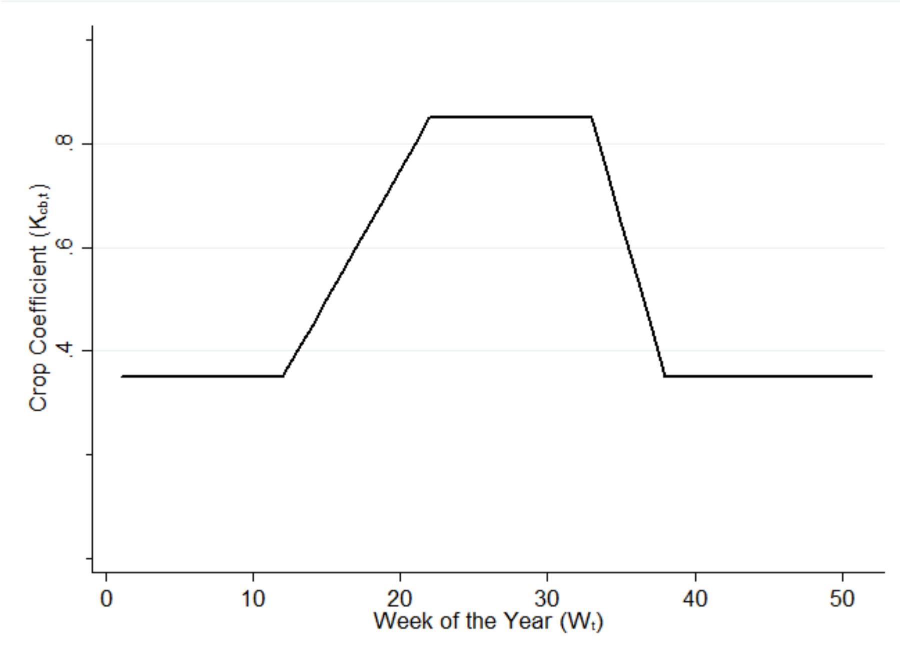
This section follows closely Allen *et al.* (2006). Trees are traditionally positioned in a square grid, each trunk 9 meters ($m$) from each other. Hence, there is a tree for every 81 $m^2$. This corresponds to our data in which, for apricot trees, the average ratio of trees per $m^2$ is 79.96 $m^2/tree$ and the ratio between total number of trees and total area is 78.25 $m^2/tree$. These numbers are slightly smaller than 81 $m^2/tree$ because some farmers place some trees very close to the edge of their plot.

Evapotranspiration (ET) is the loss of water suffered by the trees due to both Evaporation (E) of the water stored underground and Transpiration (T) of the water stored within the plant through the surface of the leaves. We use the method recommended by the Food and Agriculture Organization (FAO) to compute the evolution of the moisture due to ET:

$$ET_c = K_c \cdot ET_0,$$

where $ET_c$ is the weekly ET of crop $c$, $ET_0$ is the weekly reference ET and $K_c$ is the crop coefficient. Both $ET_c$ and $ET_0$ are measured as millimeters per week ($mm/week$). ET is affected by climatic factors: radiation, air temperature, atmospheric humidity, and wind speed. The effect of those parameters is summarized in $ET_0$. We use the estimations of $ET_0$

Figure A5: Evolution of the crop coefficient $K_{cb,t}$ over a year.




*Notes*: This figure represents the relationship between the crop coefficient $K_{cb,t}$ and the weekly calendar, for apricot trees in southern Spain. The parameters are taken from Allen *et al.* (2006). The graph is similar to Figure 37 in Allen *et al.* (2006).

in Franco *at al.* (2000).

ET would also change depending on the phase of the growing cycle:

$$ET_{cb,t} = K_{cb,t} \cdot ET_0.$$

We can then distinguish four phases (initial, development, median, and final) in the growing season. Following (Allen *et al.*, 2006, p. 107) we have that $L_{ini} = 20$, $L_{dev} = 70$, $L_{med} = 120$, and $L_{fin} = 60$; 270 days in total, finishing at the critical season. The coefficient $K_{cb,t}$ is flat during the initial period (with $K_{cb,ini} = 0.35$). It is linearly increasing during the development period until it reaches the median period. It is flat during the median period (with $K_{cb,med} = 0.85$). It is linearly decreasing during the final period (with $K_{cb,fin} = 0.60$ on average). It is then linear during the no-growth period until it reaches the initial period during the next year at $K_{cb,ini}$.[1] Figure A5 displays the evolution of the coefficient $K_{cb,t}$ over a year.

[1] Allen *et al.* (2006) formula (66). Coefficients for apricot trees without soil cover, with potential frosts, Table 17 (page 140).

### A.2.1 Evapotranspiration Under Hydric Stress.

$ET_c$ refers to the ET of crop $c$ under standard conditions. We should nonetheless adjust the value of $ET_c$ ($ET_{c,adj}$) when those conditions are not met. When the soil is wet, the water has a high potential energy, meaning that it can be easily absorbed by the roots of the tree. When the soil is dry, water is not so easily absorbed by the roots. When the moisture of the plot falls below a certain threshold, we say that the crop is under Hydric Stress (HS). The effects of HS are incorporated by multiplying $K_{cb}$ by the Hydric Stress coefficient $KS$:

$$ET_{c,adj} = KS \cdot K_{cb} \cdot ET_0.$$

Water availability refers to the ability of soil to keep water available for plants. After a heavy rain or irrigation, the soil will drain water until the full capacity is reached. The Full Capacity (FC) of a soil represents the moisture that a well drained soil keeps against gravitational forces, *i.e.*, the moisture of a soil when the downward vertical drainage has decreased substantially. In our case:

$$FC = 1000 \cdot \theta_{FC} \cdot Z_r,$$

where $\theta_{FC}$ is the moisture content of the soil at Full Capacity ($m^3\ m^{-3}$) and $Z_r$ is the depth of the tree's roots ($m$).

In absence of a source of water, the moisture in the soil will decrease due to the water consumption of the tree. As this consumption increases, the moisture level will go down, making it harder for the tree to absorb the remaining water. Eventually, a point will be reached beyond which the tree can no longer absorb any water: the Permanent Wilting (PW) point. The PW point is the moisture level of the soil at which the tree will permanently die. In our case:

$$PW = 1000 \cdot \theta_{PW} \cdot Z_r,$$

where $\theta_{WP}$ is the moisture content of the soil at the Permanent Wilting point ($m^3\ m^{-3}$) and $Z_r$ is the depth of the tree's roots ($m$).

Moisture levels above FC cannot be sustained, given the effect of gravity. Moisture levels below PW cannot be extracted by the roots of the trees. Hence, the Total Available Water (TAW) will be the difference between both:

$$TAW = FC - PW,$$

$Z_r = 4m$ in the case of apricot trees irrigated with traditional flooding methods. The soil in Murcia is limestone, hence $(\theta_{FC} - \theta_{PW}) \in [0.13, 0.19]$ and $\theta_{PW} \in [0.09, 0.21]$. For our estimation we take the middle point, *i.e.*, $FC = 1240$, $PW = 600$ and $TAW = 640$.

In practice, the tree will absorb water from the soil at a slower rate, even before reaching the PW point. When the tree is under HS, the tree is not absorbing water at the proper rate. The fraction of water that the tree can absorb without suffering HS is the Easily Absorbed Water (EAW):

$$EAW = p_c TAW,$$

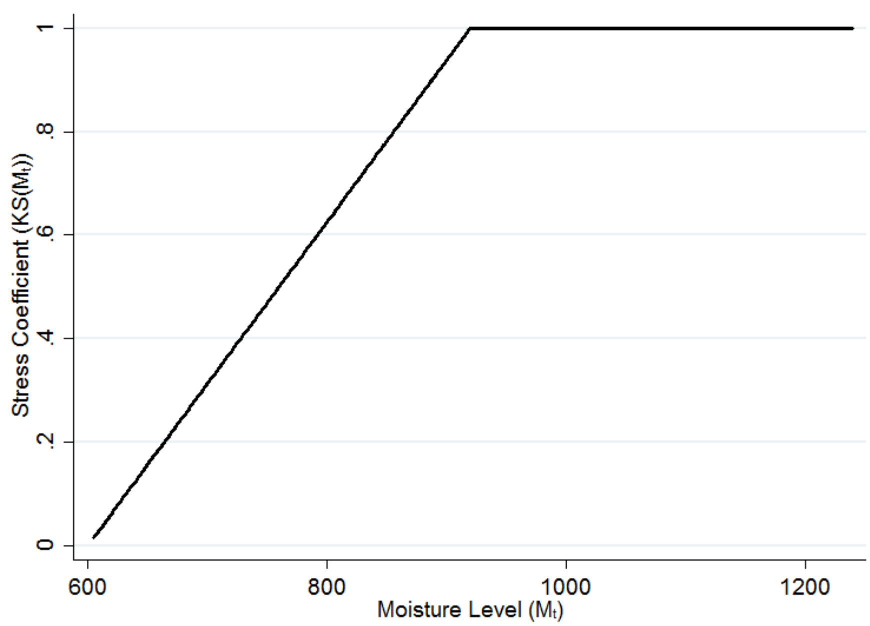
where $p_c \in [0, 1]$. For the case of the apricot tree $p_c = 0.5$, thus $EAW = 320$. The Hydric Stress coefficient $KS \equiv KS\left(M_t\right)$ is a function of the moistness of the plot $M_t$:

$$KS\left(M_t\right) = \begin{cases} 1 & \text{if} \quad M_t > FC - TAW(1 - p_c) \\ \frac{M_t - PW}{EAW} & \text{if} \quad FC - TAW(1 - p_c) \geq M_t > PW \\ 0 & \text{if} \quad M_t \leq PW \end{cases} . \tag{A.1}$$

Figure A6 shows the evolution of the coefficient of hydric stress for apricot trees, according to equation A.1. When the moisture level in the soil is below the PW point (600 $mm$), the plant dies and there is no transpiration. When the moisture level is sufficiently high (920 $mm$), the tree does not suffer from hydric stress and therefore the transpiration is maximal. When the soil has enough moisture for the tree to survive ($M_t > PW$), but not enough for the tree to function normally ($M_t < FC - TAW(1 - p_c)$), the tree suffers from hydric stress. Hydric stress makes the tree transpire less that it would otherwise.

Adding the subscripts for the periods we can write:

Figure A6: Relation between the hydric stress coefficient $KS$ and the moisture level for apricots.




*Notes*: This figure represents the relation between the hydric stress coefficient, $KS$, and the level of moisture in the soil, for apricot trees in southern Spain. The parameters are taken from Allen *et al.* (2006). The graph is similar to Figure 42 in Allen *et al.* (2006).

$$ET_{c,adj,t}\left(M_t\right) = KS\left(M_t\right) \cdot K_{cb,t} \cdot ET_0. \tag{A.2}$$

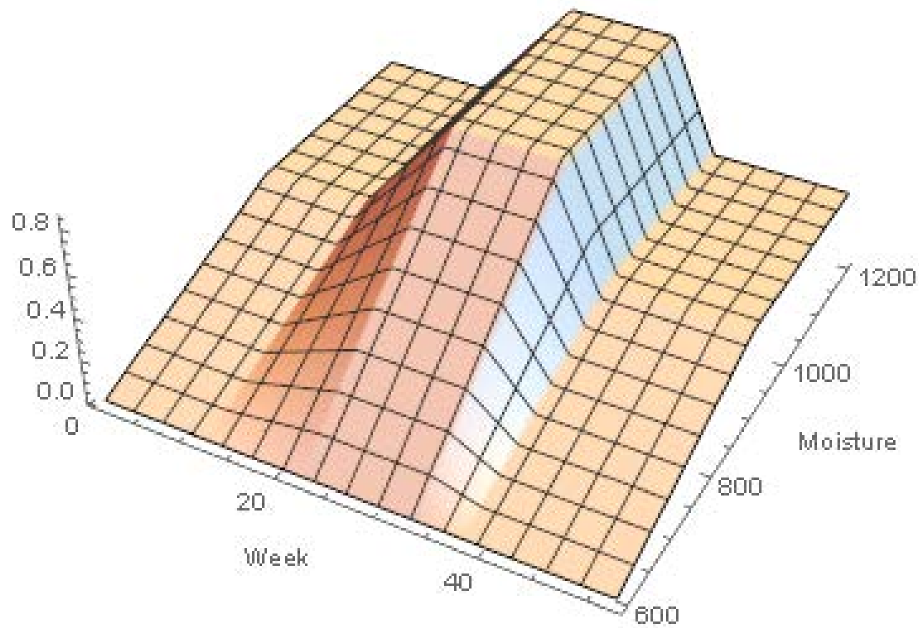
Figure A7 shows the combined effects of seasonality and hydric stress on the evapotranspiration coefficient, following equation A.2.

Finally, we have to take into account that, regardless of the amount of rain or irrigation, the moistness of the soil can never get beyond the $FC$. The evolution of the moisture $M_t$ over time is:

$$M_t = min\left\{M_{t-1} + rain_{t-1} + irrigation_{t-1} - ET_{c,adj,t-1}\left(M_{t-1}\right), FC\right\}.$$

We get an average value for $ET_c$ of 8.77, which is smaller than Franco *et al.* (2000) who find values of 23.1-30.8 $mm$ per week (3.3-4.3 $mm$ per day) for almond trees in the same region. Pérez-Pastor *et al.* (2009) report an Evapotranspiration of 1,476 $mm$ per year (28.4 $mm$ per week). This difference is due to the fact that recent studies are done using intensive dripping irrigation. Because the level of moisture of the land is greater, so is the level of

Figure A7: Relation between the hydric stress and seasonality, and the moisture level for apricots




*Notes*: This figure represents the relation between the hydric stress coefficient, $KS$, and seasonality, and the level of moisture in the soil, for apricot trees in southern Spain. The parameters are taken from Allen *et al.* (2006). The graph is a combination of Figures 37 and 42 in Allen et al (2006).

Evapotranspiration.

### A.2.2 Details about the Apricot Production Function.

Following Torrecillas *et al.* (2000) we can specify the weeks of the year in which irrigation is "critical" for apricot trees, as shown in Figure 1 in the paper. The critical weeks include the second rapid fruit growth period (Stage III) and two months after the critical, *i.e.*, Early Post-Harvest (EPH). Both periods are located before and after the harvest season.

Stage III corresponds to the period of high growth before the critical season. This stage is critical because it is the stage at which the trees "transform" water into fruit at the highest rate. The EPH period is also important because of the stress that the trees suffer during the summer after the critical season. Before and during the critical season the trees use the water at a high rate. Hence, the levels of moisture in the trees are very low after the critical season. In order for the trees to survive the summer, they need to be irrigated. Failure to do so will result in a lower output during the next season (see Pérez-Pastor *et al.,* 2009).

Table A2: Demand for Water and Urban Real Estate.

| # units bought | (1) | (2) | (3) | (4) |
|---|---|---|---|---|
| Wealthy | 0.0255*** | 0.0235*** | 0.0133** | 0.0126* |
| | (0.0063) | (0.0062) | (0.0066) | (0.0066) |
| (Wealthy) x | | | 0.0702*** | 0.0602*** |
| x (Critical Season) | | | (0.0117) | (0.0122) |
| Covariates | No | Yes | No | Yes |
| Number of observations | 14,448 | 14,448 | 14,448 | 14,448 |

*Notes:* All regressions are OLS specifications. The sample is restricted to farmers who only grow apricots. In panel A, the dependent variable is the total number of units bought by each individual farmer during a given week. In panel B, he dependent variable is the number of units bought by each individual farmer during a given week per tree. "Wealthy" is a dummy variable that equals 1 if the value of urban real estate of the farmer is above the median, and 0 otherwise. "Critical season" is a dummy that equals 1 if the observation belongs to a week during the critical season, and 0 otherwise. "Covariates" are the price paid by farmers in the auction, the amount of rainfall during the week of the irrigation, the farmer's soil moisture level, and the farmer's number of trees. Standard errors in parentheses. * p<0.10; ** p<0.05; *** p<0.01.

## A.3 Additional Preliminary Analysis.

Table A2 is similar to Table 2 in the paper, but uses the number of units bought per farmer as the dependent variable instead of the number of units bought per farmer per tree for robustness. As in Table 2 in the paper, Table A2 restricts attention to farmers who grow only apricot trees and displays OLS regressions of the number of units bought by each farmer in a given week on several covariates. The variable "wealthy" is a variable that equals 1 if the value of the urban real estate owned by the farmer is greater than the sample median, and 0 otherwise. The idea behind this dummy variable is that wealthy, elite farmers are never liquidity constrained. They do not have to pay rent for their houses, and they can collect rent from their urban real estate to obtain cash during the critical season, financing their purchases. Consider two farmers who grow apricots, who have the same number of trees, and who are not liquidity constrained. Water demand is governed by the water need of the tree, as determined by the apricot production function. These two farmers should have the same demand for water up to an idiosyncratic shock. Therefore, the demand for water and the monetary value of urban real estate should be uncorrelated for farmers who are not liquidity constrained. As expected, Table A2 shows that wealthy farmers demand more water than

Table A3: Demand for Water per tree and Urban Real Estate in dry and regular years.

| | Regular Years | | Dry Years | |
|---|---|---|---|---|
| # units bought per tree | (1) | (2) | (3) | (4) |
| Wealthy | 0.0138* | 0.010139* | -0.0016 | -0.0125 |
| | (0.0051) | (0.0054) | (0.0087) | (0.0092) |
| (Wealthy) | | 0.0209 | | 0.0630*** |
| × (Critical season) | | (0.0107) | | (0.0183) |
| Covariates | Yes | Yes | Yes | Yes |
| Number of observations | 8,736 | 8,736 | 4,992 | 4,992 |

*Notes*: All regressions are OLS specifications. The sample is restricted to farmers who only grow apricots. Dry years are defined to be the four driest harvest years in the sample, i.e., lowest rain from week 25 in the previous calendar year to week 24 in the current calendar year. Dry years are: 1960, 1962, 1964 and 1966. The remaining years are regular years. The dependent variable is the number of units bought per tree by each individual farmer during a given week. "Wealthy" is a dummy variable that equals 1 if the value of urban real estate of the farmer is above the median and 0 otherwise. "Critical season" is a dummy variable that equals 1 if the observation belongs to a week during the critical season and 0 otherwise. "Covariates" are the price paid by farmers in the auction, the amount of rainfall during the week of the irrigation, and the farmer's soil moisture level. Standard errors in parentheses. * $p<0.10$; ** $p<0.05$; *** $p<0.01$.

non-wealthy farmers throughout the year because they own larger plots than non-wealthy farmers. More importantly, and consistently with Table 2 in the paper, wealthy farmers also demand more water during the critical season because they are never liquidity constrained.

Table A3 shows that the behavior of poor and wealthy farmers during the critical season is similar in regular years (*i.e.* years without droughts). We expect poor farmers to be less likely to be liquidity constrained during regular years, when the price of water in the auction does not increase substantially. The rainfall during the harvest year (*i.e.*, from week 25 in the previous calendar year until week 24 in the present calendar year) is an exogenous shock that shifts both the demand for water and the water prices down. Table A3 shows that the purchase decisions of the wealthy and poor farmers are different for dry and regular years. In Table 2 column 4 in the paper, the interaction between wealth and the critical season is 0.0315. This is a weighted average between 0.0209, in regular years, and 0.0630, in dry years. This interaction is smaller and insignificant in regular years. This result indicates that different purchase behavior during the critical season comes from the dry years, when both the demand for water and the prices of water in the critical season are high. In other words, the purchase behavior of poor and wealthy farmers is similar during regular years, when poor farmers are less likely to be liquidity constrained.

# B Welfare Comparison of Markets *vs.* Quotas

In Donna and Espín-Sánchez (2016) we present a two period model, where some agents face liquidity constraints (LC) and their preferences are characterized by storability and seasonality. Agents differ in two unobserved dimensions, their valuation of water (productivity) and the amount of cash they have to pay for water (wealth). This is the simplest model that can accommodate the main elements analyzed in this paper: heterogeneity in productivity, liquidity constraints, storability, and seasonality. Productivity during the first period (*off season*) is lower for all farmers than during the second period (*critical season*). The model is useful to understand both positive and normative questions. The results regarding the dynamics of purchases, the inefficiency of the market equilibrium, and the relative efficiency of market *vs.* quotas apply directly to the general infinity-horizon model. The notion of first-best efficiency does not translate into the general model. When farmers are heterogeneous in productivity and the horizon is infinite, the notion of first-best is cumbersome because it has to accommodate all possible future combinations of allocations.

**Market Equilibrium.**

The model predicts that, as long as LC are *binding (i.e.* some farmers do not have enough cash to pay for the water in a given period), then the timing of purchases matters. Poor farmers will not be able to buy water in the second period, when water is more valuable. Wealthy farmers will buy all the water in the second period and some of the water in the first period. If LC are *severe (i.e.* some farmers do not have enough cash to pay for water in any period), poor farmers will not buy water in any period, and some wealthy farmers will buy several units of water. The above results imply that as long as LC are *binding,* poor farmers will not be able to buy water during the critical season. The model also predicts that farmers with intermediate wealth (LC are *binding* but not *severe*) will buy more water than wealthy farmers *off season,* while farmer with *severe* LC will buy less water than the wealthy farmer *off season.* When LC are *binding*, prices could be at their *unconstrained* level or could both be lower than their *unconstrained* level. When LC are *severe*, prices will always be lower than their *unconstrained* levels.

**Market Efficiency.**

There are two sources of market inefficiency. If LC are *binding* and farmers are homogeneous in productivity, this will not reduce welfare. As long as all units of water are allocated and no farmer buys more than two units (while others get zero units) the allocation would be efficient. If farmers are heterogeneous in productivity, efficiency requires assortative matching between units of water and farmers. In other words, efficiency requires that farmers with high productivity buy water in the second period (and never end up not buying water) and farmers with low productivity buy water in the first period (and some do not buy any water). If LC are *binding*, farmers with high productivity and low wealth might not be able to buy water in the second period when it is more productive. This is the first source of market inefficiency: inefficiency in the intensive margin or *mismatching*. Mismatching will happen only when there are differences in productivity among farmers and it will be important when those differences are large.[2]

There could still be market inefficiency even if there are no differences in productivity among farmers. If LC are *severe*, then the market allocation will be inefficient even if farmers are homogeneous in productivity. If LC are *severe*, poor farmers will not be able to buy water in any period and some wealthy farmers will buy more than one unit of water. Since the production function is concave in unit of water bought, the gains from the second unit of water of a wealthy farmer are less valuable that the gains from the first unit of water of a poor farmer. This is the second source of market inefficiency: inefficiency in the extensive margin or *overallocation*. *Overallocation* will happen only when LC are *severe*.

**Welfare: Market *vs.* Quotas.**

The results above describe the conditions under which the market will be inefficient. However, they do not compare the relative efficiency of the market mechanism *vs.* the quota mechanism when both are inefficient, *i.e.*, when neither achieves the first-best allocation.

[2] If differences in productivity are large compared to the concavity of the production function in units of water bought, then the gains from the second unit of water of a high productivity farmer could be greater than the gains from the first unit of water of a low productivity farmer. If there is perfect correlation between wealth and productivity, and differences in productivity are large compared to the concavity of the production function, then the market is efficient.

Both the quotas and the market could achieve the first-best allocation under some conditions. When farmers are heterogeneous in productivity but LC are not *binding*, then markets are efficient but the quotas are not. When LC are not *binding,* then the model is similar to the standard neoclassical model, and markets achieve efficiency by allocating more valuable units to those who value them more. Quotas allocate the units uniformly. Because units are discrete here, uniformly means randomly. This means that farmers with high productivity and farmers with low productivity have the same probability of receiving one unit in each period. Quotas will produce *mismatching* by allocating units during the *critical season* to farmers with low productivity; however all units will be allocated and no farmer will get more than one unit, thus quotas will never produce *overallocation* of units.

When farmers are homogeneous in productivity, but LC are *severe*, then the quotas are efficient but the markets are not. When farmers are homogeneous in productivity there is only one potential source of inefficiency: *overallocation*. Since all farmers have the same valuation for each unit, matching is irrelevant. Any mechanism that allocates all the units and at most one unit to each farmer will be efficient. However, the markets will not pass this test when LC are *severe*. The markets will allocate all the units to wealthy farmers and no units to poor farmers. Some wealthy farmers will receive two units *overallocation* and some high productivity poor farmers will not receive any unit in the *critical season*, thus generating *mismatching*.

In the intermediate case, when farmers are heterogeneous in productivity and LC are *binding*, the relative efficiency of markets and quotas is ambiguous. Both mechanisms suffer from *mismatching* but neither suffer from *overallocation*. Both mechanisms assign units in the *critical season* to farmers with low productivity and some high productivity farmers do not receive any units. In general, quotas are more efficient than markets when the heterogeneity in wealth is relatively large and the heterogeneity in productivity is relatively small, while markets are more efficient than quotas when the heterogeneity in wealth is relatively small and the heterogeneity in productivity is relatively large.

# C Details about the Estimation Procedure and the Structural Model

## C.1 Demand Estimates.

We construct a two-step conditional choice probability (CCP) estimator to estimate the parameters that characterize demand (Hotz and Miller 1993; Hotz *et al.* 1994).

**Step 1.** In the first step we compute transition probability matrices for the following observable state variables: moisture, week, price, and rain. As described in the paper, the productivity shocks, $\varepsilon_{ijt}$, are assumed to be *i.i.d.* Gumbel, so they can be integrated analytically (see Subsection C.3). Moisture is a continuous variable and its evolution over time depends on both the farmers' decisions to buy water and rainfall. Therefore, certain values of moisture are never reached in the sample, even when their probability of occurrence is nonzero. To estimate demand, however, we need to integrate the value function for each possible combination of the state variables in the state space. Thus, we first estimate the CCP using the values (of the state space) reached in the sample. Then we use the CCP estimator to predict the CCP on the values (of the state space) unreached in the sample.

We estimate the CCP both parametrically (using a logistic distribution) and nonparametrically (using kernel methods to smooth both discrete and continuous variables). There are four observable state variables in the structural model when the liquidity constraint is not binding (*i.e.* for the wealthy farmers): moisture, week of the year, price of water, and rain. Moisture is a deterministic continuous variable that represents the amount of water accumulated in the farmers' plot; it goes from 300 to 1200. Week of the year is a deterministic discrete variable; it goes from 1 to 52. Price of water and rain are random variables. As explained in the paper, we assume that, holding fixed the week of the year, farmers jointly draw a price-rain pair, $(p_t, r_t)$, among the 12 pairs (*i.e.* the 12 years of the same week) available in the data with equal probability (see Subsection C.4 for details). Each week, prices may take three discrete values: low, high, or no-auction. Each week, rain may take two discrete values: no rain (*i.e.* zero $mm$, which is the median of the rain distribution) or positive rain (in this case, we assign the median of the rain distribution, conditional on rain being positive, that is $31mm$). For each week, *low price* is the mean price below the median of the same week across years; *high price* is the mean price above the median of the same week across years. We estimate the joint distribution of prices and rain non-parametrically using a frequency estimator. To estimate the CCP we use the actual realizations from the data of price and rainfall.

We estimate the CCP parametrically using a logistic model (*i.e.* multinomial logit). Let $S_{it} = (M_{it}, w_t, p_t, r_t) \,\epsilon\, \mathbb{R}^{\mathbf{4}}$ be the vector of state variables, where $M_{it}$ is the moisture of the soil of farmer $i$ in week $t$, $w_t$ is the week of year, $p_t$ is the price of water in the auction in week $t$, and $r_t$ is the rainfall in week $t$. Let $j_{ikt} = 1$ if farmer $i$ bought $k$ units in period $t$ and 0 otherwise. We estimate the CCP by maximizing the following log-likelihood function $lnL = \sum_{t=1}^{T} \sum_{i=1}^{I} \sum_{k=0}^{J} j_{ikt} lnP_{ikt}$, where $P_{ikt} = \left( \frac{exp(\kappa' S_{it})}{\sum_j exp(\kappa' S_{it})} \right)$ and $\kappa$ is the parameter vector.

We also estimate CCP nonparametrically. Rather than using a traditional frequency-based approach in the presence of discrete variables, to compute the CCP we smooth both discrete and continuous variables. There are two reasons for this. First, it allows us to extend the reach of the nonparametric methods to our empirical model. It is well known that nonparametric frequency methods are useful only when the sample size is large and the discrete variables take a limited number of values: this allows the number of discrete cells to be smaller than the sample size.[3] Second, moisture is a continuous variable and its evolution over time depends on both the decisions to buy water of the farmers and the realizations of rain. Therefore, certain values of moisture are never reached in the sample even when their probability of occurrence is not zero.[4] To estimate the demand, however, we need to integrate the value function over the relevant combination of the state variables associated with the nodes of some simulated future path of the farmer.[5] Thus, we estimate the CCP

[3] The frequency approach is not feasible in our setting, even if we discretize the continuous variable moisture in a reasonable number of values. See Subsection C.8 for a thorough discussion.

[4] For example, for week 23 the joint probability of *no rain* and *low price* conditional on this week is 9.1%. This is because only in 1 out of 11 years was registered *low rain* and *low price* in the week 23 ($1/11 = .0909$). The observed different values of inventories for the 12 unconstrained farmers are (at most) $12 \times 1 = 12$. In the simulation, however, a value of moisture different from, although close to, these 12 observed values may be reached. But the frequency estimator would not be defined for any value of moisture different from those 12 values.

[5] That is, we also need to integrate over values of moisture discussed in previous footnote where the frequency estimator is not defined. These values of moisture are never reached in our sample.

nonparametrically using kernel methods to smooth both discrete and continuous variables.

We define now the nonparametric CCP estimator. Following Li and Racine (2003) we use generalized product kernels for a mix of continuous and discrete variables. Let $S_{it} = \left(M_{it}, S_t^d\right) \epsilon\, \mathbb{R} \times \mathbb{R}^{\mathbf{3}}$ be the vector of state variables, where $M_{it} \epsilon\, \mathbb{R}$ is again moisture and $S_t^d = (w_t, p_t, r_t) \epsilon\, \mathbb{R}^{\mathbf{3}}$ is the vector of discrete state variables: week, price, and rain. Let $s_k^d$ be the k$th$ component of $s^d$ and $S_t^d = (t = 1, \ldots, T)$. For $S_{tk}^d, s_k^d \,\epsilon \left\{0, 1 \ldots, c_k - 1\right\}$ (the support of each discrete variable) define the univariate kernel (Aitchison and Aitken, 1976):

$$l^u\left(S_{tk}^d, s_k^d, \lambda_k\right) = \begin{cases} 1 - \lambda_k & if\;\; S_{tk}^d = s_k^d \\ \frac{\lambda_k}{c_k - 1} & if\;\; S_{tk}^d \neq s_k^d \end{cases} .$$

We use the above kernel for prices and rain. For the ordered discrete variable week we use the kernel function (Wang and van Ryzin, 1981) $l^o\left(w_t, v, \lambda_1\right) = \lambda_1^{|w-v|}$, where $\lambda_1 \epsilon\, [0, 1]$. Therefore, for the multivariate vector of discrete state variables we use the product kernel:

$$L\left(S_{it}^d, s^d, \lambda\right) = l^o\left(w_t, v, \lambda_1\right) \prod_{k=2}^{3} l^u\left(S_{tk}^d, s_k^d, \lambda_k\right) = \lambda_1^{|w-v|} \prod_{k=2}^{3} \left(\frac{\lambda_k}{c_k - 1}\right)^{N_{tk}(s)} \left(1 - \lambda_k\right)^{1 - N_{tk}(s)}, \tag{C.1}$$

where $\lambda = (\lambda_2, \lambda_3)$ and $N_{tk}(s) = \mathbf{1}\left(S_{tk}^d \neq s_k^d\right)$ is an indicator function that equals 1 if $S_{tk}^d \neq s_k^d$ and 0 otherwise. Let $f(s) = f\left(m, s^d\right)$ be the joint probability density function (PDF) of $S_{it} = \left(M_{it}, S_t^d\right)$. We use the following kernel estimator $\hat{f}(s) = \frac{1}{T}\sum_{t=1}^{T} L_{ts^d} W_{h,tM}$, where $W_{h,tM} = h^{-1} w\left(\frac{M_t - m}{h}\right)$, $w(\cdot)$ is a standard univariate second order Gaussian kernel, and $L_{ts^d} = L\left(S_t^d, s^d, \lambda\right)$ is given by equation C.1. We select bandwidth using likelihood cross-validation. We estimate the $\hat{f}(s)$ using the observed values of the variables in the state space (in the sample). We then use the estimated density and evaluate it at the unobserved values of the state space needed to integrate the value function (out of the sample).

We extrapolate CCP on unobserved states using the estimated PDF from the sample. These are the *smooth CCP*.

**Step 2.** We restrict the sample to the twelve farmers that are not liquidity constrained. We estimate the vector of structural parameters, $\Theta \equiv (\gamma, \sigma_\varepsilon, \zeta)$, of the model in Section 4 using the conditional choice simulation estimator proposed by Hotz *et al.* (1994) which is based on the inversion theorem by Hotz and Miller (1993). We integrate the value function using the smoothed CCP as computed in the previous subsection. We set the discount factor $\beta$ equal to 0.99. Prices and rain are simulated using the joint distribution of prices and rain estimated with the procedure described in the second paragraph of step 1. We normalize the number of trees of the farmers using the median number of apricot trees for the unconstrained farmers, 76 trees. We let moisture follow the evolution described by equations in Appendix A.2 with the following values: $TAW = 1200$, $PW = 300$, $EAW = 0.5 \cdot TAW$, $E = 4$ (see A.2 for details). The observed number of units that farmers buy varies from 0 to 4 units per week.

For the initial condition of the moisture we follow Hendel and Nevo (2006, p. 1,647), and use the estimated distribution of moisture to generate its initial distribution. We do this by starting at an arbitrary initial level of moisture of: $1/2 \times (\mathrm{TAW} + \mathrm{PW})$, where TAW and PW are the Total Available Water and Permanent Welting Point described above (see Appendix A.2 for details). In practice, in our application the initial condition of the moisture has no impact on its evolution after a couple of weeks due to the evotranspiration rate and rainfall. We have experimented with different initial conditions and obtained almost identical results.

We estimate the parameter vector $\Theta$ using a GMM estimator based on moment conditions proposed by Hotz *et al.* We use 200 simulations, each with 12 individuals, and $T = 572$ weeks per individual per simulation. That is, for each simulation and each individual, we use a total of $T = 11\, years \times 52\, weeks$ periods, which is the length of our panel, leaving a total of $6,864$ observations in each simulation. We perform the estimation using KNITRO, a solver for non-linear optimization, with tolerance level of 1.0e-12. With the estimated demand we recover the annual revenue for all farmers, constrained and unconstrained.

## C.2 Properties of the Demand Estimator.

Following Aguirregabiria and Mira (2010) we now establish some properties of the demand estimator. Time is discrete and indexed by $t$. Each period represents a week. We index farmers by $i$. Farmers have preferences defined over a sequence of states from period $t = 0$ until period $t = \infty$. The state at period $t$ for wealthy farmer $i$ has two components, a vector of state variables $s_{it} = (M_{it}, w_t, p_t, r_t, \epsilon_{it})$ that is known at period $t$; and a decision vector $j_{it}$ chosen at period $t$ that belongs to the discrete set $J \in \{0, ..., J\}$. The vector of state variables, $s_{it}$, also includes the error vector $\epsilon_{it} \equiv (\epsilon_{i1t}, ..., \epsilon_{iJt})$. The time index, $t$, can be a component of the state vector, $s_{it}$, which may also contain time-invariant individual characteristics. Farmer's preferences over possible sequences of states is represented by a utility function $\sum_{\tau=0}^{\infty} \beta^{\tau} U\left(j_{i,t+\tau}, s_{i,t+\tau}\right)$, where $\beta \in (0,1)$ is the discount factor and $U\left(j_{it}, s_{it}\right) = h\left(j_{it}, M_{it}, w_t; \gamma\right) + \epsilon_{ijt} - \zeta_j - j_{it}p_t$ is the current utility function. The decision at period $t$ affects the evolution of future values of the state variables, and the farmer faces uncertainty about these future values. Liquidity constraints are never binding for wealthy farmers. Also note that the welting point constraint affects only the utility function as defined above, but not the choice set. The farmer's beliefs about future states can be represented by a Markov transition distribution function, $F\left(s_{i,t+1} | j_{it}, s_{it}\right)$. These beliefs are rational in that they are the true transition probabilities of the state variables. Every period $t$ the farmer observes the vector of state variables, $s_{it}$, and chooses an action $j_{it} \in J$ to maximize the expected utility:

$$\mathbb{E}\left(\sum_{\tau=0}^{\infty} \beta^{\tau} U\left(j_{i,t+\tau}, s_{i,t+\tau}\right) | j_{it}, s_{i,t}\right).$$

This is the farmer's dynamic programming (DP) problem. Let $\alpha\left(s_{i,t}\right)$ and $V\left(s_{i,t}\right)$ be the optimal decision rule and the value function of the DP problem, respectively. By Bellman's principle of optimality the value function can be obtained using the recursive expression:

$$V\left(s_{it}\right) = \underset{j \in J}{max}\left\{h\left(j_{it}, M_{it}, w_t; \gamma\right) + \epsilon_{jt} - \zeta_j - j_{it}p_t + \beta \int V\left(s_{i,t+1}\right) dF\left(s_{i,t+1} | j, s_{it}\right)\right\},$$

and the optimal decision rule is then $\alpha\left(s_{i,t}\right) = \underset{j \in J}{argmax}\left\{\bar{v}\left(j, s_{i,t}\right)\right\}$ where, for every $j \in J$,

$$\bar{v}\left(j, s_{i,t}\right) \equiv h\left(j_{it}, M_{it}, w_t; \gamma\right) + \epsilon_{jt} - \zeta_j - j_{it}p_t + \beta \int V\left(s_{i,t+1}\right) dF\left(s_{i,t+1} | j, s_{it}\right).$$

We are interested in the estimation of the structural parameters in preferences and transition probabilities. Suppose that a researcher has a panel of $N$ individuals who behave according to this decision model. For every observation, $(i,t)$, in this panel, the researcher observes the individual's action, $j_{it}$, and a subvector, $(M_{it}, w_t, p_t, r_t)$, of the state vector, $s_{it}$. In summary, the researcher's data set is:

$$Data = \{j_{it}, M_{it}, w_t, p_t, r_t : i = 1, 2, ..., N; t = 1, 2, ..., \infty\}.$$

We now discuss the assumptions regarding the relationship between observable and unobservable variables following Aguirregabiria and Mira (2010).

**Assumption AS (*Additive Separability*).** The utility function is additively separable in the observable and unobservable components: $U(j_{it}, s_{it}) \equiv U(j_{it}, M_{it}, w_t, p_t, r_t, \epsilon_{it}) = h(j_{it}, M_{it}, w_t; \gamma) + \epsilon_{ijt} - \zeta_j - j_{it} p_t$, where $\epsilon_{ijt}$ is a real random variable with unbounded support.

**Assumption IID (*i.i.d. Unobservables*).** The unobserved state variables in $\epsilon_{it}$ are independently and identically distributed over agents and over time with cumulative density function (CDF) $G_\varepsilon(\epsilon_{it})$, which has finite first moments and is continuous and twice differentiable in $\epsilon_{it}$.

**Assumption CI-X (*Conditional Independence of Future x*).** Conditional on the current values of the decision and the observable state variables, next period observable state variables do not depend on current $\epsilon_{it}$. This assumption holds trivially for $w_t$. It also holds trivially for $x_{it}$, because the covariates are constant for a given individual, for calendar effects, and also rain, which is exogenous. This assumption also holds for prices as discussed in Section 3 in the paper. Finally, the law of motion of the moisture is independent of $\epsilon_{it}$.

**Assumption CLOGIT.** The unobserved state variables $\epsilon_{ijt}$ are independent and have an extreme value type 1 distribution.

**Assumption DIS (*Discrete Support of x*).** The support of $(M_{it}, w_t, p_t, r_t)$ is discrete and finite: $(M_{it}, w_t, p_t, r_t) \in X = \left\{x^{(1)}, ..., x^{(|X|)}\right\}$ with $|X| < \infty$.

This model satisfies the assumptions in Hotz and Miller (1993) and Hotz *et al.* (1994), thus establishing the legitimacy of our simulation-based CCP estimator described in Section 4 in the paper.

## C.3 Specification of the Productivity Shock.

We introduce the notation with the specification in Subsection C.3.1 for the case of *i.i.d.* shocks across choice alternatives, where the result below is well known in the industrial organization literature. In Subsection C.3.2 we extend these results to the case of shocks non *i.i.d.* across choice alternatives, which is the main specification used in the paper. For robustness, we also present the demand estimates using the specification in Subsection C.3.1. Estimation results of the structural model using both specifications are in Subsection C.5 in this appendix.

### C.3.1 Specification with *i.i.d.* shocks across choice alternatives.

Under this specification the productivity shocks, $\varepsilon_{ijt}$, are drawn *i.i.d.* across choice alternatives $j \in \{0, 1, \ldots, J\}$, where each choice alternative involves the purchase of a different number of units, and across individuals and over time. Let the value function be:

$$\begin{gathered} V\left(M_{it}, w_t, p_t, r_t, \mu_{it}, \epsilon_{ijt}\right) \equiv \underset{j_{it}\in\{0,1,\ldots,J\}}{max} \left\{h\left(j_{it}, M_{it}, w_t; \gamma\right) + \varepsilon_{ijt} - \zeta_j - p_t j_{it} + \right. \\ \left. + \beta\mathbb{E}\left[V\left(M_{i,t+1}, \mu_{i,t+1}, w_{t+1}, p_{t+1}, r_{t+1}, \epsilon_{i,j,t+1}\right) | M_{it}, w_t, p_t, r_t, \mu_{it}, j_{it}\right]\right\} \\ \text{s.t.: } M_{it} \geq PW \\ \text{s.t. : } j_{it} p_t \leq \mu_{it}, \forall j_{it} > 0 \end{gathered},$$

Rewrite it as (we omit the constraints in what follows):

$$V\left(X_{it}, \epsilon_{ijt}\right) \equiv \underset{j_{it}\in\{0,1,\ldots,J\}}{max} \left\{H\left(\centerdot\right) + \varepsilon_{ijt} + \beta\mathbb{E}_{\varepsilon}\left[V\left(X_{it+1}, \epsilon_{ijt+1}\right) F_X \mathrm{d}X \left| X_{it}, j_{it}\right.\right]\right\},$$

where $X_{it} \equiv \left(M_{it}, w_t, p_t, r_t\right)$, $H\left(\centerdot\right) \equiv h\left(j_{it}, M_{it}, w_t; \gamma\right) - p_t j_{it} - \zeta_j$.

Now rewrite the previous expression as:

$$V\left(X_{it}, \epsilon_{ijt}\right) \equiv \underset{j_{it}\in\{0,1,\ldots,J\}}{max} \left\{v(j_{it}, X_{it}) + \varepsilon_{ijt}\right\},$$

where $v(j_{it}, X_{it})$ is the choice specific value function.

Then if the productivity shocks $\varepsilon_{ijt}$ follow a Gumbel distribution with CDF $F_\varepsilon\left(\varepsilon_{it};\sigma_\varepsilon\right) = e^{-e^{-\varepsilon it/\sigma_\varepsilon}}$:

$$\mathbb{E}_\varepsilon V\left(X_{it+1},\epsilon_{ijt+1}\right) = \int_\varepsilon \underset{j_{it}\in\{0,1,\dots,J\}}{max}\left\{v(j_{it},X_{it})+\varepsilon_{ijt}\right\}\mathrm{d}F_\varepsilon = log\left(\sum_{j=0}^{J} exp\left(v(j_{it},X_{it})\right)\right)+\bar{\gamma}\sigma_\varepsilon, \tag{C.2}$$

where the last equality follows from the properties of the Gumbel distribution and $\bar{\gamma}=0.5772$ is the Euler's constant.

Then:

$$\mathbb{P}\left(j_{it}=d\right)=\mathbb{P}\left(\varepsilon_{idt}-\varepsilon_{ikt}>v(j_{it}=k,X_{it})-v(j_{it}=d,X_{it}),\ \forall k\neq d\right).$$

Using the properties of the Gumbel distribution:

$$\mathbb{P}\left(j_{it}=d\right)=\frac{exp\left(v^d\right)}{\sum_{k=0}^{J} exp\left(v^k\right)}, \tag{C.3}$$

where $v^r \equiv v(j_{it}=r,X_{it})$.

Then replacing the equation in (C.3) into (C.2):

$$\mathbb{E}_\varepsilon V\left(X_{it+1},\epsilon_{ijt+1}\right)=v^d - log\left(\mathbb{P}\left(j_{it}=d\right)\right)+\bar{\gamma}\sigma_\varepsilon.$$

### C.3.2 Specification with shocks non *i.i.d.* across choice alternatives.

We follow the same steps as in Subsection C.3.1. Now $\hat{j}\in\{0,1\}$, where $\hat{j}=0$ if $j=0$ and $\hat{j}=1$ if $j>0$. Then the productivity shocks, $\varepsilon_{i\hat{j}t}$, are drawn *i.i.d.* across $\hat{j}\in\{0,1\}$ and the shock is the same for every unit, so $\varepsilon_{ijt}=\varepsilon_{i\hat{j}t}$ for $j=0$ and $\varepsilon_{ijt}=\varepsilon_{i\hat{j}t}$ for $j>0$. Let the value function be:

$$\begin{aligned} V\left(X_{it},\epsilon_{i\hat{j}t}\right) \equiv \underset{j_{it}\in\{0,1,\dots,J\}}{max}\Big\{ & h\left(j_{it},M_{it},w_t;\gamma\right)+\varepsilon_{i,\hat{j}=0,t}1\left\{j_{it}=0\right\}+\varepsilon_{i,\hat{j}=1,t}1\left\{j_{it}>0\right\}-\zeta_j-p_t j_{it}+ \\ & +\beta\mathbb{E}\left[V\left(X_{it+1},\epsilon_{i,\hat{j},t+1}\right)|X_{it},j_{it}\right]\Big\}. \end{aligned}$$

Rewrite the previous expression as:[6]

$$V\left(X_{it},\epsilon_{i\hat{j}t}\right)\equiv max\left\{v(j_{it}=0,X_{it})+\varepsilon_{i,\hat{j}=0,t},\ \underset{j_{it}^{+}\in\{1,\ldots,J\}}{max}\left\{v(j_{it}=j_{it}^{+},X_{it})+\varepsilon_{i,\hat{j}=1,t}\right\}\right\}.$$

Using the properties of the Gumbel distribution:

$$\mathbb{E}_{\varepsilon}V\left(X_{it+1},\epsilon_{i\hat{j}t+1}\right)=log\left(exp\left(v^{0}\right)+exp\left(v^{j^{+}}\right)\right)+\bar{\gamma}\sigma_{\varepsilon}, \tag{C.4}$$

and:

$$\mathbb{P}\left(j_{it}=0\right)=\frac{exp\left(v^{0}\right)}{exp\left(v^{0}\right)+exp\left(v^{j^{+}}\right)}, \tag{C.5}$$

$$\mathbb{P}\left(j_{it}=j_{it}^{+}\right)=\frac{exp\left(v^{j^{+}}\right)}{exp\left(v^{0}\right)+exp\left(v^{j^{+}}\right)}. \tag{C.6}$$

Replacing the equations in (C.5) and (C.6) into (C.4) we obtain a the closed-form solution for:

$$\mathbb{E}_{\varepsilon}V\left(X_{it+1},\epsilon_{i\hat{j}t+1}\right).$$

## C.4 Serial Correlation Over Time in Rainfall and Auction Water Price.

In this subsection we discuss our modeling assumption for the evolution of the state variables for auction water price and rainfall, and their fit to the data. In particular, we discuss the role of serial correlation over time in rainfall and the price of water.

We model the evolution of prices and rainfall to capture two main empirical regularities from our setting. First, the major determinant of the price of water is weather seasonality. Second, the variation of prices and rainfall across years is low, conditional on the week of

[6] Note that the inner maximization process is deterministic. That is, conditional on buying, there is only one shock. Therefore, no integration is needed for the inner maximization process.

the year (which captures seasonality).[7] Our data in this paper span 12 years. We model the joint evolution of the price of water in period $t$, that denotes a week, and rainfall in period $t-1$ assuming that, holding fixed the week of the year, farmers jointly draw a price-rain pair, $(p_t, r_{t-1})$, *i.i.d.* among the 12 pairs (*i.e.* the 12 years of the same week) available in the data with equal probability.

Serial correlation in the price of water arises because weather seasonality is its main determinant, and dry weather in a given week is usually followed by dry weather in a subsequent week. During summers, for example, prices are systematically higher for several weeks. Indeed, in the tables described below we find that the price of water displays serial correlation across the weeks of the year. Accounting for such serial correlation is important because it affects the dynamics of prices ,and farmers' decisions, if such dynamics are taken into account when the farmers bid in the auction.

One potential concern with our specification for the evolution of price and rainfall is that it may not account for the serial correlation in prices. We investigate this concern in Tables A4 and A5, that display OLS regressions using the price of water from the data and the price of water simulated for the structural model as dependent variables, along with two tests for the disturbance from these regressions: (1) a Breusch-Godfrey test for serial correlation, and (2) a Durbin's alternative test for serial correlation.[8]

Table A4 shows that under our specification for the evolution of prices, as described above, the predicted prices closely follow the prices in the data, and that the residual part of the price of water from the data that is not explained by the simulated price displays no serial correlation. Column 1 shows that, as expected, the Breusch-Godfrey and Durbin's alternative tests strongly reject the null of no first-order serial correlation in the disturbance of a regression of price of water on a constant. Column 2 shows the regression of the price of water from the data on the simulated price of water using the specification described in the first paragraph, which corresponds to the one used for the structural model in Section 3 in the paper. As expected, the simulated price tracks the price from the data quite well as indicated by an estimated coefficient of 0.71, which is statistically significant, and a goodness of the fit

[7]See Donna and Espín-Sánchez (2018) for details.

[8]For details about these tests see, *e.g.*, Davidson and MacKinnon (1993).

of $R^2 = 47$ percent. Importantly, the tests for the disturbance in this regression show that we cannot reject the null hypotheses of no first-order serial correlation (the p-value of the Breusch-Godfrey test is 32.9 percent and the p-value of the Durbin's alternative test is 33.3 percent).[9] That is, the residual part of the price of water from the data that is not explained by the simulated price displays no serial correlation. This indicates that the specification used to simulate the price of water for the structural model accounts for the serial correlation in the price of water from the data. Intuitively, this occurs because our specification models price and rainfall as a *joint distribution conditional* on the week of the year, using the price of water from the current week and the rainfall from the previous week, $(p_t, r_{t-1})$. The main determinants of the price of water in a given week are the rainfall from the previous week, and the calendar week. As shown in A4, these two variables are sufficient to account for the serial correlation observed in the data. Column 3 performs a similar analysis as column 1, but using the price of water simulated for the structural model. Similar to column 1, column 3 shows that for the simulated prices we also reject the null of no first-order serial correlation in the disturbance. This shows that the simulated price for the structural model displays serial correlation similar to that in the price of water in the data.

Table A5 focuses on the alternative modeling assumptions for the evolution of the state variables for price of water and rainfall. For example, an alternative assumption commonly used in the industrial organization literature is to use a Markov process with lagged price of water, or rainfall, or both (recall that rainfall corresponds to the rain during the previous week). Column 1 displays a regression of price of water on the lagged price of water, and shows that we cannot reject the null of no serial correlation in the disturbance from such regression. Columns 2 and 3 show similar results when we include lagged rainfall, or both, lagged price of water and lagged rainfall: the null of no serial correlation cannot be rejected. Columns 4 to 6, repeat the analysis in the previous columns including week fixed effects, a set of 51 dummy variables each corresponding to one week of the year. Again, the null of no serial correlation cannot be rejected.[10] We interpret this as evidence that using a Markov process with either lagged price of water, or lagged rainfall, or both as an alternative modeling

[9] Similar results are obtained if one additionally includes the simulated rain as dependent variable, and/or if one includes additional lags.

[10] Similar results are obtained using additional lags for the price of water and rainfall.

assumption does not account for the serial correlation in the prices from the data.

We conclude this subsection comparing the correlation patterns between price of water and rainfall from the data, and the simulation of these variables using the specification from the structural model. Table A6 displays these correlations. The first column, labeled as "Data," shows that there is a negative correlation between price of water in week $t$ and rainfall in previous week(s), as expected. The correlation tends to decrease with higher lags because the effect of rain on demand dissipates as times passes due to the evapotranspiration of water. Again, this negative correlation is the main reason why the price of water displays serial correlation across weeks of the year. It is also the main reason why we use the conditional joint process described above to model price of water and rainfall. The second column, labeled as "Simulation," shows the same patterns from the simulated variables (price of water and rainfall) used in the structural model. The correlation patterns from the simulated variables are similar to the ones from the data.

Table A4: Price regressions and tests for serial correlation in the disturbance using specification from the structural model

| | (1)<br>Price data | (2)<br>Price data | (3)<br>Price simulated |
|---|---|---|---|
| Price simulated | | 0.7060***<br>(0.0681) | |
| Constant | 310.2234***<br>(20.0981) | 103.0802***<br>(32.3308) | 304.5125***<br>(18.5680) |
| $R^2$ | 0.000 | 0.473 | 0.000 |
| Breusch-Godfrey test | | | |
| $\chi^2$ statistic: | 33.21 | 0.952 | 10.03 |
| Prob. $> \chi^2$: | 8.27e-09 | 0.329 | 0.00154 |
| Durbin's alternative test | | | |
| $\chi^2$ statistic: | 38.22 | 0.936 | 10.38 |
| Prob. $> \chi^2$: | 6.31e-10 | 0.333 | 0.00127 |

*Notes:* The table displays OLS regressions using the variable displayed in each column as dependent variable. "Price data" is the price of water from the data. "Price simulated" is the price of water that was simulated for the structural model as described in Section 3 in the paper, using the procedure described in this subsection. All regresions include a constant that is reported in the table, and are performed with the sample used in the structural model from Section 3, when the auction was run. Standard errors are in parenthesis. The bottom part of the table displays two tests for the disturbance from the regressions in each column: (1) a Breusch-Godfrey test for serial correlation in the disturbance, and (2) a Durbin's alternative test for serial correlation in the disturbances. For each test and for each column, the table displays the value of the $\chi^2$ statistic of the test and the corresponding p-value, denoted by "Prob.$>\chi^2$." *** p<0.01, ** p<0.05, * p<0.1.

Table A5: Price regressions and tests for serial correlation in the disturbance using alternative specifications

| | (1) Price data | (2) Price data | (3) Price data | (4) Price data | (5) Price data | (6) Price data |
|---|---|---|---|---|---|---|
| Lagged price data | 0.5496*** | | 0.5426*** | 0.4978*** | | 0.5002*** |
| | (0.0699) | | (0.0695) | (0.0924) | | (0.0923) |
| Rain data | | -3.9101** | -2.3622* | | -1.5531 | -1.5063 |
| | | (1.7812) | (1.2845) | | (1.2561) | (1.2961) |
| Constant | 106.6150*** | 328.9105*** | 122.6566*** | 61.4336 | 132.7750* | 61.2521 |
| | (22.4892) | (21.6805) | (23.9456) | (156.0353) | (79.5221) | (155.7288) |
| Week fixed effects | No | No | No | Yes | Yes | Yes |
| $R^2$ | 0.306 | 0.020 | 0.323 | 0.703 | 0.694 | 0.708 |
| Breusch-Godfrey test | | | | | | |
| $\chi^2$ statistic: chi2 | 27 | 30.45 | 25.99 | 29.54 | 56.79 | 31.16 |
| Prob. $> \chi^2$: | 2.03e-07 | 3.42e-08 | 3.42e-07 | 5.48e-08 | 0 | 2.38e-08 |
| Durbin's alternative test | | | | | | |
| $\chi^2$ statistic: chi2 | 32.64 | 34.44 | 30.92 | 23.12 | 57.66 | 24.45 |
| Prob. $> \chi^2$: | 1.11e-08 | 4.39e-09 | 2.68e-08 | 1.53e-06 | 0 | 7.61e-07 |

*Notes:* The table displays OLS regressions using the variable "Price data" as dependent variable, which is the price of water from the data. The variable "Lagged price data" is the lag of "price data." The variable "Rain data" is the rain from the data. All regresions include a constant that is reported in the table, and are performed with the sample used in the structural model from Section 3, when the auction was run. Standard errors are in parenthesis. The bottom part of the table displays two tests for the disturbance from the regressions in each column: (1) a Breusch-Godfrey test for serial correlation in the disturbance, and (2) a Durbin's alternative test for serial correlation in the disturbance. For each test and for each column, the table displays the value of the $\chi^2$ statistic of the test and the corresponding p-value, denoted by "Prob.$>\chi^2$." *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Correlation between Price and Rain: Data and Simulation.

| Data | | Simulation | |
|---|---|---|---|
| | $Price_t$ | | $Price_t$ |
| $Price_t$ | 1.000 | $Price_t$ | 1.000 |
| $Rain_t$ | -0.152 | $Rain_t$ | -0.141 |
| $Rain_{t-1}$ | -0.101 | $Rain_{t-1}$ | -0.105 |
| $Rain_{t-2}$ | -0.074 | $Rain_{t-2}$ | -0.058 |
| $Rain_{t-3}$ | -0.116 | $Rain_{t-3}$ | -0.126 |
| $Rain_{t-4}$ | -0.093 | $Rain_{t-4}$ | -0.112 |
| $Rain_{t-5}$ | -0.061 | $Rain_{t-5}$ | -0.083 |
| $Rain_{t-6}$ | -0.071 | $Rain_{t-6}$ | -0.030 |

*Notes*: The table displays the correlation between the price of water in week $t$, denoted by $Price_t$, and rain in period $\hat{t}$, denoted by $Rain_{\hat{t}}$, for $\hat{t} = t, t-1, \ldots, t-6$. The column labeled as "Data," displays these correlations using the data (with the same sample as the one used in the structural model from Section 3, when the auction was run). The column labeled as "Simulation," displays these correlations using the simulated prices and rainfall used in the structural model, as described in Section 3 in the paper.

## C.5 Demand Estimates: Estimates of Additional Specifications of the Model.

For robustness, we also present the demand estimates using the specification in Subsection C.3.1. Table A7 displays the demand estimates using the specifications in Subsections C.3.1 and C.3.2, the latter being the same as the one presented in the paper in Section 5, which is our preferred specification.

Table A7: Structural Estimates

| Estimates | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
|---|---|---|---|---|---|---|---|---|
| Transformation rate ($18 \leq week \leq 32$): | | | | | | | | |
| – Linear term: $\hat{\gamma}_L$ | 0.1584<br>(0.0067) | 0.1790<br>(0.0094) | 0.2124<br>(0.0026) | 0.0734<br>(0.0064) | 0.1449<br>(0.0051) | 0.1804<br>(0.0099) | 0.2040<br>(0.0183) | 0.10494<br>(0.0057) |
| – Quadratic term: $\hat{\gamma}_Q$ | – | 1.36e-04<br>(1.59e-05) | – | 6.19e-05<br>(8.82e-06) | – | 1.36e-04<br>(1.57e-05) | – | 1.53e-04<br>(8.41e-06) |
| Irrigating cost: $\hat{\zeta}$ | 24.3755<br>(6.9848) | 182.174<br>(13.2357) | 78.8924<br>(3.1994) | 34.3495<br>(6.76245) | 11.3193<br>(5.4146) | 183.5007<br>(2.2251) | 69.6141<br>(8.6176) | 201.8714<br>(8.1738) |
| Scale parameter of Gumbel distribution: $\hat{\beta}$ | 1.0100<br>(0.2568) | 1.0854<br>(0.1286) | 0.9361<br>(0.0393) | 1.0144<br>(0.1048) | 1.0252<br>(0.0116) | 1.0321<br>(0.0601) | 0.9923<br>(0.0786) | 1.1987<br>(0.0612) |
| Area heterogeneity | No | No | Yes | Yes | No | No | Yes | Yes |
| Specification of the choice specific error term | C.3.2 | C.3.2 | C.3.2 | C.3.2 | C.3.1 | C.3.1 | C.3.1 | C.3.1 |

*Notes*: Standard errors in parenthesis. See Subsection 6.1 for details about this table.

## C.6 Probability of Being Liquidity Constrained.

In this subsection we discuss how we estimate a lower bound on the probability of being liquidity constrained for the poor farmers during the critical season. The demand estimates (Sections 4 and 5 in the paper) and welfare analysis (Section 6 in the paper) are unaffected by the estimates in this subsection. The estimates in this subsection provide additional evidence about the presence of the liquidity constraints. They complement the evidence in Table A2.

In the data we only observe whether a poor farmer buys water or not, in addition to the number of units they purchase. When a farmer does not buy water, we do not know whether it is because the farmer does not demand water at the equilibrium price and is not liquidity constrained, or whether the farmer is liquidity constrained. That is, for the poor farmers the decision variable is censored. We compute a lower bound on the probability that a poor farmer is liquidity constrained during the critical season using the demand estimates from Section 4, and assuming that the CCPs of the poor and wealthy farmers would coincide during the critical season if the poor farmers were not liquidity constrained. We provide evidence about the latter behavior in Table A.3 in Appendix A.3. The intuition is that farmers are heterogeneous in two dimensions, their productivity and their ability to pay for the water or cash holdings. However, during the critical season, the purchase decisions are determined by the production function of apricots. So the purchase behavior of a potentially constrained farmer who is in the unconstrained state is the same as the purchase behavior of a permanently unconstrained farmer. The "exclusion restriction" for the analysis in this subsection is that poor and wealthy farmers have the same production function (*i.e.* no persistent unobserved heterogeneity). See Subsection C.6 for details.[11]

The cash of farmer $i$ in period $t$, denoted by $\mu_{it}$, evolves according to:

$$\mu_{it} = \mu_{i,t-1} - p_{t-1} j_{i,t-1} + \Phi_t \left( re_i; \phi \right) + \eta_{it} + \nu_{it}, \tag{C.7}$$

[11] This subsection provides a simple procedure to obtain approximate estimates (*i.e.* lower bounds), on the probability of being liquidity constrained for the poor farmers. In Appendix C.8 we discuss two alternatives to estimate the model, and their fit to our data. The first alternative consists of implementing the analysis in this subsection using also the decisions of the poor farmers outside the critical season. The second alternative consists of implementing simultaneously the demand estimation and the analysis in this subsection, using the decisions from all the farmers, wealthy and poor. These alternatives are considerably more complex than our approach. In Appendix C.8 we discuss the complications and additional assumptions needed to implement them.

$$i = 1, \ldots, I, \qquad t = 1, \ldots, T.$$

where $\Phi_t (re_i; \phi) = \phi_{i0} + \phi_1 re_i$ captures the weekly cash flow function derived from the real estate, $\phi re_i$, minus the weekly consumption of individual $i$ that is constant over time, $\phi_{i0}$; $\eta_{it}$ is the farmer's revenue from selling the harvest that we define below; and $\nu_{it}$ is an idiosyncratic financial shock that we specify below. The farmer collects the revenue after the harvest, in week 24. Thus, the revenue, $\eta_{it}$, is:

$$\eta_{it} = \begin{cases} 0 & \{t : w_t \neq 24\} \\ R_{it} & \{t : w_t = 24\} \end{cases}, \tag{C.8}$$

where $R_{it} = \sum_{w_t=1}^{52} h(M_{t-1}, w_t) = \sum_{w_t=18}^{32} \gamma \times (M_{t-1} - PW) \times KS(M_t)$.

Given that farmers buy more than one unit very seldom, we assume $J = 1$ in this subsection, and focus only on the decision of buying *vs.* not buying. The probability that farmer $i$ is liquidity constrained in period $t$, denoted by $\mathbb{P}(p_t j_{it} > \mu_{it})$, is given by:

$$\begin{aligned} \mathbb{P}(p_t j_{it} > \mu_{it}) &= \mathbb{P}(p_t j_{it} > \mu_{i,t-1} - p_{t-1} j_{i,t-1} + \Phi_t (re_i; \phi) + \eta_{it} + \nu_{it}) \\ &= \mathbb{P}(\nu_{it} < -C_{it}) \\ &= \mathbb{F}_\nu (-C_{it}) \end{aligned} \tag{C.9}$$

$$i = 1, \ldots, I, \qquad t = 1, \ldots, T,$$

where the first line follows from the equation in (C.7); $C_{it} \equiv \mu_{i,t-1} - p_{t-1} j_{i,t-1} + \Phi_t (re_i; \phi) + \eta_{it} - p_t j_{i,t}$; and $\mathbb{F}_\nu(\cdot)$ denotes the cumulative distribution function of $\nu_{it}$.

Similarly, the probability that farmer $i$ is not liquidity constrained in period $t$, denoted by $\mathbb{P}(p_t j_{it} \leq \mu_{it})$, is given by:

$$\begin{aligned} \mathbb{P}(p_t j_{it} \leq \mu_{it}) &= \mathbb{P}(\nu_{it} \geq -C_t) \\ &= 1 - \mathbb{F}_\nu (-C_t) \end{aligned} \tag{C.10}$$

$$i = 1, \ldots, I, \qquad t = 1, \ldots, T,$$

where the second line uses the symmetry of the distribution of $\nu_{it}$.

In the data we only observe whether a farmer buys water. When a farmer does not buy water, we do not know whether it is because they do not need the water and have a low valuation or because they are liquidity constrained and have a high valuation. That is, for the liquidity-constrained farmers, the decision variable is censored. An additional complication is that we know that the wealthy farmers are not liquidity constrained, but we do not know which of the poor farmers are liquidity constrained. In this subsection, we estimate a lower bound on the probability of being liquidity constrained for each of the poor farmers during the critical season using the procedure described below.

There are three main difficulties in estimating the probabilities in the equations in (C.9) and (C.10) for the poor farmers who are potentially liquidity constrained. First, the revenue, $\eta_{it}$, from the equation in (C.7) is unobserved in the data. Recovering the revenue requires an estimate of the production function. We estimate the parameters that characterize the production function in Section 4 in the paper using only the purchase decisions of the wealthy (unconstrained) farmers, and the moisture level resulting from the actual purchase decisions of the poor (potentially constrained) farmers.

Second, the conditional choice probabilities of the poor farmers and the wealthy farmers may differ outside the critical season, but will coincide during the critical season if the poor farmers were not liquidity constrained. Outside the critical season, the purchase behavior of a potentially constrained farmer who is in the unconstrained state may differ from the behavior of a potentially constrained who is in the constrained state. For example, a potentially constrained farmer may abstain from purchasing water outside of the critical season, even when feasible, to make sure that they are not constrained during the critical season when the marginal return on water is higher. Similarly, a potentially constrained farmer who believes that they may be constrained during the critical season will purchase water before the critical season and "store" it within the trees and soil by increasing their moisture. Part of the water will evaporate during the critical season, but this is the best the farmer can do when they are constrained during the critical season. The latter behavior can be seen in the paper in Figure 2 and Table 2, which show that potentially constrained farmers buy more water before the critical season. For these reasons, using the smooth conditional choice probabilities from the wealthy farmers would underestimate the probability of being liquidity constrained for a potentially constrained farmer who believes that may be constrained during the critical season. During the critical season, however, the purchase decision is determined by the tree's need for water (stages II, III, and the early post harvest as depicted in Figure 1 in the paper), conditional on the moisture. Thus, during the critical season, the purchase behavior of a potentially constrained farmer who is in the unconstrained state is the same as the purchase behavior of a permanently unconstrained farmer. This is captured by their conditional choice probabilities, as discussed below. In turn, these purchase behaviors are the same as the purchase behavior that would be observed for a potentially constrained farmer who is in the constrained state had the farmer not been constrained. Table A3 provides evidence about this. During regular years (*i.e.* years without droughts), potentially constrained farmers who believe that they may be constrained during the critical season, purchase more water than unconstrained farmers before the critical season (*i.e.* before the uncertainty about rain is realized). However, their purchases are not statistically different from the purchases of the wealthy farmers during the critical season (*i.e.* after uncertainty about rain is realized) in regular years (Table A3 column 2). In a dry year, however, when poor farmers are likely to be liquidity constrained, wealthy farmers do buy more water during the critical season (Table A3 column 4). We interpret this as evidence that, during the critical season, and conditional on moisture, the smooth conditional choice probabilities of the wealthy (unconstrained) farmers can be used to infer the purchase behavior of the poor (potentially constrained) farmers in the counterfactual, unobserved scenario that the latter were unconstrained. For these reasons, in the procedure below we estimate the probability of being liquidity constrained for the poor farmers only during the critical season.

Finally, weekly water consumption is also unobserved. In principle, the weekly consumption can be estimated using the procedure described below under the additional assumption that, during the critical season, the weekly consumption for a potentially constrained farmer who is in the unconstrained state is the same as the weekly consumption of a potentially constrained farmer who is in the constrained state. This assumption may be violated if, for example, one of the poor farmers is permanently unconstrained. For this reason, rather than estimating the weekly consumption of the poor farmers, we set it equal to zero for all farmers, and estimate an upper bound on the probability that the farmer is liquidity constrained. That is, with positive weekly consumption, the probability of being liquidity constrained will be higher than the one we estimate, but it will still be contained within our bounds.

To summarize, we estimate a lower bound on the probability of being liquidity constrained during the critical season for potentially liquidity constrained farmers as follows. First, we generate the actual revenue of the farmer using the estimated demand system and the moisture level resulting from the actual purchase decision of the poor farmers. Second, we focus on the decisions under the critical season and exploit the fact that, during the critical season, purchase decisions are determined by the production function of apricots. Finally, we set the consumption of the poor farmers equal to zero and obtain an upper bound on these probabilities.

We now describe a simple procedure to estimate a lower bound on the probability of being liquidity constrained for the poor farmers during the critical season. To simplify notation, in what follows we omit conditioning on the state variables. Everything is, however, conditional on the state. Let the estimated smooth conditional choice probability (CCP) of not buying water, *i.e.*, $j_{it} = 0$, for a wealthy farmer be $\hat{\mathbb{P}}_{CCP}(j_{it} = 0)$. Similarly, let the estimated smooth CCP of buying water, *i.e.*, $j_{it} = 1$, for a wealthy farmer be $\hat{\mathbb{P}}_{CCP}(j_{it} = 1)$. For potentially liquidity constrained farmers, define the following variable:

$$\tilde{j}_{it} = \begin{cases} j_{it} & if \;\; p_t j_{it} < \mu_{it} \\ 0 & if \;\; p_t j_{it} \geq \mu_{it} \end{cases} .$$

Then, during the critical season:

$$\begin{aligned} \mathbb{P}(\tilde{j}_{it_c} = 0) &= \mathbb{P}[(j_{it_c} = 0 \wedge p_{t_c} j_{it_c} < \mu_{it_c}) \vee (p_{t_c} j_{it_c} \geq \mu_{it_c})], \\ &= \mathbb{P}(j_{it_c} = 0)\mathbb{P}(p_{t_c} j_{it_c} < \mu_{it_c}) + \mathbb{P}(p_{t_c} j_{it_c} \geq \mu_{it_c}), \end{aligned}$$

$$i = 1, \ldots, I, \qquad t_c = \{t : 18 \leq w_t \leq 32\},$$

where the second equality follows because during the critical season, purchase decisions are determined by the production function of apricots, thus making $j_{it}$ independent of $\mu_{it}$ due to the former being determined by the observable states and the *i.i.d.* productivity shocks.

Thus:

$$\hat{\mathbb{P}}\left(\tilde{j}_{it_c}=0;\chi\right)=\hat{\mathbb{P}}_{CCP}\left(j_{it_c}=0\right)\left[1-\mathbb{F}_{\nu}\left(-C_{t_c};\chi\right)\right]+\mathbb{F}_{\nu}\left(-C_{t_c};\chi\right), \tag{C.11}$$

$$i=1,\ldots,I, \qquad t_c=\{t:18\leq w_t\leq 32\}.$$

where $\chi\equiv(\phi_{i0},\phi_1,\sigma_{\nu}^2)$ is a parameter vector.

Using the same argument as above:

$$\begin{aligned}\mathbb{P}(\tilde{j}_{it_c}=1)&=\mathbb{P}[(j_{it_c}=1\wedge p_{t_c}j_{it_c}<\mu_{it_c})],\\&=\mathbb{P}(j_{it_c}=1)\mathbb{P}(p_t j_{it_c}<\mu_{it_c}),\end{aligned}$$

$$i=1,\ldots,I, \qquad t_c=\{t:18\leq w_t\leq 32\}.$$

Thus:

$$\hat{\mathbb{P}}\left(\tilde{j}_{it_C}=1;\chi\right)=\hat{\mathbb{P}}_{CCP}\left(j_{it_C}=1\right)\left[1-\mathbb{F}_{\nu}\left(-C_{t_C};\chi\right)\right]. \tag{C.12}$$

Note that $\mathbb{P}\left(\tilde{j}_{it_C}=0\right)+\mathbb{P}\left(\tilde{j}_{it_C}=1\right)=1$.

We estimate the parameter vector by maximizing the log-likelihood function:

$$\chi=\underset{\chi}{arg\,max}\sum_{i=1}^{I}\sum_{\tilde{t}_c}\mathbf{1}\left(\tilde{j}_{i\tilde{t}_c}=0\right)log\hat{\mathbb{P}}\left(\tilde{j}_{i\tilde{t}_c}=0;\chi\right)+\mathbf{1}\left(\tilde{j}_{i\tilde{t}_c}=1\right)log\hat{\mathbb{P}}\left(\tilde{j}_{i\tilde{t}_c}=1;\chi\right),$$

$$i=1,\ldots,12, \qquad \tilde{t}_c=\{t:18\leq w_t\leq 32\},$$

where $\hat{\mathbb{P}}\left(\tilde{j}_{it_C}=0;\chi\right)$ and $\hat{\mathbb{P}}\left(\tilde{j}_{it_C}=1;\chi\right)$ are given by the equations in (C.11) and (C.12), respectively; and $i$ index poor farmers.

As discussed above, for the estimation we set the weekly consumption to zero, $\phi_{i0}=0$, and obtain a lower bound on poor farmers' probability of being liquidity constrained during the critical season. We also set the cash flow derived from farmers' real state to zero, $\phi_1=0$, because there is no variation in this variable in the data with 10 out of the 12 poor farmers having no urban real estate (*i.e.* $re_i=0$ for $i=1,\ldots,10$ among the poor farmers). We let $\nu_{it}\sim\mathbf{N}\left(0,\sigma_{\nu}^2\right)$. Thus, $\chi=\sigma_{\nu}^2$. Finally, we use the estimated distribution of cash holdings to generate the initial distribution as follows: (i) start the cash holdings at an arbitrary initial level of zero in 1955; (ii) use the first three years in the sample (years 1955, 1956, and

1957) to generate an initial distribution of cash determined by the model conditional on the parameter vector; and (iii) use the remaining years in the data (years 1958 to 1966) and the generated initial condition in (ii) to perform the estimation.[12]

### Estimation Results

The estimated parameter from this subsection is $\hat{\sigma}_{\nu}^{2} = 184.32$ with a standard error of 58.29 for the poor farmers.[13] The distribution of yearly estimated lower bounds of poor farmers' probability of being liquidity constrained during the critical season, defined as the mean across $\tilde{t}_c$ by farmer, where $\tilde{t}_c = \{t : 18 \leq w_t \leq 32\}$, are given in Figure A8. They correspond to $\hat{\mathbb{P}}\left(\tilde{j}_{itC} = 0; \hat{\chi}\right)$ and $\hat{\mathbb{P}}\left(\tilde{j}_{itC} = 1; \hat{\chi}\right)$, given by the equations in C.11 and C.12, with $\hat{\chi} = (0, 0, \hat{\sigma}_{\nu}^{2})$. In all years the estimated probabilities range from a minimum of zero to a maximum of one, indicating substantial heterogeneities in the probabilities of being liquidity constraints during the critical season for the poor farmers.[14] Some of the poor farmers are liquidity constrained with probability one, while others are not liquidity constrained. As expected, the mean probability increases during the dry years of 1960 and 1964.

## C.7 Welfare Measures.

In this subsection we describe how we construct the welfare measures. Given rainfall and the allocation of water among farmers, the yearly average revenue per tree for farmer $i$ is given by:

$$Revenue_i = \frac{1}{\#\ trees_i} \frac{1}{T} \sum_{t=1}^{T} \left[ \sum_{w_t=1}^{52} h\left(M_{i,t-1}, w_t\right) - \zeta_j \right]. \tag{C.13}$$

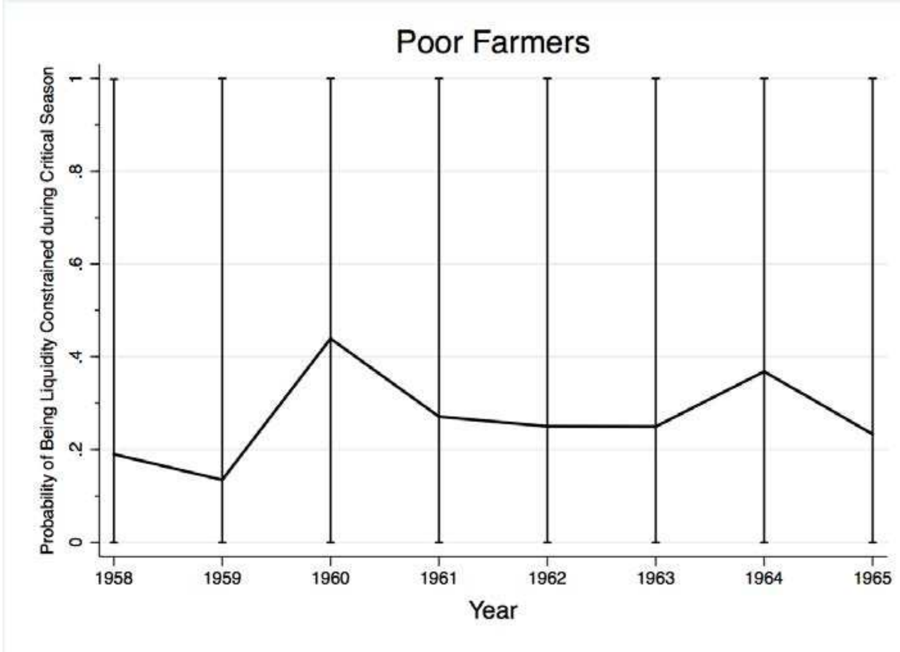
Note that we do not take into account water expenses because we are interested in welfare measures (*i.e.* transfers are not taken into account). We define welfare as follows:

[12] This is the standard approach in the industrial organization literature to deal with the unobserved initial condition of the inventory (see *e.g.* Hendel and Nevo, 2006, p. 1647, where our unobserved initial cash holding is analogous to Hendel and Nevo's unobserved initial inventory). We have experimented with different initial conditions and obtained similar results.

[13] The standard error is computed by bootstrapping from the asymptotic distribution of the parameters in Section 4 in the paper.

[14] Note that the figure is informative about these heterogeneities, because its displays the distribution of estimates *across* poor farmers.

Figure A8: Lower Bound on the Probability of Being Liquidity Constrained




*Notes:* The figure displays the estimated lower bounds of the probability of being liquidity constrained (PLC) during the critical season for the poor farmers using the procedure described in Subsection C.6. Each vertical line displays the distribution of the mean PLC across farmers, defined as the the mean across $\tilde{t}_c$ by farmer with $\tilde{t}_c = \{t : 18 \leq w_t \leq 32\}$. Each vertical line displays the maximum PLC (upper whisker), mean (solid line), and minimum PLC (lower whisker). The figure shows the distribution across poor farmers.

$$Welfare_i = \frac{1}{\#\ trees_i} \frac{1}{T} \sum_{t=1}^{T} \left[ \sum_{w_t=1}^{52} h\left(M_{i,t-1}, w_t\right) - \zeta_j + \epsilon_{ijt} \right]. \tag{C.14}$$

**Auctions using Complete Units ($Ac$).**

We compute both revenue and welfare.

- **Poor farmers.** We compute revenue using the estimated demand system, $\hat{\Theta} \equiv \left(\hat{\gamma}, \hat{\sigma}_\varepsilon, \hat{\zeta}\right)$. and actual purchases made by poor farmers. We use equations C.13 (revenue) and C.14 (welfare), and the moisture level in the farmers' plots (*i.e.* the moisture resulting from their actual purchase decisions).

- **Wealthy farmers.** We compute the revenue using the estimated demand system, $\hat{\Theta} \equiv \left(\hat{\gamma}, \hat{\sigma}_\varepsilon, \hat{\zeta}\right)$, and the actual purchases made by wealthy farmers. We use equations C.13 (revenue) and C.14 (welfare), and moisture level in the farmers' plots (*i.e.* the moisture resulting from their actual purchase decisions). Note that the revenue for wealthy farmers can be greater than the $HVc$ average revenue. This is because poor farmers are sometimes liquidity constrained, so wealthy farmers buy more water than the amount required by the $HVc$ allocation.

**Quotas ($Q$).**

Revenue and welfare coincide under the quota system because farmers do not choose when to irrigate. We only report one measure that we call "welfare." As explained in Section 2 in the paper, in this paper we focus on the 24 farmers who only grow apricot trees. These farmers bought 633 units of water under the auction system over the sample period. Under the quota system, we allocate the same number of units of water (633 units) in each week when these units were bought under the auction. In the empirical application the quota implemented was closest to $Qc25\%$. We also compute the welfare under other quota configurations, where we allocate units among the farmers as follows:

- **Quotas with random assignment of complete units, $Qc$.** Every time we observe that a farmer bought a unit of water during the auction on a particular date, the complete unit of water is assigned uniformly at random, proportional to their amount of land, among all farmers.

- **Quotas with non-random assignment of complete units, $QcX\%$.** Every time we observe that a farmer bought a unit of water during the auction on a particular date, the complete unit of water is assigned uniformly at random, proportional to their amount of land, among the X percent of farmers who had not received irrigation for the longest amount of time, on the same date. That is, we keep track of when the last time was that each farmer irrigated under the quota system. Then, to allocate a unit of water on week $t$, we only consider the subset of farmers whose last irrigation period was furthest away from $t$. This is the subset of farmers who value water most. Then we allocate the unit of water uniformly at random, proportional to farmers' amount of land, among this subset of farmers. The value of $X$ defines how large this set is. For example, if $X = 100\%$, then all farmers are included in the set and the unit of water is allocated uniformly at random, proportional to their amount of land, among *all* farmers. Formally, the subset is defined as follows. Let $t_i^{Last} < t$ be the last week farmer $i$ was allocated a unit of water under the quota system. Let $I$ be the total number of farmers and let $\mathcal{I}$ be the set of all farmers. Let us index the farmers according to the last time that each farmer irrigated, being farmer $I$ the one who irrigated in the week closest to $t$ and being farmer 1 the farmer who irrigated in the week farthest away from $t$. Then $t_1^{Last} \leq t_2^{Last} \leq t_3^{Last} \leq \cdots \leq t_I^{Last}$. (Note that such ranking can always be done and, typically, can be done using several strict inequalities, depending on how many units have been allocated in the past.) Let $X = {}^{x}\!/_{I} \times 100$ for $x \in 1, 2, \ldots, I$. So given $X$, we can compute $x = {}^{X}\!/_{100} \times I$. Then, under $QcX\%$ we allocate the unit of water uniformly at random, proportional to farmers' amount of land, among the subset of farmers $\tilde{\mathcal{I}}_{X\%} \equiv \{i \in \mathcal{I} : i \leq x, \ with \ x = {}^{X}\!/_{100} \times I\}$. For example, if $I = 10$, $t_1^{Last} \leq t_2^{Last} \leq t_3^{Last} < t_4^{Last} \leq \cdots \leq t_{10}^{Last}$, and $X = 30\%$, then $x = {}^{30}\!/_{100} \times 10 = 3$ and $\tilde{\mathcal{I}}_{30\%} = \{1, 2, 3\}$. So, we allocate the unit of water uniformly at random, proportional to their amount of land, among farmers indexed as 1, 2, and 3. These are the three farmers whose last irrigation was farthest away from $t$. In case of ties, we include all tied farmers in the subset $\tilde{\mathcal{I}}$. In the previous example, if if $t_1^{Last} \leq t_2^{Last} \leq t_3^{Last} = t_4^{Last} = t_5^{Last} < t_6^{Last} \leq \cdots \leq t_{10}^{Last}$, then $\tilde{\mathcal{I}}_{30\%} = \{1, 2, 3, 4, 5\}$. For example, in $Qc50\%$, complete units of water are allocated among the 50 percent of farmers who did not receive irrigation the longest; in $Qc25\%$, complete units of water are allocated among the 25 percent of farmers who did not receive irrigation the longest; and so on. As indicated before, under $QcX\%$ we need to keep track of when the last time was that each farmer irrigated under the quota system. We do not have this information for the initial weeks in the sample. So, under $QcX\%$, we allocate units uniformly at random, proportional to farmers' amount of land, at the beginning of the sample as described in the procedure above. In $Qc$ and $QX\%$ units are allocated uniformly at random, proportional to their amount of land, among the corresponding set of farmers. We simulate the allocation $S = 1,000$ times under $Qc$ and $QX\%$. In Table 4 in the paper we report the mean welfare measures across simulations.

**Highest Valuation using Complete Units ($HVc$).**

We compute the highest-valuation allocation using complete units, denoted by $HVc$, as follows. Every time we observe that a farmer bought a unit of water during the auctions on a particular date, the complete unit of water is assigned to the farmer who values water the most on that date.

## C.8 Alternative Estimation Methods and Fit to the Data.

Next we discuss two alternatives to estimate the model, and their fit to the data in our setting. (a) The first alternative consists of estimating probability of being liquidity constrained from Subsection C.6, using also the decisions of the poor farmers outside the critical season, by replacing the smooth conditional choice probabilities with the choice probabilities from the structural model. (b) The second alternative consists of implementing simultaneously the demand estimation in Section 4 in the paper and the estimation of the probability of being liquidity constrained in Subsection C.6, using the decisions from all the farmers, wealthy and poor, and estimating simultaneously both the parameters that characterize demand and the probability of being liquidity constrained for all the farmers.

Both alternatives are considerably more complex than our approach because they would require including: (i) two additional state variables relative to the demand estimation as we implemented it in the paper, the cash holdings, $\mu_{it}$, and the financial shock, $\nu_{it}$; and (ii) an unobserved state variable, the cash holdings. The first point increases the dimension of the state space to five, plus the random shocks, $\epsilon_{it}$ and $\nu_{it}$, thus increasing the computational complexity and, most importantly, the data requirements for identification and to obtain precise estimates, as discussed below. The second point precludes estimating the conditional choice probabilities as we did in the first step of the demand estimation in the paper, thus making no longer possible the "simple" application of the procedure from Hotz and Miller (1993) and Hotz *et al.* (1994).

In principle, one could implement the alternatives in the previous paragraphs using one of the following two approaches. First, a full solution or nested fixed point procedure to solve the dynamic programming problem of the farmers.[15] In the case of alternative (a), one would solve the dynamic programming problem of the poor farmers with two additional state variables relative to the demand estimation as implemented in the paper. In the case of alternative (b), one would solve the dynamic programming problem of all the farmers. Second, an iterative procedure like, *e.g.*, Arcidiacono and Miller (2011) could be used to calculate the conditional choice probabilities with the unobserved state variable $\mu_{it}$, using the expectation–maximization (EM) algorithm. The EM algorithm approach does not require solving the dynamic programming problem of the farmers, but it requires solving the maximization step multiple times, which can be computationally intensive if the unobserved state does not follow a Markov process, or if one needs to discretize a continuous state in small bins, as it is in our case for the cash holdings, $\mu_{it}$, as discussed next.[16]

The data requirements to obtain precise estimates when implementing the two approaches described in the first paragraph are substantially higher than in using our approach. Alternatively, one would need to make additional assumptions that are difficult to justify in the empirical context of Mula, such as, using larger discrete bins for the continuous state variable moisture or the unobserved state variable cash holdings, collapsing purchase decisions of the farmers into a dummy variable for buying instead of modeling the decision of how many units to buy, assuming a Markov process for the unobserved state cash holdings or for the prices, *etc*. Consider, for example, the discretization of the state variables moisture and cash holdings. With five discrete variables, and assuming we discretize the continuous variable moisture into just 11 values, the cash holdings into just 2 values (*e.g.* high and low cash), and the additional assumption that the cash holdings follow a first order Markov process, the number of discrete cells that would arise is $11\times52\times3\times2\times2 = 6,864$. Thus, the average number of observations in each cell, the effective sample size for the conditional choice

[15]See, *e.g.*, Miller (1984), Pakes (1986), Rust (1987), and Wolpin (1984), and Keane and Wolpin (1994, 1997).

[16]A third approach would consist of using a mathematical program with equilibrium constraints, MPEC, (*e.g.* Luo *et al.* 1996; Su and Judd 2012; Dubé *et al.* 2012). The MPEC approach would only require solving for the equilibrium at the final estimate of the structural parameters, avoiding repeatedly solving for an equilibrium at each candidate of the parameter vector, thus reducing the computational burden (see, *e.g.*, Su and Judd 2012 for an application of the MPEC approach to estimate the single agent dynamic discrete choice model by Rust 1987). The main difficulty to apply the MPEC approach to our case is the presence of the latent states. We are not aware of any paper that uses an MPEC approach to estimate dynamic discrete choice models in the presence of latent states. See, *e.g.*, Connault (2016) for a discussion about dynamic discrete choice models with unobserved dynamics.

probabilities, would be $T/6{,}864 = 6{,}864/6{,}864 = 1$ , where $T = 6,864$ is our sample size.[17] But discretizing moisture into 11 values and the cash holdings into two values would be too low because it will not capture the variability in the purchase decisions of the farmers. Importantly, the revenue of the farmers, $\eta_{it}$, that enters into the unobserved state cash holdings in the equation in C.7, depends on the seasonal effect, thus violating the simple Markov assumption. Although in principle one could discretize the continuous variable moisture in smaller bins, and keep track of the cash holdings throughout the year, the computational and identification burden would increase as discussed below. Similarly, collapsing purchase decisions of the farmers into a dummy variable for buying would eliminate, by construction, any variation in the number of units bought when farmers do buy. Such variation is important in the data because poor farmers are more likely than wealthy farmers to buy multiple units before the critical season to "store" such water in their moisture in case they cannot buy water during the critical season due to being liquidity constrained (*e.g.* see Figure 2 in the paper). Not accounting for such variation would artificially generate lower revenue for the poor farmers relative to the wealthy farmers, because the moisture of poor farmers would be artificially lower before the critical season.

There are also computational and identification limitations in implementing the alternative approaches presented above. The full solution approach requires solving the dynamic programming problem of each farmer at each candidate value of the parameter vector. The computational requirements of the full solution approach increase substantially when there is an unobservable, time varying state variable like cash holdings in our setting. Keane and Wolpin (2010) is one of the few papers we are aware of that performs a full solution approach with unobservable, time-varying state variables. The EM algorithm approach requires repeating the maximization step multiple times. Integrating over the unobserved state complicates forming the likelihood of the data required to compute the conditional choice probabilities of being in particular values of the unobserved state. The computational requirements can be mitigated, to a certain extent, using some of the assumptions discussed in the previous paragraph. However, these assumptions are problematic in our case due to their fit to the empirical setting, as discussed above. Finally, identification of the unobserved state variables

[17] Note that: $T = 6,864 = 12$ unconstrained farmers $\times$ 52 weeks per year $\times$ 11 years.

in both approaches is limited by the length of the panel and the variation in the observable state variables (see, *e.g.*, Arcidiacono and Miller 2011).[18]

The specificities of our setting allow us to separate the estimation into two parts, which in turn enables us to exploit variation in our data as described below. For the demand estimation, wealthy farmers are never liquidity constrained, as discussed in the paper. So we do not use the model to estimate their probability of being liquidity constrained. We determine their unconstrained status directly from the data by looking at the value of their urban real estate. Thus, we estimate the parameters that characterize demand without incorporating the unobserved cash holdings. This simplifies the dimension of the state space to four observed state variables, plus the random shocks, $\epsilon_{it}$, that we integrate analytically as described in Subsection C.3. This approach also allows us to use smaller (relative to the alternative approaches presented above) discrete bins for the continuous state variable moisture, and to use the variability in the moisture level resulting from the number of units bought by the farmers rather than a dummy variable for whether to buy. Because we do not have the unobserved cash holdings as a state variable, we apply the procedure from Hotz and Miller (1993) and Hotz *et al.* (1994). In our case in particular, the forward simulation procedure from Hotz *et al.* (1994) has two advantages. First, it allows us to incorporate the continuous variable moisture in a straight manner by using Monte Carlo simulations to approximate continuation values at states not observed in the data. Second, it requires only estimating future choice and transitions probabilities associated with the nodes of some simulated future path of the farmer. Thus, by exploiting the representation by Hotz and Miller (1993), we avoid the full solution that would require solving the dynamic programming problem of each farmer at each candidate value of the parameter vector. By using the forward simulation procedure from Hotz *et al.* (1994), we avoid integration of the value functions over all future paths. By determining the unconstrained status directly from the data, we avoid the EM algorithm that would require solving the maximization step multiple times. For the estimation of the probability of being liquidity constrained, we exploit that in our setting the purchase decision during the critical season is determined by the technological

[18]For further details about advances in the estimation of dynamic discrete choice models see Arcidiacono and Ellickson (2011).

constrained imposed by the production function of apricots. Thus, we use the conditional choice probabilities of the wealthy farmers during the critical season to obtain a lower bound on the probability that the poor farmers are liquidity constrained. The lower bound results from setting the unobserved consumption to zero in Subsection C.6.[19] The assumptions required to implement the estimation of the demand and the probability of being liquidity constrained are, respectively, that wealthy farmers are never liquidity constrained, and that poor and wealthy farmer have the same production function for apricots. We provide evidence to support these assumptions, and discuss their validity in our empirical context in Section 7 in the paper and in Section D in this appendix.

## C.9 Dynamic Demand System.

In this paper we do not model the auction game and, thus, we abstract from the within-week variation in prices. We translate the auction mechanism into a simpler dynamic demand system, whereby individual farmers take prices as exogenous. To that end, we compute the average price per week and assume that farmers can purchase water at this price. This allows us to focus on the dynamic behavior of farmers across weeks in the presence of liquidity constraints, which is the main focus of the paper. In theory, one could nest the full dynamic auction game within the dynamic framework currently used, and estimate the game like in Jofre-Bonet and Pesendorfer (2003). In practice, for the case of Mula, however, it is unfeasible to perform such estimation because of the small variation in average prices per farmer within weeks. Nonetheless, we provide the full dynamic auction model, in Subsection C.10 in this appendix. See Donna and Espín-Sánchez (2018) for details about the dynamics and strategic behavior within four-unit auctions.

In our simulations, the market clears by setting the total number of available units in a given week equal to the total demand from the farmers who buy water in that week. Each week there are 40 units to be assigned among all farmers. Apricot farmers decide whether to buy at the weekly price. In addition, the total number of units that apricot farmers can buy throughout the sample is fixed to 633 units, the actual number of units the apricot farmers

[19] Note that the estimation of the probability of being liquidity constrained in Subsection C.6 is not needed to perform the welfare analysis from Section 6 in the paper.

bought during the sample under the auction.

## C.10 Dynamic Auction Model with Within-Week Price Variation.

In this section, we consider the dynamic bidding case in which the farmer can buy several units of water each week. The purchase is sequential. The farmer is offered a price for the first unit and decides whether to purchase the unit or not. After this decision, the farmer is offered a price for the second unit, and so forth. The prices offered to the farmer follow a stochastic Markov process. The farmer knows the parameters governing this process.

There are 40 units auctioned every week. Before the first price is offered, the farmer observes the rain in the previous week and a $10 \times 1$ vector containing the shocks to the utility for the next week: 5 days, Monday to Friday; 2 schedules, day and night; $\epsilon_{it} = (\epsilon_{it1}, ..., \epsilon_{it10})$. Each value of the shock represents a shock to the utility for all four units in a four-unit auction. We abstract here from the equilibrium played within each four-unit auction, see Donna and Espín-Sánchez (2018) for details.

We index each of the ten units by $k$. Let us denote a purchase of $j$ units of water by farmer $i$, during period $t$ and within the $k^{th}$ four-unit auction by $j_{itk}$, with $0 \leq j_{itk} \leq 4$, and $\sum_{k=1}^{k=10} j_{itk} = j_{it} \leq 40$. Also denote by $p_{tk}$ the price associated with buying any unit within the $k^{th}$ four-unit auction in period $t$, and by $V(M_{it}, x_{it}, \epsilon_{it})$ the value of a farmer to participate in a 40-unit auction at week $t$, where $\epsilon_{it} \equiv (\epsilon_{it1}, ..., \epsilon_{it10})$ is an unobserved state and $x_t \equiv (w_t, p_t, r_t, \mu_{it})$ is an observed state. . Let $\varphi_{itk}$ be a state variable in the within-period game. The variable $\varphi_{itk}$ includes the units of water already purchased by the farmer in period $t$ up to auction $k-1$, thus $\varphi_{it1} = 0$ and $\varphi_{itk} = \sum_{l=1}^{l=k-1} j_{itl}$. Hence, for the within week auction game we have:

$$V(M_{it}, x_{it}, \epsilon_{it}) = h(M_{it}, w_t; \gamma) - \sum_{k=1}^{k=10} \left(j^*_{itk} p_{tk} + \zeta_{jk} + \epsilon_{ijtk}\right) + \beta V(M_{it+1}, x_{it+1}, \epsilon_{it+1}),$$

where $j^*_{itk}$ are the elements of the solution to the game below. We define the value of the farmer of entering the within-week game as:

$$V\left(M_{it},x_{it},\epsilon_{it}\right)\equiv W_1\left(M_{it},x_{it},\epsilon_{it};0\right).$$

Then the (finite) within-week game is:

$$\begin{aligned}
W_1\left(M_{it},x_{it},\epsilon_{it};0\right) &= \underset{j_{it1}\in\{0,\ldots,4\}}{Max}\left\{h\left(M_{it},w_t;\gamma\right)-\left(j_{it1}p_{t1}+\zeta_{j1}+\epsilon_{ijt1}\right)+W_2\left(M_{it},x_{it},\epsilon_{it};\varphi_{it2}\right)\right\},\\
&\vdots\\
W_k\left(M_{it},x_{it},\epsilon_{it};\varphi_{itk}\right) &= \underset{j_{itk}\in\{0,\ldots,4\}}{Max}\left\{-\left(j_{itk}p_{tk}+\zeta_{jk}+\epsilon_{ijtk}\right)+W_{k+1}\left(M_{it},x_{it},\epsilon_{it};\varphi_{itk+1}\right)\right\},\\
&\vdots\\
W_{10}\left(M_{it},x_{it},\epsilon_{it};\varphi_{it10}\right) &= \underset{j_{i,t,10}\in\{0,\ldots,4\}}{Max}\left\{-\left(j_{it,10}p_{t,10}+\zeta_{j,10}+\epsilon_{ijt,10}\right)+W_{11}\left(M_{it},x_{it},\epsilon_{it};\varphi_{it11}\right)\right\}.
\end{aligned}$$

Note that:

$$\beta V\left(M_{it+1},x_{it+1},\epsilon_{it+1}\right)\equiv W_{11}\left(M_{it},x_{it},\epsilon_{it};\varphi_{it11}\right).$$

If we do not assume that all prices are learned at the beginning of the week, but rather, that prices are learned at the beginning of each four-unit auction, then we have:

$$\begin{aligned}
W_1\left(M_{it},x_{it},\epsilon_{it};0\right) &= \underset{j_{it1}\in\{0,\ldots,4\}}{Max}\left\{h\left(M_{it},w_t;\gamma\right)-\left(j_{it1}p_{t1}+\zeta_{j1}+\epsilon_{ijt1}\right)+\mathbb{E}\left[W_2\left(M_{it},x_{it},\epsilon_{it};\varphi_{it2}\right)\right]\right\},\\
&\vdots\\
W_k\left(M_{it},x_{it},\epsilon_{it};\varphi_{itk}\right) &= \underset{j_{itk}\in\{0,\ldots,4\}}{Max}\left\{-\left(j_{itk}p_{tk}+\zeta_{jk}+\epsilon_{ijtk}\right)+\mathbb{E}\left[W_{k+1}\left(M_{it},x_{it},\epsilon_{it};\varphi_{itk+1}\right)\right]\right\},\\
&\vdots\\
W_{10}\left(M_{it},x_{it},\epsilon_{it};\varphi_{it10}\right) &= \underset{j_{i,t,10}\in\{0,\ldots,4\}}{Max}\left\{-\left(j_{it,10}p_{t,10}+\zeta_{j,10}+\epsilon_{ijt,10}\right)+W_{11}\left(M_{it},x_{it},\epsilon_{it};\varphi_{it11}\right)\right\},
\end{aligned}$$

where the expectation is taken with respect to the remaining prices to be disclosed in the current week, and the price sequence follows a Markov chain.

We do not include a discount factor within week because the time from one auction to the following is just a few minutes and the discount factor in this case is virtually 1 (*i.e.* we only include a weekly discount factor, $\beta$). Instead, we can solve this game by backward induction if we know the value of $V_{i,t+1}(\cdot)$. The solution concept is Subgame Perfect Equilibrium in the first case, and Perfect Bayesian Equilibrium if price are learned at every step.

# D Extended Discussion about the Fit of the Model to the Data

In this section we thoroughly discuss the fit of our model to the empirical setting in Mula. To facilitate the reading, we repeat the discussion in Section 7 in the paper.

## D.1 Unobserved Heterogeneity.

The differences in production as a measure of welfare in Table 4 are due to differences in soil moisture levels (*i.e.* some farmers irrigate more than others) because our specification assumes that all farmers are equally productive, up to an idiosyncratic productivity shock. An alternative explanation would be that differences in production are due to unobserved differences in productivity. For example, it could be that wealthy farmers used additional productive inputs (*e.g.* fertilizers, hired labor, manure, *etc.*) in greater quantities than did poor farmers. Thus, poor farmers' production would be lower than wealthy farmers' production due to both differences in soil moisture levels and greater use of these additional productive inputs.

Although we cannot rule out this argument explicitly, it does not affect our counterfactual results from Table 4 in the paper. We cannot rule it out explicitly because we have no data about the relative use of these additional productive inputs and our econometric specification does not allow for persistent differences in productivity among farmers. However, it does not affect our counterfactual results in the historical context of Mula. Artificial fertilizers were not introduced in Mula until the 1970s. Farmers did use manure and mules when farming the land. If poor farmers faced liquidity constraints when buying water, it is reasonable that they also faced LC when buying these other inputs or capital. So if wealthy farmers used additional productive inputs in greater quantities than did poor farmers under the auction, the transition from auctions to quotas would increase the production of poor farmers more than we predicted in the counterfactual from Table 4 in the paper. This is because, under quotas, farmers do not have to make large payments for water (maintenance costs are substantially lower than the prices of water under the auction), leaving them extra cash to buy additional productive inputs. In other words, under quotas poor farmers are less likely to be liquidity constrained to buy additional inputs. Even if poor farmers were less productive than wealthy farmers during the auctions period due to underuse of inputs/capital, they would be as productive as wealthy farmers under the quotas.[20] Next we further explore this issue by generalizing the model and allowing correlation between the wealth and productivity of the farmers.[21]

## D.2 Correlation Between Wealth and Productivity.

Throughout the paper we assume that there are no persistent differences in productivity between wealthy and poor farmers. Although this hypothesis is untestable, we believe it is reasonable in the historical context of Mula. All farmers' plots are located in a small, relatively flat area spanning less than 2 km; thus, weather conditions are the same. To the best of our knowledge, there are no historical sources mentioning (explicitly or implicitly) differences in productivity among farmers, or between wealthy and poor farmers. Table A8, Panel A, shows that although wealthy farmers have larger plots (column 1), when considering all agricultural products there are no differences in revenue per tree (column 5) between poor and wealthy farmers in 1954 (the only year where revenues are observed). Interviews with surviving farmers confirm this. The differences between poor and wealthy farmers (columns 2, 3, and 4) are attributable to the larger plots of wealthy farmers. Note that the year responsible for 1954 revenue was particularly dry (water prices were substantially higher than other years in the sample). So we would expect large differences in revenue per tree if differences in productivity were large. However, Table A8, Panel B, shows that there are only large differences in revenue per tree for farmers who grow only apricot trees. These differences are accounted for by moisture differences (lower water purchases) of poor farmers relative to wealthy farmers during the critical season in 1954.

[20] In terms of the model, this can be interpreted as a weaker assumption required for the welfare results to hold. The welfare analysis only requires that poor farmers are as productive as wealthy farmers under quotas (not under auctions), which is a credible assumption in the historical context of Mula as explained above.

[21] Although our model does not allow for persistent unobserved heterogeneity, we do estimate the parameter $\sigma_\epsilon^2$, which determines the variance of the idiosyncratic shock, $\sigma_\epsilon^2\pi^2/6$. The higher the value of $\sigma_\epsilon^2$, the more heterogeneous the distribution of productivity. If $\sigma_\epsilon^2$ is large enough, auctions are more efficient than quotas because, under quotas, there is no decision nor gains from trade. See Section 6 in the paper.

Moreover, Table A8, Panel B, shows that solely among farmers who grow only apricot trees, do wealthy farmers obtain greater revenue than poor farmers. However, if a farmer grows another agricultural product in addition to apricot trees (*e.g.* oranges), then there are no substantial differences between wealthy and poor farmers. Moreover, revenue for oranges is not correlated with the wealth of the farmer either. This is because oranges are harvested in winter, unlike apricots, which are harvested in the summer when the prices of water in the auction are high. Prices of water during the orange harvest season are low; thus LC plays no role. Farmers who grow both apricots and oranges use the cash obtained in winter from the orange harvest to buy water for apricots in the summer. Similarly, farmers use cash obtained from the apricot harvest to buy water for oranges in winter. Hence, these multi-crop farmers are not affected by LC. Farmers who grow only apricots do not have access to this "cash smoothing mechanism" and are therefore affected by LC. Results for other agricultural products harvested in the summer such as lemons and peaches are similar to those for apricots. The results in Table A8, Panel B, provide evidence of both LC and low-productivity heterogeneity. Column (1) shows that the average revenue per apricot tree for farmers growing only apricots is substantially lower for poor farmers. Column (2) shows that the revenue per orange tree is similar for poor and wealthy farmers. Column (3) shows that the same is true among the farmers who grow apricots and other crops, as well as for lemons and peaches. We interpret these results as evidence that the differences in revenue observed among the farmers who grow only apricots are due to differences in input utilization (*e.g.* water) used by wealthy and poor farmers, and not due to differences in their production function.[22]

The evidence presented above suggests that the correlation between wealth and productivity is small. (The actual correlation coefficient between urban real estate and revenue per tree in 1954 is -0.06.) Nonetheless we performed a sensitivity analysis to examine how large the correlation should be to revert the welfare results from Table 4.[23] We explore this by

[22]When looking at the revenue per tree for wealthy farmers, farmers growing only apricot trees have a greater revenue than farmers growing also other crops. The reason behind this result is that wealthy farmers growing only apricot trees have a lower average number of trees (72 trees) than farmers growing also other crops (109 trees). This is due to disseconomies of scale. The number of trees for poor farmers growing only apricot trees is 73, thus disseconomies of scale play no role when comparing poor and wealthy farmers.

[23]We thank a co-editor, Fabrizio Zilibotti, for this suggestion.

Table A8: Farmers characteristics and wealth.

**Panel A: Size and Composition of Plots and Wealth, for all agricultural products.**

| | (1) Area Total (Ha) | (2) Area with trees (Ha) | (3) Fraction with trees | (4) Revenue (*pesetas*) | (5) Revenue/ area ($pesetas/m^2$) |
|---|---|---|---|---|---|
| Urban real estate | 34,023*** | 22,069*** | -0.0355 | 23,894*** | -0.1797 |
| | (9,747) | (7,031) | (0.0320) | (4,024) | (0.7543) |
| Number of observations | 388 | 388 | 388 | 388 | 388 |

*Notes*: All regressions are OLS specifications. The dependent variable is the variable in each column. "Urban real estate" measures the value of a farmer's urban real estate in pesetas. Standard errors in parentheses. * p<0.10; ** p<0.05; *** p<0.01.

**Panel B: Revenue per tree in 1954 for each agricultural products.**

| | | Apricot (only) (1) | Orange (other) (2) | Apricot (other) (3) | Lemon (other) (4) | Peach (other) (5) |
|---|---|---|---|---|---|---|
| Total | Rev. per tree | 134.21 | 125.13 | 124.70 | 112.92 | 51.81 |
| | # trees | 73.0 | 152.0 | 90.0 | 102.4 | 93.1 |
| Poor | Rev. per tree | 105.47 | 131.65 | 129.69 | 120.37 | 47.09 |
| | # trees | 73.6 | 137.8 | 71.8 | 97.0 | 81.6 |
| Wealthy | Rev. per tree | 162.94 | 119.48 | 119.23 | 105.93 | 55.58 |
| | # trees | 73.4 | 164.6 | 109.0 | 107.4 | 102.3 |
| # farmers | | 24 | 322 | 239 | 64 | 45 |

*Notes*: Own elaboration from the 1954 Agricultural census. "CROP (only)" refers to the revenue generated by CROP trees for farmers that only grow CROP trees. "CROP (other)" refers to the revenue generated by CROP trees for farmers who grow CROP and other trees. (CROP represents Apricot, Orange, Lemon, and Peach.) "Wealthy" ("Poor") is a dummy variable that equals 1 if the value of urban real estate of the farmer is above (below) the median, and 0 otherwise.

generalizing the model and allowing for correlation between wealth and land quality (or the use of additional inputs). One way to do this is to allow the apricot production function, $h(j_{it}, M_t, w_t; \gamma)$, to shift with wealth. Let $\Phi_i$ be a factor multiplying the apricot production function of farmer $i$ and be given by:

$$\Phi_i \equiv 1 + \rho_{w,p}\mathrm{NW}_i + (1 - \rho_{w,p})\vartheta_i \qquad \forall t, \tag{D.1}$$

where $\rho_{w,p} \in [0, 1]$ is the correlation between wealth and productivity, $\mathrm{NW}_i$ is the normalized wealth of farmer $i$ such that $\mathbb{E}(\mathrm{NW}_i) = 0$ and $\mathbb{V}(\mathrm{NW}_i) = 1$, and $\vartheta_i$ is an *i.i.d.* random shock to farmer $i$ such that $\mathbb{E}(\vartheta_i) = 0$ and $\mathbb{V}(\vartheta_i) = 1$. Note that $\mathbb{E}(\Phi_i) = 1$. Also note that if $\rho_{w,p} = 0$, there is no correlation between wealth and productivity but it has permanent heterogeneity unlike the original model. When the variance of the random shock goes to zero, we are back in our original model.

Data about land quality or the use of additional inputs is not available, so it is impossible to pin down the correlation parameter, $\rho_{w,p}$, from the data. To perform the sensitivity analysis we simulate the model for different values of $\rho_{w,p}$ using equation (D.1) as follows. In each simulation $s \in S = 1,000$, each farmer $i \in \{1, \dots, 24\}$ has always the *same* normalized wealth, $\mathrm{NW}_i$, obtained from the data. To avoid arbitrarities choosing the distribution of the white noise, we let $\vartheta_i$ to be a random draw from the normalized empirical wealth distribution (*i.e.* a random draw from $\mathrm{NW}_i$.).[24] Thus, in each simulation $s$, each farmer $i$ has a *different* random draw, $\vartheta_i$ (this introduces noise to the simulation that vanishes progressively as $\rho_{w,p} \to 1$). For each simulation $s$ we obtain $\Phi_i^s$ for farmer farmer $i$. Then we use the same procedure as described in baseline model in the paper. The results displayed in Figure 6 in the paper represents the average across simulations.

Figure A9 shows the sensitivity of the welfare results from Table 4 to the correlation between wealth and productivity, $\rho_{w,p}$, for $\rho_{w,p} \in [0, 1]$. The figure displays the welfare difference between quotas minus auctions as function of $\rho_{w,p}$ and as percentage of the welfare under auctions with $\rho_{w,p} = 0$ (the baseline in Table 4). The top panel displays the welfare of quotas $Qc25\%$ minus the welfare of auctions $Ac$. In our base line case in Table 4 $\rho_{w,p} = 0$ and

[24] Results are almost identical with other distributions, such as a standard normal.

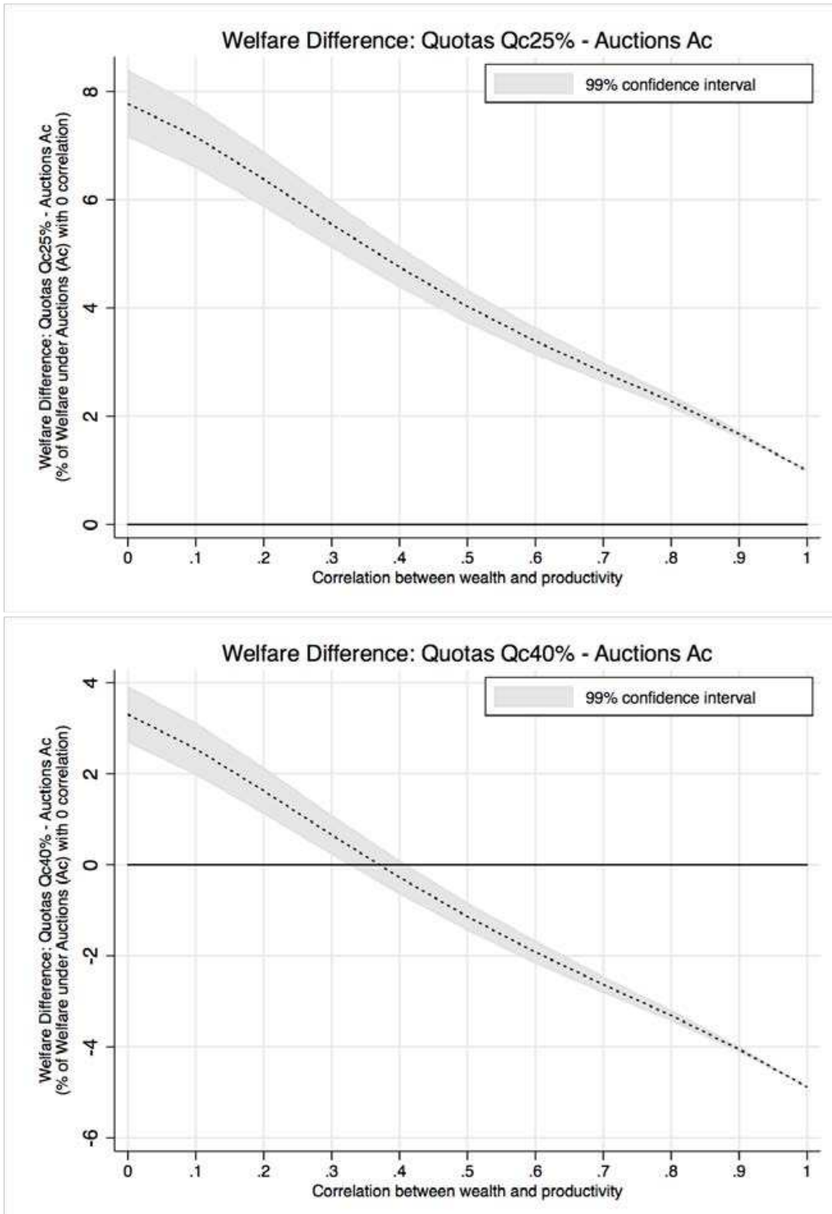
the quotas $Qc25\%$ produce 7.6% $(^{(1{,}480.47-1{,}375.67)}/_{1{,}375.67})$ more output per tree than auctions $Ac$. As expected, as the correlation increases, quotas are relatively less efficient than auctions. (When $\rho_{w,p} \in [-1, 0]$ the welfare difference of quotas minus auctions is larger.) In the extreme case where $\rho_{w,p} = 1$ (*i.e.* wealthy farmers are always more productive than poor farmers with the same soil moisture level), the welfare difference between quotas $Qc25\%$ and auctions $Ac$ is minimal because under auctions wealthy farmers buy more water during the critical season than do poor farmers (Figure 2 in the paper).

The top panel in Figure A9 shows that quotas $Qc25\%$ are more efficient than auctions $Ac$ even when wealth and productivity are perfectly correlated (*i.e.* when $\rho_{w,p} = 1$). This may seem counterintuitive because by moving from quotas $Qc25\%$ to auctions $Ac$ there is a transfer of water from wealthy (more productive) to poor (less productive) farmers according to equation D.1. However, note that equation D.1 defines a shift in productivity (*i.e.* wealthy farmers are more productive than poor farmers) for farmers with the same soil moisture level. Under auctions $Ac$, wealthy farmers have substantially higher levels of moisture than do poor farmers and, thus, wealthy farmers are less productive than poor farmers (due to the concavity of the apricot production function, created by the cap on the amount of moisture absorbed by the soil, even when wealthy farmers are more productive than poor farmers for the same level of moisture). Hence, a redistribution of water from wealthy to poor farmers under quotas $Qc25\%$ results in a net increase in efficiency because the increase in efficiency due to the concavity of the production function is greater than the decrease in efficiency due to the water being used by "less productive" poor farmers (the latter effected captured by equation D.1).

The bottom panel in Figure A9 displays the welfare of quotas $Qc40\%$ minus the welfare of auctions $Ac$.[25] In our base line case in Table 4 $\rho_{w,p} = 0$ and the welfare difference of auctions $Qc40\%$ minus auctions $Ac$ is approximately 3 percent. Now, as $\rho_{w,p}$ increases, quotas $Qc40\%$ are less productive than auctions $Ac$ in contrast to the top panel, where quotas $Qc25\%$ are always more efficient than auctions. Note, however, that both panels in Figure A9 show that auctions are relatively more efficient than quotas as $\rho_{w,p}$ increases (downward slope). This

[25] Notice that Figure A9 shows $Qc40\%$ instead of $Qc50\%$. We do this because $Qc50\%$ is not statistically different than the auction, even when productivity is not correlated with wealth.

Figure A9: Efficiency gains as a function of the correlation between wealth and productivity.




*Notes*: See Subsection C.7 for a discussion about the computation of the welfare measures in this figure. Confidence intervals account for uncertainty about the estimated parameters (by drawing from the asymptotic distribution) and across simulations.

is due to the shifter in productivity from equation D.1. In each panel the mechanisms to allocate water are fixed ($Qc25\%$ and $Ac$ in the top panel, and $Qc40\%$ and $Ac$ in the bottom panel), so there is no increase in efficiency as $\rho_{w,p}$ varies. The increase in efficiency due to the concavity in the production function can be seen in Figure 4 in the paper for a given value of correlation between wealth and productivity, $\rho_{w,p} = 0$.[26]

## D.3 Strategic Supply.

The president of the *Heredamiento de Aguas* decided whether to run the auction or not. There is no evidence that running the auction or not was a strategic decision. If there was enough water in the dam, the auction was held. However, the president could stop the auction at any time, and used to do so if the price fell considerably, usually to less than 1 *peseta*. This uncommon situation happened only after an extraordinarily rainy season.

## D.4 Strategic Size and Sunk Cost.

The results obtained when comparing revenue from quotas and auctions suggest that the choice of the unit size in the auction (*i.e.* three hours of irrigation) was not innocuous. In particular, the fact that in some years poor farmers under the quota system produced higher revenue than wealthy farmers under the auction system suggests that the size of the units sold at the auction might be too large. The size of the units sold at auction had not changed since the middle ages. This could be due to institutional persistence or due to other reasons, *i.e.*, three hours could be the size that maximizes revenue. It could be the case that three hours maximizes profits, but not welfare.

As shown in Donna and Espín-Sánchez (2018), there is a sunk cost to the first unit of water allocated to a plot because the dry channel absorbs some water. Subsequent units associated with the same channel flow through a wet channel, thus, the loss is negligible for subsequent units. In the auction system, subsequent units are allocated to different farmers, depending on who has won each unit. The optimal size of the unit (*i.e.* the size of the unit that maximizes welfare) would be determined by a trade-off between the sunk cost incurred every time a farmer irrigates, due to the loss of water flowing through a dry channel, and

[26] In principle, one could argue that the shifter in productivity from equation D.1 could be large enough such that the slope of the lines in Figure A9 were steeper and, hence, auctions $Ac$ outperformed quotas $Qc25\%$ for large values of $\rho_{w,p}$. We believe this is not the case in the historical context of Mula given the information presented in Table A8, where the correlation coefficient between wealth and revenue per tree in 1954 is $\rho_{w,p} = -0.06$. Also as emphasized above, data about land quality or the use of additional inputs is not available, so it is impossible to pin down the contribution of the shifter in productivity, $\rho_{w,p}$, from the data. Note that if the production function is linear, and wealth and productivity are perfectly correlated, auctions are always more efficient than any mechanism of quotas.

the diminishing return of water. In the quota system, units are allocated to each farmer in geographical order (*i.e.* every unit is allocated to a neighbor farmer down the channel with respect to the previous farmer). Therefore, the sunk cost due to the sunk water is minimal.

## D.5 Optimal Mix of Crops and Size.

Our analysis only considers the case in which farmers grow apricot trees. Since different agricultural products have different irrigation needs in different seasons, the optimal crop mix involves diversifying among several agricultural products with different irrigation needs. For example, oranges are harvested in winter, and their need for water peaks in December. Apricots are harvested in summer, and their need for water peaks in May-June. Hence, a mix of crops with apricot and orange trees would outperform one with just apricot trees, in terms of spending smoothing. We observe this optimal mix in the data. Many farmers have orange trees and either apricot, peach, or lemon trees, all three of them are harvested during summer. In this paper, we focus on the set of farmers who only grow apricot trees because they have the same production function. This allows us to account for unobserved heterogeneity without modeling it, as discussed in Section 7 in the paper.

If farmers with a summer-winter crop mix can smooth spending, and avoid being liquidity constrained, then we would observe all, not just most, farmers with such a crop mix. However, for a variety of reasons, Mula farmers did not all grow mixed crops. First, the farmers in Mula inherited their land from their fathers who, in turn, inherited their land from their fathers, who, in turn, inherited their land from their fathers, and so on. From the point of view of the farmers studied in this paper, this means that there was no initial "choice" for the type of crops. Moreover, if there are "fixed costs" per crop, the farmers would be better off with only one crop. Second, if some farmers are liquidity constrained to buy water, they *might* also be liquidity constrained to buy more land to plant the second crop (*e.g.*, oranges). A critique of the previous arguments is that wealthier farmers could just buy the land of the poor farmers, and consolidate it with their own land. However, as shown in, *e.g.*, Hoffman (1996), growing fruit trees exhibits "decreasing marginal returns" in terms of the size of the land with trees. This means that there is an optimal number of trees. This is because the owner of a large plot cannot work all the land by himself, but rather has to hire workers, who he would have to supervise and/or accept a lower quality of work from. This is indeed what we see in the data when more than 90% of the parcels in Mula were smaller than one hectare.

## D.6 Trees.

Quotas are desirable during a drought because they allocate a certain amount of water periodically to each farmer. Quotas also function as insurance for farmers, who have less uncertainty when carrying out risky investments, such as trees. A tree takes several years to be fully productive, but will die if it does not get enough water in any given year. On the other hand, vegetables grow more quickly than trees, and can be harvested within a year of planting. Hence, a farmer with a secure supply of water is more likely to plant trees and receive a higher expected profit from them.

## D.7 Collusion.

The presence of a centuries-stable market alongside repeated interaction among farmers raises a concern about bidding collusion in the auction. Historical sources (González Castaño and Llamas Ruiz, 1991) and personal interviews with surviving farmers point in the opposite direction.[27]

Rather than a system in which farmers colluded to pay a price lower than what would have prevailed without collusion, there seemed to be bitterness among farmers competing for water, and between farmers and employees of the cartel. Fights, loud arguments, and complaints were common. In many instances the police intervened during the auction to guarantee its normal development. See Donna and Espín-Sánchez (2018) for a detailed investigation of bidding collusion in this setting.

[27] A summary of the interview is available upon request.

## D.8 Liquidity Constraints *vs.* Risk Aversion or Impatience.

One concern to identify LC is that some empirical implications of markets where agents face liquidity constraints are similar to those of markets where agents are risk averse. In particular, poor farmers buying water before the critical season (*i.e.* before uncertainty about rain is realized) is consistent with both LC and risk aversion. We now use the response of poor farmers to their purchase timing to investigate this concern.

The main difference in farmers' behavior under LC and risk aversion occurs during the summer, when prices are high. On the one hand, if poor farmers are liquidity constrained, they would not be able to buy water when the price is high, even if the moisture level in their plots is low. On the other hand, if farmers are unconstrained but risk averse, they would have the same demand for water as wealthy farmers during the summer, conditional on soil moisture levels. In Table 2, Panel B, column 4 in the paper we show that holding the moisture level fixed, poor farmers buy less water than do wealthy farmers. Following the results in this table, along with the opinions presented above, we conclude that poor farmers were liquidity constrained.[28]

The same argument rules out the possibility that the results are driven by poor farmers being more impatient (lower discount factor) than wealthy farmers. If poor farmers were more impatient, their soil moisture level would be always lower than that of wealthy farmers: an extra peseta spent on water has an immediate cost and a future reward. However, poor farmers have higher moisture levels than wealthy farmers before the critical season, lower moisture levels during the critical season, and again high moisture levels right after the critical season (Figure 2 in the paper). This behavior rules out differences in discount factors between wealthy and poor farmers.

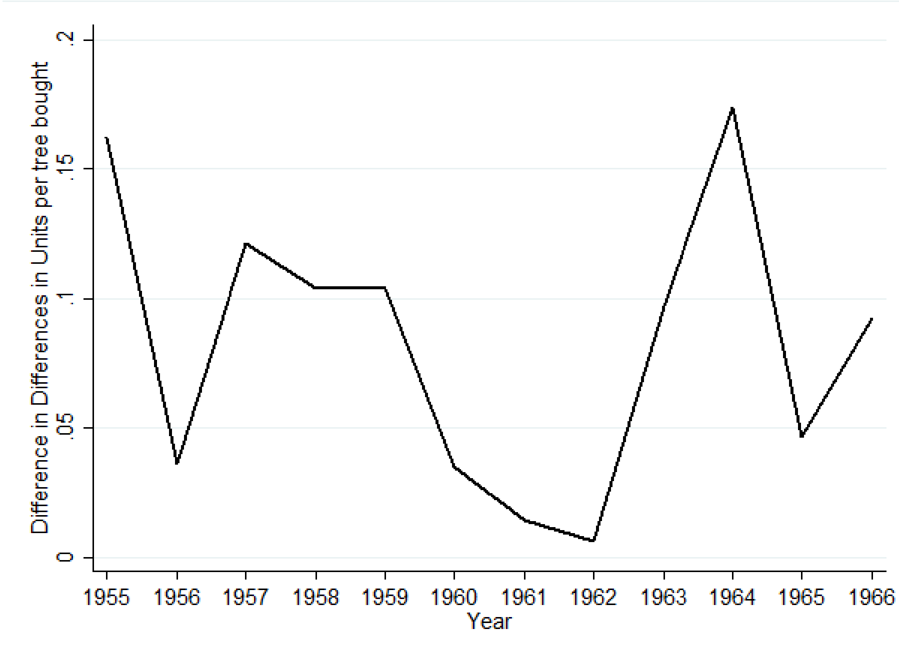
[28]In this paper we abstract from differences in prices within the week (*i.e.* Monday to Friday, and Day to Night). However, differences in prices within the week can also be used to asses the importance of liquidity constraints. As shown in Donna and Espín-Sánchez (2018) prices are higher for night-time irrigation and higher earlier in the week (prices on Mondays are higher than on Fridays). Although not reported here, we find that poor farmers are more likely to buy water during nights and later in the week.

## D.9 Attrition.

While we have weekly panel data about water purchases, our data only contains only one cross-sectional observation of the characteristics of the farmers. The cross-sectional characteristics of the farmers were obtained from a detailed agricultural census carried by the Franco regime in 1955. This agricultural census took place only once, in 1955, to estimate the national capacity to produce agricultural products. One concern about observing cross-sectional characteristics only once is potential attrition in the data. For example, it could be that some of the farmers who only grew apricots in 1955 switched to growing apricots and oranges during the following decade. The incentives to plant other trees, in particular orange trees, would be greater for poor farmers facing liquidity constraints than for wealthy farmers. If poor farmers switched, then we should expect a change in poor and wealthy farmers' relative purchase of water during the critical season versus the rest of the year (difference in differences).

We investigate this issue in Figure A10, which displays the difference in liters of water bought per tree during the critical *vs.* non-critical season between wealthy and poor farmers over time. If poor farmers were switching, we would expect a downward trend in Figure A10. That is, over time, poor farmers growing only apricots will disappear. This is not what we see in Figure A10. There are large differences between wealthy and poor farmers from year to year. During dry years (1955, 1957, and 1964) price differences in summer are large, so differences in water purchases are also large. During rainy years (1956, 1960, 1961, and 1962) price differences in summer are small, so differences in water purchases are small. Importantly, there is no trend in the difference in differences data, suggesting that attrition is not a concern in the case of Mula. This is consistent with the notion that switching costs are high, especially for poor farmers.

Figure A10: Differences in liters of water bought per tree during the critical vs non-critical season between wealthy and poor farmers (*i.e.* difference in differences).




*Notes:* For each year, we compute the amount of water per tree bought: *(i) on* and *off* season, and (ii) by wealthy and poor farmers. The figure shows the evolution of the difference in differences for these groups: wealthy *on-season*, wealthy *off-season*, poor *on-season,* and poor *off-season*. The unit of the vertical axis are liters of water per tree.

# References


[1] Aguirregabiria, V. and Mira, P., 2010, "Dynamic discrete choice structural models: A survey," *Journal of Econometrics*, Vol. 156, 38-67.

[2] Arcidiacono, P. and Ellickson, P. B, .2011, "Practical methods for estimation of dynamic discrete choice models." *Annual Review of Economics* Vol. 3, No. 1, 363-394.

[3] Aitchison, J. and Aitken, C. G. G., 1976, "Multivariate binary discrimination by the kernel method," *Biometrika*, Vol. 63, 413-420.

[4] Allen, R. G., Pereira, L. S., Raes, D. and Smith, M., 2006, *Evapotranspiración del cultivo: Guías para la determinación de los requerimientos de agua de los cultivos*. Food and Agriculture Organization (FAO) 2006.

[5] Connault, B., 2016, "Hidden rust models," Working Paper, University of Pennsylvania.

[6] Davidson, R., and J. G. MacKinnon. 1993. *Estimation and Inference in Econometrics*. New York: Oxford University Press.

[7] Donna, J. D. and Espín-Sánchez, J., 2018 "Complements and Substitutes in Sequential Auctions: The Case of Water Auctions," *RAND Journal of Economics,* Vol. 49, No. 1, pp. 87–127.

[8] Donna, J. D. and Espín-Sánchez, J., 2016, "A note on Resource Allocation with Liquidity Constrained Agents," working paper. Available online at: https://www.dropbox.com/s/k7fqvz0micathlp/DE-Note-LC.pdf?dl=1

[9] Dubé, J. P., Fox, J. T., & Su, C. L. 2012, "Improving the numerical performance of static and dynamic aggregate discrete choice random coefficients demand estimation," *Econometrica*, Vol. 80, No. 5, 2231-2267.

[10] Franco, J. A., Abrisqueta, J. M., Hernansáez, A. and Moreno, F., 2000, "Water balance in a young almond orchard under drip irrigation with water of low quality," *Agricultural Water Management*, Vol. 42, 75-98.

[11] Hendel, I., and Nevo, A., 2006, "Measuring the implications of sales and consumer inventory behavior," *Econometrica*, 74(6), 1637-1673.

[12] Hoffman, P. T., 1996, *Growth in a Traditional Society: The French Countryside 1450– 1815*. Princeton, NJ: Princeton University Press.

[13] Hotz, J. and Miller, R., 1993, "Conditional Choice Probabilities and the Estimation of Dynamic Models," *Review of Economic Studies*, Vol. 60, No. 3, 497-529.

[14] Hotz, J., Miller, R., Sanders, S. and Smith, J., 1994, "A Simulation Estimator for Dynamic Models of Discrete Choice," *Review of Economic Studies*, Vol. 61, No. 2, 265-289.

[15] Jofre-Bonet, M., and Pesendorfer, M., 2003, "Estimation of a Dynamic Auction Game," *Econometrica*, Vol. 71, No. 5, 1443-1489.

[16] Keane, M. P. and Wolpin, K. I., 1994, "The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence," *Review of Economics and Statistics*, Vol. 76, No. 4, 648-672.

[17] Keane, M. P. and Wolpin, K. I., 1997, "The career decisions of young men," *Journal of Political Economy*, Vol. 105, No. 3, 473-522.

[18] Keane, M. P. and Wolpin, K. I., 2010, "The role of labor and marriage markets, preference heterogeneity, and the welfare system in the life cycle decisions of black, hispanic, and white women," *International Economic Review*, Vol. 51, No. 3, 851-892.

[19] Li, Q. and Racine, J., 2003, "Nonparametric estimation of distributions with categorical and continuous data," *Journal of Multivariate Analysis*, Vol. 86, 266-292.

[20] Luo, Z. Q., Pang, J. S., and Ralph, D., 1996, *Mathematical programs with equilibrium constraints*. Cambridge University Press.

[21] Miller, R. A., 1984, "Job matching and occupational choice," *Journal of Political Economy*, Vol. 92, No. 6, 1086-1120.

[22] Pakes, A., 1986, "Patents as options: Some estimates of the value of holding European patent stocks," *Econometrica*, Vol. 54, No. 4, 755-784.

[23] Pérez-Pastor, A., Domingo, R. and Torrecillas, A., 2009, "Response of apricot trees to deficit irrigation strategies," *Irrigation Science*, Vol. 27, 231-242.

[24] Rust, J., 1987, "Optimal replacement of GMC bus engines: An empirical model of Harold Zurcher," *Econometrica*, Vol. 55, No. 5, 999-1033.

[25] Su, C. L., and Judd, K. L., 2012, "Constrained optimization approaches to estimation of structural models," *Econometrica*, Vol. 80, No. 5, 2213-2230.

[26] Torrecillas, A., Domingo, R., Galego, R. and Ruiz-Sánchez, M.C., 2000, "Apricot tree response to withholding irrigation at different phenological periods," *Scientia Horticulturae*, Vol. 85, 201-205.

[27] Wang, M. and van Ryzin, J., 1981, "A Class of Smooth Estimators for Discrete Distributions," *Biometrika*, Vol. 68, 301-309.

[28] Wolpin, K. I., 1984, "An estimable dynamic stochastic model of fertility and child mortality," *Journal of Political Economy*, Vol. 92, No. 5, 852-874.