Matching, Competition, and Auction Design for Differentiated Goods*

Nicolò Generoso† Miho Hong‡

October 25, 2026

Please click here for the latest version

Abstract

Allocation of differentiated goods to buyers with heterogeneous tastes introduces a tradeoff between efficient product-buyer matching and competition between buyers. Separate auctions improve matching by letting buyers target preferred items but thin competition, while pooling differentiated objects increases competition at the expense of match quality. We quantify this tradeoff in wholesale peach auctions in South Korea, where hundreds of lots are auctioned sequentially each day. We document bidder-variety sorting in auction participation and estimate a model of sequential auctions with multi-unit bidder demand, heterogeneous latent tastes for varieties, and selective entry. The estimates reveal heterogeneity in both relative valuations and willingness to bid across varieties. Revenue-maximizing schemes pool varieties that attract different bidders while preserving the horizontal differentiation most valuable for matching. These schemes would raise revenue by 4.8–7.5%, with effects on efficiency ranging from a 2.2% loss to a 0.5% gain. The status quo is typically dominated by pooling schemes that would improve both revenue and match value. Supply composition changes the balance between participation gains and matching losses, and an index of these forces closely predicts the resulting revenue ranking of pooling schemes.

*Many thanks to Phil Haile, Charles Hodgson, and Adam Rosenberg for helpful conversations; and to the Yale Center for Research Computing for assistance in computation run on the Bouchet cluster.

†Yale University, nicolo.generoso@yale.edu.

‡Yale University, miho.hong@yale.edu.

1 Introduction

Selling heterogeneous goods to buyers with heterogeneous tastes forces a choice about how finely to differentiate among the objects. Treating similar but distinct objects as identical thickens markets and intensifies competition, but prevents buyers from sorting towards their preferred goods. This basic market design tradeoff is well understood (Levin and Milgrom, 2010), but its magnitude depends on how tastes are distributed across goods, which goods are available to sell, and which buyers are willing to participate in the market. A designer faces a frontier whose shape changes with these primitives.

We study this tradeoff in wholesale auctions for fresh peaches in South Korea. A single intermediary repeatedly allocates a supply whose varietal composition shifts with the season to a set of wholesale buyers with heterogeneous tastes across varieties. The intermediary earns ad valorem fees on transaction values and controls the key design margins: how supply is partitioned into lots, what buyers observe about each lot, and whether lots from different farmers or varieties are pooled.

If peaches traded in a frictionless competitive market where prices fully reflect fruit heterogeneity and buyer tastes, market-clearing prices would coordinate trade and implement an efficient allocation. Relative to this benchmark, design choices matter because they determine which dimensions of heterogeneity are priced and which ones are averaged away, thereby reshaping substitution patterns and trading off efficiency with revenues. In our setting, trade is organized through a sequential auction mechanism. The auctioneer partitions the daily supply into small lots based on fruit variety and sells them sequentially. Individual auctions typically last only a few seconds, leaving buyers little time to decide whether to bid and leading to "thin" auctions: while daily attendance is between 20–30 bidders, the median auction receives only three bids. This limited participation leaves scope for pooling to attract buyers who would otherwise enter different auctions, broadening competition at the cost of restricting sorting. Because buyers demand multiple units, competition is inherently dynamic: winning early affects a buyer's willingness to bid later, and expectations about future lots feed back into current bids.

We begin by documenting some descriptive facts that inform our modeling choices. First, bidders selectively sort into auctions. We test and reject the hypothesis that entry occurs at random across lots and independently of variety, corroborating the intuition that bidders are more likely to bid on lots that belong to a preferred variety. We also find that bidders are heterogeneous along different margins – how likely and how aggressive they are in bidding on given varieties. The two margins are weakly correlated, so that bidders that tend to bid on a given variety on a given day need not submit systematically higher bids on that variety. Second, controlling for lot features and bidder-day fixed effects, bids tend to decline with accumulated inventory and become more aggressive towards the end of each auction session.

Third, supply composition changes considerably over the course of a trading season, with different peach varieties becoming available in a staggered fashion. This implies that different distributions of varieties are matched to the same set of bidders from day to day.

Motivated by these facts, we formalize a model with persistent idiosyncratic tastes for varieties, multi-unit declining demand, selective entry based on lot features, and dynamic considerations. Bidders draw persistent types every day, which determine both their valuations and their propensity to bid on different varieties. Their willingness to pay for different varieties depends on such types and declines in the quantity already purchased. When a lot is called, they decide whether to place a bid based on the lot's features and the realization of a stochastic participation cost, whose mean depends on the bidder's type. Entrants submit bids to maximize their payoffs, which include both their valuation for the current lot and the expected continuation value from optimal entry and bidding in subsequent auctions. A typical high-season trading day features hundreds of auctions and dozens of bidders in attendance, leading to an intractably large number of variables for bidders to track. In a spirit similar to oblivious equilibrium (Weintraub et al., 2008), we assume bidders only keep track of their own purchases and of a coarse subset of the publicly available information.

The main empirical challenge is to recover heterogeneity in bidder demand from auction histories in which participation is endogenous and bidding incentives depend on both past purchases and expectations about future auctions. We model persistent taste heterogeneity using a finite number of latent bidder-day types and exploit repeated choices of the same bidder within a day to recover their distribution (Kasahara and Shimotsu, 2009). Persistent differences in entry and bidding across fruit varieties reveal heterogeneity in types' tastes and participation costs, while changes in bidding behavior as purchases accumulate discipline the decline of marginal values with inventory. Seasonal variation in the supply composition provides additional information by exposing the same population of bidder types to different distributions of lot features. Together, the sequence of entry decisions, bids, and purchases across supply regimes identifies persistent heterogeneity in demand, the effect of accumulated purchases on willingness to pay, and the distribution of participation costs.

We estimate the model in a sequence of steps. First, we homogenize bids to remove shifts due to observed lot heterogeneity over which bidders are assumed to have common tastes, such as fruit size and packaging format (Haile et al., 2003; Athey and Haile, 2007). Second, we estimate the distribution of bids conditional on different states. The usual first-price bid inversion (Guerre et al., 2000) then maps observed bids into "pseudo-valuations," which combine the bidder's current willingness to pay for the lot with the effect that winning has on future bidding opportunities. Third, we estimate the different types' tastes and their demand slope using an adaptation of the EM algorithm for dynamic models with latent heterogeneity developed by Arcidiacono and Miller (2011). The algorithm combines posterior probabilities over latent bidder types with conditional choice probabilities (CCPs) for the dynamic entry

decision. Given current posterior type probabilities and continuation values, we update the type-specific mean participation costs so that posterior-weighted optimal entry matches observed participation, which in turn determines a new set of type-specific entry probabilities. Holding these CCPs fixed, we evaluate bidders' continuation values and update the taste parameters and the decline of marginal values with inventory by maximizing the likelihood of submitted bids. The observed entry and bidding history then updates the posterior probability that each bidder-day belongs to each type.

The estimates reveal substantial heterogeneity in both tastes and participation. Higher-valuation types account for about 30% of bidder-days and have baseline willingness to pay 44–68% above lower-valuation types. Horizontal tastes differ most strongly between Nectarines and Yellow peaches: bidders who favor Yellows value them 30–50% above Nectarines, while those who favor Nectarines value them 18–28% above Yellows. By contrast, valuations for White peaches differ by less than 3% across horizontal types with the same vertical component. Participation costs divide the product space differently, and the main difference in entry propensities separates bidders more inclined to bid on White peaches from those more inclined to bid on Nectarines and Yellows. These participation differences arise even between bidders with identical tastes, so differentiation can separate potential competitors along product distinctions that matter relatively little for horizontal sorting.

We ask how pooling different varieties affects revenue and efficiency, and how these effects depend on the economic environment. We compare the status quo with schemes that combine selected varieties into mixed lots while retaining the sequential auction mechanism. We find that the status quo is typically dominated by pooling schemes which would improve both revenue and efficiency. While the magnitudes of such effects depend on economic primitives such as the number of attending bidders and the composition of supply, this result shows that the underlying distribution of tastes and participation costs would allow a designer to broaden participation while preserving the horizontal product differentiation most valuable for matching. The revenue-maximizing schemes involve partial pooling, leaving some product distinctions intact. These schemes would achieve revenue gains relative to the status quo ranging from 4.8–7.5%, with corresponding changes in efficiency ranging from a loss of 2.2% to a gain of 0.5%. Bidder surplus would fall at these revenue maxima in every case. Full pooling would generate smaller revenue gains and substantially larger matching losses.

The revenue and efficiency effects of pooling depend on how differences in tastes and willingness to participate interact with the composition of supply. Once varieties attracting different bidders have been combined, further pooling would add little potential participation while suppressing additional sorting opportunities. We construct and validate an index of potential participation gains and horizontal matching losses based on the estimated primitives. The index closely tracks the revenue ranking of pooling schemes across supply compositions. The quantities of each variety available for sale weight participation gains and matching

losses differently, so changes in supply composition can reverse the revenue ranking of pooling schemes even with bidder heterogeneity held fixed.

Related literature We combine first-price bid inversion (Guerre et al., 2000) with conditional choice probability methods (Hotz and Miller, 1993) and EM estimation with latent heterogeneity (Arcidiacono and Jones, 2003; Arcidiacono and Miller, 2011). Our alternating updates of choice probabilities and parameters relate to the iterative estimators of Aguirregabiria and Mira (2002; 2007). Identification of persistent taste types draws on results for dynamic finite mixtures (Kasahara and Shimotsu, 2009; Higgins and Jochmans, 2023). Our model is related to Gentry and Li (2014) in that we allow for selective entry; instead of shifters, we exploit repeated entry decisions to recover the persistent tastes that drive selection before lot-specific shocks are learned.

We also contribute to the empirical literature on dynamic auctions in procurement (Jofre-Bonet and Pesendorfer, 2003; Altmann, 2023; De Silva et al., 2023), differentiated-product markets (Bodoh-Creed et al., 2021; Backus and Lewis, 2025), oil and gas (Kong, 2021), and wholesale markets (Lu et al., 2019; Choi, 2024). Choi (2024) studies the same market as us but focuses on synergies across lots. Like Cui (2025), we use EM to recover persistent heterogeneity, but our setting features endogenous entry and bid-determined winners, avoiding the additional latent-type dependence of outcomes that motivates her partial-information likelihood. Like Kobayashi and Alcobendas (2024), we model forward-looking entry and bidding, but dynamic incentives arise from declining marginal values rather than budget constraints. Altmann (2023) identifies both unobserved preferences and inventory dynamics, while we leverage both accumulated purchases and repeated entry and bidding to separate persistent latent tastes from inventory-dependent demand. In advertising auctions, Sousa (2026) estimates how targeting improves matches but reduces participation and competition, while Tunuguntla et al. (2026) empirically evaluate information-bundling mechanisms.

Our design question builds on the tradeoff between differentiation and thickness emphasized by Levin and Milgrom (2010), which has inspired theory research. Previous works studied sellers' incentives to coarsen information (Ganuza and Penalva, 2010; Bergemann et al., 2022), bundle goods (Jehiel et al., 2007), and offer lotteries over differentiated goods (Loertscher and Muir, 2025). In matching settings, recent work has illustrated how profit motives might lead a designer to provide matching technologies that reduce match quality (Gieselmann, 2025), delay matches (Antler et al., 2025), or otherwise reshape how agents are matched to objects (Fershtman and Pavan, 2022).

Roadmap Section 2 describes the institutional details and the data. Section 3 presents descriptive facts that inform our modeling choices. Sections 4 and 5 discuss our model and estimation. Section 6 illustrates counterfactuals.

2 Institutions and data

This section presents the key institutional details and introduces the data. A more comprehensive description of the setting is provided by Generoso and Hong (2026).

2.1 Institutional details

Wholesale auctions for fruits and vegetables in South Korea are run by a set of licensed auction houses. Licensed wholesale buyers can participate in the auctions run by a given house upon registration at that house. On a trading day, auction houses accept consignments from farmers to be auctioned on the same day. Farmers are free to choose which auction house to supply. Auction house employees inspect the supplied produce, categorize and label it based on standardized quality and fruit-size scales, divide it into lots, and display it in an open area. Each lot is a subset of the fruit supplied by a farmer with homogeneous characteristics such as variety, quality grade, size, and packaging format. Lot sizes are typically between 10 and 100 kilograms. Bidders can inspect lots before the start of the auction session, both visually, by tasting samples, and by measuring sweetness.

An auction session comprises a sequence of auctions for individual lots. A licensed auctioneer, an employee of the house, conducts auctions whose format resembles an ascending auction. Interested bidders submit their bids using an electronic device. The auctioneer sees all standing bids, and closes the auction when they deem appropriate. As long as the auction is open, bidders can raise, decrease, or cancel their bids, but they cannot observe the bids submitted by others. The highest bidder at the time of closure wins the lot and pays the bid amount. Auctions are typically very fast, with most auctions closing within seconds and only a negligible share dragging for as long as a minute. Since peaches are highly seasonal, hundreds of lots are auctioned on a typical high-season day (e.g., any day in July or August). Low-season days (e.g., in early June or late September) can feature several dozen auctions.[1] Most auction houses operate from 5 to 6 days a week.

Peaches come in many varieties, differentiated primarily by skin type, flesh color, and texture. The most botanically salient distinction is between nectarines and fuzzy peaches: the former are smooth-skinned, whereas the latter have the characteristic fuzzy peel. Varieties also differ in flavor and texture. White-fleshed peaches are typically lower in acidity and are therefore perceived as sweeter or more floral, whereas yellow-fleshed peaches tend to have a tangier balance (Petruccelli et al., 2023). Texture ranges from melting to non-melting to stony-hard: non-melting fruit retain firmness and are commonly used for processing and canning,

[1] The amount of auctions of course depends on the size of the auction house and on the available supply. The auction house we will focus on for most of our estimation exercise typically runs 150-380 auctions per day between mid June and the end of August. Table 1 reports some moments of the distribution of the number of daily auctions in our sample.

while melting fruit soften rapidly and are mostly consumed fresh (Christofi et al., 2022). For brevity, we use the term "peach" to refer to both nectarines and fuzzy peaches.

Bidders are wholesale traders that purchase large quantities of fruit on a daily basis. They resell the produce to both institutional buyers (e.g., grocery stores and food processors) and final consumers.[2] Auctions are run early in the morning, and bidders typically collect orders from their clients the night prior. Accordingly, a variety's attractiveness to a bidder depends on the particular needs of their clients.

Table 1 Summary statistics

Sample size				
Days				192
Auctions				27,018
Bids				97,207
Bidders				33
Bidder-days				4,712
Auction environment				
	25th	50th	75th	Mean
Daily auctions	44	149	227	140.72
Lot size (kg)	20	45	104	77.25
Daily supply (t)	3.15	10.42	16.84	10.87
Bids / auction	2	3	5	3.60
Daily winners	23	27	28	24.54
Bid (KRW/kg)	1,400	2,100	2,888	2,197
Winning bid (KRW/kg)	1,377	2,150	3,000	2,259
Won quantity (kg)	130	315	644	496.20

Notes: Table reports summary statistics for our main dataset. The top panel reports the sample size. *Bidders* reports the number of unique observed bidders. The bottom panel reports distributional summaries of the auction environment. Daily supply is reported in metric tonnes, or thousands of kilograms. *Daily winners* is the number of bidders who won at least one lot during an auction day. *Won quantity* is the total number of kilograms purchased by a bidder on a given day, and its distribution is over bidder-days. In 2017, KRW 1,000 was worth approximately USD 0.9.

[2] Most bidders operate small shops in the same marketplace where the auctions are conducted, and consumers are free to shop there.

2.2 Data

Our main data covers the universe of bids submitted for all peach lots auctioned in 2016 and 2017 from one of the two auction houses operating in the Noeun marketplace, in the city of Daejeon. The data includes lot-level information on (i) the supplier: farmer name, consignor name, and origin; (ii) the auction outcome: all submitted bids and bidder ID; and (iii) lot characteristics: lot size, variety, packaging, and standardized quality and fruit-size levels. The latter are assessed by auction house inspectors prior to the auction. The quality grade is based on visual inspection of whether the fruit included in the lot are homogeneous in terms of appearance and display good skin integrity. The fruit size grade depends on measurement of the circumference of fruits in the lot; bidders seem to attach a premium to larger fruits. Table 1 presents summary statistics on the size of the dataset and the auction environment bidders face.

We also make limited use of data from the Korea Agro-Fisheries & Food Trade Corporation (hereafter AT), the public agency that monitors the operations of auction houses and collects data on auction outcomes. The data provided by AT covers the universe of traded lots in all auction houses, includes the same lot characteristic information as our main dataset, but only reports winning bids. This is the same data used in Generoso and Hong (2026): more details can be found there.

Lot varieties are reported in a non-uniform way, with different auction houses following different conventions on the level of detail at which varieties are recorded. We manually classify the observed variety names into labels that trace differentiation along the dimensions discussed in Section 2.1. For our analysis, we group varieties into five families: Hard White, Soft White, Yellow, Nectarine, and Niche, the latter collecting less common varieties.

3 Descriptives

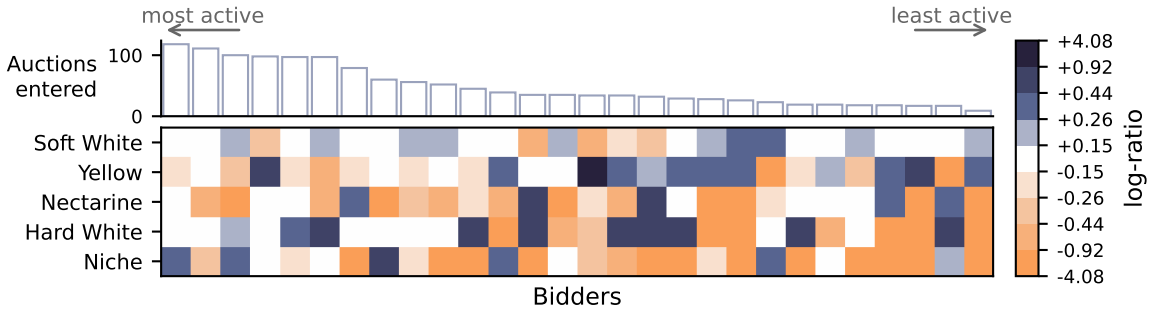
This section presents some descriptive facts that motivate our research questions and modeling choices.

Fact 1 *Bidders selectively sort into auctions.*

Within-bidder entry patterns can be informative about bidders' tastes. If a bidder has strong taste for some variety, one would expect them to be relatively more likely to enter auctions for that variety. Because the composition of available supply varies across days, we benchmark each bidder's entry pattern against the day's variety mix. Indexing bidders by i, varieties by j, and days by d, denote bidder i's empirical distribution of entered auctions across varieties and the distribution of available varieties by $\boldsymbol{q}_{id} = (q_{ijd})_j$ and $\boldsymbol{p}_d = (p_{jd})_j$, respectively. Within-day selectiveness can be measured by the log-ratio $\log(q_{ijd} / p_{jd})$. Positive values indicate that bidder i enters variety j more often than implied by its availability on

day d. Figure 1 plots this object for the day with the largest number of auctions. While these patterns are suggestive, bidder-day counts are often small, making it difficult to draw sharp conclusions from raw entry shares alone. In Appendix A.1, we formally test and strongly reject the null hypothesis that, conditional on the observed number of entered auctions and the day's variety mix, a bidder's entries are randomly allocated across varieties.[3] Our model will allow for selective participation into auctions.

Figure 1 Log-ratio of entry probabilities to available variety shares



Notes: Figure plots the log-ratio of each bidder's shares of bid-on varieties to the shares of available varieties on August 8, 2016. This is the day recording the most auctions in our sample. Rows correspond to varieties, columns to bidders. Blue cells mark families a bidder enters more often than their availability, and orange cells less often. Bidders are sorted by the number of auctions entered, shown in the top panel. On this day 28 bidders entered at least one of 382 auctions; the mean bidder entered 48.0 auctions (25th percentile: 22; median: 34.5; 75th percentile: 64.8).

Fact 2 *Entry and bidding reflect distinct dimensions of bidder heterogeneity.*

Selective entry could arise because of (i) differences in bidders' willingness to pay across varieties, reflecting the value of downstream demand when it is present, or (ii) differences in how often bidders face downstream demand for each variety. We distinguish these margins by comparing bidders' relative entry rates and relative bids across varieties. Figure 2 shows that bidders who bid relatively more for a variety are only slightly more likely to enter its auctions. Table 2 shows that both entry and bidding patterns persist across separate subsets of auctions within the same bidder-day, while their cross-correlation is small. Appendix A.2 further shows that the same pattern holds across individual pairs of varieties. These results suggest that bidder heterogeneity operates along two dimensions: how aggressively a bidder bids and how likely they are to bid across different varieties. Our model will allow bidders to differ both in their willingness to pay and in their variety-specific participation decisions.

Fact 3 *Dynamic considerations affect intra-day bidding patterns.*

[3] For the reader's convenience, appendix figures expose hyperlinks to return to the relevant sections of the main text.

Figure 2 Relative bids and entry across varieties

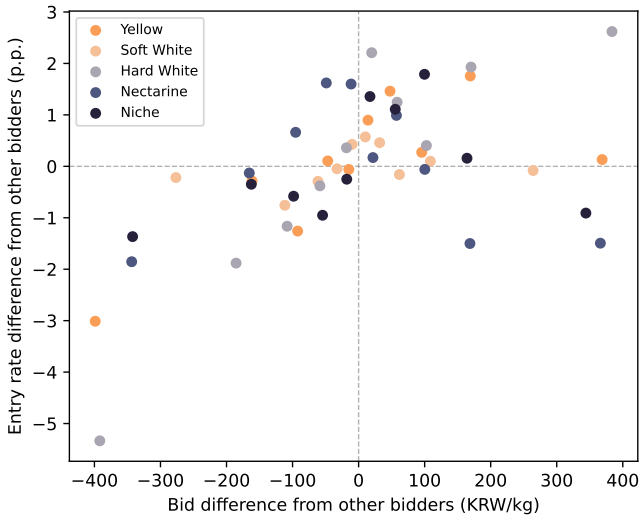


Table 2 Within-day correlations

	Correlation
Entry across samples	0.63
Bids across samples	0.29
Entry vs. bids	0.08

Notes: Figure 2 plots (i) a bidder's average bid residuals after controlling for lot, bidder-day, and inventory bin fixed effects, against (ii) the difference between the bidder's entry rate for that variety and the corresponding entry rate of other bidders. Observations are at the bidder-day-variety level. Within each fruit family, we bin them by their average bid residual and plot bin averages. Table 2 reports correlations across bidder-day-variety observations constructed from two non-overlapping samples of auctions. We divide each day's auction sequence into four equal-sized time blocks and, within each block and variety, assign successive auctions alternately to samples A and B. For each bidder-day-variety, we separately compute the average bid residual and entry-rate difference in each sample. The first row reports the correlation between entry-rate differences in A and B, the second the correlation between bid residuals in A and B, and the third the average of the correlation between bids in A and entry in B and the correlation between bids in B and entry in A. In 2017, KRW 1,000 was worth approximately USD 0.9. The mean per-kilogram bid across our sample is KRW 2,100, or approximately USD 1.89.

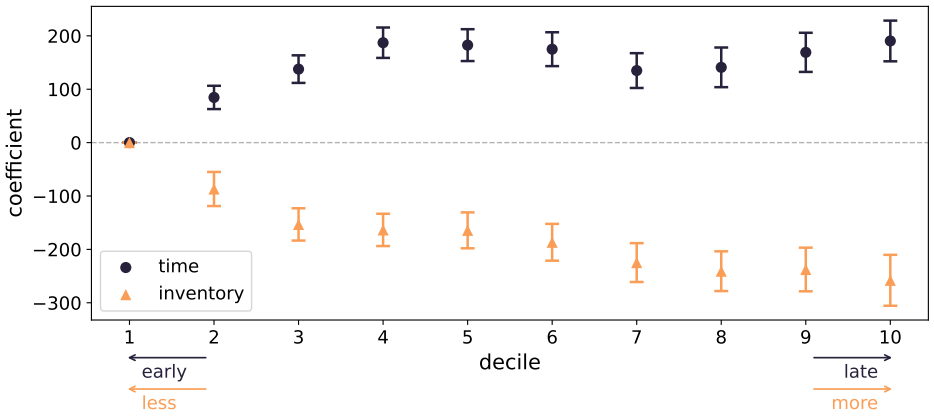
Bidding should be less aggressive when there are more auctions left. If marginal valuations decline with accumulated inventory, bidders should also become less aggressive after securing large amounts of fruit. We leverage the panel structure of our data to assess the magnitude of these effects. We refer to the total amount (in kilograms) of peaches won by a bidder when a given lot is about to be auctioned as *inventory* and to the share of auctions already occurred when a given lot is auctioned as *time*. Dividing time and inventory in deciles, we estimate the following regression

$$\text{bid}_{id\ell} = \sum_{m=2}^{10} \mathbf{1}(\text{time}_\ell = m) \times \gamma_m^{\text{time}} + \sum_{n=2}^{10} \mathbf{1}(\text{inventory}_{i\ell} = n) \times \gamma_n^{\text{inv}} + \boldsymbol{x}_\ell' \boldsymbol{\beta} + \alpha_{id} + \varepsilon_{id\ell}, \qquad (1)$$

where bidders are indexed by i, days by d, lots by ℓ. The vector $\boldsymbol{x}_\ell$ includes lot variety, quality grade, fruit size, and binned lot size. Bidder-day fixed effects are denoted by α_{id}.

Figure 3 reports the estimated coefficients $\gamma^{\text{time}}, \gamma^{\text{inv}}$. Controlling for lot observables and bidder-day fixed effects, higher levels of inventory are associated with lower bids, while a lower

Figure 3 Bid dynamics



Notes: Figure reports coefficients γ_{mn} from regression (1) and confidence intervals from standard errors clustered at the bidder-day level. Full estimated coefficients are available in Table A2. All amounts are in Korean won. In 2017, KRW 1,000 was worth approximately USD 0.9. The mean per-kilogram bid across our sample is KRW 2,100, or approximately USD 1.89.

share of remaining auction is associated with higher bids.[4] Our dynamic auction model will explicitly incorporate the option value of losing, which is greater when more auctions remain and leads to softer bidding early in the sequence. To capture the decline in bids with inventory, we will allow bidders to have multi-unit demand with decreasing marginal valuations for each additional kilogram of peaches.

Fact 4 *Supply follows markedly seasonal patterns.*

Peaches are traded between June and September and different varieties become available at different times within the season. Successive days tend to have "similar" supply compositions, which motivates us to classify days into a finite set of what we label *supply regimes*. These seasonal patterns provide variation in the environment bidders operate in, and our model will allow bidders to have different beliefs on the supply composition of future lots based on the auction day's supply regime. Moreover, when varying supply configurations need to be matched to the same set of bidders, the revenue incentives to distort the matching could vary with the underlying regime. Appendix A.4 discusses our classification of days into regimes, Figure A3 plots the cosine distance between variety share vectors for all pairs of days in the 2017 season, and Figure A4 shows the supply composition of different regimes and their temporal variation.

[4] Of course having a high inventory is typically correlated with later auctions, so that the two effects partially offset each other. In the data, raw bids tend on net to decline during the day.

4 Model

We model a trading session as a sequence of first-price auctions. Motivated by the facts presented in Section 3, we assume that bidders have latent type capturing their willingness to pay and propensity to bid on different varieties. When a lot is called, bidders observe its features and decide whether to incur a stochastic cost to participate in the auction. Participants incur a shock to their willingness to pay for the lot. Bidders have declining demands and optimize participation and bidding accounting for future opportunities.

Preferences A lot of peaches ℓ is characterized by $(\boldsymbol{x}_\ell, q_\ell, \xi_\ell)$, where $\boldsymbol{x}_\ell$ is a vector of observable characteristics such as variety and packaging format, q_ℓ its size in kilograms, and $\xi_\ell \in \mathbb{R}$ its quality, determined by features such as sweetness and ripeness.

Bidders serve different clients and hence have idiosyncratic tastes for lot features. Their marginal valuation declines linearly in purchased stock, since each bidder moves down the willingness-to-pay curve of downstream buyers to sell larger quantities. A bidder i who owns w_i kilograms of peaches values an additional kilogram with features $(\boldsymbol{x}, \xi)$ at

$$\boldsymbol{\gamma}_i' \boldsymbol{x} - \alpha w_i + \xi + \varepsilon_i,$$

where $\boldsymbol{\gamma}_i$ is the vector of bidder's *tastes*, and ε_i is a bidder-lot *taste shock*.

Integrating marginal valuations, bidder i's willingness to pay for lot ℓ is

$$\int_{w_i}^{w_i+q_\ell} (\boldsymbol{\gamma}_i' \boldsymbol{x}_\ell - \alpha q + \xi_\ell + \varepsilon_{i\ell}) \mathrm{d}q = \left(\boldsymbol{\gamma}_i' \boldsymbol{x}_\ell - \alpha w_i - \frac{1}{2} \alpha q_\ell + \xi_\ell + \varepsilon_{i\ell} \right) q_\ell.$$

Game structure Each day is associated with a public information state z collecting commonly known facts: the number of lots available, the number of attending bidders, and features of the daily supply. At the beginning of each auction day, bidder i draws a *type* m_i, which determines both their tastes $\boldsymbol{\gamma}_i = \boldsymbol{\gamma}_{m_i}$ and the distribution of participation costs they face across lots. When exactly one private bidder enters, we represent the auctioneer's discretion over auction closure by an unobserved competing bid R_ℓ. An offer below R_ℓ is not accepted. Conditional on the state, this bid is independent of the private bidder's offer. For the theoretical bidding problem, bids are unrestricted in sign.[5]

The cumulated quantity won by a bidder at a given point in time is called *inventory*. All bidders start the day with empty inventory $w_{i0} = 0$. Indexing lots by their position in the daily sequence of auctions, $w_{i\ell}$ denotes the cumulated inventory secured by bidder i in the auctions preceding lot ℓ. When lot ℓ is called, bidders observe its features $(\boldsymbol{x}_\ell, q_\ell, \xi_\ell)$ and

[5] A similar device is used by Krasnokutskaya and Seim (2011).

decide whether to incur idiosyncratic cost $c_{i\ell}$ to participate in the auction.[6] Participating bidders learn their lot-specific taste shock $\varepsilon_{i\ell}$, a first-price auction is run, and the winner's identity i^* and winning bid are observed. Inventories are updated to $w_{i^*\,\ell+1} = w_{i^*\ell} + q_\ell$ and $w_{i\,\ell+1} = w_{i\ell}$ for $i \neq i^*$.

We assume the following: (i) bidder types $m_i \sim F_m$ are iid across bidders and days; (ii) taste shocks $\varepsilon_{i\ell} \,|\, \boldsymbol{x} \sim F_{\varepsilon|\boldsymbol{x}}$ are independent across bidders and lots conditional on lot observables, and satisfy $\mathbb{E}[\varepsilon \,|\, \boldsymbol{x}] = 0$; (iii) participation costs $c_{i\ell} \,|\, \boldsymbol{x}, q_\ell, m_i, z \sim F_{c|\boldsymbol{x},q,m,z}$ are independent across bidders and lots conditional on bidder types, lot observables, and public information; and (iv) lot features $(\boldsymbol{x}_\ell, q_\ell, \xi_\ell) \,|\, z \sim F_{\boldsymbol{x},q,\xi|z}$ are independent of bidder types and observable characteristics $(\boldsymbol{x}_\ell, q_\ell)$ are independent across lots conditional on z.

Coarse beliefs The information available to bidders throughout the game includes the full sequence of past auction outcomes and the features of past lots. In a spirit similar to oblivious equilibrium (Weintraub et al., 2008), we assume bidders only keep track of a subset of the information available to them.[7] We assume that (i) they know their type m_i and public information z; (ii) they know the current lot's features $(\boldsymbol{x}_\ell, q_\ell, \xi_\ell)$ and the lot's position ℓ in the sequence of auctions; (iii) they keep track only of their own inventory $w_{i\ell}$. We call a *state* a vector $s = (z, \ell, \boldsymbol{x}_\ell, q_\ell, \xi_\ell)$, and we assume that bidders best respond to the state-contingent distribution of highest opposing bids, denoted $(\psi(\cdot \,|\, s))_s$. This allows bidders to form expectations on competitors that depend on a rich set of features, capturing the first-order effects on bidding of aggregate intra-day bidding patterns, seasonality, and lot desirability. Equilibrium will require that this distribution be consistent with the opponents' average bidding behavior conditional on each state. We focus on symmetric equilibria, so that such distribution is common to all players and not indexed by i.

Optimal entry and bidding A bidder optimizes their participation and bidding decisions forming expectations over the realizations of future participation costs, lot features, and taste shocks. The value function for this single-agent dynamic problem is denoted $\mathcal{V}_{i\ell}(s, c_i, w_i)$, and captures bidder i's expected value from optimal entry and bidding in lots $\ell, \ell+1, \ldots, L$

[6] We use heterogeneity in participation costs to capture the variety-specific entry propensities documented in Section 3. These costs are not primitives of downstream preferences, but rather a reduced-form representation of the factors determining whether a bidder is active for a particular lot in this auction environment. Given the fast pace of the market, these factors may include transitory attention as well as whether the bidder currently has downstream resale opportunities for the auctioned variety. We load this variation in entry propensities on participation costs rather than on the slope of demand because Figure A2 shows that these differences do not meaningfully change with inventory.

[7] Oblivious equilibrium does not literally apply to our finite-horizon setting. Despite imposing a similar restriction, our assumption is best interpreted as describing the steady state of a learning process across repeated auction days, in which bidders learn the distribution of competition conditional on coarse states without retaining the full history of past play. Section B.1 elaborates.

when the current state is s, and i's current participation cost and inventory are c_i and w_i, respectively. It is indexed by i because the bidder's type determines both their tastes and the distribution of future participation costs.

Starting from $\mathcal{V}_{i\,L+1} = 0$, it is defined recursively as

$$\mathcal{V}_{i\ell}(s_\ell, c_{i\ell}, w_{i\ell}) = \max_{e\in\{0,1\}} \underbrace{e}_{\text{entry}} \Bigg(\mathbb{E}_{\varepsilon_{i\ell}}\Bigg[\max_b \underbrace{q_\ell\left(\boldsymbol{\gamma}_{m_i}'\boldsymbol{x}_\ell - \alpha w_{i\ell} - \frac{1}{2}\alpha q_\ell + \xi_\ell + \varepsilon_{i\ell} - b\right)\psi(b \mid s_\ell)}_{\text{current auction's payoff}}$$
$$+ \underbrace{\psi(b \mid s_\ell)\mathbb{E}_{c',s'\mid s,m_i}[\mathcal{V}_{i\,\ell+1}(s', c', w_{i\ell} + q_\ell)]}_{\text{continuation value from winning}} + \underbrace{(1 - \psi(b \mid s_\ell))\mathbb{E}_{c',s'\mid s,m_i}[\mathcal{V}_{i\,\ell+1}(s', c', w_{i\ell})]}_{\text{continuation value from losing}}\Bigg]$$
$$- \underbrace{c_{i\ell}}_{\text{entry cost}} \Bigg) + \underbrace{(1-e)\mathbb{E}_{c',s'\mid s,m_i}[\mathcal{V}_{i\,\ell+1}(s', c', w_{i\ell})]}_{\text{value from not entering}}. \tag{2}$$

Bidder i's lot-ℓ value can be unpacked as the expected payoff from optimal entry and bidding in auction ℓ and the expected continuation values of moving to the next lot's auction. The expected payoff from entry integrates over the yet-unknown taste shock, anticipating the implied optimal bid. Inventory updates upon winning, so the expectation over continuation values depends on win probabilities at the optimal bid. The payoff from not entering is the continuation value holding inventory fixed. The computation of continuation values integrates over states and participation costs.

Denote the difference in continuation values from winning versus losing by

$$\Delta\mathcal{V}_{i\ell}(w_{i\ell}, s_\ell) := \mathbb{E}_{c',s'\mid s_\ell,m_i}[\mathcal{V}_{i\,\ell+1}(s', c', w_{i\ell} + q_\ell) - \mathcal{V}_{i\,\ell+1}(s', c', w_{i\ell})],$$

and define bidder i's *pseudo-valuation* and *mean pseudo-valuation* for lot ℓ as

$$\bar{v}_{i\ell} := \boldsymbol{\gamma}_i'\boldsymbol{x}_\ell - \alpha w_{i\ell} - \frac{1}{2}\alpha q_\ell + \xi_\ell + \frac{1}{q_\ell}\Delta\mathcal{V}_{i\ell}(w_{i\ell}, s_\ell), \qquad v_{i\ell} := \bar{v}_{i\ell} + \varepsilon_{i\ell}.$$

Both depend on both a bidder's features and the state.

By inspection of (2), an entrant submits a bid that maximizes $(v_{i\ell} - b)\psi(b \mid s_\ell)$, so that a bidder's pseudo-valuation is sufficient for optimal bidding, given a state. Denote a symmetric optimal bidding strategy by $\beta : (v, s) \mapsto b$. Dynamic considerations are captured by $\Delta\mathcal{V}_{i\ell}(w_{i\ell}, s_\ell)$: such term is negative because losing entails an option value, and bidding is less aggressive when such term is large in magnitude (e.g., when there are many remaining auctions). Assuming $\psi(\cdot \mid s)$ is strictly increasing and differentiable for all s, the optimal bid

$b^*_{i\ell} = \beta(v_{i\ell}, s_\ell)$ satisfies the necessary first-order condition

$$v_{i\ell} = b^*_{i\ell} + \frac{\psi(b^*_{i\ell} \mid s_\ell)}{\psi'(b^*_{i\ell} \mid s_\ell)}. \tag{3}$$

Taste shocks only realize upon entry, so that expected profits from bidding optimally are

$$\Pi(\bar{v}_{i\ell}, s_\ell) := \mathbb{E}_{\varepsilon_{i\ell}|\boldsymbol{x}_\ell}\left[q_\ell(\bar{v}_{i\ell} + \varepsilon_{i\ell} - b^*_{i\ell})\,\psi(b^*_{i\ell} \mid s_\ell)\right],$$

where $b^*_{i\ell} = \beta(\bar{v}_{i\ell} + \varepsilon_{i\ell}, s_\ell)$. A bidder bids on a lot if and only if their expected profits exceed the participation costs. An optimal symmetric entry strategy $\eta : (\bar{v}, c, s) \mapsto e \in \{0, 1\}$ satisfies

$$\eta(\bar{v}_{i\ell}, c_{i\ell}, s_\ell) = \mathbf{1}\big(\Pi(\bar{v}_{i\ell}, s_\ell) \geqslant c_{i\ell}\big), \tag{4}$$

Given a state, expected profits from bidding depend on bidders' types and inventory only through their mean pseudo-valuations. Participation additionally depends on the realized participation cost, whose distribution varies with bidder type. Under mild conditions, expected profits are strictly increasing in mean pseudo-valuations and bidders are more likely to place bids on lots of their preferred varieties.[8]

Equilibrium We close the model by requiring that the bid distributions ψ facing a given bidder are generated by optimal entry and bidding behavior.

Assume bidders follow symmetric entry and bidding strategies (η, β). These strategies imply a probability distribution $F_{s,E,\bar{\boldsymbol{v}}\,|\,\eta,\beta}$ over states, entrant sets, and mean pseudo-valuations of entrants.[9] Denote the induced distribution over entrant sets and mean pseudo-valuations conditional on states as $F_{E,\bar{\boldsymbol{v}}\,|\,s;\,\eta,\beta}$.

With endogenous entry, a bidder can be the only entrant with positive probability. Following Krasnokutskaya and Seim (2011), we assume that the auctioneer submits a competing bid r_ℓ whenever exactly one private bidder enters. Conditional on s, this bid is independent of the entrant's private information and has the same distribution as the highest opposing private bid when at least one opponent enters. If $E \setminus \{i\}$ is empty, the highest opposing bid faced by i is therefore r_ℓ.

We say that distributions of opposing bids $\psi = (\psi(\,\cdot \mid s))_s$ are *consistent* with (η, β) if

$$\psi_i(b \mid s) = \mathbb{E}_{E,\bar{\boldsymbol{v}}\,|\,s;\,\eta,\beta}\left[\mathbb{E}_{\boldsymbol{\varepsilon}}\left[\mathbf{1}\left(\max_{j \in E \setminus \{i\}} \beta(\bar{v}_j + \varepsilon_j, s) \leqslant b\right)\right] \Big|\, s;\,\eta,\beta\right]. \tag{5}$$

[8] The claim follows if $\beta(v, s) > 0$ almost surely and $\psi(\,\cdot \mid s)$ is strictly increasing for every s. Under such assumptions, $\psi(\beta(v, s) \mid s) > 0$ for all (v, s), and the objective of the bidding problem, $(v - b)\psi(b \mid s)$, is strictly increasing in v. This implies that the optimized value $\pi(v, s) := \max_b (v - b)\psi(b \mid s)$ is also increasing in v. Since this holds pointwise for v, expected profits $\Pi(\bar{v}, s) = \mathbb{E}_\varepsilon[\pi(v + \varepsilon, s)]$ are strictly increasing in $\bar{v}$.

[9] An explicit derivation is in Appendix B.3.

Under the assumption on r_ℓ, $\psi(\cdot \mid s)$ coincides with the distribution of the highest opposing private bid conditional on at least one opponent entering. We assume that the continuous game admits a symmetric equilibrium for any distribution of lot characteristics satisfying the assumptions above.

A *symmetric equilibrium* is a profile of strategies and distributions of opposing bids (η, β, ψ) such that η and β are optimal, and ψ is consistent with (η, β). Existence of a symmetric equilibrium can be established for a discretized version of the game.[10]

The next result extends a standard homogenization argument (Haile et al., 2003; Krasnokutskaya, 2011) from static settings to our dynamic model.[11] In the homogenization spirit, we compare two states that differ only in lot features over which all bidders agree. Call two states $s^0 = (z, \ell, \boldsymbol{x}^0, q, \xi^0)$ and $s^1 = (z, \ell, \boldsymbol{x}^1, q, \xi^1)$ *translation-equivalent with shift* δ if (i) $\boldsymbol{\gamma}_m'(\boldsymbol{x}^1 - \boldsymbol{x}^0) + \xi^1 - \xi^0 = \delta$ for every type m, (ii) $F_{\varepsilon \mid \boldsymbol{x}^1} = F_{\varepsilon \mid \boldsymbol{x}^0}$, and (iii) $F_{c \mid \boldsymbol{x}^1, q, m, z} = F_{c \mid \boldsymbol{x}^0, q, m, z}$ for every type m.

Proposition 1 (Homogenization)
There exists a symmetric equilibrium (η, β, ψ) *such that* $\beta(v + \delta, s^1) = \beta(v, s^0) + \delta$ *for all* $v \in \mathbb{R}$ *and all translation-equivalent* s^0, s^1 *with shift* δ.

Proposition 1 proves that additive shifts in valuations common across all bidders shift all equilibrium bids linearly. An immediate implication is that lot quality ξ_ℓ is fully passed through to bids. A further implication is that the expected profits from participation do not depend on the lot's quality, which is therefore irrelevant to the entry decision.

Parameterization We parameterize the model as follows:

1. Bidder types differ in tastes and participation costs.

   Tastes have a vertical and a horizontal component: $\boldsymbol{\gamma}_m = \boldsymbol{\gamma}_m^{\updownarrow} + \boldsymbol{\gamma}_m^{\leftrightarrow}$. We assume two possible values of each component. *Vertical tastes* take values in $\Gamma^{\updownarrow} = \{\boldsymbol{\gamma}^{\updownarrow 1}, \boldsymbol{\gamma}^{\updownarrow 2}\}$, with $\boldsymbol{\gamma}^{\updownarrow 1} < \boldsymbol{\gamma}^{\updownarrow 2}$. *Horizontal tastes* take values in $\Gamma^{\leftrightarrow} = \{\boldsymbol{\gamma}^{\leftrightarrow 1}, \boldsymbol{\gamma}^{\leftrightarrow 2}\}$, with $\sum_j \gamma_j^{\leftrightarrow k} = 0$ for $k = 1, 2$. These generate four possible taste-types, $\boldsymbol{\gamma} \in \Gamma = \{\boldsymbol{\gamma}^{\updownarrow r} + \boldsymbol{\gamma}^{\leftrightarrow k} : r, k \in \{1, 2\}\}$.

   Entry-types take values in $\{\mathbf{0}, \boldsymbol{a}^1\}$, where $\boldsymbol{a}^1 = (a_j^1)_j$ is a vector of variety-specific entry-cost deviations.

   Bidder types m are identified with combinations of taste- and entry-types $(\boldsymbol{\gamma}, \boldsymbol{a}) \in \Gamma \times \{\mathbf{0}, \boldsymbol{a}^1\}$. This gives 8 possible types.[12]

[10] In this game, there are finitely many states and lot features, inventories are discrete, feasible bids take values on a grid, and ties are broken at random. This is the game we use to simulate our counterfactuals. See Section F.1 for details.

[11] The proof is in Appendix B.2.

[12] Their population probabilities are unrestricted, and correlations between horizontal tastes and entry costs is allowed.

2. *Taste shocks* are mean-zero normally distributed conditional on lot observables: $\varepsilon \,|\, \boldsymbol{x} \sim N(0, \sigma^2_{\boldsymbol{x}})$. The variance depends on fruit size.

3. *Entry costs* are logistic with location depending on states and types: $c_{i\ell} = c_{s_\ell} + a_{j_\ell m_i} + \tau\nu_{i\ell}$, with $\nu_{i\ell} \sim \text{Logistic}(0,1)$. Here j_ℓ denotes the variety of lot ℓ, and the shocks $\nu_{i\ell}$ are iid across bidders and lots, independently of types and lot characteristics. The common location c_s depends on the state only through (i) public information, (ii) fruit variety, and (iii) lot size bin. The term a_{jm} captures type m's variety-specific deviation from the common location. We normalize one of the two participation cost profiles to zero for every variety. We further restrict the second entry-type profile to have zero mean across varieties under the empirical distribution of varieties across auctions.[13] This removes an entry-type-wide cost level, so entry types differ only in their relative propensity to enter auctions across varieties.

The taste decomposition accommodates differences in overall willingness to pay and in relative preferences across varieties. The additional differences in entry-types captures a bidder's client base and allows bidders with identical tastes to differ in their propensity to enter auctions across varieties.[14] Finally, the dependence of taste shocks on fruit size accommodates differences in bid dispersion across size categories.

5 Estimation

For estimation, we first homogenize bids on observed lot characteristics. We then invert bids through the distribution of opposing bids. Finally, we estimate the dynamic model by EM, treating bidder types as latent.

5.1 Homogenization

Motivated by the observation that a farmer's yield is supplied on multiple occasions within a trading season and is often split in multiple lots on each trading day, we impose that the same unobserved quality ξ be shared by all lots produced by the same farmer in a given year. This gives a reasonable approximation to lots' unobserved features such as sweetness and ripeness.

We also divide lot features in two groups, based on whether bidders have common versus idiosyncratic tastes over them. Partition lot observable and taste vectors as $\boldsymbol{x}_\ell = (\boldsymbol{x}^{\text{com}}_\ell, \boldsymbol{x}^{\text{idio}}_\ell)$

[13] That is, $\sum_j \omega_j a^1_j = 0$, where ω_j is variety j's share of auctions in the estimation sample.

[14] At the low entry probabilities observed in the data, the logistic entry rule implies that a_{jm} approximately scales type m's entry probability by $\exp(-a_{jm}/\tau)$ relative to the common participation margin. That is, $p_{jm} \approx \exp(-a_{jm}/\tau)p_{j0}$, where p_{j0} is the entry probability under the baseline participation profile. We interpret this relative propensity as capturing differences in downstream clienteles that go beyond a bidder's resale values implied by their taste-type. Types with lower a_{jm} are more likely to have a client for variety j and therefore to consider its auctions.

and $\gamma_i = (\gamma_i^{\text{com}}, \gamma_i^{\text{idio}})$, with $\gamma_i^{\text{com}} = \gamma^{\text{com}}$ almost surely. We normalize the location of common tastes by assuming there is a baseline combination of common-taste lot features $\tilde{x}^{\text{com}}$ such that $\gamma^{\text{com}\prime}\tilde{x}^{\text{com}} = 0$. Then, by Proposition 1,

$$b_{hi\ell} = \xi_{\text{farm}_\ell,\text{year}_\ell} + \gamma^{\text{com}\prime}\boldsymbol{x}_\ell^{\text{com}} + b_{i\ell}^0, \tag{6}$$

where $b_{i\ell}^0$ is the bid i would have submitted for a lot with unobserved quality $\xi = 0$ and common-taste features $\tilde{x}^{\text{com}}$. We let bidders have common tastes over the packaging format, fruit size, and the quality grade.[15] This leaves fruit variety as the single feature over which bidders have idiosyncratic tastes. Henceforth, we work with homogenized bids $b_{i\ell}^0$ and we identify x_ℓ with x_ℓ^{idio}, i.e., with an indicator of the lot's variety.

Equation (6) directly implies identification of γ^{com} and $\xi_{\text{farm,year}}$ from observed bids (Athey and Haile, 2007). Identifying a fully nonparametric distribution of lots' unobserved (to the econometrician) quality ξ_ℓ at the lot level is challenging for two reasons. First, Proposition 1 implies ξ_ℓ shifts bids linearly and does not affect entry decisions, preventing us from recovering its distribution from entry outcomes – a strategy exploited, e.g., by Compiani et al. (2020). Second, deconvolution arguments that exploit the joint distribution of bids (e.g., Krasnokutskaya, 2011 and Krasnokutskaya and Seim, 2011) only give the latent quality's distribution, not of the auction-level realizations. While integration over such distribution can be feasible in static settings, it seems computationally unattractive in our dynamic environment.

5.2 Bid inversion

State-dependent bid distributions are identified from observation of all bids. While the model presented in Section 4 does not restrict the dimensionality or the cardinality of the set of states, we discretize the state variables for tractability. In particular, each discretized state s_ℓ collects (i) public information: supply regime and indicators for above-median attending bidders and aggregate supply; (ii) lot observables: variety, fruit size bin, and an indicator for above-median lot size; (iii) within-day timing: an indicator for each quartile of remaining auctions. Henceforth, we refer to these discretized states simply as "states."

For each state s, we use the observed highest opposing bids in auctions with at least two bidders[16] to estimate

$$\psi(b \mid s) = \Pr\left(\max_{i' \neq i} b_{i'\ell} \leqslant b \mid s_\ell = s\right).$$

[15] Unobserved quality ξ captures features such as sweetness and ripeness, while quality grade is an observable label attached to each lots. It is based on visual inspection and assess features such as the fruit integrity.

[16] Section 4 introduced the assumption that the auctioneer submits a bid following this distribution whenever there is a unique entrant. As argued there, this implies that the restricting attention to auctions with more than one bidder to estimate ψ is inconsequential.

We estimate distributions $\psi(\cdot \mid s)$ nonparametrically via cross-validated Bernstein polynomials, which allow a convenient analytic expression for the densities $\psi'(\cdot \mid s)$. Since Bernstein polynomials approximate functions defined over closed intervals, we restrict the range of the state-specific distribution of bids to be between 0 and the maximum bid observed in that state. Appendix C.1 provides details on the implementation.

Using (3), each observed bid gives the pseudo-valuation

$$\hat{v}_{i\ell} = b_{i\ell} + \frac{\hat{\psi}(b_{i\ell} \mid s_\ell)}{\hat{\psi}'(b_{i\ell} \mid s_\ell)}. \tag{7}$$

This inversion is ill behaved in regions where the density is flat, e.g., the boundaries of the support. We therefore restrict attention to observations for which the inversion is sufficiently far from the boundaries of the estimated opposing-bid distribution.

5.3 Dynamics

We estimate the model using CCPs and a variant of the EM algorithm (Hotz and Miller, 1993; Arcidiacono and Jones, 2003; Arcidiacono and Miller, 2011). The CCPs fully summarize bidders' dynamic entry behavior conditional on their types and allow for a convenient representation of continuation values. Given parameters and current posterior type probabilities, we update the participation-cost profile from the entry likelihood, then jointly solve for common cost locations and entry probabilities to reproduce observed entry frequencies. We then hold the resulting entry probabilities fixed while updating the remaining parameters from the bid likelihood, before updating posterior type probabilities.

Since the EM algorithm assigns a posterior distribution over latent types to each bidder-day, we index bidder-specific objects by both bidder-day i and candidate type m. In particular, $\bar{v}_{im\ell}$ and $p_{im\ell}$ denote the mean pseudo-valuation and entry probability that bidder-day i would have on lot ℓ if assigned type m. Before the EM loop, we estimate the empirical distribution of future lot characteristics, which is needed to compute bidders' expectations throughout the game. We model variety and fruit size as independent draws conditional on the day's public information,[17] and draw lot quantity separately from its empirical distribution.

Partition parameters in $\boldsymbol{\theta} = ((\gamma_m)_m, \alpha, \tau, (\sigma_x)_x)$, $\boldsymbol{c} = (c_s)_s$, and $\boldsymbol{a} = (a_j)_j$. We initialize $\boldsymbol{\theta}^{(0)}$, $\boldsymbol{c}^{(0)}$, $\boldsymbol{a}^{(0)}$, mixture probabilities $\boldsymbol{\lambda}^{(0)} = (\lambda_m^{(0)})_m$, posterior type probabilities $\boldsymbol{\rho}^{(0)} = (\rho_{im}^{(0)})_{im}$, entry probabilities $\boldsymbol{p}^{(0)} = (p_{im\ell}^{(0)})_{im\ell}$. We set $\rho_{im}^{(0)} = \lambda_m^{(0)}$ and iterate the following updates for $n \geqslant 0$:

1. *Update entry probabilities and participation costs*

[17] In particular, public information includes information on the "supply regime" of a given day, i.e., the distribution of available varieties. Appendix A.4 describes how days are classified into different regimes. Appendix C.2 gives details on the implementation of our estimator.

Given $\boldsymbol{\theta}^{(n)}$ and $\boldsymbol{\rho}^{(n)}$, we first compute expected profits, then update the participation-cost profile and solve jointly for common cost locations and entry probabilities.

(a) For any mean pseudo-valuation $\bar{v}$ and state s, the first-order condition (3) implies that expected profits from optimal bidding can be written as

$$\Pi(\bar{v}, s) = \mathbb{E}_{\varepsilon \mid \boldsymbol{x}}\left[q \frac{\psi(b^* \mid s)^2}{\psi'(b^* \mid s)} \right],$$

where b^* solves $\bar{v} + \varepsilon = b^* + \frac{\psi(b^*|s)}{\psi'(b^*|s)}$. We obtain $\Pi(\bar{v}, s)$ directly from the estimated ψ and ψ', using Gauss-Hermite quadrature to integrate over the normally-distributed taste shocks using variances $\boldsymbol{\sigma}^{(n)}$.

Given the current entry probabilities, we compute continuation values backwards from the last auction. At each auction, the expected value associated with the entry decision is $-\tau^{(n)} \log(1 - p_{im\ell})$, and we integrate this value over the distribution of the next lot's characteristics. Starting from zero continuation value after the final auction and repeating this recursion backwards gives the continuation values for every type and inventory, and hence $\Delta\mathcal{V}_{m\ell}$, $\bar{v}_{im\ell}$, and $\Pi(\bar{v}_{im\ell}, s_\ell)$.

(b) Holding expected profits and common cost locations fixed, we update the participation-cost profile $\boldsymbol{a}$ by maximizing the posterior-weighted entry log-likelihood:

$$\boldsymbol{a}^{(n+1)} = \arg\max_{\boldsymbol{a}} \sum_i \sum_{m=1}^{M} \rho_{im}^{(n)} \sum_\ell \left[e_{i\ell} \log p_{im\ell}(\boldsymbol{a}) + (1 - e_{i\ell}) \log(1 - p_{im\ell}(\boldsymbol{a})) \right],$$

where $e_{i\ell} \in \{0, 1\}$ is the observed entry decision and

$$p_{im\ell}(\boldsymbol{a}) = \frac{\exp\left(\Pi(\bar{v}_{im\ell}, s_\ell)/\tau^{(n)}\right)}{\exp\left(\Pi(\bar{v}_{im\ell}, s_\ell)/\tau^{(n)}\right) + \exp\left((c_{s_\ell}^{(n)} + a_{j_\ell m})/\tau^{(n)}\right)}.$$

Holding $\boldsymbol{a}^{(n+1)}$ fixed, we update the common cost locations to reproduce observed entry frequencies. Let L_s denote the lots with the same public information, variety, and lot-size bin as s. Each c_s solves

$$\sum_i \sum_{\ell \in L_s} \sum_{m=1}^{M} \rho_{im}^{(n)} \frac{\exp\left(\Pi(\bar{v}_{im\ell}, s_\ell)/\tau^{(n)}\right)}{\exp\left(\Pi(\bar{v}_{im\ell}, s_\ell)/\tau^{(n)}\right) + \exp\left((c_s + a_{j_\ell m}^{(n+1)})/\tau^{(n)}\right)} = \sum_i \sum_{\ell \in L_s} e_{i\ell}.$$

Because predicted entry is decreasing in c_s, this is a one-dimensional root-finding problem for each cost location.

(c) Using the updated cost locations, we update entry probabilities according to

$$p_{im\ell} = \frac{\exp\left(\Pi(\bar{v}_{im\ell}, s_\ell)/\tau^{(n)}\right)}{\exp\left(\Pi(\bar{v}_{im\ell}, s_\ell)/\tau^{(n)}\right) + \exp\left((c_{s_\ell} + a_{j_\ell m}^{(n+1)})/\tau^{(n)}\right)}.$$

The updated entry probabilities change continuation values and expected profits. Holding $\boldsymbol{a}^{(n+1)}$ fixed, we repeat the continuation-value calculation, common-cost inversion, and entry-probability update until entry probabilities converge.

2. *M-step*
Given $\boldsymbol{p}^{(n+1)}$, the mean pseudo-valuation implied by candidate parameters $\boldsymbol{\theta}$ is

$$\bar{v}_{im\ell}\big(\boldsymbol{\theta}; \boldsymbol{p}^{(n+1)}\big) = \boldsymbol{\gamma}_m' \boldsymbol{x}_\ell - \alpha w_{i\ell} - \frac{1}{2}\alpha q_\ell + \frac{1}{q_\ell}\Delta\mathcal{V}_{m\ell}\big(w_{i\ell}, q_\ell, s_\ell; \boldsymbol{\theta}, \boldsymbol{p}^{(n+1)}\big),$$

where dependence of continuation term $\Delta\mathcal{V}$ on parameters and CCPs is made explicit in the notation. We update $\boldsymbol{\theta}^{(n+1)}$ as

$$\boldsymbol{\theta}^{(n+1)} = \arg\max_{\boldsymbol{\theta}} \sum_i \sum_{m=1}^{M} \rho_{im}^{(n)} \sum_{\ell: e_{i\ell}=1} \log f_{\varepsilon|\boldsymbol{x}_\ell}\big(\hat{v}_{i\ell} - \bar{v}_{im\ell}\big(\boldsymbol{\theta}; \boldsymbol{p}^{(n+1)}\big)\big),$$

where $\hat{v}_{i\ell}$ is obtained from inversion (7).

3. *E-step*
Using $\boldsymbol{\theta}^{(n+1)}$ and the entry probabilities $\boldsymbol{p}^{(n+1)}$, define the likelihood of bidder-day i conditional on type m as

$$\mathcal{L}_{im}^{(n+1)} = \prod_\ell \left(1 - p_{im\ell}^{(n+1)}\right)^{1-e_{i\ell}} \left[p_{im\ell}^{(n+1)} f_{\varepsilon|\boldsymbol{x}_\ell}\big(\hat{v}_{i\ell} - \bar{v}_{im\ell}\big(\boldsymbol{\theta}^{(n+1)}; \boldsymbol{p}^{(n+1)}\big)\big)\right]^{e_{i\ell}}.$$

We update posterior type probabilities using Bayes' rule

$$\rho_{im}^{(n+1)} = \frac{\lambda_m^{(n)} \mathcal{L}_{im}^{(n+1)}}{\sum_{m'=1}^{M} \lambda_{m'}^{(n)} \mathcal{L}_{im'}^{(n+1)}},$$

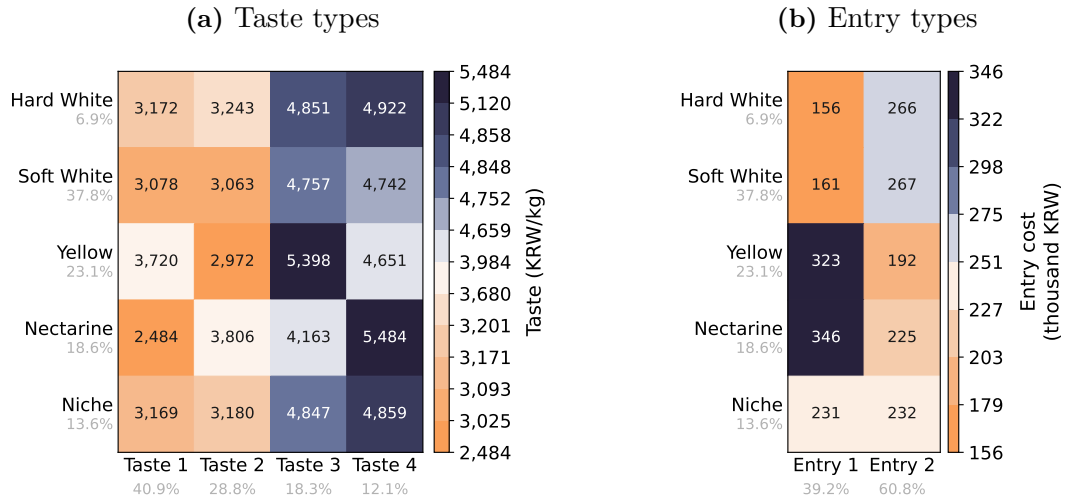
and population probabilities by averaging the posteriors over bidder-days

$$\lambda_m^{(n+1)} = \frac{1}{N} \sum_i \rho_{im}^{(n+1)}.$$

The next iteration starts from $(\boldsymbol{\theta}^{(n+1)}, \boldsymbol{c}^{(n+1)}, \boldsymbol{a}^{(n+1)}, \boldsymbol{\lambda}^{(n+1)}, \boldsymbol{\rho}^{(n+1)}, \boldsymbol{p}^{(n+1)})$. We iterate until convergence. Implementation and computation details are in Appendix C.2.

Identification Identification combines the panel structure of entry decisions with the cardinal information contained in bids. Under standard rank and support conditions, repeated within-day outcomes and variation in states identify the latent bidder-day types, their probabilities, and their type-specific CCPs up to relabeling (Kasahara and Shimotsu, 2009; Higgins and Jochmans, 2023). Unlike a dynamic discrete-choice model identified from choices alone, bid inversion reveals pseudo-valuations in monetary units and therefore provides a cardinal scale for payoffs. Conditional on the type-specific CCPs, the logit representation implies $\Delta\mathcal{V}_{m\ell} = \tau\Delta\mathcal{V}^1_{m\ell}$, where $\Delta\mathcal{V}^1_{m\ell}$ is determined by the CCPs and the distribution of future lot characteristics. Since inverted pseudo-valuations have conditional mean $\boldsymbol{\gamma}_m'\boldsymbol{x}_\ell - \alpha(w_{i\ell} + q_\ell/2) + \tau\Delta\mathcal{V}^1_{m\ell}/q_\ell$, the parameters $(\boldsymbol{\gamma}_m, \alpha, \tau)$ are therefore identified from three sources of variation: (i) lot characteristics identify tastes; (ii) accumulated inventory and lot size identify the demand slope; and (iii) variation in future opportunities that is excluded from current valuation (e.g., the number of remaining auctions or the daily supply regime) identifies the scale of dynamic incentives. Conditional on the valuation parameters and type-specific entry probabilities, the logit entry condition identifies type-specific cost locations. The restriction that there are only two possible entry-types separates costs into common cost locations c_s and variety-specific participation-cost deviations. Finally, $(\sigma_x)_x$ is identified from the conditional dispersion of inverted pseudo-valuations.

Figure 4 Estimates



Notes: Figure (a) reports estimated taste coefficients for each of the four latent taste-type and variety. Panel (b) reports the estimated participation costs for each of the two latent entry-types and variety. Each variety's and type's share are reported below their labels. Type shares are estimated. Variety shares are computed as the share of transacted kilograms belonging to each variety.

5.4 Results

Figure 4 gives a visualization of the estimated type-specific tastes and entry costs. The most salient dimension of horizontal taste differentiation is a relative preference for Yellow peaches compared to Nectarines. Vertical differentiation is quantitatively significant, with "strong" taste-types valuing any variety at roughly 1,700 KRW/kg more[18] than "weak" types. The latter are more common in the population. Along the entry dimension, types are mostly differentiated by their propensity to bid on either White peaches or Yellows and Nectarines. Figure D1 reports the average probabilities of each type bidding on each variety. Summary statistics on the distributions of state-specific entry costs are reported in Table D3.

The full estimates are reported in Tables D1–D2. The estimated demand slope implies that moving from the 25th to the 75th percentile of cumulative purchases lowers the direct component of willingness to pay by about 360 KRW/kg. This magnitude is consistent with the decline of bids in inventory discussed in Section 3. Estimates of regression (1) documented that bids in the highest inventory decile are about 215 KRW/kg lower than in the lowest decile. Figure D2 plots the observed and model-predicted homogenized bids in the inventory-remaining auctions space and shows that our estimated model replicates the observed patterns well, although at a slightly lower magnitude.

6 Counterfactuals

The estimates reveal differences in horizontal tastes and entry behavior that divide the product space differently. Horizontal tastes differ most strongly between Nectarines and Yellow peaches, while entry types separate bidders relatively likely to participate in auctions for White peaches from those relatively likely to participate in auctions for Nectarines and Yellows. This multidimensional heterogeneity suggests that differentiation can improve matching while also separating bidders who would otherwise compete for the same supply. We use pooling counterfactuals to examine the resulting tradeoff between matching and competition.

6.1 Pooling and the revenue-efficiency frontier

We first ask how pooling varieties affects revenue and efficiency when differentiation improves matching but also separates potential competitors. Selling varieties separately allows bidders to target the goods they value most, but selective entry can leave each variety contested by a narrower set of bidders. We compare schemes that combine selected varieties into mixed lots, with their composition dictated by available supply, while retaining the sequential first-price auction mechanism. This allows us to completely enumerate the set of feasible pooling

[18] That amounts to approximately 1.53 USD/kg at the 2017 exchange rates.

schemes, solve for the equilibrium, and evaluate the revenue and efficiency implications of each of them.[19] Efficiency is measured by expected match value, the total value of the goods to the bidders who receive them, accounting for their tastes and declining marginal willingness to pay as purchases accumulate.[20] We focus on five *environments* chosen to span contrasting supply compositions: Nectarine-heavy, Soft White-heavy, Soft White-led mix, Yellow and Soft White, and Yellow and Niche.[21]

Figure 5 reports the revenue-efficiency frontier. Across the five environments, the revenue-maximizing schemes pair each White variety with either Nectarine or Yellow while leaving Niche separate. The two schemes, labeled *S1* and *S2*, preserve the Nectarine-Yellow distinction while combining varieties favored by different entry types. The revenue-maximizing scheme in each environment would raise seller revenue by 4.8–7.5% (171–233 KRW/kg) relative to separate sales. Expected match value would fall by 0.8–2.2% (38–119 KRW/kg) in four environments but rise by 0.5% (29 KRW/kg) in Yellow and Niche. These revenue gains would entail a 7.8–18.6% decline in bidder surplus. Table F1 reports the levels and changes in revenue, expected match value, and bidder surplus for the status quo, all frontier schemes, *S1*, *S2*, and full pooling.

These schemes exploit the different product distinctions governing tastes and entry. Pooling a White variety with Nectarine or Yellow would bring into competition bidders who otherwise tend to participate in different auctions. At the same time, retaining separate pools for Nectarines and Yellows would preserve sorting along the main dimension of horizontal taste differentiation. The profitable schemes therefore eliminate distinctions that separate entry types while preserving those along which bidders have different relative tastes.

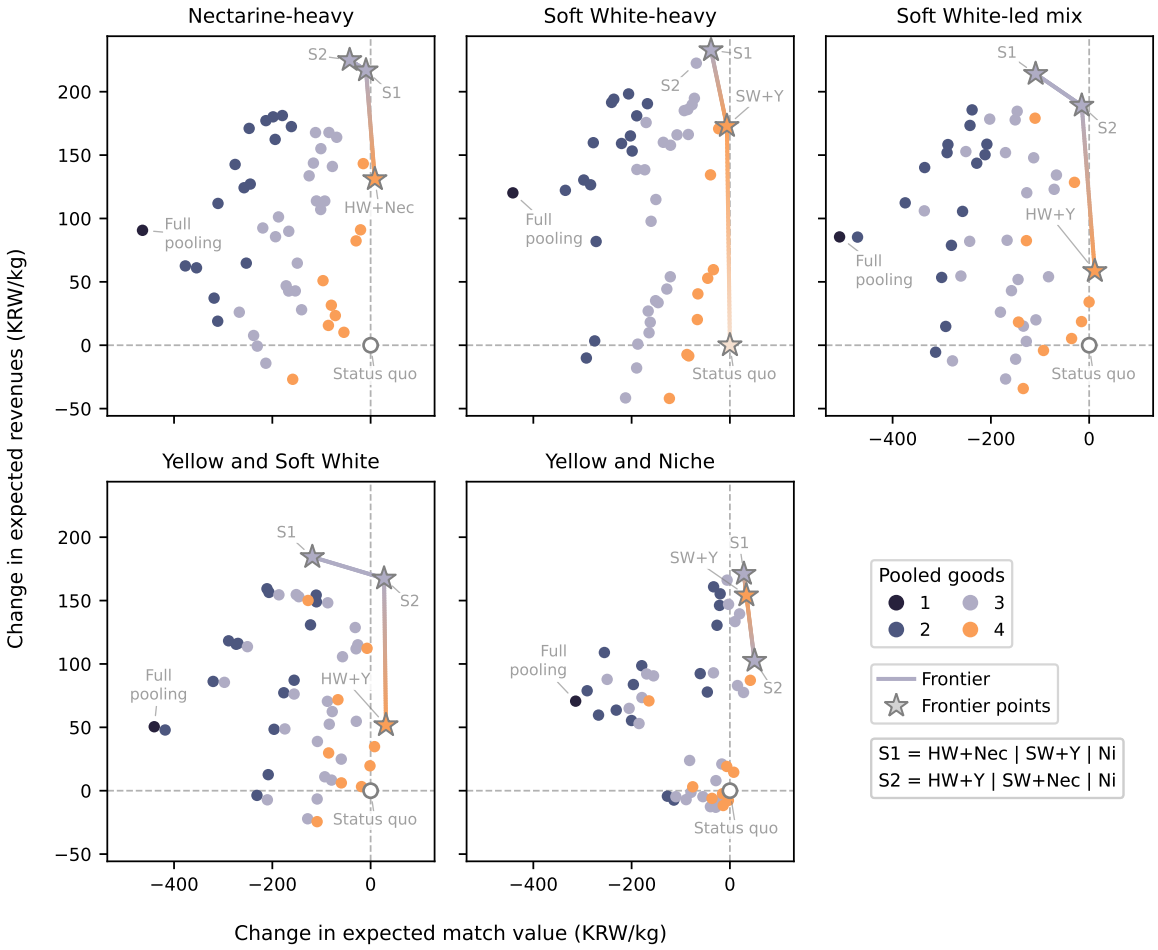
Revenue gains need not always require a loss in match value. In the *Nectarine-heavy* environment, pooling Hard White with Nectarine alone would raise revenue by 131.0 KRW/kg (4.3%) and expected match value by 8.6 KRW/kg (0.17%). In the *Soft White-led mix* environment, pooling Hard White with Yellow alone would raise revenue by 58.6 KRW/kg (1.9%) and expected match value by 11.5 KRW/kg (0.22%). In *Yellow and Niche*, revenue maximization would *improve* efficiency, with revenues increasing by 171.1 KRW/kg (4.8%) and expected match value by 28.5 KRW/kg (0.53%). Although the matching gains would be small, these

[19] Note that our restriction rules out "partial" pooling scheme. For instance, an auctioneer with Hard White and Yellow peaches could in principle decide to sell some White-only lots, some Yellow-only lots, and some 50-50 lots split between Hard Whites and Yellows. The set of such schemes is too high dimensional to allow any useful characterization, either empirically or theoretically. Under our restriction, there are 52 possible pooling configurations, ranging from the status quo to full pooling.

[20] Expected match value is $\mathbb{E}[\sum_{\ell=1}^{L}\sum_{i=1}^{N}\mathbf{1}\{i \text{ wins lot } \ell\}q_\ell(\boldsymbol{\gamma}_i'\boldsymbol{x}_\ell + \varepsilon_{i\ell} - \alpha w_{i\ell} - \alpha q_\ell/2)]$, where $w_{i\ell}$ is bidder i's cumulative purchases before lot ℓ. The expression integrates the winner's marginal willingness to pay over the quantity purchased, with the expectation taken over equilibrium outcomes. Participation costs are interpreted as capturing imperfect attention across the sequence of auctions and are assumed to not contribute to welfare.

[21] An environment is a realization of the public state z, which pins down the daily supply composition and an indicator of whether the number of attending bidders is above the median.

Figure 5 Revenue-efficiency frontier for selected environments



Notes: Figure reports the revenue-efficiency frontier for selected environments. Changes are relative to the status quo. Levels and changes are measured per kilogram sold under the corresponding environment's status quo. Across the five environments, the status quo expected revenues are 3,038, 3,106, 3,122, 3,655, and 3,575 KRW/kg and the expected match values are 5,181, 5,036, 5,176, 5,281, and 5,402 KRW/kg. Based on their supply composition, environments are labeled as *Nectarine-heavy*, *Soft White-heavy*, *Soft White-led mix*, *Yellow and Soft White*, and *Yellow and Niche*. Dots represent pooling schemes. Their color represents how many goods remain available after pooling. For instance, full pooling corresponds to one pooled good and a scheme that pools two varieties while leaving the other ones untouched corresponds to four pooled goods. Stars identify frontier points, i.e., those for which no other scheme yields more revenues and higher match values. Labels *S1* and *S2* are used to denote two focal pooling schemes. Scheme *S1* pools Hard Whites with Nectarines and Soft Whites with Yellows. Scheme *S2* pools Hard Whites with Yellows and Soft Whites with Nectarines. Both leave Niche unpooled.

results show that broader participation can offset the loss from preventing the constituent varieties from being assigned separately. Selective entry limits the matching benefits of differentiation because bidders who would value a lot highly may not participate in its auction.

More extensive pooling would generally entail larger matching losses without delivering

higher revenues. Full pooling would raise revenue by only 50–120 KRW/kg (1.4–3.9%) while reducing expected match value by approximately 314–508 KRW/kg (5.8–9.8%). Schemes that combine Nectarines and Yellows would also be dominated. These varieties attract the same entry type but lie on opposite sides of the main horizontal taste divide, so combining them would sacrifice valuable sorting while doing relatively little to broaden competition.

6.2 Which fruits should be pooled?

We next ask which features of demand and supply explain why the revenue gains from pooling vary across schemes and environments. Pooling varieties that attract different bidders could broaden competition, but assigning them together would sacrifice sorting opportunities when bidders disagree about their relative values. The supply mix weights these potential participation gains and horizontal matching losses differently. We construct measures of both forces from the estimated primitives and examine how their balance accounts for the revenue effects of different pooling schemes.

Our first measure captures the potential participation gain from combining varieties favored by different entry types. Denote the set of varieties sold as one good by $P \subseteq J$, the quantity of variety j by Q_j and the quantity contained in good P by $Q_P = \sum_{j \in P} Q_j$. We use estimated participation cost differences to construct type m's relative entry propensity for variety j, denoted by r_{jm}.[22] Interest in any constituent variety can make the pooled good relevant to a bidder, but the strength of this effect should depend on how much of that variety the lot contains. We allow its relevance to increase with quantity at a diminishing rate, using $h(q) := 1 - \exp(-q/\kappa)$. For a pooled lot of q kilograms, we weight variety j's entry propensity by $h(qQ_j/Q_P)/h(q)$, comparing the relevance of its quantity within the pooled lot with that of a full lot of the same variety. The potential *entry gain* is

$$G(P) = \frac{Q_P}{\sum_j Q_j} \sum_m \lambda_m \Bigg\{ \underbrace{\mathbb{E}_q\left[1 - \prod_{j \in P}\left(1 - r_{jm}\frac{h(qQ_j/Q_P)}{h(q)}\right)\right]}_{\substack{\text{entry propensity} \\ \text{after pooling}}} - \underbrace{\sum_{j \in P}\frac{Q_j}{Q_P} r_{jm}}_{\substack{\text{status quo} \\ \text{entry propensity}}} \Bigg\},$$

where λ_m is the frequency of bidder type m and the expectation averages over lot quantities. The first term combines the constituent varieties' entry propensities, discounting each according to its quantity within the pooled lot. The second term is their supply-weighted average under separate sales. The measure captures the scope for bringing additional entry types into

[22] We define $r_{jm} := \exp\{-[a_{jm} - \min_{m'} a_{jm'}]/\tau\}$, where a_{jm} is the variety-specific participation cost deviation for type m and τ is the scale of the participation cost distribution. The normalization assigns propensity one to the type facing the lowest participation cost for variety j and lower propensities to types facing higher costs. The common cost location cancels from this comparison. Equilibrium entry probabilities also depend on expected profits, whereas r_{jm} isolates the relative participation cost component.

competition for the same supply.

Our second measure captures the sorting opportunities lost when varieties must be assigned together. For a fixed number of bidders drawn from the estimated type distribution, the *matching loss* is

$$L(P) = \frac{Q_P}{\sum_j Q_j}\mathbb{E}\Bigg[\underbrace{\sum_{j\in P}\frac{Q_j}{Q_P}\max_i\left\{\gamma_{m_i j}+\varepsilon_{ij}\right\}}_{\substack{\text{status quo}\\ \text{horizontal match value}}} - \underbrace{\max_i\left\{\sum_{j\in P}\frac{Q_j}{Q_P}\left(\gamma_{m_i j}+\varepsilon_{ij}\right)\right\}}_{\substack{\text{horizontal match value}\\ \text{under pooled assignment}}}\Bigg].$$

The first term assigns each variety to the bidder who values it most; the second assigns the entire pooled good to one bidder. The expectation is over bidder types and variety-specific taste shocks. This comparison measures the loss from suppressing sorting while holding participation fixed. Appendix F describes the construction of entry propensities and the calibration and numerical choices used to evaluate both measures.

A pooling scheme $\mathcal{P}$ is a partition of the available varieties, with each element $P \in \mathcal{P}$ specifying the varieties sold together as one good. Singleton elements correspond to varieties left separate. We combine the participation gains and matching losses across pooled goods into the following index of the scheme's revenue change:

$$\Delta\hat{R}(\mathcal{P}) = \sum_{P\in\mathcal{P}:|P|>1}\Bigg[\underbrace{\zeta_{\text{entry}}G(P)}_{\substack{\text{predicted gain from}\\ \text{broader entry}}} - \underbrace{\zeta_{\text{match}}L(P)}_{\substack{\text{predicted loss from}\\ \text{coarser matching}}}\Bigg]. \tag{8}$$

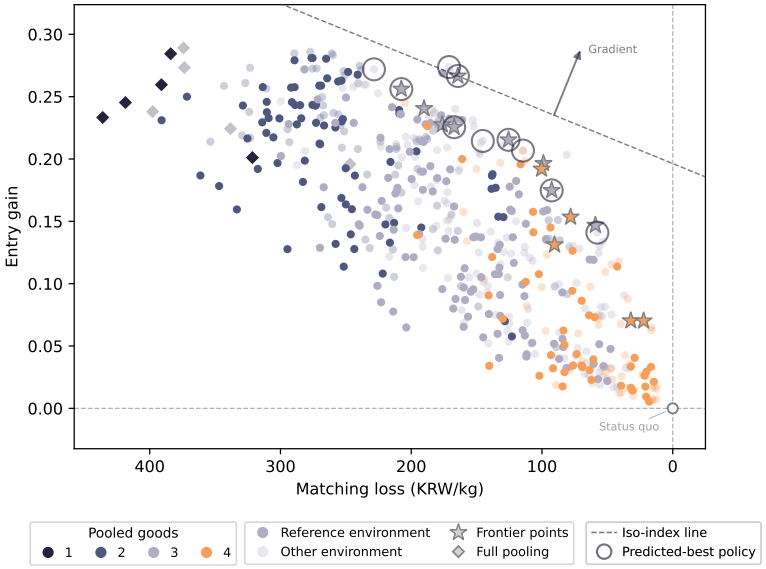
We estimate the two coefficients using only the 50 counterfactuals that pool a single pair of varieties while leaving all remaining varieties separate, comprising ten schemes in each of the five environments. We then hold these coefficients fixed and evaluate the index across all pooling schemes. Each pooled good is evaluated using its own participation gain and matching loss.

Figure 6a compares the potential participation gains and matching losses associated with different pooling schemes across environments. The schemes with the highest predicted revenues differ across environments in both dimensions, implying different losses of horizontal match value even when their pooling structures are similar. Which product distinctions remain profitable to preserve depends on how much additional competition pooling could attract relative to the sorting opportunities it would remove. More extensive pooling can offer little additional participation while substantially increasing matching losses.

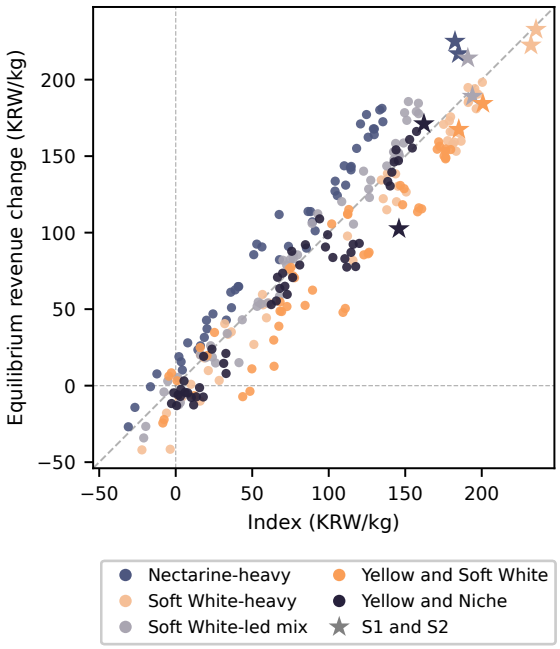
Figure 6b compares the index with equilibrium revenue changes over the full policy space, for the five environments considered in Figure 5. Schemes containing larger pooled goods or several pooled goods were excluded from fitting and provide a test of the index's predictive performance. Across the full policy space, within-environment rank correlations range from

Figure 6 Predicting revenue changes and preferred pooling schemes

(a) Participation gains and matching losses



(b) Predictive performance



Notes: Panel (a) plots each pooling scheme's potential entry gain against its matching loss. Each point represents a scheme-environment combination. The horizontal axis is reversed. Colors indicate the number of goods remaining after pooling, as in Figure 5 Solid markers correspond to the five environments considered in Figure 5, translucent markers correspond to other environments. Stars identify schemes on the equilibrium revenue-efficiency frontiers of the five selected environments. Diamonds identify full pooling. Large outlined circles identify the scheme with the highest index value in each environment. The dashed line joins combinations with equal values of the index in (8), with the arrow pointing toward increasing index values. Panel (b) plots equilibrium revenue changes against the index, both relative to the status quo and measured in KRW/kg. Each point represents a pooling scheme. Colors distinguish environments, and stars identify *S1* and *S2*. The index coefficients are fitted using single-pair pooling counterfactuals and held fixed for all other schemes. Scheme *S1* pools Hard White with Nectarine and Soft White with Yellow. Scheme *S2* pools Hard White with Yellow and Soft White with Nectarine. Both schemes leave Niche separate.

0.955 to 0.989, with a pooled rank correlation of 0.947 across all five environments. The relationship estimated from single-pair counterfactuals remains highly predictive of the ranking of more complex pooling schemes. The remaining errors include systematic differences across environments in the level of revenue gains. The index understates gains in the Nectarine-heavy environment by 25.3 KRW/kg on average and overstates them in Yellow and Soft White by 20.9 KRW/kg.

The close fit is striking because the index does not explicitly account for how pooling changes inventories and subsequent entry and bidding. It suggests that much of the variation in profitability reflects the separation between entry and taste heterogeneity. Pooling a White

variety with Nectarine or Yellow could therefore broaden competition while preserving the Nectarine-Yellow distinction. In the index, these combinations generate substantial potential entry gains with comparatively modest matching losses. Pooling Nectarine with Yellow instead offers little additional participation while suppressing the strongest horizontal taste distinction. The two measures capture this separation between the product distinctions that fragment competition and those that support valuable sorting.

The index also provides a simple way to construct a pooling scheme. Starting from the status quo's separate sales, we combine the two goods whose merger produces the largest positive increase in the index. We then reevaluate the remaining mergers and continue until none would increase it. Figure 7 illustrates this greedy procedure for three environments. In the Nectarine-heavy environment, it first combines Soft White with Nectarine and then Hard White with Yellow. In the other four environments, it first combines Soft White with Yellow and then Hard White with Nectarine. In all five environments, the procedure stops after two poolings and selects the revenue-maximizing scheme identified by the full equilibrium comparisons. Using equilibrium revenue rather than the index to choose each merger yields the same sequence in every environment. The stopping point has a direct economic interpretation. Once each White variety has been paired with Nectarine or Yellow, further pooling would add little potential participation while increasing matching losses. The index therefore favors stopping after two pairings and leaving Niche separate. The preferred pairings change with supply, but follow the same economic principle: combine varieties that attract different entry types while preserving the distinctions that matter most for horizontal sorting.

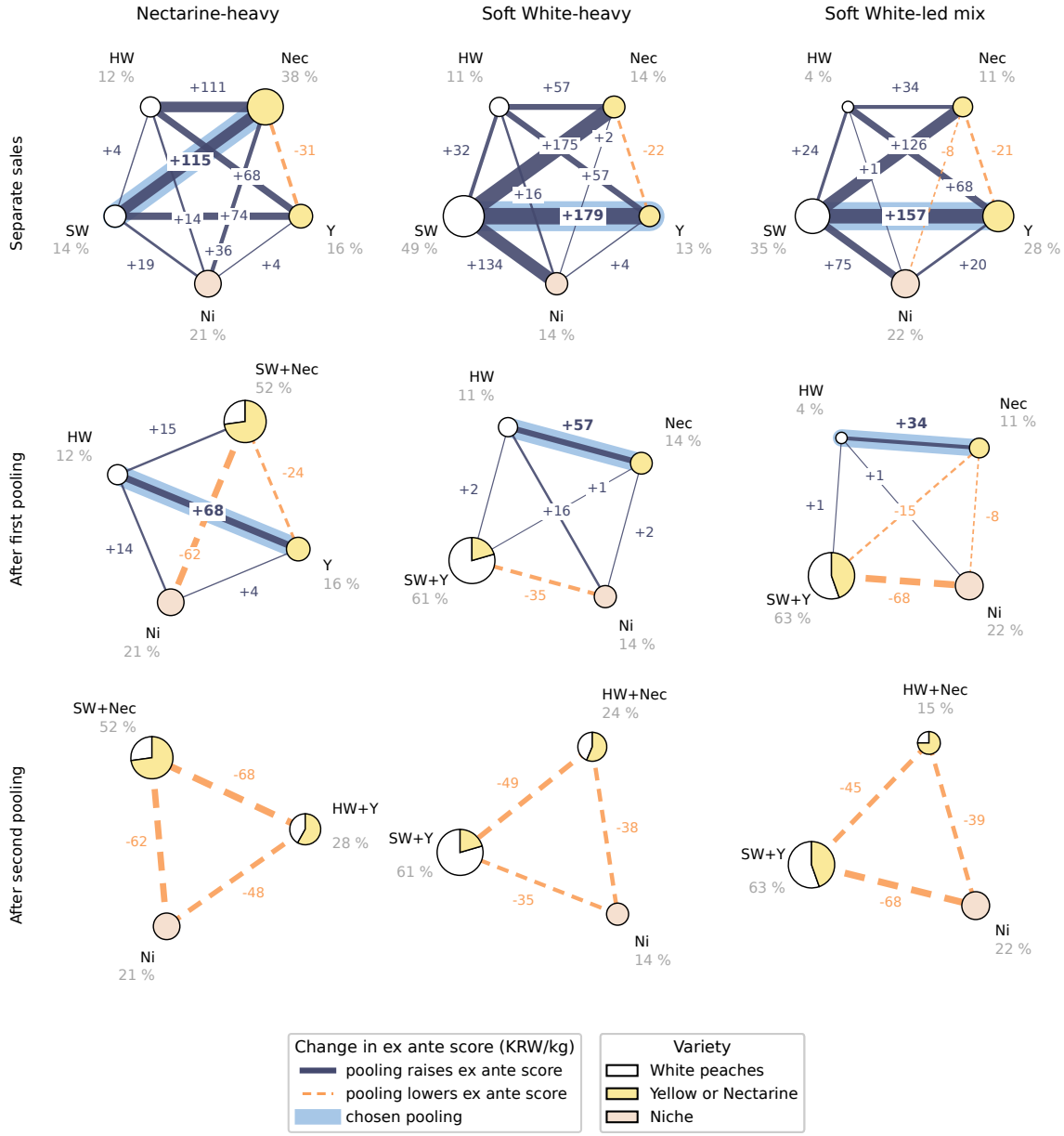
6.3 How does the tradeoff depend on the environment?

We then use the index to examine how the predicted revenue-maximizing pooling scheme changes with supply composition. Figure 8 varies the relative supply of Nectarines and Yellows and the share of White peaches in non-Niche supply.[23] At each supply composition, the color reports the predicted revenue advantage, in KRW/kg, of the better of *S1* and *S2* over the best of all remaining pooling schemes. Positive values therefore identify compositions at which one of these two schemes has the highest index value. The dashed curve marks compositions at which *S1* and *S2* have equal index values.

One of the two schemes that pair Whites with Nectarines and Yellows is preferred over a broad range of supply compositions. Changes in relative supply can reverse which pairing is preferred, while sufficiently different compositions can favor another pooling structure. These

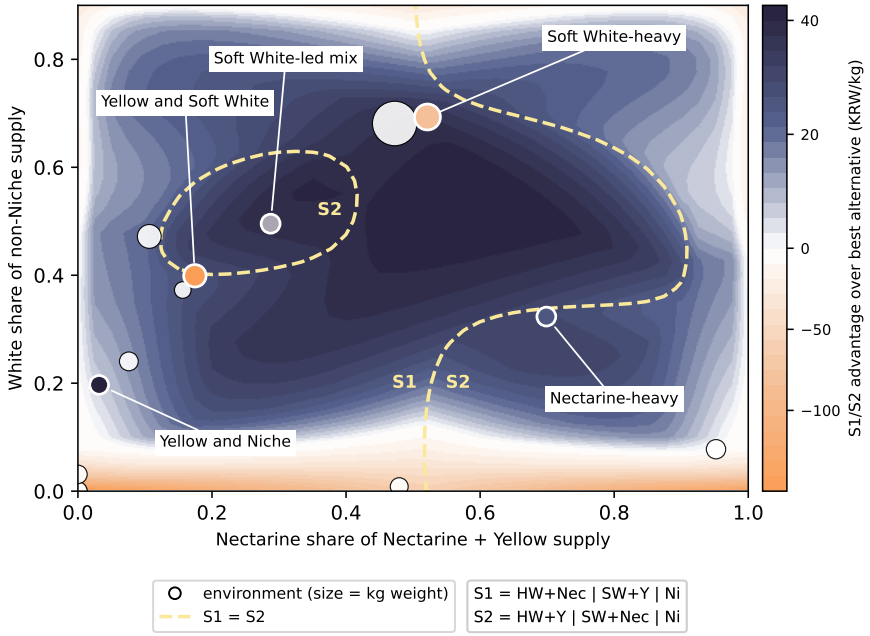
[23]These axes vary supply along the two economically distinct divisions identified in the estimates. The horizontal axis changes the balance between Nectarines and Yellows, which attract the same entry type but differ sharply in horizontal tastes. The vertical axis changes the balance between Whites and the combined supply of Nectarines and Yellows, which attract different entry types. Niche's supply share and the Hard White share within Whites are held fixed, as described in the figure notes.

Figure 7 Choosing which varieties to pool



Notes: Figure shows a sequential, greedy procedure for choosing a pooling scheme using predicted revenue gains for selected environments. Columns represent three supply environments. Each column should be read downward, with the first row corresponding to the status quo. Each node represents a variety. *HW* denotes Hard White, *SW* Soft White, *Nec* Nectarine, *Y* Yellow, and *Ni* Niche. A plus sign denotes varieties sold together as one pooled good. Node areas and gray percentages indicate shares of total supply in kilograms. Pie slices show the composition of pooled goods. Each edge represents a possible pooling of two varieties and its label reports the incremental predicted revenue change in KRW/kg relative to the scheme shown in that row. Predictions use the index defined in (8). Solid purple edges indicate predicted gains and dashed orange edges indicate predicted losses. Thicker edges represent larger absolute changes. The light-blue highlight identifies the pooling operation yielding the largest positive predicted gain. Pooling such varieties produces the scheme in the next row, where all remaining pooling possibilities are re-evaluated. In the bottom row, every remaining pooling lowers the index, so the procedure stops. The final scheme maximizes equilibrium revenue among all feasible schemes in each of the three environments.

Figure 8 Supply composition and revenue-maximizing pooling schemes



Notes: Figure compares the index values of *S1* and *S2* with those of other pooling schemes across supply compositions. The color reports the index value of the better of *S1* and *S2* minus the highest index value among all other pooling schemes, measured in KRW/kg. Positive values indicate that *S1* or *S2* is preferred. In *S1*, Hard White is pooled with Nectarine and Soft White with Yellow. In *S2*, Hard White is pooled with Yellow and Soft White with Nectarine. Both schemes leave Niche separate. The horizontal axis is Nectarine's share of combined Nectarine and Yellow supply. The vertical axis is White peaches' share of non-Niche supply. The dashed line marks equal index values for *S1* and *S2*. Niche's supply share and the Hard White share within Whites are held at their kilogram-weighted empirical values. Other primitives are held at the reference environment's values. Circles locate observed supply compositions, with area proportional to their share of total kilograms. Colored circles identify the five environments considered in Figure 5.

departures occur mainly when Whites account for a very small or very large share of non-Niche supply. In these cases, the potential participation gain from a pairing is limited by the small quantity of the constituent variety that would attract the other entry type. Other pooling structures can then achieve a more favorable balance between participation gains and matching losses.

7 Conclusion

We study how pooling differentiated goods changes revenue and matching when bidders participate selectively. Broader participation could offset the loss of sorting opportunities, allowing some pooling schemes to improve both revenue and match value relative to the status quo. Revenue would be maximized under partial pooling, but bidder surplus would fall at these maxima in every case. More extensive pooling would sacrifice additional sorting opportunities

without attracting enough additional competition to raise revenue further.

The revenue ranking of pooling schemes changes with supply composition, even when tastes and participation costs remain fixed. The quantities available determine how participation gains and matching losses are weighted across varieties. The effects of pooling two goods therefore depend on their place in the broader supply mix, rather than on their similarity alone.

References

Aguirregabiria, V., and P. Mira (2002): "Swapping the Nested Fixed Point Algorithm: A Class of Estimators for Discrete Markov Decision Models," *Econometrica*, 70 (4), 1519–1543.

——— (2007): "Sequential Estimation of Dynamic Discrete Games," *Econometrica*, 75 (1), 1–53.

Altmann, S. (2023): "Identification and Estimation of a Dynamic Multi-Object Auction Model," *Mimeo*.

Altmann, S. M. (2023): "Choice by Design: Evidence from Feeding America's Food Allocation Problem," *Mimeo*.

Antler, Y., D. Bird, and D. Fershtman (2025): "Search, Matching, and Dating Apps: Designed to be Deleted, or Designed to be Repeated?" *Mimeo*.

Arcidiacono, P., and J. B. Jones (2003): "Finite Mixture Distributions, Sequential Likelihood and the EM Algorithm," *Econometrica*, 71 (3), 933–946.

Arcidiacono, P., and R. A. Miller (2011): "Conditional Choice Probability Estimation of Dynamic Discrete Choice Models With Unobserved Heterogeneity," *Econometrica*, 79 (6), 1823–1867.

Athey, S., and P. A. Haile (2007): "Nonparametric Approaches to Auctions," in *Handbook of Econometrics* Volume 6, Amsterdam: Elsevier, 3847–3965.

Backus, M., and G. Lewis (2025): "Dynamic Demand Estimation in Auction Markets," *The Review of Economic Studies*, 92 (2), 837–872.

Bergemann, D., T. Heumann, S. Morris, C. Sorokin, and E. Winter (2022): "Optimal Information Disclosure in Classic Auctions," *American Economic Review: Insights*, 4 (3), 371–388.

Bodoh-Creed, A. L., J. Boehnke, and B. Hickman (2021): "How Efficient are Decentralized Auction Platforms?" *The Review of Economic Studies*, 88 (1), 91–125.

Choi, J. J. (2024): "Sequential First-Price Auctions Under Partial Disclosure: An Application to Korean Fruit Auction," *Mimeo*.

Christofi, M., A. Pavlou, D. Lantzouraki et al. (2022): "Profiling Carotenoid and Phenolic Compounds in Fresh and Canned Fruit of Peach Cultivars: Impact of Genotype and Canning on Their Concentration," *Journal of Food Composition and Analysis*, 114, 104734.

Compiani, G., P. Haile, and M. Sant'Anna (2020): "Common Values, Unobserved Heterogeneity, and Endogenous Entry in US Offshore Oil Lease Auctions," *Journal of Political Economy*, 128 (10), 3872–3912.

Cui, J. (2025): "Bidding for Reputation," *Mimeo*.

De Silva, D. G., M. L. Gentry, and P. Schiraldi (2023): "Dynamic Oligopoly in Procurement Auction Markets," *Mimeo*.

Fershtman, D., and A. Pavan (2022): "Matching Auctions," *The RAND Journal of Economics*, 53 (1), 32–62.

Ganuza, J.-J., and J. S. Penalva (2010): "Signal Orderings Based on Dispersion and the Supply of Private Information in Auctions," *Econometrica*, 78 (3), 1007–1030.

Generoso, N., and M. Hong (2026): "Who Gains from Thicker Markets? Access and Market Power in Agricultural Trade," *Mimeo*.

Gentry, M., and T. Li (2014): "Identification in Auctions With Selective Entry," *Econometrica*, 82 (1), 315–344.

Gieselmann, J. (2025): "(Mis-)Matchmaker," *Mimeo*.

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-price Auctions," *Econometrica*, 68 (3), 525–574.

Haile, P. A., H. Hong, and Shum (2003): "Nonparametric Tests for Common Values In First-Price Sealed-Bid Auctions," *Mimeo*.

Higgins, A., and K. Jochmans (2023): "Identification of Mixtures of Dynamic Discrete Choices," *Journal of Econometrics*, 237 (1), 105462.

Hotz, V. J., and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60 (3), 497–529.

Jehiel, P., M. Meyer-ter Vehn, and B. Moldovanu (2007): "Mixed bundling auctions," *Journal of Economic Theory*, 134 (1), 494–512.

Jofre-Bonet, M., and M. Pesendorfer (2003): "Estimation of a Dynamic Auction Game," *Econometrica*, 71 (5), 1443–1489.

Kasahara, H., and K. Shimotsu (2009): "Nonparametric Identification of Finite Mixture Models of Dynamic Discrete Choices," *Econometrica*, 77 (1), 135–175.

Kobayashi, S. J., and M. Alcobendas (2024): "Dynamic Auctions with Budget-Constrained Bidders: Evidence from the Online Advertising Market," *Mimeo*.

Kong, Y. (2021): "Sequential Auctions with Synergy and Affiliation across Auctions," *Journal of Political Economy*, 129 (1), 148–181.

Krasnokutskaya, E. (2011): "Identification and Estimation of Auction Models with Unobserved Heterogeneity," *The Review of Economic Studies*, 78 (1), 293–327.

Krasnokutskaya, E., and K. Seim (2011): "Bid Preference Programs and Participation in Highway Procurement Auctions," *American Economic Review*, 101 (6), 2653–2686.

Levin, J., and P. Milgrom (2010): "Online Advertising: Heterogeneity and Conflation in Market Design," *American Economic Review*, 100 (2), 603–607.

Loertscher, S., and E. V. Muir (2025): "Optimal Hotelling Auctions."

Lu, Y., A. Gupta, W. Ketter, and E. van Heck (2019): "Dynamic Decision Making in Sequential Business-to-Business Auctions," *Management Science*.

Petruccelli, R., A. Bonetti, L. Ciaccheri, F. Ieri, T. Ganino, and C. Faraloni (2023): "Evaluation of the Fruit Quality and Phytochemical Compounds in Peach and Nectarine Cultivars," *Plants*, 12 (8).

Sousa, T. (2026): "Competition and Targeting in Auctions for Online Ads," *Mimeo*.

Tunuguntla, S., C. F. Mela, and J. Pratt (2026): "Targeting Information in Ad Auction Mechanisms," *Mimeo*.

Weintraub, G. Y., C. L. Benkard, and B. Van Roy (2008): "Markov Perfect Industry Dynamics With Many Firms," *Econometrica*, 76 (6), 1375–1411.

A Additional descriptive evidence

A.1 Testing for selective entry

We test whether bidders selectively enter auctions of different varieties. A bidder is classified as an entrant into an auction if they submit at least one bid in that auction.

For bidder i, day d, and variety j, let E_{ijd} denote the number of distinct auctions of variety j entered by bidder i, and let $E_{id} = \sum_j E_{ijd}$ denote the bidder's total number of entered auctions that day. Define the bidder's empirical entry shares by

$$q_{ijd} := \frac{E_{ijd}}{E_{id}}.$$

Denote the number of lots of variety j auctioned on day d by L_{jd}, and the total number of lots on day d as $L_d = \sum_j L_{jd}$. Define variety lot shares as

$$p_{jd} := \frac{L_{jd}}{L_d}.$$

We test the null hypothesis that, conditional on bidder i's observed number of entered auctions and the day's variety mix, the bidder's entries are randomly allocated across varieties:

$$(E_{ijd})_j \,\big|\, E_{id}, \boldsymbol{p}_d \sim \text{Multinomial}(E_{id}, \boldsymbol{p}_d).$$

For each bidder-day, we measure the departure from the null using the multinomial likelihood-ratio statistic

$$G^2_{id} = 2 \sum_j E_{ijd} \log\left(\frac{E_{ijd}}{E_{id}\, p_{jd}}\right) = 2 E_{id} D_{\text{KL}}(\boldsymbol{q}_{id} \,\|\, \boldsymbol{p}_d).$$

where terms with $E_{ijd} = 0$ are defined to be zero and D_{KL} denotes the Kullback-Leibler divergence. We aggregate these bidder-day terms to construct the observed test statistic

$$T_{\text{obs}} = \sum_{i,d} G^2_{id}.$$

Since many bidder-days contain relatively few entered auctions, we construct the null distribution of T by simulation. For simulation s, we draw

$$\left(E^{(s)}_{ijd}\right)_j \sim \text{Multinomial}(E_{id}, \boldsymbol{p}_d),$$

independently across bidder-days, holding fixed both the observed number of auctions entered by each bidder and the varieties available on that day. We then construct the corresponding

aggregate statistic $T^{(s)}$. The Monte Carlo p-value is

$$\hat{p} = \frac{1 + \sum_{s=1}^{S} \mathbf{1}\{T^{(s)} \geq T_{\text{obs}}\}}{S + 1}.$$

The upper panel of Table A1 reports the results. None of the 2,000 simulated statistics is as large as the observed statistic. To assess the magnitude of selective entry, the top panel of Table A1 reports the KL divergence between the profiles of entered and available varieties. Even under random entry, realized bidder-day entry shares generally differ from the available variety shares because bidders enter only finitely many auctions, so the relevant benchmark for the divergence is its value under the simulated null rather than zero. The comparison of observed values to the null-implied ones suggests that the excess selectiveness is not confined to a small number of extreme bidder-days: the divergence for the median bidder-day is about 60 percent larger than would be expected from random entry alone.

Table A1 Within-day selectiveness in entry across varieties

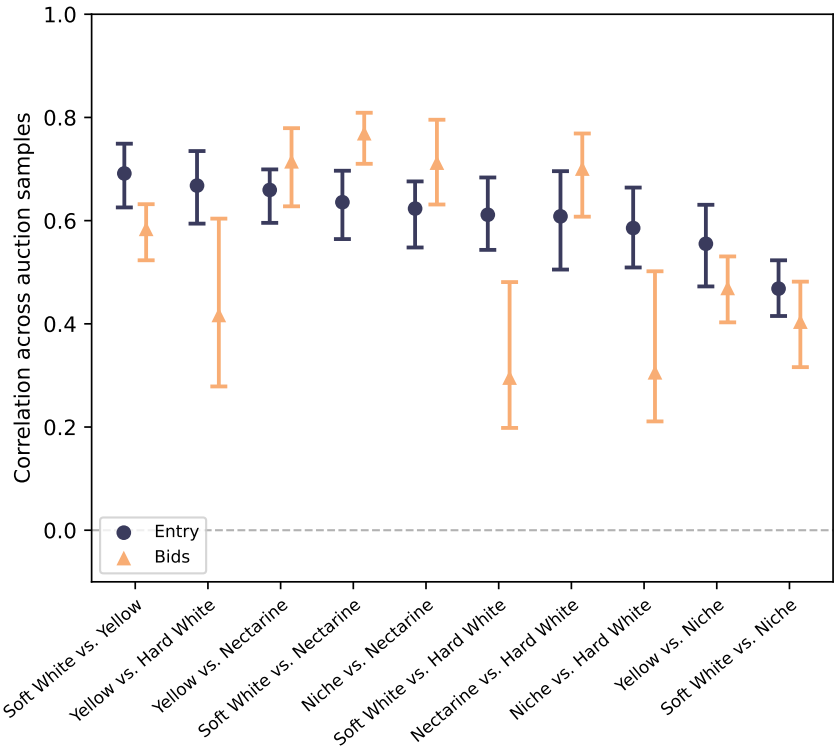
Test of random entry across varieties		
Bidder-days in test sample		4,529
Observed statistic T_{obs}		31,905
Simulated null mean		15,701
Simulated null median		15,706
Simulated null 99th percentile		16,070
Number of simulations		2,000
Monte Carlo p-value		< 0.001
Bidder-day sorting magnitudes		
	Observed	Null
Mean $D_{\text{KL}}(\boldsymbol{q}_{id} \,\|\, \boldsymbol{p}_d)$	0.333	0.210
Median $D_{\text{KL}}(\boldsymbol{q}_{id} \,\|\, \boldsymbol{p}_d)$	0.208	0.126

Notes: The top panel reports the result of our test for random entry across varieties. The bottom panel reports the observed and simulated (under the null of random entry across varieties) mean and median Kullback-Leibler divergence of bidder-day entry shares $\boldsymbol{q}_{id}$ from the day's variety mix $\boldsymbol{p}_d$.
(Return to the descriptives section in the main text.)

A.2 Entry and bidding heterogeneity across varieties

This section provides additional evidence on heterogeneity in bidders' entry and bidding patterns across varieties.

Figure A1 Persistence of entry and bidding patterns across varieties



Notes: Figure reports, for each pair of varieties, correlations across samples A and B in relative entry differences, relative bid differences, and the relationship between the two. Each row corresponds to one variety pair. *Entry* and *Bids* report correlations of the same pairwise difference across samples A and B. Horizontal lines report 95% confidence intervals from 999 bootstrap draws over bidders. Only variety pairs with at least 50 bidder-days are included.
(Return to the descriptives section in the main text.)

Fact A1 *Relative entry and bidding patterns across varieties are persistent within bidder-days, but weakly related to each other.*

We use pairwise comparisons across varieties to isolate horizontal heterogeneity from differences in bidders' overall entry and bid levels. For each pair of varieties (j, j'), we ask whether (i) bidder-days that disproportionately enter variety j rather than variety j' exhibit the same relative entry pattern in a separate set of auctions, and (ii) bidder-days that bid relatively more for variety j than for variety j' exhibit the same relative bidding pattern in a separate set of auctions. We then compare the two patterns to assess whether relative entry and relative bidding reflect the same dimension of bidder heterogeneity.

For each trading day, we divide the auction sequence into four blocks containing approximately equal numbers of auctions. Within each time block and variety, successive auctions are assigned alternately to samples A and B. This produces two non-overlapping samples contain-

ing auctions of the same varieties at similar points in the auction day. All objects below are computed separately within each time block and sample, and we suppress these two indices to simplify notation.

Index bidder-days by i and varieties by j. Denote the number of auctions of variety j bidder-day i is exposed to in the relevant time block and sample by L_{ij}, and the number of other bidders who submit at least one bid on bidder-day i's auction day by N_i. We define bidder-day i's *relative entry rate* for variety j as the difference between bidder-day i's entry rate into variety j and the average entry rate of the other bidders over the same auctions:

$$\bar{e}_{ij} \coloneqq \frac{1}{L_{ij}} \sum_{\ell : j(\ell)=j} \left[\mathbf{1}(i \text{ bids on } \ell) - \frac{1}{N_i} \sum_{i' \neq i} \mathbf{1}(i' \text{ bids on } \ell) \right].$$

For a pair of varieties (j, j'), the difference $\bar{e}_{ij} - \bar{e}_{ij'}$ measures whether bidder-day i disproportionately enters variety j rather than variety j'. We compute this difference within each time block and average it across blocks, so that varieties are compared at similar points in the auction day. We perform this calculation separately in samples A and B, giving two measurements of the same bidder-day's relative entry pattern.

For bids, let $\tilde{b}_{i\ell}$ denote bidder-day i's residualized bid for lot ℓ, obtained as the residual of a regression of all bids on observed lot characteristics, inventory, time of day, and the bidder-day fixed effects. Denote by $\bar{b}_{ij}$ the average of such residuals over lots of variety j bid on by i. For a pair of varieties (j, j'), the difference $\bar{b}_{ij} - \bar{b}_{ij'}$ measures whether bidder-day i bids relatively more for j than for j'. As for entry, we compute these objects separately in samples A and B and within each time block before averaging across blocks.

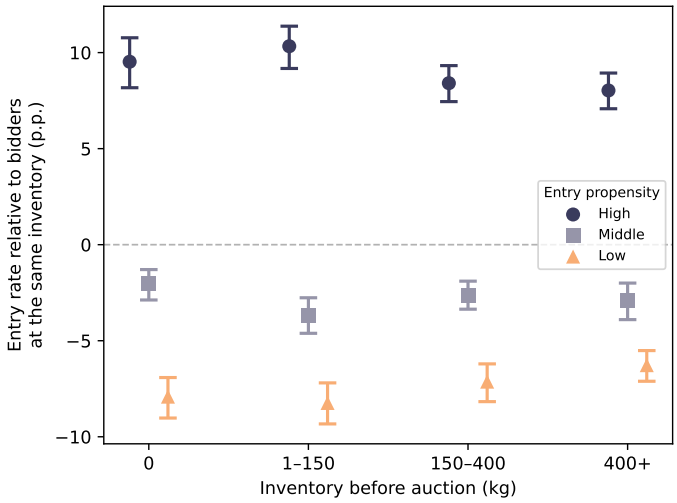
Figure A1 reports the persistence of these pairwise entry and bidding differences across samples A and B. For each variety pair, the "Entry" series reports the correlation across bidder-days between the relative entry difference measured in A and the corresponding difference measured in B. The "Bids" series reports the analogous correlation for relative bids. We retain variety pairs for which both correlations can be constructed in both samples for at least 50 bidder-days.

Relative entry patterns are strongly persistent across the two auction samples, with a median correlation of 0.62 across the supported variety pairs. Relative bidding patterns are also persistent, with a median correlation of 0.53. In an additional cross-sample comparison, the median correlation between relative entry and relative bids is only 0.05. Thus, bidder-days differ systematically in which varieties they favor in both entry and bidding, but these two dimensions of horizontal heterogeneity are only weakly related.

Fact A2 *Variety-specific differences in entry are present at zero inventory and persist as inventory accumulates.*

We examine whether the variety-specific entry patterns documented above emerge only as

Figure A2 Variety-specific entry heterogeneity over inventory



Notes: Figure reports average relative entry across inventory bins for bidder-day-variety observations classified into low, middle, and high variety-specific entry groups. Entry groups are constructed at the bidder-day-variety level using sample A. Relative entry is residualized on bidder-day, inventory-bin, and time-block fixed effects, averaged within bidder-day-variety, and divided into terciles separately within each variety. The plotted outcomes use sample B and are demeaned within inventory-bin-variety cells before pooling varieties. Points report group means, vertical bars report 95% bidder-bootstrap confidence intervals, and the procedure is repeated with samples A and B reversed before pooling both directions.
(Return to the descriptives or model sections in the main text.)

bidders accumulate inventory. We use the same auction samples A and B and use one sample to classify bidder-day-variety observations and the other to measure their subsequent entry behavior.

Using sample A, we residualize relative entry on bidder-day, inventory-bin, and time-block fixed effects. For each bidder-day-variety, we average these residuals across auctions of that variety and, separately within each variety, divide observations into Low, Middle, and High terciles. Because bidder-day fixed effects absorb each bidder-day's overall entry rate, this classification captures which varieties a bidder enters unusually often rather than how active the bidder is overall. The same bidder-day can therefore be classified as High for one variety and Low for another.

Using sample B, we compute relative entry for each group at zero inventory, 1–150 kilograms, 150–400 kilograms, and more than 400 kilograms. Within each inventory category and variety, we demean relative entry before pooling varieties. The plotted differences therefore remove both average differences in entry across varieties and the common relationship between entry and inventory. We repeat the procedure with A and B reversed and pool the two directions.

Figure A2 shows that the separation between groups is already present at zero inventory and remains similar throughout the inventory distribution. At zero inventory, the High group enters about 18 percentage points more often than the Low group. The corresponding split-sample correlation in variety-specific entry is 0.47 at zero inventory and 0.52 at positive inventory.

A.3 Dynamics

Tables A2 reports the complete coefficient estimates underlying Figure 3. The dependent variable is the bidder's final homogenized per-kilogram bid on each entered lot. The time variable is divided into ten equal-width bins of the fraction of auctions remaining, while inventory is divided into ten approximately equal-sized groups based on the empirical distribution of kilograms previously won by the bidder that day.

A.4 Seasonal patterns

We construct daily supply compositions using the universe of lots observed across auction houses. This makes use of data on auction houses other than the one we focus on. As described in Section 2, that includes information on the universe of transacted lots, therefore allowing us to construct supply vectors on a larger sample of auction house-day combinations. This augmented sample adds 58 auction houses and 11,280 house-day combinations.

For auction house h, day d, and variety j, let Q_{jhd} denote the total quantity in kilograms supplied in lots belonging to variety j and define

$$s_{jhd} = \frac{Q_{jhd}}{\sum_{j'} Q_{j'hd}}.$$

Figure A3 illustrates the persistence of supply composition at the auction house in our main samples. For each pair of trading days d and d' in the 2017 season, we compute the cosine distance between their supply-share vectors,

$$1 - \frac{\boldsymbol{s}_{hd}' \boldsymbol{s}_{hd'}}{\|\boldsymbol{s}_{hd}\|_2 \|\boldsymbol{s}_{hd'}\|_2}.$$

Nearby trading days tend to have similar supply compositions, reflecting the gradual availability and disappearance of different peach varieties over the season.

We summarize variation in daily supply composition by classifying house-days into seven supply regimes. We use seven clusters, which provides a parsimonious discretization of the pronounced changes in supply composition over the season while retaining economically distinct variety mixes. One of our variety labels, *General*, collects lots for which the original description does not permit a more informative classification. We exclude this category when construct-

Table A2 Bid dynamics

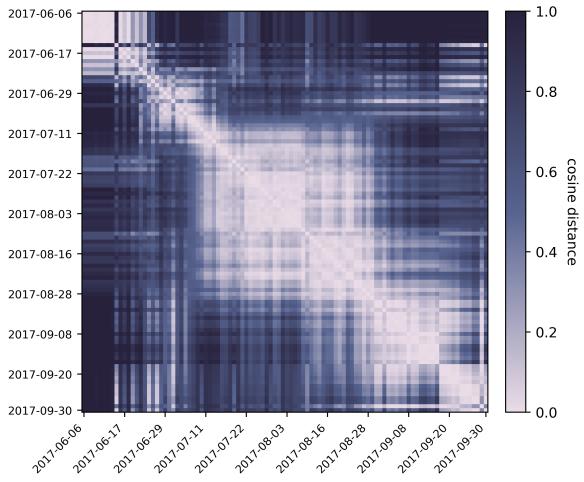
Dynamic effects					
Auctions remaining	Coefficient		Inventory	Coefficient	
80–90%	84.64	(11.13)	Decile 2	-86.90	(16.28)
70–80%	137.66	(13.18)	Decile 3	-153.26	(15.44)
60–70%	187.14	(14.52)	Decile 4	-163.52	(15.43)
50–60%	182.61	(15.22)	Decile 5	-164.34	(17.17)
40–50%	174.99	(16.18)	Decile 6	-186.71	(17.59)
30–40%	134.95	(16.63)	Decile 7	-224.73	(18.57)
20–30%	141.00	(18.96)	Decile 8	-240.85	(18.98)
10–20%	169.08	(18.71)	Decile 9	-237.66	(20.86)
0–10%	190.40	(19.45)	Decile 10	-257.91	(24.34)
Lot observables					
Variety	Coefficient		Other observables	Coefficient	
Nectarine	-102.56	(18.45)	Grade: Off-grade	97.58	(91.62)
Niche	-0.40	(12.45)	Grade: Standard	-509.52	(110.92)
Soft White	-42.13	(10.59)	Grade: High	134.55	(91.52)
Yellow	203.00	(13.47)	Grade: No grade	-874.59	(93.48)
			Grade: Special	78.81	(91.02)
			Packaging: Box	-187.04	(147.42)
			Fruit-size: bin 2	392.73	(8.73)
			Fruit-size: bin 3	918.63	(11.12)
			Fruit-size: bin 4	1556.03	(12.76)
			Lot size: above median	37.47	(5.74)
Observations					96,889
Bidder-day fixed effects					Yes

Notes: Table reports the coefficients from regression (1). Standard errors clustered at the bidder-day level are reported in parentheses. The omitted categories are 90–100% of auctions remaining, the first inventory decile, Hard White, the reference categories for quality grade and packaging, fruit-size bin 1, and the below-median bin of the lot-size distribution. All coefficients are in KRW/kg.
(Return to the estimation section in the main text.)

ing the vectors used for clustering. Denote the share of supply belonging to the remaining categories as $m_{hd} = 1 - s_{\text{General},hd}$ and rescale non-*General* varieties' shares as $\tilde{s}_{jhd} = s_{jhd} / m_{hd}$.

We apply Euclidean k-means to the renormalized supply-share vectors of the retained auction houses. Before estimating the regime centroids, we exclude auction houses whose observed supply composition contains essentially no variation across informative variety categories. For

Figure A3 Similarity of supply across days



Notes: Figure reports the cosine distance between the vectors of variety shares of any pair of days in the 2017 season. The share of each variety is computed as the share of the total quantity (in kilograms) of peaches of that variety among the supply available on a given day.
(Return to the descriptives section in the main text.)

each auction-house day, define the effective number of supplied varieties by

$$N_{hd}^{\text{eff}} = \exp\left(-\sum_{j} s_{jhd} \log s_{jhd} \right).$$

This measure accounts for varieties' relative supply shares: varieties representing only a negligible fraction of daily supply contribute little, whereas a more balanced composition yields a larger effective count. For example, an equal split between two varieties gives $N_{hd}^{\text{eff}} = 2$, while a highly concentrated two-variety composition yields a value only slightly above one. We retain auction houses for which the tenth percentile of N_{hd}^{eff} across trading days exceeds one. This screen effectively exclude house-days combinations where *General* acts as a catch-all label that covers most of the supply. Denoting the set of retained house-days as $\mathcal{X}$ and indexing regimes by r, the estimated centroids satisfy

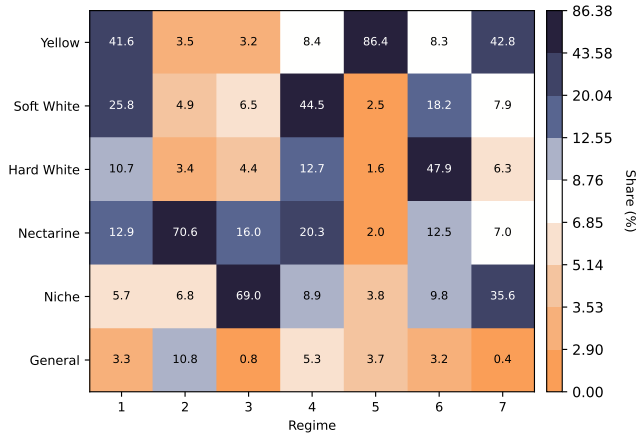
$$(\hat{\boldsymbol{\mu}}_r)_{r=1}^{7} \in \arg\min_{(\boldsymbol{\mu}_r)_{r=1}^{7}} \sum_{hd \in \mathcal{X}} \min_{r \in \{1,\ldots,7\}} \|\tilde{\boldsymbol{s}}_{hd} - \boldsymbol{\mu}_r\|_2^2.$$

The estimated centroids are then used to classify every auction-house day in the data, including days from auction houses not used to estimate the centroids. For each house-day, we use renormalized shares $\tilde{\boldsymbol{s}}_{hd}$, and assign the resulting vector to the closest centroid. We leave a day unclassified if more than 90% of its supply is classified as *General*. All 192 trading days at the auction house featuring in our main sample receive a regime assignment.

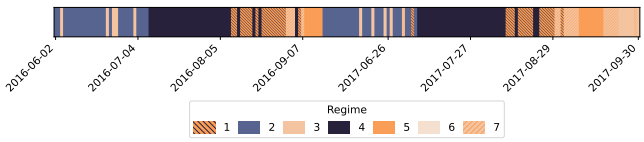
Panel A4a reports average daily variety shares within each regime. The averages use all classified auction-house days. The *General* category is included when describing the resulting regimes even though it is excluded from the clustering variables. For each regime, we average these original daily variety shares over all classified auction-house days assigned to that regime. The *General* category therefore appears in the figure even though it is excluded from the vectors used to assign regimes. Unclassified days are omitted from these averages.

Figure A4 Supply regimes

(a) Variety shares by regime



(b) Regimes timeline



Notes: Figure reports properties of the supply regimes classified by the k-means algorithm. Panel (a) reports the distribution of varieties across regimes. Panel (b) reports the classification of days in the 2016 and 2017 seasons for the auction house in our main dataset.
(Return to the descriptives section in the main text.)

B Theory

B.1 Modeling assumptions

Our equilibrium is reminiscent of oblivious equilibrium (Weintraub et al., 2008), but its interpretation is slightly different. Oblivious equilibrium is typically invoked in infinitely-repeated games where a Markovian structure leads to a stationary distribution of states and actions. Obliviousness maintains that players only keep track of a subset of the relevant state variables, and is particularly sensible in settings with large player sets.

The finite-horizon nature of the game we model makes the interaction inherently non-stationary, therefore preventing us from applying the exact oblivious equilibrium's motivation and interpretation. The relevant repetition in our setting occurs across days, as bidders participate in many auctions each day and return to the same market over months and years. We interpret bidders as learning the distribution of competition associated with relatively coarse states from this repeated experience, while not retaining or conditioning on the complete history of every previous auction day. In particular, bidders can learn how strong competing bids tend to be for a given position in the daily sequence, supply environment, and set of lot characteristics without remembering the identities, inventories, bids, and preceding outcomes associated with every occurrence of that state. The consistency condition requires these coarse beliefs to be correct in the long run.

This motivation is not extraneous to the theory literature. Self-confirming equilibrium concepts are often interpreted as steady states of learning processes, in which repeated experience no longer contradicts the beliefs on which players base their decisions. In this sense, our solution concept has a self-confirming interpretation in the spirit of Battigalli (1987) and Fudenberg and Levine (1993): bidders need not hold or update a complete model of opponents' strategies at every possible history, but the payoff-relevant beliefs they retain from repeated market experience are confirmed by the outcomes observed within each coarse state.

B.2 Proof of Proposition 1

Write $s^0 \sim s^1$ if s^0 and s^1 are translation-equivalent, and denote the associated shift by $\delta(s^0, s^1) := \boldsymbol{\gamma}_m'(\boldsymbol{x}^1 - \boldsymbol{x}^0) + \xi^1 - \xi^0$. The relation $\sim$ is reflexive and symmetric, with $\delta(s, s) = 0$ and $\delta(s^1, s^0) = -\delta(s^0, s^1)$. If $s^0 \sim s^1$ and $s^1 \sim s^2$, then for every type m

$$\boldsymbol{\gamma}_m'(\boldsymbol{x}^2 - \boldsymbol{x}^0) + \xi^2 - \xi^0 = \delta(s^0, s^1) + \delta(s^1, s^2),$$

and the distributions of taste shocks and participation costs at s^0, s^1, and s^2 coincide. This implies that $s^0 \sim s^2$, proving that $\sim$ is an equivalence relation.

For each equivalence class of $\sim$, fix a reference state s° with $\xi = 0$. Since $\boldsymbol{x}$ takes finitely many values, $s \mapsto s^\circ$ is measurable. Let $d(s) = \delta(s^\circ, s)$. If $s^0 \sim s^1$ with shift δ, they have the same reference, and

$$d(s^1) = d(s^0) + \delta. \tag{9}$$

Reference game Denote a generic game as defined in Section 4 as Γ. Let Γ° be constructed from Γ by assigning lot ℓ characteristics $(\boldsymbol{x}_\ell^\circ, q_\ell, 0)$ whenever the original lot has characteristics $(\boldsymbol{x}_\ell, q_\ell, \xi_\ell)$, with all other primitives unchanged. The replacement is applied lot by lot and depends only on the lot's characteristics and public information, so Γ° satisfies the maintained

distributional assumptions (i)–(iv) from Section 4. Since Section 4 maintains that a symmetric equilibrium exists, we conclude that a symmetric equilibrium $(\eta^\circ, \beta^\circ, \psi^\circ)$ also exists for Γ°. Define

$$\beta(v, s) \coloneqq \beta^\circ(v - d(s), s^\circ) + d(s), \tag{10}$$

$$\eta(\bar{v}, c, s) \coloneqq \eta^\circ(\bar{v} - d(s), c, s^\circ), \tag{11}$$

$$\psi(b \mid s) \coloneqq \psi^\circ(b - d(s) \mid s^\circ). \tag{12}$$

We show that (η, β, ψ) is an equilibrium of Γ.

Optimality Given ψ, a bidder's problem at s is the problem at s° with valuation and bid both shifted by $d(s)$. Bidding and entry in (10)–(11) are then optimal because their reference counterparts are. For every v and every $b \in \mathbb{R}$, (12) gives

$$(v - b)\,\psi(b \mid s) = \big(v - d(s) - (b - d(s))\big)\,\psi^\circ(b - d(s) \mid s^\circ). \tag{13}$$

Since $b \mapsto b - d(s)$ is a bijection of $\mathbb{R}$, $\beta(\,\cdot\,, s)$ in (10) maximizes the left side because $\beta^\circ(\,\cdot\,, s^\circ)$ maximizes the right side. Taste shocks have the same distribution at s and s°. Evaluating (13) at $v = \bar{v} + \varepsilon$ and the optimal bids, multiplying by q, and integrating over ε,

$$\Pi(\bar{v}, s) = \Pi^\circ(\bar{v} - d(s), s^\circ). \tag{14}$$

Participation costs have the same distribution at s and s°, so η satisfies (4) because η° does. Thus η and β are optimal given ψ.

Continuation values Consistency concerns the play that (η, β) generates, and the pseudo-valuations in that play contain $\Delta\mathcal{V}_{i\ell}$. We therefore need the continuation values of the two games to coincide, which requires a backward induction. Let $\mathcal{V}^\circ$ be the value function of Γ°. We show by backward induction on ℓ that

$$\forall\,(s, c, w), \quad \mathcal{V}_{i\ell}(s, c, w) = \mathcal{V}^\circ_{i\ell}(s^\circ, c, w). \tag{15}$$

At $\ell = L + 1$ both sides are zero. Suppose (15) holds at $\ell + 1$. Since the next lot's characteristics in Γ map to those in Γ°, we have

$$\mathbb{E}_{c', s' \mid s, m_i}[\mathcal{V}_{i\,\ell+1}(s', c', w')] = \mathbb{E}_{c', s' \mid s^\circ, m_i}[\mathcal{V}^\circ_{i\,\ell+1}(s', c', w')]$$

for $w' \in \{w,\, w + q\}$. This implies that $\Delta\mathcal{V}_{i\ell}(w, s) = \Delta\mathcal{V}^\circ_{i\ell}(w, s^\circ)$ and

$$\bar{v}_{i\ell} = \bar{v}^\circ_{i\ell} + d(s). \tag{16}$$

By (2), the value at ℓ is the maximum over entry of the optimized current payoff plus expected

continuation values. The optimized current payoff coincides across the two games by (13) and (16), the expected continuation values by the display above, and the participation cost is the same draw. This proves (15) at ℓ.

Consistency Consistency requires $\psi(\cdot \mid s)$ to be the distribution of the highest opposing bid that (η, β) generates at s. That distribution depends on the types and inventories bidders carry into auction ℓ, hence on play in auctions $1, \dots, \ell-1$. The argument has two parts. First, we show that, for all possible realizations of primitives, playing (η, β) in Γ and $(\eta^\circ, \beta^\circ)$ in Γ° generate the same inventories and bids that differ by $d(s_\ell)$ at every auction. This is an induction on ℓ. Second, we show that this implies identity of the conditional distributions required by consistency.

Fix a realization of the primitives: bidder types m_i, lot characteristics $(\boldsymbol{x}_\ell, q_\ell, \xi_\ell)$, participation costs $c_{i\ell}$, and taste shocks $\varepsilon_{i\ell}$. The realized $c_{i\ell}$ and $\varepsilon_{i\ell}$ are admissible in Γ° because their distributions at s_ℓ and at s°_ℓ coincide. Denote the inventories, entry decisions, bids, auctioneer's bid, and winner generated in Γ° by $w^\circ_{i\ell}$, $e^\circ_{i\ell}$, $b^\circ_{i\ell}$, r°_ℓ, and $i^{\circ *}_\ell$, and those generated in Γ by the same symbols without the superscript. Let $d_\ell := d(s_\ell)$ and $b^{\max}_{-i\ell} := \max_{j \in E_\ell \setminus \{i\}} b_{j\ell}$, the highest bid opposing participant i. We show by induction on ℓ that inventories coincide:

$$\forall\, i, \forall\, \ell, \quad w_{i\ell} = w^\circ_{i\ell}. \tag{17}$$

At $\ell = 1$ all inventories are zero. Suppose (17) holds at ℓ. Then (16) gives $\bar{v}_{i\ell} = \bar{v}^\circ_{i\ell} + d_\ell$, and (11)–(10) imply

$$e_{i\ell} = \eta(\bar{v}_{i\ell}, c_{i\ell}, s_\ell) = \eta^\circ(\bar{v}^\circ_{i\ell}, c_{i\ell}, s^\circ_\ell) = e^\circ_{i\ell}, \tag{18}$$

$$b_{i\ell} = \beta(\bar{v}_{i\ell} + \varepsilon_{i\ell}, s_\ell) = \beta^\circ(\bar{v}^\circ_{i\ell} + \varepsilon_{i\ell}, s^\circ_\ell) + d_\ell = b^\circ_{i\ell} + d_\ell. \tag{19}$$

The realization of the primitives does not pin down r_ℓ, which Section 4 characterizes only through its distribution conditional on s_ℓ. To pass between conditioning on s_ℓ in Γ and on s°_ℓ in Γ° we use the following identity. Let y° be any variable determined by $(m_j, w^\circ_{j\ell})_j$, s°_ℓ, $(c_{j\ell}, \varepsilon_{j\ell})_j$, and r°_ℓ. By (iv), $(m_j, w^\circ_{j\ell})_j$ is independent of lot ℓ's characteristics. By translation-equivalence, $(c_{j\ell}, \varepsilon_{j\ell})_j$ has the same distribution given $s_\ell = s$ as given $s^\circ_\ell = s^\circ$, and r°_ℓ depends on the lot only through s°_ℓ. Hence

$$\forall\, A \text{ measurable}, \quad \Pr(y^\circ \in A \mid s_\ell = s) = \Pr(y^\circ \in A \mid s^\circ_\ell = s^\circ). \tag{20}$$

Set $r_\ell := r^\circ_\ell + d_\ell$. Consider the highest private bid opposing i. By (18)–(19), in Γ it equals its value in Γ° plus d_ℓ. By definition of the auctioneer's bid in Γ°, conditional on $s^\circ_\ell = s^\circ$ and on at least two private entrants, it has the same distribution as r°_ℓ By (20), this remains true conditional on $s_\ell = s$ and on at least two private entrants. Adding d_ℓ to both sides, in Γ it has the same distribution as r_ℓ conditional on $s_\ell = s$ and on at least two private entrants. This is

the distribution Section 4 requires of r_ℓ. Every bid facing a participant therefore shifts by d_ℓ:

$$\forall\, i \in E_\ell, \quad b^{\max}_{-i\ell} = b^{\max\circ}_{-i\ell} + d_\ell. \tag{21}$$

By (19) and (21), $b_{i\ell} > b^{\max}_{-i\ell}$ if and only if $b^\circ_{i\ell} > b^{\max\circ}_{-i\ell}$, so $i^*_\ell = i^{\circ *}_\ell$ almost surely. This implies

$$w_{i\,\ell+1} = w_{i\ell} + q_\ell\, \mathbf{1}(i = i^*_\ell) = w^\circ_{i\ell} + q_\ell\, \mathbf{1}(i = i^{\circ *}_\ell) = w^\circ_{i\,\ell+1},$$

which is (17) at $\ell + 1$.

This completes the first part. For the second, fix ℓ, a state s at position ℓ, a participant i, and $b \in \mathbb{R}$ and note that

$$\begin{aligned}
\Pr(b^{\max}_{-i\ell} \leqslant b \mid s_\ell = s) &= \Pr(b^{\max\circ}_{-i\ell} \leqslant b - d(s) \mid s_\ell = s) \\
&= \Pr(b^{\max\circ}_{-i\ell} \leqslant b - d(s) \mid s^\circ_\ell = s^\circ) \\
&= \psi^\circ(b - d(s) \mid s^\circ) = \psi(b \mid s).
\end{aligned}$$

The first equality follows from (21), the second from (20) with $y^\circ = b^{\max\circ}_{-i\ell}$, the third is implied by consistency of ψ° in Γ°, and the fourth is the definition (12). This shows that ψ is consistent with (η, β), and (η, β, ψ) is a symmetric equilibrium of Γ.

Translation Let $s^0 \sim s^1$ with shift δ. They have the same reference, so (10) and (9) give

$$\forall\, v \in \mathbb{R}, \quad \beta(v + \delta, s^1) = \beta^\circ(v + \delta - d(s^1), s^\circ) + d(s^1) = \beta^\circ(v - d(s^0), s^\circ) + d(s^0) + \delta = \beta(v, s^0) + \delta.$$

This establishes the desired result. ∎

B.3 Equilibrium entry and consistency

For a bidder with mean pseudo-valuation $\bar v$ in state s, optimal entry in (4) implies $e = \mathbf{1}(\Pi(\bar v, s) \geq c)$. Since participation costs are independently distributed according to F_c, the probability of entry conditional on $(\bar v, s)$ is

$$p(\bar v, s) = F_{c\,|\,s}(\Pi(\bar v, s)).$$

Conditional on state s_ℓ and the vector of mean pseudo-valuations $\bar{\boldsymbol{v}}_\ell$, independence of participation costs implies that the probability of entrant set A is

$$\Pr(E_\ell = A \mid \bar{\boldsymbol{v}}_\ell, s_\ell) = \prod_{j \in A} F_{c\,|\,s}\big(\Pi(\bar v_{j\ell}, s_\ell)\big) \prod_{j \notin A} \Big[1 - F_{c\,|\,s}\big(\Pi(\bar v_{j\ell}, s_\ell)\big)\Big].$$

C Estimation details

C.1 Estimation of bid distributions

Since the estimation of bid distribution is done state by state, fix s and drop it from notation. We retain each bidder's final bid in an auction and discard auctions with fewer than two participating bidders. For every bidder in the remaining auctions, we compute the highest bid submitted by another participant. We weight these observations by the inverse of the number of participating bidders, so that every auction contributes equal total weight to the empirical distribution.

We approximate the distribution of highest opposing bids on the interval $[0, b^{\max}]$, where $b^{\max}$ is the largest opposing bid observed. For a polynomial of degree d, the Bernstein basis is defined for $t \in [0,1]$ as

$$B_{k,d}(t) = \binom{d}{k} t^k (1-t)^{d-k}.$$

We approximate the state-specific bid distribution as

$$\hat{\psi}(b) = \sum_{k=0}^{d} a_k B_{k,d}\left(\frac{b}{b^{\max}}\right).$$

Its density is available analytically:

$$\hat{\psi}'(b) = \frac{d}{b^{\max}} \sum_{k=0}^{d-1} (a_{k+1} - a_k) B_{k,d-1}\left(\frac{b}{b^{\max}}\right). \tag{22}$$

The Bernstein representation is convenient because economically relevant shape restrictions can be imposed directly on the fitted distribution and its derivatives. We fix $a_0 = 0$ and $a_d = 1$ and restrict the coefficients to satisfy $0 = a_0 \leqslant a_1 \leqslant \cdots \leqslant a_d = 1$, which guarantees that $\hat{\psi}$ is a valid CDF.

Since bid inversion can be sensitive to noise in the estimated density, we impose a small roughness penalty to limit local oscillations in $\hat{\psi}'$. Equation (22) clarifies that adjacent coefficient differences $a_{k+1} - a_k$ determine the shape of the density, so we use a roughness penalty of $10^{-3} \sum_{k=1}^{d-1} (a_{k+1} - 2a_k + a_{k-1})^2$.

Finally, we restrict the fitted distribution so that the inversion implied by the bidding first-order condition is weakly increasing. Since the derivative of inverted bids is

$$\frac{\partial}{\partial b}\left[b + \frac{\hat{\psi}(b)}{\hat{\psi}'(b)}\right] = \frac{2\hat{\psi}'(b)^2 - \hat{\psi}(b)\hat{\psi}''(b)}{\hat{\psi}'(b)^2},$$

we impose

$$2\hat{\psi}'^2 - \hat{\psi}\hat{\psi}'' \geqslant 0.$$

This rules out fitted bid distributions for which a higher observed bid would imply a lower pseudo-valuation.

We choose the polynomial degree by cross-validation over $d = 2, \ldots, 30$. For a given degree, we estimate the Bernstein coefficients by minimizing the mean squared distance between the fitted and the empirical distributions, plus the roughness penalty and subject to the restrictions described above. To select the degree, we randomly partition the distinct empirical distribution evaluation points into K subsets. For each candidate degree, we estimate the coefficients using all but one subset and compute the mean squared difference between the fitted and empirical distribution values on the omitted subset. We repeat this procedure until each subset has been omitted once and average the resulting out-of-sample errors. We select the degree with the smallest average out-of-sample error and re-estimate its coefficients using all evaluation points of the empirical distribution. Depending on the number of distinct opposing bids observed in the state, we use between two and five subsets.

C.2 Implementation of the EM algorithm

State transitions The backward recursion requires a transition matrix over the discretized states s defined in Section 5.2. Recall that each state contains three groups of variables: public information, consisting of the supply regime and indicators for above-median attending bidders and aggregate supply; current-lot characteristics, consisting of variety, fruit-size bin, and an indicator for above-median lot size; and within-day timing, represented by the quartile of remaining auctions. The public-information component is fixed throughout a day. The timing component evolves deterministically with the position of the next auction in the daily sequence. Conditional on public information, the next lot's variety and fruit-size bin are drawn independently from their empirical distribution. Lot quantity is drawn separately from its empirical distribution in the estimation sample, and the corresponding lot-size indicator is determined by this draw. The state transition matrix is computed in the natural way, by taking products of these distributions over state components just described.

Expected profits For each state s, we construct a grid of bids and use the estimated $\hat{\psi}$ and $\hat{\psi}'$ to compute the corresponding pseudo-valuations

$$v(b, s) = b + \frac{\hat{\psi}(b \mid s)}{\hat{\psi}'(b \mid s)}. \tag{23}$$

The first-order condition implies that gross profit conditional on the realized pseudo-valuation equal

$$\pi(v,s) := q\frac{\hat{\psi}(\beta(v,s)\mid s)^2}{\hat{\psi}'(\beta(v,s)\mid s)} \tag{24}$$

at all grid points. These two arrays provide a numerical representation of optimal gross profit as a function of the realized pseudo-valuation. For a candidate mean pseudo-valuation $\bar{v}$ and taste-shock scale $\sigma_{\boldsymbol{x}}$, Gauss-Hermite quadrature evaluates this function at the shifted nodes $\bar{v}+\sigma_{\boldsymbol{x}}\varepsilon_u$ and averages the resulting profits using the quadrature weights ω_u:

$$\Pi(\bar{v},s) \approx \sum_u \pi(\bar{v}+\sigma_{\boldsymbol{x}}\varepsilon_u, s)\omega_u. \tag{25}$$

The bid-to-value and realized-profit grids pinned down by (23) and (24) can be computed once and cached. Because $\sigma_{\boldsymbol{x}}$ is estimated, the quadrature approximation (25) is recomputed as the parameters change inside the optimizer.

Backward induction We compute continuation values using the same recursion as in Section 4, evaluated on a discrete inventory grid. We discretize inventory on 137 equally spaced points, corresponding to $0, 20, \ldots, 2720$ kilograms. Lot quantities are rounded and discretized using the same 20-kilogram unit, so each quantity is associated with an integer inventory increment d_q. If the bidder loses or does not enter an auction, their inventory index is unchanged. If the bidder wins, an inventory index k transitions to $k' = \min\{k+d_q, 136\}$. Continuation values conditional on winning versus losing are obtained by comparing the next-auction value at grid points $k+d_q$ and k, without interpolation over inventory.

For type m, inventory w, and current state s_ℓ, the difference in expected continuation values between winning and losing the auction is

$$\Delta\mathcal{V}_{m\ell}(w,s_\ell) = \mathbb{E}_{c',s'|s_\ell}\big[\mathcal{V}_{m\,\ell+1}(s',c',w+q_\ell) - \mathcal{V}_{m\,\ell+1}(s',c',w)\big].$$

Since the game has finite horizon, the terminal value is $\mathcal{V}_{m\,L+1} = 0$. Proceeding backwards, we use the next-auction value function to compute $\Delta\mathcal{V}_{m\ell}(w,s_\ell)$ and hence each type's mean pseudo-valuation

$$\bar{v}_{m\ell}(w,s_\ell) = \boldsymbol{\gamma}_m'\boldsymbol{x}_\ell - \alpha w - \frac{1}{2}\alpha q_\ell + \frac{1}{q_\ell}\Delta\mathcal{V}_{m\ell}(w,s_\ell).$$

Expected profits from entering $\Pi(\bar{v}_{m\ell}(w,s_\ell), s_\ell)$ are obtained via interpolation and quadrature as described above.

The corresponding type-specific entry probability is

$$p_{m\ell}(w, s_\ell) = \frac{\exp\big(\Pi(\bar{v}_{m\ell}(w, s_\ell), s_\ell)/\tau\big)}{\exp\big(\Pi(\bar{v}_{m\ell}(w, s_\ell), s_\ell)/\tau\big) + \exp\big((c_{s_\ell} + a_{j_\ell m})/\tau\big)}.$$

The logit distributional assumption implies that, after integrating over the current entry cost, the value function is updated according to (Hotz and Miller, 1993)

$$\mathbb{E}_{c|s_\ell}\big[\mathcal{V}_{m\ell}(s_\ell, c, w)\big] = \mathbb{E}_{c',s'|s_\ell}\big[\mathcal{V}_{m\,\ell+1}(s', c', w)\big] - \tau \log\big[1 - p_{m\ell}(w, s_\ell)\big].$$

Thus, expected profits determine entry probabilities, and entry probabilities determine the current value used in the next step of the backward recursion. Expectations over future states are evaluated by direct summation using the empirical transition matrix described above.

Holding $\boldsymbol{p}$ fixed, every current-auction inclusive-value term is proportional to τ, the terminal value is zero, and the transition expectation is linear. A simple induction then shows that continuation value differences are homogeneous of degree 1 in τ:

$$\Delta\mathcal{V}_{m\ell}(w, s; \tau, \boldsymbol{p}) = \tau \Delta\mathcal{V}_{m\ell}(w, s; 1, \boldsymbol{p}),$$

where dependence on entry probabilities and the entry cost scale is made explicit in notation. (Dependence on the other parameters is left implicit.) In light of this, we solve the backward recursion imposing $\tau = 1$ for fixed entry probabilities. The mean pseudo-valuations for candidate parameters are then evaluated as

$$\bar{v}_{im\ell}\big(\boldsymbol{\theta}; \boldsymbol{p}^{(n+1)}\big) = \boldsymbol{\gamma}_m' \boldsymbol{x}_\ell - \alpha w_{i\ell} - \frac{1}{2}\alpha q_\ell + \frac{\tau}{q_\ell}\Delta\mathcal{V}_{m\ell}\big(w_{i\ell}, s_\ell; 1, \boldsymbol{p}^{(n+1)}\big).$$

Candidate values of τ therefore only rescale the unit-scale continuation-value wedge and do not require another backward recursion.

Entry probabilities and participation costs update Given the continuation values and expected profits from the current-parameter backward-induction step, we first update the participation cost vector $\boldsymbol{a}$. Holding common cost locations $\boldsymbol{c}^{(n)}$ and expected profits fixed, we maximize the posterior-weighted entry log-likelihood in Section 5.3 over $\boldsymbol{a} = (a_j)_j$. We solve this concave maximization problem by Newton iteration. The resulting $\boldsymbol{a}^{(n+1)}$ is held fixed throughout the subsequent computation of common cost locations and entry probabilities.

We update each common cost location c_s by solving the entry-frequency equation in Section 5.3:

$$\sum_i \sum_{\ell \in L_s} \sum_{m=1}^{M} \rho_{im}^{(n)} \frac{\exp\big(\Pi(\bar{v}_{im\ell}, s_\ell)/\tau^{(n)}\big)}{\exp\big(\Pi(\bar{v}_{im\ell}, s_\ell)/\tau^{(n)}\big) + \exp\big((c_s + a_{j_\ell m}^{(n+1)})/\tau^{(n)}\big)} = \sum_i \sum_{\ell \in L_s} e_{i\ell}.$$

With expected profits and $\boldsymbol{a}^{(n+1)}$ fixed, the left-hand side is strictly decreasing in c_s. We solve these equations using vectorized bisection.

Let h index iterations of the inner fixed point, and let $\tilde{\boldsymbol{c}}^{(h)}$ denote the cost locations returned by the bisection given entry probabilities $\boldsymbol{p}^{(h)}$. Before updating entry probabilities, we damp the cost update according to $\boldsymbol{c}_{\text{map}}^{(h)} = (1-\delta)\boldsymbol{c}^{(h)} + \delta\tilde{\boldsymbol{c}}^{(h)}$, for some $\delta \in (0,1]$. Using expected profits from the same backward-induction pass, the entry-probability update is

$$p_{im\ell,\text{map}}^{(h)} = \frac{\exp\left(\Pi(\bar{v}_{im\ell}^{(h)}, s_\ell)/\tau^{(n)}\right)}{\exp\left(\Pi(\bar{v}_{im\ell}^{(h)}, s_\ell)/\tau^{(n)}\right) + \exp\left((c_{s_\ell,\text{map}}^{(h)} + a_{j_\ell m}^{(n+1)})/\tau^{(n)}\right)}. \tag{26}$$

The pair $(\boldsymbol{p}_{\text{map}}^{(h)}, \boldsymbol{c}_{\text{map}}^{(h)})$ represents one update of the inner fixed point.

Updating common cost locations changes entry probabilities, which in turn change continuation values and expected profits. We therefore iterate the cost inversion, backward recursion, and entry-probability update to a joint fixed point, holding $\boldsymbol{a}^{(n+1)}$ fixed. We Anderson-accelerate the sequence of common cost updates using the five most recent iterates. Entry probabilities are instead updated directly from (26). Convergence is checked by applying one ordinary fixed-point update to the current candidate and requiring the maximum discrepancy between predicted and observed entry shares to be below 10^{-6} and the maximum resulting change in entry probabilities at observed states to be below 10^{-5}. Across iterations of the estimation algorithm, we warm start this step from the common cost locations and entry probabilities obtained in the preceding iteration.

M-step The inverted pseudo-valuations and their Jacobians with respect to bids are computed once before the EM iterations and held fixed throughout estimation. We retain an entered observation only if $\hat{v}_{i\ell}$ and its derivative with respect to $b_{i\ell}$ are finite, $\hat{\psi}(b_{i\ell} \mid s_\ell) \leqslant 0.9$, and the implied markup is below KRW 5,000 (which is more than double the median winning bid). These restrictions prevent observations in the thin upper tail of the estimated opposing-bid distribution from generating arbitrarily large pseudo-valuations through a very small value of $\hat{\psi}'$.

During the M-step, the posterior type probabilities $\boldsymbol{\rho}^{(n)}$ and entry probabilities $\boldsymbol{p}^{(n+1)}$ are held fixed. As described above, we compute $\Delta\mathcal{V}_{m\ell}(w, s; 1, \boldsymbol{p}^{(n+1)})$ once before optimization, so evaluating a candidate $\boldsymbol{\theta}$ does not require another backward recursion. For each retained observation and candidate type, we construct $\bar{v}_{im\ell}(\boldsymbol{\theta}; \boldsymbol{p}^{(n+1)})$ and evaluate the normal pseudo-valuation log-likelihood

$$\log f_{im\ell}(\boldsymbol{\theta}) = \log\phi\left((\hat{v}_{i\ell} - \bar{v}_{im\ell})/\sigma_{\boldsymbol{x}_\ell}\right) - \log\sigma_{\boldsymbol{x}_\ell} + \log\frac{\partial\hat{v}_{i\ell}}{\partial b_{i\ell}},$$

where the last term accounts for the bid-to-values change of variables. The objective is the sum

of these contributions over entered observations and types, weighted by the current posterior probabilities $\rho_{im}^{(n)}$, as reported in Section 5.3.

We maximize this objective using L-BFGS-B. The likelihood and its gradient are evaluated jointly in JAX, with gradients obtained by automatic differentiation. Candidate mean pseudo-valuations and likelihood contributions are evaluated as array operations over observations and latent types, while the inverted pseudo-valuations, Jacobians, and unit-scale continuation wedges are cached. Positive parameters α, τ, and $\sigma_{\boldsymbol{x}}$ are optimized in logs, while the taste coefficients are optimized directly in the vertical-horizontal parameterization described in Section 4. Each M-step is warm started from the parameter estimates obtained in the preceding EM iteration.

E-step The products defining $\mathcal{L}_{im}^{(n+1)}$ in Section 5.3 can underflow because a bidder-day contains many auctions. We therefore compute $\log \mathcal{L}_{im}^{(n+1)}$ by summing the lot-level log contributions. Posterior type probabilities are then computed from $\log \lambda_m^{(n)} + \log \mathcal{L}_{im}^{(n+1)}$ using the log-sum-exp identity. This calculation is numerically equivalent to the expressions in Section 5.3 but avoids forming products of many small probabilities.

Computational details The three most expensive operations in estimation are repeated evaluation of expected profits over the dynamic state space, backward induction over auctions and inventories, and repeated evaluation of the M-step objective and its gradient. We organize each calculation to exploit array parallelism and GPU computation.

First, evaluation of expected profits requires interpolation of the state-specific profit function at Gauss-Hermite nodes for every combination of state, inventory, and latent type. We compute the bid-to-value and realized-profit curves on a fixed grid of 512 points once and cache them. For each evaluation of expected profits, we form all shifted pseudo-valuations $\bar{v} + \sigma_{\boldsymbol{x}} \varepsilon_u$ simultaneously and evaluate the corresponding state-specific profit curves in parallel. Because each profit curve is evaluated on the same 512-point grid, locating the two grid points surrounding each query requires exactly nine comparisons. We perform these searches and the resulting linear interpolations simultaneously for all states, inventories, types, and quadrature nodes, allowing the GPU to evaluate a very large number of independent interpolation problems in parallel. In our data there are 716 discretized state. On an AMD Ryzen AI 9 HX 370 CPU, evaluating these interpolations across the state space takes approximately one minute, compared with less than one second on an NVIDIA RTX 5000 Ada 32 GB GPU.

Second, the backward recursion updates continuation values over states, inventories, latent types, and possible lot quantities. At each auction position, quantity nodes are evaluated in parallel and expectations over future states are computed by array operations using the transition probabilities described above. The auction sequence itself is traversed using a compiled scan, so only the time dimension remains sequential. The unit-scale continuation-value wedges

described above are cached once entry probabilities have converged and reused throughout the subsequent M-step.

Third, the M-step repeatedly evaluates the pseudo-valuation likelihood and its gradient. The inverted pseudo-valuations, their Jacobians, and the unit-scale continuation-value wedges are fixed before optimization, so candidate likelihoods require only construction of mean pseudo-valuations and evaluation of the normal densities over observations and latent types. We evaluate these likelihood contributions as array operations and obtain the objective and its gradient jointly by automatic differentiation. This avoids both numerical differentiation and repeated solution of the dynamic problem inside L-BFGS-B.

Several smaller computations are organized similarly. The participation cost deviations $\boldsymbol{a}$ are updated once per EM iteration by maximizing the posterior-weighted entry log-likelihood using Newton iteration. Holding these deviations fixed, the common cost equations are solved simultaneously using 96 vectorized bisection steps, and the resulting entry probabilities are evaluated over the full state-inventory-type grid in parallel. Across EM iterations, the participation costs and entry probabilities are warm started from the preceding iteration. Anderson acceleration is applied only to the common cost locations, while entry probabilities are generated directly from the model.

D Estimation results

D.1 Additional outputs

Figure D1 reports the average entry probabilities by type and variety. Recall that entry-type 2 has a higher propensity to bid on nectarines or yellow peaches (cf. Figure 4b) while taste-types 1 and 3 prefer yellow peaches to nectarines (cf. Figure 4a).

D.2 Model fit

We assess the model's ability to reproduce the decline in bids over the auction day by comparing bid patterns in the data and the model across combinations of inventory and time. As in Section 3, we refer to the total amount of peaches won by bidder i when lot ℓ is about to be auctioned as *inventory*, and to the share of auctions already occurred when lot ℓ is auctioned as *time*. We divide both variables into deciles.

To remove variation in bids associated with bidder-day heterogeneity and observed lot characteristics, we estimate

$$\text{bid}_{id\ell} = x_\ell'\beta + \alpha_{id} + u_{id\ell},$$

where x_ℓ contains lot variety, fruit size, and binned lot size, and α_{id} denotes bidder-day fixed

Table D1 Estimates

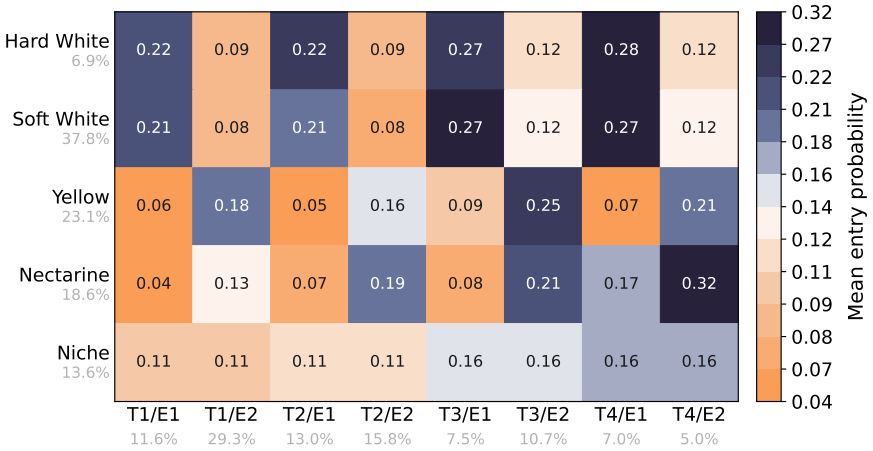
Tastes (KRW/kg)	Taste type			
	1	2	3	4
Hard White	3,172 (182)	3,243 (219)	4,851 (417)	4,922 (450)
Nectarine	2,484 (163)	3,806 (274)	4,163 (346)	5,484 (329)
Niche	3,169 (225)	3,180 (476)	4,847 (435)	4,859 (713)
Soft White	3,078 (197)	3,063 (243)	4,757 (429)	4,742 (465)
Yellow	3,720 (379)	2,972 (189)	5,398 (602)	4,651 (419)

Entry cost deviations (1,000 KRW)	Entry type	
	1	2
Hard White	–	110.2 (9.7)
Nectarine	–	-120.9 (34.0)
Niche	–	1.1 (13.9)
Soft White	–	105.9 (7.9)
Yellow	–	-131.1 (17.8)

Other parameters	
Slope (α)	0.680 (0.142)
Taste shock s.d. (σ_x)	
Fruit size bin 1	1,357 (34)
Fruit size bin 2	1,098 (36)
Fruit size bin 3	1,205 (39)
Fruit size bin 4	1,416 (53)

Notes: Table reports estimates of the parameters governing tastes, entry costs, and demand. Entry type 1 is the reference type, with all entry-cost deviations normalized to zero. Entry type 2 deviations average to zero across varieties, weighted by their empirical shares of auctions in the estimation sample. Taste-shock standard deviations are common across latent types and vary by fruit-size bin. The estimation sample contains 3,304 bidder-days observed over 140 auction days, corresponding to 466,293 bidder-day-lots. Standard errors clustered by auction day are reported in parentheses. Reported standard errors do not yet account for estimation error in bid homogenization and bid distribution estimation.
(Return to the estimation section in the main text.)

Figure D1 Entry probabilities by type and variety



Notes: Figure reports the average entry probabilities by variety and type. Each cell is the mean fitted probability that a bidder of the given type enters an auction of the given variety. It is averaged over the observed entry opportunities for that variety. Columns are the eight latent types: T`p`/E`e` is taste profile `p` combined with entry type `e`. Gray figures under the variety names are each variety's share of kilograms sold in the estimation sample. Gray figures under the types are their estimated population shares. Colors use twelve equal-count bins. *(Return to the estimation section in the main text.)*

effects. Let $\hat{u}_{id\ell}$ denote the resulting bid residual.

For each time decile m and inventory decile n, we compute the average residual

$$\bar{u}_{mn} = \frac{\sum_{id\ell} \mathbf{1}(\text{time}_\ell = m)\mathbf{1}(\text{inventory}_{i\ell} = n)\hat{u}_{id\ell}}{\sum_{id\ell} \mathbf{1}(\text{time}_\ell = m)\mathbf{1}(\text{inventory}_{i\ell} = n)}.$$

We recenter these averages by subtracting their empirical-frequency-weighted mean across inventory-time combinations. The resulting objects measure how much bids for a given combination of inventory and time lie above or below the conditional average bid, in KRW/kg.

Figure D2 reports these bid deviations for the data and for model-implied bids. The model-implied bids are posterior-weighted expected bids evaluated at the states observed in the data. They are constructed as follows. Holding fixed the observed lot characteristics, auction position, and bidder inventory, we compute each type's mean pseudo-valuation $\hat{\bar{v}}_{im\ell}$ implied by the estimated model and integrate the optimal bidding rule over the taste shock to obtain $\hat{b}_{im\ell} = \mathbb{E}_{\varepsilon|\boldsymbol{x}_\ell}[\beta(\hat{\bar{v}}_{im\ell} + \varepsilon, s_\ell)]$. We then average across taste types using bidder-day i's posterior type probabilities from the EM estimator, $\hat{b}_{i\ell} = \sum_{m=1}^{M} \rho_{im}\hat{b}_{im\ell}$.

E Inference

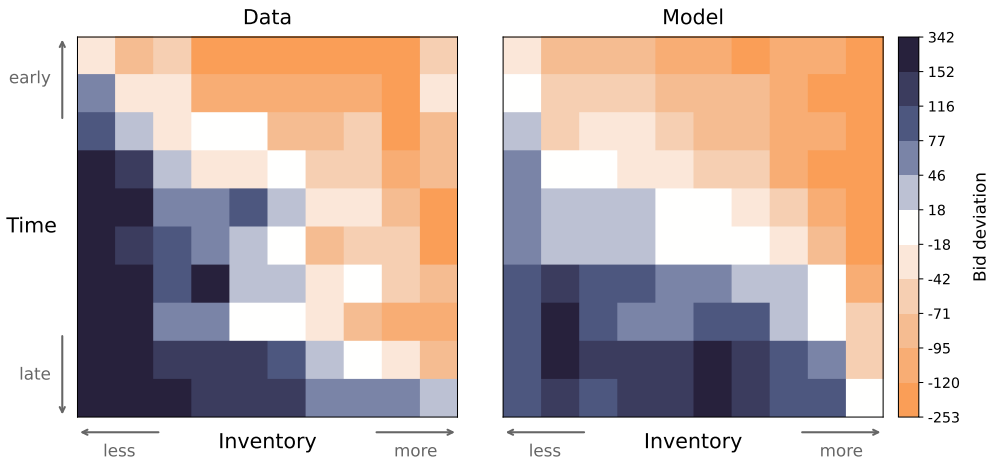
Inference must account for two features of the estimation procedure. First, the parameters estimated in the EM and cost update block are jointly determined with the latent-type proba-

Table D2 Type distribution

	Entry type 1	Entry type 2
Taste type 1	0.116 (0.041)	0.293 (0.064)
Taste type 2	0.130 (0.062)	0.158 (0.075)
Taste type 3	0.075 (0.031)	0.107 (0.034)
Taste type 4	0.070 (0.020)	0.050 (0.019)

Notes: Table reports the estimated joint population distribution of latent taste and entry types. Rows index the four taste-types governing willingness to pay across fruit families, while columns index the two entry-types governing variety-specific participation costs. Each cell reports the estimated probability of the corresponding taste-type and entry-type combination. The estimates use 3,304 bidder-days observed over 140 auction days. Standard errors clustered by auction day are reported in parentheses. Reported standard errors do not yet account for estimation error in bid homogenization and bid distribution estimation.
(Return to the estimation section in the main text.)

Figure D2 Model fit



Notes: Figure reports bid deviations in KRW/kg across combinations of cumulative quantity won and remaining auction time. Bids are regressed on bidder-day fixed effects and the observed lot characteristics used in the bid homogenization. We then compute the average bid residual for each inventory-time combination and recenter them by subtracting the weighted mean residual. The model panel applies the same calculation to model-implied bids using the empirical distribution of observations across inventory-time combinations.
(Return to the estimation section in the main text.)

bilities, state-specific entry-cost locations, and equilibrium entry probabilities, so uncertainty in these objects propagates through continuation values and posterior type probabilities. Second, this block depends on several objects estimated in preceding steps. We first characterize the converged EM and cost update estimator as a system of estimating equations and derive its variance holding the preceding estimates fixed. We then account for estimation of the pre-

Table D3 Entry cost estimates

	Common cost location c_s (KRW)	Entry type 2 deviation a_j (KRW)
Distribution of common cost locations		
Mean	235,610	
25th percentile	167,411	
Median	213,159	
75th percentile	321,828	
Logit scale τ	97,749	
Variety		
Hard White	155,585	110,204
Soft White	160,927	105,948
Yellow	323,133	−131,078
Nectarine	346,021	−120,860
Niche	230,956	1,078
Lot size		
Below median	220,990	
Above median	254,217	

Notes: Table reports estimated participation cost parameters. The common cost location c_s applies to entry type 1, while entry type 2 faces location $c_s + a_j$ for variety j. The deviations a_j are normalized to have zero mean across varieties, weighted by their shares of auctions in the estimation sample. Distributional statistics weight fitted cost locations by the number of observed bidder-lot entry opportunities they represent. The parameter τ is the scale of the logistic participation cost shocks.
(Return to the estimation section in the main text.)

ceding objects using the fast bootstrap of Armstrong et al. (2014), which re-estimates them in each bootstrap sample but avoids rerunning the EM and cost update block to convergence.

E.1 Estimating equations

Let $\mathcal{N}$ denote the set of bidder-days and $N = |\mathcal{N}|$. The parameters are $\boldsymbol{\theta} = ((\boldsymbol{\gamma}_m)_m, \alpha, \tau, (\sigma_x)_x)$, the population probabilities of the bidder types are $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_M)$, the state-specific entry-cost locations are $\boldsymbol{c} = (c_s)_s$, and the variety-specific participation-cost deviations are $\boldsymbol{a} = (a_j)_j$. We use $\lambda_1, \ldots, \lambda_{M-1}$ as free coordinates and recover $\lambda_M = 1 - \sum_{m=1}^{M-1} \lambda_m$. As in Section 4, the

first entry type has zero deviations, while the second satisfies $\sum_j \omega_j a_j = 0$, where ω_j is variety j's share of auctions. We use a_j for $j \neq j_0$ as free coordinates, with $a_{j_0} = -\sum_{j \neq j_0} \omega_j a_j / \omega_{j_0}$. Let $\boldsymbol{\delta}$ collect the continuous objects estimated before the EM and cost update block: the coefficients used to homogenize bids, the Bernstein coefficients defining state-specific bid distributions, and the empirical distributions of future fruit variety, fruit size, and lot quantity. Objects obtained from these estimates (e.g., inverted pseudo-valuations) are treated as functions of $\boldsymbol{\delta}$.

For given $(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{c}, \boldsymbol{a}, \boldsymbol{\delta})$, the entry probabilities are obtained by solving the fixed point described in Section 5.3. Posterior type probabilities are then obtained from the E-step. The EM iterations and the cost update are computational procedures for finding the joint fixed point, while inference depends only on the equations satisfied at convergence.

For bidder-day i and bidder type m, let $\mathcal{L}^b_{im}$ denote the bid log likelihood conditional on type m. The M-step implies

$$\boldsymbol{G}_\theta(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{c}, \boldsymbol{a}, \boldsymbol{\delta}) = \frac{1}{N} \sum_{i \in \mathcal{N}} \sum_{m=1}^{M} \rho_{im} \frac{\partial \mathcal{L}^b_{im}}{\partial \boldsymbol{\theta}} = \mathbf{0}.$$

The update of the population type probabilities implies

$$G_{\lambda_m}(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{c}, \boldsymbol{a}, \boldsymbol{\delta}) = \frac{1}{N} \sum_{i \in \mathcal{N}} (\rho_{im} - \lambda_m) = 0.$$

Only $M - 1$ of these equations are independent because both $\boldsymbol{\rho}_i$ and $\boldsymbol{\lambda}$ sum to one, and inference is computed using the same $M - 1$ free coordinates used in estimation.

Each entry-cost location is chosen to equate posterior-weighted predicted entry with observed entry, giving

$$G_{c_s}(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{c}, \boldsymbol{a}, \boldsymbol{\delta}) = \frac{1}{N} \sum_{i \in \mathcal{N}} \sum_{\ell : s_\ell = s} \left[\sum_{m=1}^{M} \rho_{im} p_{im\ell} - e_{i\ell} \right] = 0.$$

The participation-cost deviations maximize the posterior-weighted entry likelihood subject to their normalization. Let $\mathcal{M}_2$ denote the set of bidder types associated with the second entry type. The first-order conditions for the free deviations are

$$G_{a_j}(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{c}, \boldsymbol{a}, \boldsymbol{\delta}) = \frac{1}{N\tau} \sum_{i \in \mathcal{N}} \sum_{\ell} \left(\mathbf{1}\{j_\ell = j\} - \frac{\omega_j}{\omega_{j_0}} \mathbf{1}\{j_\ell = j_0\} \right) \sum_{m \in \mathcal{M}_2} \rho_{im} (p_{im\ell} - e_{i\ell}) = 0,$$

for every $j \neq j_0$. The second term accounts for the adjustment of the reference variety's deviation required by the normalization.

Stacking the independent equations gives $\boldsymbol{G}(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{c}, \boldsymbol{a}, \boldsymbol{\delta}) = \mathbf{0}$. Conditional on the specification and the preceding estimates, the converged EM and cost update estimator can therefore be treated as an exactly identified Z-estimator.

The derivative relevant for inference is the total derivative of this stacked system rather than the Hessian of the M-step objective. A change in $\boldsymbol{\theta}$, $\boldsymbol{c}$, or $\boldsymbol{a}$ changes equilibrium entry probabilities, continuation values, and posterior type probabilities, which in turn change all blocks of the estimating equations. Writing the entry probability fixed point as $\boldsymbol{p} = \mathcal{P}(\boldsymbol{p}; \boldsymbol{\theta}, \boldsymbol{c}, \boldsymbol{a}, \boldsymbol{\delta})$, the implicit function theorem gives

$$\left[I - \frac{\partial \mathcal{P}}{\partial \boldsymbol{p}'}\right] \frac{\partial \boldsymbol{p}}{\partial(\boldsymbol{\theta}', \boldsymbol{c}', \boldsymbol{a}')} = \frac{\partial \mathcal{P}}{\partial(\boldsymbol{\theta}', \boldsymbol{c}', \boldsymbol{a}')}.$$

We incorporate this response when differentiating the estimating equations.

E.2 Conditional variance

Let $\boldsymbol{\zeta}$ collect the free coordinates of $(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{a}, \boldsymbol{c})$, and let $\boldsymbol{g}_i(\boldsymbol{\zeta}, \boldsymbol{\delta})$ denote bidder-day i's contribution to the stacked estimating equations. The sample equations are

$$\boldsymbol{G}(\boldsymbol{\zeta}, \boldsymbol{\delta}) = \frac{1}{N} \sum_{i \in \mathcal{N}} \boldsymbol{g}_i(\boldsymbol{\zeta}, \boldsymbol{\delta}).$$

Holding the preceding estimates fixed at $\hat{\boldsymbol{\delta}}$, define

$$\hat{A} = \frac{\partial \boldsymbol{G}(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\delta}})}{\partial \boldsymbol{\zeta}'},$$
$$\hat{\Omega} = \frac{1}{N} \sum_{i \in \mathcal{N}} \boldsymbol{g}_i(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\delta}}) \boldsymbol{g}_i(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\delta}})'.$$

The corresponding covariance matrix is

$$\hat{V}_{\text{cond}} = \frac{1}{N} \hat{A}^{-1} \hat{\Omega} \hat{A}^{-T}.$$

The bidder-day is the natural observation unit because taste types are drawn at the bidder-day level and the mixture likelihood is formed at that level. To allow dependence among bidder-days exposed to the same sequence of auctions, we also construct a day-clustered covariance by summing $\boldsymbol{g}_i$ within auction days before forming $\hat{\Omega}$. We additionally report a two-way day-by-bidder covariance to allow for dependence generated by the same bidder appearing on different days.

The state-specific entry-cost locations are nuisance parameters for the results reported in the paper, but they cannot be held fixed because their estimation affects the remaining parameters. Let $\boldsymbol{r}$ collect the free coordinates of $(\boldsymbol{\theta}, \boldsymbol{\lambda}, \boldsymbol{a})$ and partition the Jacobian as

$$\hat{A} = \begin{pmatrix} \hat{A}_{rr} & \hat{A}_{rc} \\ \hat{A}_{cr} & \hat{A}_{cc} \end{pmatrix}.$$

Define

$$\hat{B} = \hat{A}_{rr} - \hat{A}_{rc}\hat{A}_{cc}^{-1}\hat{A}_{cr},$$
$$\boldsymbol{f}_i = \boldsymbol{g}_{r,i} - \hat{A}_{rc}\hat{A}_{cc}^{-1}\boldsymbol{g}_{c,i},$$
$$\hat{\Omega}_r = \frac{1}{N}\sum_{i\in\mathcal{N}} \boldsymbol{f}_i\boldsymbol{f}_i'.$$

The covariance matrix for the reported parameters is

$$\hat{V}_{r,\text{cond}} = \frac{1}{N}\hat{B}^{-1}\hat{\Omega}_r\hat{B}^{-T}.$$

This covariance is exactly the block of the complete joint covariance corresponding to the reported parameters, so profiling the entry-cost locations does not ignore their estimation uncertainty.

The conditional covariance captures sampling uncertainty generated by the EM and cost update block while holding $\hat{\boldsymbol{\delta}}$ fixed. It does not capture sampling variation in bid homogenization, unobserved lot quality, opposing-bid distributions, or the empirical distribution of future lot characteristics. Our preferred inference procedure also accounts for estimation of these objects.

E.3 Fast bootstrap

A conventional bootstrap would resample the data and repeat the complete estimation procedure in every draw, including bid homogenization, estimation of unobserved lot quality, estimation of the opposing-bid distributions, reconstruction of the dynamic problem, and the full EM and cost update procedure. Repeating the EM and cost update procedure hundreds of times is computationally prohibitive in our application. We instead use the fast bootstrap of Armstrong et al. (2014), which repeatedly re-estimates the preceding objects but replaces repetition of the EM and cost update block by one local update around the original estimate.

We resample entire auction days with replacement because bidders within a day face the same sequence of lots and because construction of the opposing-bid distributions requires the joint set of bids observed in each auction. For each bootstrap draw t, we sample with replacement the same number of auction days as in the original sample. Each sampled day contributes all auctions, lots, bids, and bidder-days observed on that day. If a day is sampled more than once, each copy is assigned a distinct day identifier and treated as a separate bootstrap day. The estimating equations remain indexed by bidder-day; day resampling only determines which bidder-days and auctions are resampled together.

Let $\hat{\boldsymbol{\delta}}^{(t)}$ denote the estimates obtained by repeating the preceding estimation blocks on this resampled data. We reconstruct the bid inversion and dynamic-program objects implied

by $\hat{\boldsymbol{\delta}}^{(t)}$ and resolve the entry-probability fixed point at the original values of $\boldsymbol{\theta}$, $\boldsymbol{c}$, and $\boldsymbol{a}$. We then evaluate the stacked estimating equations on the bootstrap sample without rerunning the EM and cost-location iterations.

Let $\boldsymbol{G}^{(t)}$ denote the bootstrap-sample estimating equations evaluated at $(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\delta}}^{(t)})$. The fast local bootstrap draw is

$$\hat{\boldsymbol{\zeta}}^{(t)}_{\text{fast}} = \hat{\boldsymbol{\zeta}} - \hat{A}^{-1}\left[\boldsymbol{G}^{(t)}(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\delta}}^{(t)}) - \boldsymbol{G}(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\delta}})\right].$$

Under the regularity conditions in Armstrong et al. (2014), the distribution of these local draws provides a first-order approximation to the sampling distribution obtained by rerunning the complete estimator in every bootstrap sample.

The same profiling used above is applied before constructing the reported bootstrap draws. Define

$$\begin{aligned}\boldsymbol{d}_r^{(t)} &= \boldsymbol{G}_r^{(t)}(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\delta}}^{(t)}) - \boldsymbol{G}_r(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\delta}}) \\ \boldsymbol{d}_c^{(t)} &= \boldsymbol{G}_c^{(t)}(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\delta}}^{(t)}) - \boldsymbol{G}_c(\hat{\boldsymbol{\zeta}}, \hat{\boldsymbol{\delta}}).\end{aligned}$$

The local bootstrap draw for the reported parameters is

$$\hat{\boldsymbol{r}}^{(t)}_{\text{fast}} = \hat{\boldsymbol{r}} - \hat{B}^{-1}\left[\boldsymbol{d}_r^{(t)} - \hat{A}_{rc}\hat{A}_{cc}^{-1}\boldsymbol{d}_c^{(t)}\right].$$

We compute bootstrap standard errors from the dispersion of $\hat{\boldsymbol{r}}^{(t)}_{\text{fast}}$ across draws and confidence intervals from the corresponding empirical quantiles. For transformed parameters, including the reported vertical and horizontal taste components, the normalized participation-cost deviations, and λ_M, we apply the transformation separately to every bootstrap draw.

The procedure provides a first-order approximation to the sampling distribution of the full estimator. We assess the accuracy of the approximation by comparing the fast-bootstrap draws with complete re-estimation on a small number of bootstrap samples and in a smaller simulated version of the model where repeated full estimation is feasible.

F Counterfactuals

F.1 Computation of the equilibrium

Overview Computing the equilibrium requires making bidders' entry and bidding behavior consistent with the distribution of opponents' bids and with the endogenous evolution of inventories over the auction sequence. Given candidate distributions of inventories and continuation values, we first solve for state-contingent entry and bidding behavior. We then use the resulting behavior to propagate the inventory distribution forward through the sequence

of auctions and recompute continuation values by backward induction. These updates generate new mean pseudo-valuation distributions and hence new entry and bidding behavior. We iterate this procedure until the induced inventory distributions, continuation values, entry probabilities, and bid distributions are mutually consistent.

The forward and backward updates ensure consistency of inventories and continuation values across auctions. Equilibrium additionally requires that state-contingent optimal bids be consistent with the distribution of opposing bids generated in the same state. We exploit the sufficiency of pseudo-valuations for bidding, discussed in Section 4: conditional on the discretized state s, an entrant's optimal bid depends on taste, inventory, exact auction position, and quantity only through the realized pseudo-valuation v. This allows us to solve for a single symmetric bidding strategy $\beta(\cdot, s) : v \mapsto b$ for each state s. Conditional on the current distribution of pseudo-valuations, we compute this strategy using Algorithm 2 of Kastl et al. (2026, henceforth KRR), while jointly imposing consistency of entry probabilities with the profits generated by the resulting strategy.

The computation therefore has two nested levels. An inner fixed point solves entry and symmetric bidding conditional on the current distributions of mean pseudo-valuations. An outer iteration then propagates inventories forward and continuation values backward, producing updated distributions of mean pseudo-valuations and continuation-value wedges. We iterate these forward and backward updates until the dynamic objects stabilize.

Solution procedure The state s is the discretized state defined in Section 5.2. It contains timing and lot-size bins, but not the exact auction position ℓ or quantity q. These variables remain relevant for valuations, participation, and inventory transitions, so we retain them separately when describing those objects. Write $\boldsymbol{x}_s$ for the lot observables associated with state s. For a bidder with taste $\boldsymbol{\gamma}_m$ and inventory w, let $\Delta\mathcal{V}_{m\ell q}(w, s)$ denote the expected continuation-value difference between winning quantity q and not winning it at position ℓ.[24] The corresponding mean pseudo-valuation is

$$\bar{v}_{m\ell q}(w, s) = \boldsymbol{\gamma}_m' \boldsymbol{x}_s - \alpha w - \frac{\alpha q}{2} + \frac{\Delta\mathcal{V}_{m\ell q}(w, s)}{q}. \tag{27}$$

For every combination (ℓ, q, s), let $\mu_{\ell qs}$ denote the distribution of these mean pseudo-valuations among potential bidders before entry. Each $\mu_{\ell qs}$ is a probability distribution on $\mathbb{R}$, represented computationally by the population shares of taste-inventory combinations (m, w) and their associated values in (27). The combination labels are retained throughout the calculation because inventory transitions must distinguish bidders with different tastes and inventories, even when their mean pseudo-valuations coincide. The population shares before an

[24] The additional q index makes explicit a dependence already present in $\Delta\mathcal{V}_{i\ell}(w_{i\ell}, s_\ell)$ in Section 4, where the non-discretized state contains the exact lot quantity.

auction do not depend on its current quantity, but the associated mean pseudo-valuations generally do.

Collect the distributions as $\boldsymbol{\mu} = (\mu_{\ell qs})_{\ell,q,s}$ and the wedges as $\Delta\boldsymbol{\mathcal{V}} = (\Delta\mathcal{V}_{m\ell q}(w,s))_{m,\ell,q,w,s}$. Conditional on these mean pseudo-valuations, current bidding does not depend separately on continuation values. In particular, the sufficiency property in Section 4 implies that entrants with the same realized pseudo-valuation v and state s face the same bidding problem. We therefore compute one common strategy $\beta(\cdot, s)$ for all positions and quantities associated with s, while retaining their separate entry probabilities and population distributions.

Our iterative procedure has two levels: (i) holding $(\boldsymbol{\mu}^{(t)}, \Delta\boldsymbol{\mathcal{V}}^{(t)})$ fixed, we solve jointly for state-contingent bidding and participation; (ii) we then use the resulting behavior to propagate inventories forward and update continuation-value wedges backward, which yields an updates $(\boldsymbol{\mu}^{(t+1)}, \Delta\boldsymbol{\mathcal{V}}^{(t+1)})$.

We use Algorithm 2 of Kastl et al. (2026, henceforth KRR) to compute the common monotone bidding strategy. For each state, we use an ordered bid grid $b^0 < \cdots < b^B$ and a pseudo-valuation grid $v^1 < \cdots < v^K$, augmented by endpoints $v^0 = -\infty$ and $v^{K+1} = +\infty$. Candidate strategies are piecewise constant, with cutoffs selected from the pseudo-valuation grid.[25]

We initialize the outer iteration with $(\boldsymbol{\mu}^{(0)}, \Delta\boldsymbol{\mathcal{V}}^{(0)})$. Given $(\boldsymbol{\mu}^{(t)}, \Delta\boldsymbol{\mathcal{V}}^{(t)})$, one outer update proceeds as follows.

1. *Solve state-contingent entry and bidding*
   Fix $(\boldsymbol{\mu}^{(t)}, \Delta\boldsymbol{\mathcal{V}}^{(t)})$. For each state s, initialize entry probabilities $p^{(t,0)}_{\ell qs}(\bar{v})$ for every (ℓ, q, s). Given entry probabilities $p^{(t,h)}_{\ell qs}(\bar{v})$ at inner iteration h, perform the following operations.

   (a) *Construct the entry-adjusted pseudo-valuation distributions*
   For each (ℓ, q, s), define

   $$A^{(t,h)}_{\ell qs}(v) = \int \left[1 - p^{(t,h)}_{\ell qs}(\bar{v}) + p^{(t,h)}_{\ell qs}(\bar{v}) F_{\varepsilon|\boldsymbol{x}_s}(v - \bar{v})\right] \mathrm{d}\mu^{(t)}_{\ell qs}(\bar{v}). \tag{28}$$

   This is the distribution of the augmented pseudo-valuation $\tilde{v}$, which equals the realized pseudo-valuation of an entrant and equals $-\infty$ under non-entry.

   (b) *Compute state-contingent winning probabilities*
   For pseudo-valuation grid points $v^a < v^b$, the tie-adjusted probability of winning in

[25] This approximation leads to ties happening with positive probability, and this requires us to distinguish the probability of winning after a tie-break from the probability that all opposing bids are weakly below a given bid.

(ℓ, q, s) when the same bid is submitted by bidders with $v \in (v^a, v^b]$ is

$$\omega_{\ell qs}^{(t,h)}(v^a, v^b) = \sum_{k=0}^{N-1} \binom{N-1}{k} A_{\ell qs}^{(t,h)}(v^a)^{N-1-k} \big[A_{\ell qs}^{(t,h)}(v^b) - A_{\ell qs}^{(t,h)}(v^a)\big]^k \frac{1}{k+1}. \quad (29)$$

Aggregation over (ℓ, q) gives

$$\Omega_s^{(t,h)}(v^a, v^b) = \sum_{\ell,q} \Pr(\ell, q \mid s) \omega_{\ell qs}^{(t,h)}(v^a, v^b). \quad (30)$$

(c) *Compute the common symmetric bidding strategy*
Using $\Omega_s^{(t,h)}$, apply Algorithm 2 of Kastl et al. (2026, henceforth KRR) to compute a symmetric monotone strategy $\beta^{(t,h)}(\cdot, s)$ on the bid grid $b^0 < \cdots < b^B$. The algorithm selects pseudo-valuation grid points as thresholds separating the intervals assigned to different bid levels.
Let $a_n^{(t,h)}(s)$ denote the selected index of the upper threshold associated with b^n. The winning probability associated with bid b^n is then

$$\Omega_{s,n}^{(t,h)} = \Omega_s^{(t,h)}\big(v^{a_{n-1}^{(t,h)}(s)}, v^{a_n^{(t,h)}(s)}\big).$$

The corresponding distribution of highest opposing bids satisfies

$$\psi^{(t,h)}(b^n \mid s) = \sum_{\ell,q} \Pr(\ell, q \mid s) A_{\ell qs}^{(t,h)}\big(v^{a_n^{(t,h)}(s)}\big)^{N-1}. \quad (31)$$

(d) *Update entry probabilities*
Given $\beta^{(t,h)}$ and $\Omega_{s,n}^{(t,h)}$, expected profits from entry at (ℓ, q, s) are

$$\Pi_{\ell qs}^{(t,h)}(\bar{v}) = q \sum_{n=0}^{B} \Omega_{s,n}^{(t,h)} \int_{v^{a_{n-1}^{(t,h)}(s)}}^{v^{a_n^{(t,h)}(s)}} (v - b^n)\, \mathrm{d}F_{\varepsilon \mid \boldsymbol{x}_s}(v - \bar{v}). \quad (32)$$

Update entry probabilities according to

$$p_{\ell qs}^{(t,h+1)}(\bar{v}) = F_{c \mid s}\big(\Pi_{\ell qs}^{(t,h)}(\bar{v})\big). \quad (33)$$

If $p_{\ell qs}^{(t,h+1)}(\bar{v})$ differs from $p_{\ell qs}^{(t,h)}(\bar{v})$ beyond numerical tolerance, repeat the steps (a)–(d). Once the entry probabilities converge, denote the resulting objects by $p_{\ell qs}^{(t)}$, $\beta^{(t)}(\cdot, s)$, $\Pi_{\ell qs}^{(t)}$, and $\psi^{(t)}(\cdot \mid s)$.

2. *Update the population distributions*
Starting from zero inventory, forward-simulate outcomes through auctions $\ell = 1, \ldots, L$ using the converged objects $p^{(t)}$ and $\beta^{(t)}$ from Step 1 and the input $\Delta \boldsymbol{\mathcal{V}}^{(t)}$. Let $F_{w \mid \gamma_m, z, \ell}^{(t+1)}$

denote the distribution of inventory among type-m bidders immediately before auction ℓ generated by this forward pass. Note that such distribution only depends on the public component of the state because lot observables and size are iid across auctions. The initial condition is $F^{(t+1)}_{w\,|\,\boldsymbol{\gamma}_m,z,1}(0) = 1$ for every m.

Consider a type-m bidder entering auction ℓ with inventory w, when the realized quantity and state are (q,s). Their mean pseudo-valuation $\bar{v}^{(t)}_{m\ell q}(w,s)$ is computed from (27) using $\Delta\boldsymbol{\mathcal{V}}^{(t)}$. Let $\chi^{(t)}_{m\ell qs}(w)$ denote the unconditional probability that this bidder enters and wins the auction. Using the thresholds selected by $\beta^{(t)}(\,\cdot\,,s)$, we have

$$\begin{aligned}\chi^{(t)}_{m\ell qs}(w) = p^{(t)}_{\ell qs}\big(\bar{v}^{(t)}_{m\ell q}(w,s)\big)\sum_{n=0}^{B}\omega^{(t)}_{\ell qs}\big(v^{a^{(t)}_{n-1}(s)},v^{a^{(t)}_{n}(s)}\big)\\ \times\Big[F_{\varepsilon\,|\,\boldsymbol{x}_s}\big(v^{a^{(t)}_{n}(s)}-\bar{v}^{(t)}_{m\ell q}(w,s)\big)-F_{\varepsilon\,|\,\boldsymbol{x}_s}\big(v^{a^{(t)}_{n-1}(s)}-\bar{v}^{(t)}_{m\ell q}(w,s)\big)\Big].\end{aligned} \tag{34}$$

Equation (34) integrates over the bidder's taste shock, multiplying the probability of selecting each action by the auction-specific probability of winning that action.[26]

Inventory evolves based on wins, so for every type m,

$$\begin{aligned}F^{(t+1)}_{w\,|\,\boldsymbol{\gamma}_m,z,\ell+1}(w') = \sum_{w}F^{(t+1)}_{w\,|\,\boldsymbol{\gamma}_m,z,\ell}(w)\sum_{q,s}\Pr(q,s\mid\ell,z)\Big[\chi^{(t)}_{m\ell qs}(w)\mathbf{1}\{w'=w+q\}\\ +\big(1-\chi^{(t)}_{m\ell qs}(w)\big)\mathbf{1}\{w'=w\}\Big].\end{aligned} \tag{35}$$

Applying (35) sequentially through the day gives $F^{(t+1)}_{w\,|\,\boldsymbol{\gamma}_m,z,\ell}$ for every ℓ and m. The resulting distributions are converted into the next mean pseudo-valuation distributions $\boldsymbol{\mu}^{(t+1)}$ only after $\Delta\boldsymbol{\mathcal{V}}^{(t+1)}$ has been computed in the backward step below.

3. *Update continuation-value wedges*
   Conditional on the converged entry probabilities $p^{(t)}$ and expected profits $\Pi^{(t)}$ from Step 1, apply the backward-induction procedure described in Appendix C. Starting from the terminal condition $\mathcal{V}_{i\,L+1} = 0$, this yields the updated continuation-value wedges $\Delta\boldsymbol{\mathcal{V}}^{(t+1)}$. No entry or bidding problem is re-solved in this step.

4. *Update the mean pseudo-valuation distributions*
   Update the distribution of mean pseudo-valuations as

$$\mu^{(t+1)}_{\ell qs}(\bar{v}) = \sum_{m}\lambda_m\sum_{w}F^{(t+1)}_{w\,|\,\boldsymbol{\gamma}_m,z,\ell}(w)\mathbf{1}\{\bar{v}^{(t+1)}_{m\ell q}(w,s)=\bar{v}\}.$$

[26] Note that (34) uses $\omega^{(t)}_{\ell qs}$ rather than the state-level average $\Omega^{(t)}_{s}$ because the realized allocation is generated by the auction-specific $\omega^{(t)}_{\ell qs}$ despite the bidders optimizing against the aggregate $\Omega^{(t)}_{s}$.

Steps 1–4 define the update $(\boldsymbol{\mu}^{(t)}, \Delta\boldsymbol{\mathcal{V}}^{(t)}) \mapsto (\boldsymbol{\mu}^{(t+1)}, \Delta\boldsymbol{\mathcal{V}}^{(t+1)})$. We iterate this mapping until both objects stabilize. At convergence, entry and bidding are mutually consistent conditional on the dynamic objects and the distributions of mean pseudo-valuations and continuation-value wedges are reproduced by the resulting allocation dynamics and future payoffs, therefore giving an oblivious equilibrium as defined in Section 4.

For computational convenience, we iterate these updates on a coarse pseudo-valuation grid. We then use the converged solution to initialize a finer grid and re-run the iteration. In practice, we find that this two-step performs well: the coarse-grid iteration is reliable and relatively fast, and warm starting the more expensive finer-grid iteration allows the procedure to effectively converge in fewer steps.

F.2 Pooling counterfactuals

Lot construction A pooling scheme $\mathcal{P}$ partitions the available varieties into goods sold as mixed lots. For each $P \in \mathcal{P}$, let $Q_P = \sum_{j \in P} Q_j$ and $\omega_{jP} = Q_j / Q_P$ denote its total supply and constituent shares. Every lot of good P contains these proportions, so a lot of q kilograms contains $\omega_{jP} q$ kilograms of variety j. The construction is applied within lot-size categories, whose index we suppress together with the public state. Pooling preserves the number of auctions, the distribution of lot quantities, and the number and type distribution of attending bidders. The probability of auctioning good P equals the sum of the probabilities of auctioning its constituent varieties.

Valuations and taste shocks Type m's baseline valuation for P is $\gamma_{mP} = \sum_{j \in P} \omega_{jP} \gamma_{mj}$. Since pooling also results in different lots' taste shocks being pooled, we draw independent variety-specific taste shocks $\varepsilon_{i\ell j} \sim N(0, \sigma^2)$ and form the pooled shock as

$$\varepsilon^P_{i\ell} = \sum_{j \in P} \omega_{jP} \varepsilon_{i\ell j}, \qquad \mathrm{Var}(\varepsilon^P_{i\ell}) = \sigma^2 \sum_{j \in P} \omega^2_{jP}.$$

The component shocks are independent across bidders, lots, and varieties, and are learned upon entry. Since our pooling scheme might mix fruits of different sizes, we assume that the common variance σ^2 is the environment's frequency-weighted average of the estimated fruit-size-specific variances. In homogenized units, bidder i's willingness to pay for a pooled lot is $q_\ell(\gamma_{m_i P} + \varepsilon^P_{i\ell} - \alpha w_{i\ell} - \alpha q_\ell / 2)$. Winning increases inventory by the lot's entire quantity, regardless of its composition.

Participation We distinguish having a downstream opportunity for a lot from finding it profitable to incur the cost of bidding. A variety is relevant to a bidder when they have a downstream opportunity for it. Without such an opportunity, the bidder does not participate.

Conditional on relevance, participation depends on expected auction profits and a stochastic participation cost.

We use the estimated participation-cost differences to calibrate the probability of relevance. Write c_{jm} for the fitted participation-cost location and define $c_j^0 := \min_m c_{jm}$ and $r_{jm} := \exp\{-(c_{jm} - c_j^0)/\tau\}$. At low entry probabilities, the fitted cost difference $c_{jm} - c_j^0$ approximately scales participation by r_{jm} relative to a bidder facing location c_j^0. We interpret this proportional difference as the probability that variety j is relevant to type m, while c_j^0 governs the cost of participation conditional on relevance.

A mixed lot is relevant if at least one constituent variety provides a downstream opportunity. However, the quantity of that variety must be sufficient for the opportunity to be useful. We discount each constituent's relevance by $h(\omega_{jP}q)/h(q)$, where $h(q) = 1 - \exp(-q/\kappa)$. This ratio can be interpreted as the probability that the constituent quantity is sufficient, conditional on a full lot being sufficient, when the minimum useful quantity has distribution h. Assuming these opportunities are independent across varieties conditional on bidder type, the probability that a pooled lot is relevant is

$$r_{Pm}(q) = 1 - \prod_{j \in P} \left[1 - r_{jm} \frac{h(\omega_{jP}q)}{h(q)} \right].$$

We set $\kappa = 10$ kilograms because that is among the smallest lot sizes observed in the data. This rule preserves relevance for an unpooled lot and prevents a negligible quantity of a variety from making the entire pooled lot relevant.

Conditional on relevance, participation costs are $c_{i\ell} = c_P^0 + \tau\nu_{i\ell}$, where $c_P^0 = \sum_{j \in P} \omega_{jP} c_j^0$ and $\nu_{i\ell} \sim \text{Logistic}(0, 1)$ is independent of downstream opportunities and taste shocks. The bidder enters if expected profits $\Pi(\bar{v}_{m\ell}(w, s_\ell), s_\ell)$ exceed this cost. Integrating over relevance and the cost draw gives

$$p_{m\ell}(w, s_\ell) = r_{Pm}(q_\ell) \frac{\exp\left(\Pi(\bar{v}_{m\ell}(w, s_\ell), s_\ell)/\tau\right)}{\exp\left(\Pi(\bar{v}_{m\ell}(w, s_\ell), s_\ell)/\tau\right) + \exp(c_P^0/\tau)}.$$

Here s_ℓ contains the pooled lot's characteristics, and $\bar{v}_{m\ell}(w, s_\ell)$ incorporates its valuation and the continuation values under the pooling scheme.

Equilibrium comparisons We recompute equilibrium under every partition, allowing participation, bidding, inventories, and continuation values to adjust jointly. The solution procedure is described in Appendix F.1. The status quo is also recomputed under the same participation specification, with every variety in a singleton pool. All reported changes compare a pooling scheme with this common benchmark.

F.3 Additional counterfactual results

Table F1 reports revenues, efficiency, and bidder surplus of select pooling schemes under the five environments considered in Figure 5.

References

Armstrong, T. B., M. Bertanha, and H. Hong (2014): "A Fast Resample Method for Parametric and Semiparametric Models," *Journal of Econometrics*, 179 (2), 128–133.

Battigalli, P. (1987): "Comportamento razionale ed equilibrio nei giochi e nelle situazioni sociali," Unpublished thesis, Università Bocconi.

Fudenberg, D., and D. K. Levine (1993): "Self-Confirming Equilibrium," *Econometrica*, 61 (3), 523–545.

Hotz, V. J., and R. A. Miller (1993): "Conditional Choice Probabilities and the Estimation of Dynamic Models," *The Review of Economic Studies*, 60 (3), 497–529.

Kastl, J., E. Richert, and J. Rüdiger (2026): "Efficient Grid Search to Solve Static Games with Private Information," *Mimeo*.

Weintraub, G. Y., C. L. Benkard, and B. Van Roy (2008): "Markov Perfect Industry Dynamics With Many Firms," *Econometrica*, 76 (6), 1375–1411.

Table F1 Revenues and efficiency under selected environments and pooling schemes

	Revenue		Match value		Bidder surplus	
Policy	Level	Change	Level	Change	Level	Change
Nectarine-heavy						
Status quo	3,037.8	–	5,180.5	–	2,142.7	–
Hard White + Nectarine†	3,168.9	131.0	5,189.1	8.6	2,020.3	-122.4
S1†	3,254.7	216.9	5,171.1	-9.4	1,916.4	-226.3
S2†	3,262.8	225.0	5,137.9	-42.6	1,875.1	-267.6
Full pooling	3,128.5	90.7	4,716.3	-464.2	1,587.9	-554.8
Soft White-heavy						
Status quo†	3,105.7	–	5,035.6	–	1,930.0	–
Soft White + Yellow†	3,278.6	173.0	5,029.6	-6.1	1,750.9	-179.0
S1†	3,338.5	232.8	4,997.2	-38.4	1,658.7	-271.2
S2	3,328.2	222.5	4,967.3	-68.3	1,639.1	-290.8
Full pooling	3,225.9	120.2	4,594.0	-441.6	1,368.1	-561.8
Soft White-led mix						
Status quo	3,121.9	–	5,175.9	–	2,054.0	–
Hard White + Yellow†	3,180.4	58.6	5,187.4	11.5	2,007.0	-47.1
S1†	3,335.9	214.1	5,066.9	-108.9	1,731.0	-323.0
S2†	3,310.9	189.0	5,160.8	-15.1	1,849.9	-204.2
Full pooling	3,207.3	85.4	4,668.0	-507.9	1,460.7	-593.4
Yellow and Soft White						
Status quo	3,655.2	–	5,281.1	–	1,625.9	–
Hard White + Yellow†	3,706.9	51.7	5,312.2	31.2	1,605.4	-20.5
S1†	3,839.5	184.4	5,162.4	-118.6	1,322.9	-303.0
S2†	3,822.6	167.4	5,308.5	27.4	1,485.9	-140.0
Full pooling	3,705.6	50.5	4,840.7	-440.3	1,135.1	-490.8
Yellow and Niche						
Status quo	3,575.2	–	5,401.6	–	1,826.4	–
SW+Y†	3,729.3	154.1	5,435.1	33.5	1,705.8	-120.6
S1†	3,746.3	171.1	5,430.1	28.5	1,683.7	-142.6
S2†	3,677.7	102.5	5,451.8	50.2	1,774.0	-52.3
Full pooling	3,645.9	70.7	5,087.7	-313.9	1,441.9	-384.5

Notes: Table reports expected revenue, match value, and bidder surplus under the status quo, every scheme on the revenue-efficiency frontier, *S1*, *S2*, and full pooling. Levels and changes are measured in KRW/kg. Changes are relative to the corresponding environment's status quo. The reported environments are the same as in Figure 5. Bidder surplus equals match value minus seller revenue. In the pooling scheme labels, a plus sign denotes varieties sold together as one pooled good. Varieties omitted from a label remain separate. Scheme *S1* pools Hard White with Nectarine and Soft White with Yellow. Scheme *S2* pools Hard White with Yellow and Soft White with Nectarine. Both schemes leave Niche separate.

†Pooling scheme belongs to the revenue-efficiency frontier in the corresponding environment.

(Return to the counterfactuals section in the main text.)