Kyoto University,
Graduate School of Economics
Discussion Paper Series

Quality-Driven Price Dispersion:

Implications for Testing Market Efficiency

Hisaki Kono, Ernst David, Tsilavo Ralandison, Yutaka Arimoto

Discussion Paper No. E-25-004

Graduate School of Economics
Kyoto University
Yoshida-Hommachi, Sakyo-ku
Kyoto City, 606-8501, Japan

May, 2025

Quality-Driven Price Dispersion: Implications for Testing Market Efficiency*

Hisaki Kono† David Ernst‡ Tsilavo Ralandison§ Yutaka Arimoto¶

May 30, 2025

Abstract

Spatial price dispersion in agricultural markets is often interpreted as evidence of market inefficiencies. Yet, price differences may also reflect variations in product quality, especially where formal grading is absent. This study utilizes a novel dataset of transaction-level paddy rice sales from rural Madagascar, collected in 2022-2023, that includes laboratory-assessed grain quality indicators. By employing hedonic regressions, we construct a composite quality index to quantify how much of the observed cross-regional price variation can be explained by quality differences. We find that quality accounts for only a small share of observed dprice dispersion, and controlling for quality has little impact on conventional measures of market integration. This suggests that buyers may face challenges in accurately assessing quality at the point of sale. In such contexts, buyers appear to rely on village-level reputations as a substitute, with villages known for high-quality rice commanding price premiums that exceed what is explained by observable quality attributes. These results highlight the importance of informal reputation as a substitute for formal quality verification in rural markets and suggest that weak quality-price linkages may undermine incentives to invest in quality.

*This research was supported by the Japan Society for the Promotion of Science (JSPS) under Grants-in-Aid for Scientific Research (19H01482, 23K25510). We are grateful for the outstanding research assistance provided by Hanchen Lee. We also extend our sincere appreciation to Faliarivony Razaiarinirina for her dedicated field coordination, and to the enumerators—Diavolana Rakotondratsima, Rindrasoa Randrianavalona, Allan Rahajananahary, Nambinintsoa Ramananjoelina, Zo Rabearimanana, and Don Cabrel Miarintsoaniaina—for their invaluable contributions to data collection.

†Kyoto University. E-mail: kono@econ.kyoto-u.ac.jp

‡Kyoto University. E-mail: davidernst.321@gmail.com

§Northern Ontario School of Medicine University. E-mail: tsilavo@gmail.com

¶Metrics Work Consultants Inc. and Hitotsubashi University. Email: arimotoy@gmail.com

1 Introduction

Agricultural markets in developing countries often exhibit substantial price dispersion across locations, even for seemingly identical commodities. These spatial price differences are typically interpreted as evidence of market inefficiencies or frictions that hinder arbitrage (Ravallion, 1986; Fackler and Goodwin, 2001; Moser et al., 2009). In a frictionless market, the law of one price should hold–identical goods would trade at the same price once transport costs are accounted for. Deviations from this benchmark are commonly viewed as indicative of incomplete spatial market integration. Numerous studies from Africa and Asia have documented considerable variation in staple crop prices across regions, often exceeding what transport costs alone can explain (Baulch, 1997; Barrett and Li, 2002; Negassa and Myers, 2007).

However, an alternative explanation for price dispersion is heterogeneity in product quality. While agricultural commodities are often treated as homogeneous goods in empirical analyses, in practice they can differ considerably in attributes valued by consumers (Unnevehr, 1986; Peterson-Wilhelm et al., 2023; Twine et al., 2023). If such quality differences are not properly accounted for, observed price dispersion across markets would be wrongly attributed to market inefficiencies or high transaction costs. In contrast, such price differences may reflect quality-based price premiums, consistent with efficient market functioning. This perspective is supported by empirical evidence that improved market access can lead producers to upgrade quality in response to demand-side incentives (Bold et al., 2022). Conversely, if higher quality is not translated into higher prices, this may indicate market inefficiency in the sense that the market fails to reward quality adequately and thereby undermines producers' incentives to invest in quality improvements. Accounting for quality is thus essential when evaluating spatial market integration and efficiency.

A key empirical challenge is the lack of data containing both price and quality across locations. To address this, we use a unique dataset of transaction-level paddy rice prices from rural Madagascar, collected in 2022 and 2023. Each transaction record was accompanied by laboratory assessments of key grain characteristics such as moisture content, sterility, presence of fungi and soil, varietal identity, and variety purity, providing objective quality indicators. Using these data, we construct a composite quality index through hedonic price regressions and estimate the extent to which quality explains cross-regional price variation, controlling for time and transportation accessibility.

The study area covers a relatively small geographic region whererice varieties are broadly similar and arbitrage is relatively easy due to proximity. This context limits price variations unexplained by quality, allowing us to precisely estimate the contribution of quality differences to price dispersion.

We find that measured quality differences account for only a small portion of price variation. Consequently, controlling for quality has a limited impact on conventional market integration measures. We interpret this small magnitude of the quality premium as evidence of information frictions: quality attributes are not easily observed or verified due to the limited availability of equipment for measuring quality at the point of sale. Supporting this interpretation, we find that rice from villages with higher average quality commands a price premium beyond what is predicted by measured attributes. In the absence of reliable quality verification, buyers rely on village-level reputation. Villages known for high-quality rice consistently secure better prices because buyers use reputation as a heuristic to navigate information asymmetry. This finding aligns with models of repeated interaction and reputation as substitutes for formal quality verification (Tadelis, 1999).

Our findings contribute to several strands of literature on agricultural markets, information frictions, and informal institutions. First, we contribute to the literature on market integration and price dispersion in agriculture. Previous studies assumed a homogeneous good and attributed all inter-market price gaps to inefficiencies or transaction costs (Ravallion, 1986; Baulch, 1997; Barrett and Li, 2002), but product quality can differ, especially in developing countries. With quality differences, what has been labeled as a "malfunctioning of the market" could be a market-based quality premium, which may indicate the efficiency of the market. We found differences in product quality, but the magnitude of the quality premium is relatively modest, which in turn suggests the potential of market inefficiency in the sense that the market fails in offering a sufficiently higher price for higher quality products, which would reduce the producer's incentives to improve quality.

Second, our study speaks to the literature on information frictions and the role of reputation in market transactions. We provide quantitative evidence of how informal reputational capital can facilitate market exchange in the absence of formal grading systems. In many rural economies, formal grading systems or branding for products like rice are absent, so buyers and sellers rely on trust and informal information to gauge quality (Demont et al., 2013; Hoffmann and Gatobu, 2014; Bold et al., 2017; Minten et al., 2017). A particular village might develop a reputation for producing high-quality rice, due to favorable local agro-climatic conditions, better post-harvest practices, or a history of reliable quality, while another village is known for inferior output. Traders learn these reputations over time and may be willing to pay a premium for rice from the reputed village, effectively using the village name as a proxy for unobserved quality.

The structure of the paper is as follows. Section 2 presents the local context of our study area and explains the data, followed by Section 3, which reports the variation in price and measured quality. Section 4 shows the results of price regression measuring the quality premium. Section 5 conducts the analysis, measuring market integration after controlling for quality differences. Section 6 concludes.

2 Local context and data

2.1 Local context

We conducted a series of surveys in the Amparafaravola district of the Alaotra-Mangoro region of Madagascar, one of the country's major rice-producing areas. The region is especially renowned for its Makalioka rice, a translucent, long, and thin-grained variety widely regarded as high quality in the domestic market. "Makalioka" is a generic name encompassing several varieties derived from MK 34, a cultivar released in 1934. Within this group, Makalioka luxe refers to the high-grade type, while Makalioka ordinaire denotes the standard grade.

The main transportation route in the region is National Highway 3A, which runs north-south and serves as the primary artery for rice shipments. Most villages are accessible by 10-ton trucks, and nearly all are reachable by hand tractors–small vehicles capable of transporting approximately one ton of rice.

Rice transactions in the region generally follow two distinct modes. The first involves trade with wholesalers or rice millers, who typically deal in milled rice through long-term relationships. The second mode entails trade in paddy rice between buyers and large farmers or village-level collectors, with brokers acting as intermediaries. While some wholesalers also deal in paddy rice, their prices are usually higher. As a result, buyers seeking lower prices often rely on brokers, who actively gather information on potential sellers. Buyers frequently contact brokers in advance to inquire about the availability and price of paddy and to schedule their arrival to ensure smooth transactions. In the absence of appointments, brokers commonly wait at community meeting points for incoming trucks.

We focus our analysis on paddy rice transactions for the following two reasons. First, milled rice prices are substantially affected by the quality of milling, which is difficult to control or observe. In contrast, the quality of paddy rice can be objectively assessed using a standardized set of laboratory tests proposed by International Rice Research Institute (2013). Second, milled rice is typically traded through relational contracts, where the price is determined considering the dynamic incentive constraints, and the law of one price does not necessarily hold (Baker et al., 2002; Macchiavello, 2022). In contrast, paddy rice is traded in the spot market, making it more suitable for analyzing spatial integration based on the law of one price.

Although quality variation in paddy is generally smaller than in milled rice, buyers often expressed concerns about paddy quality (Ralandison et al., 2025). Commonly cited issues include: varietal purity, moisture content, fungal development, foreign matter (e.g., stones or soil), and the presence of sterile (empty) grains. Buyers typically assess a grain sample by visual inspection, focusing on easily observable characteristics. However, detailed assessments are rarely conducted due to the lack of quality-testing equipment at transaction points.

In response to buyer concerns about paddy quality, farmers and village-level collectors employ several practices to improve and preserve the condition of their rice. One commonly used method is winnowing, which helps reduce the presence of sterile (empty) grains. After threshing and before storage, winnowing primarily separates grains from husks and distinguishes heavier, healthy grains from lighter, sterile ones using wind or airflow. However, this method is ineffective for removing heavier contaminants, such as stones or soil particles, which may still remain in the paddy.

Another essential practice is drying, which helps control moisture levels and prevents fungal growth during storage. Farmers typically dry their paddy by spreading it on large plastic sheets laid out on roads or open spaces under the sun.

Larger farmers and village collectors often place paddy bags on wooden pallets to further protect rice during storage. This elevates the rice off the ground, helping to prevent moisture absorption and allowing for air circulation, which reduces the risk of spoilage.

2.2 Survey

We surveyed the transaction price of paddy and its quality in the Alaotra region in Madagascar in 2022 and 2023, using a different sampling procedure.

In 2022, we first identified active rice traders in the region using information from our previous surveys. These traders were then contacted regularly (at least every 10 days) for survey interviews. During each interview, we collected a 250-gram grain sample from every variety-grade of rice sold in their store for quality testing. We also recorded key transaction details, including the purchase price, the variety the trader reported, and the produce's origin. To track price outcomes, we followed up with traders to obtain the selling price of the paddy bag from which the sample had been collected. In addition, we gathered information on all sales transactions conducted by each trader, including the selling price and characteristics of the buyers. In some cases, traders facilitated direct transactions by taking buyers to producers, such as farmers or village collectors. When this occurred, we accompanied them to the transaction site, collected a 250-gram sample from the produce purchased, and recorded the corresponding transaction details. The survey ran from September 7th to November 5th, 2022. However, we continued to follow up on the sampled produce in order to obtain the eventual selling price.

By repeatedly surveying the same traders, this survey design allows us to measure the trader-specific price premium. However, it did not capture many direct transactions between sellers and buyers mediated by brokers.

To address this gap, we conducted another survey in 2023 focused specifically on broker-mediated trade. Since buyers from outside the region often rely on local brokers (known as mpanera) to locate sellers, we first identified active brokers operating in the region. We then divided all communes into six geographic zones and assigned one enumerator to each zone. Each zone included typical meeting points where brokers and external buyers regularly interacted. Every morning, each enumerator contacted a broker in their assigned zone to inquire about the time and location of upcoming buyer visits. If no appointments were scheduled, enumerators waited with the brokers at known meeting points for potential buyers to arrive. Enumerators accompanied the brokers and collected a 250-gram sample from the traded paddy for quality testing. We ensured that each enumerator covered all active brokers in their zone and visited multiple locations within the same day to collect price and transaction information across communes. This survey was conducted from June 8th to July 10th, 2023. Although the survey period was shorter than in 2022, the higher frequency of broker-mediated transactions allowed us to collect a larger volume of transaction data.

Thus, the two surveys differ along two key dimensions: trade mode and timing. The mode of trade likely affects the degree of market integration. In broker-mediated transactions, buyers can access information from a wider geographic area, which increases spatial competition and is expected to result in stronger market integration. In contrast, the timing difference influences the role of quality. Without adequate storage technology, rice quality deteriorates rapidly over time, leading to changes in taste and the development of fungi. As a result, quality concerns are likely more pronounced in the later survey period.

To illustrate the geographic features of the survey area and the spatial distribution of transactions, Figure 1 presents a map of the region, with trade locations recorded in the 2023 survey shown as black circles. The area outlined in black represents the survey site, Amparafaravola District. Lake Alaotra lies to the east, and flatlands suitable for rice cultivation stretch to its southwest. Light red lines on the map represent major roads, with the north-south route corresponding to National Highway 3A. It is evident that most transactions took place along this highway.

To assess grain quality, we partnered with a local agricultural research institute to conduct laboratory testing. According to International Rice Research Institute (2013), several characteristics determine rice quality. Poor post-harvest handling and inadequate cleaning are common causes of reduced quality.

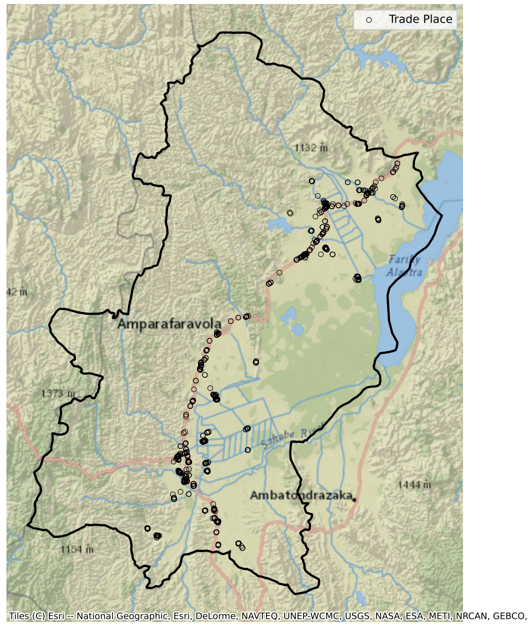
Moisture content (MC): Moisture levels above 14% can cause spoilage, fungal growth, and reduced storability. To ensure safe long-term storage, MC should remain below 13–14%.[1]

Presence of foreign matter: Contaminants such as stones, soil, and straw increase processing costs and reduce milling yield.

Variety Purity: Mixing different rice varieties reduces size uniformity and negatively affects

[1] Rice Knowledge Bank (`http://www.knowledgebank.irri.org/step-by-step-production/postharvest/milling/milling-and-quality/measuring-moisture-content-in-milling`).

Figure 1: Survey location and observed trade location in the 2023 survey



taste. Many buyers prefer to source paddy directly from large farmers to avoid mixing, frequently occurring when small collectors aggregate paddy from multiple farmers.[2]

Grain size: Variations in grain size and shape make it difficult to adjust hullers and polishers optimally, resulting in low initial de-hulling efficiency, higher rates of recirculated paddy, uneven polishing, and reduced overall quality of milled rice.

Percentage of sterile or immature grains: Sterile grains are completely empty, while immature grains contribute to excessive bran production and broken grains during milling.

Percentage of broken and cracked grains: Cracked grains often result from exposure to fluctuating temperature or moisture levels. These cracks lead to breakage during milling.

Presence of fungi and discoloration: Visual discoloration (e.g., black or yellow grains) may indicate fungal development. Fungal contamination lowers both the aesthetic and nutritional value of rice and may signal the presence of aflatoxins or other mycotoxins.

To capture these quality attributes, we agreed with the research institute to measure the following variables: (1) moisture content (MC), (2) paddy purity, (3) presence of soil, (4) variety purity, (5) weight of 1,000 sampled grains, (6) sterility, and (7) fungal presence.

MC measures the percentage of water weight in paddy rice. Levels above 14% increase the risk of fungal growth, as well as insect and rodent damage. To reflect quality penalties from excessive moisture, we also construct an indicator of inappropriate moisture content: $InappMC \equiv$

[2] Village collectors typically purchase small lots of paddy (20–30 kg bags) from individual farmers and resell them to wholesalers in larger bags (80–100 kg). During this aggregation process, different rice varieties are often mixed. Such mixing can also occur at the farm level, as smallholder farmers frequently obtain seeds from multiple sources, leading to varietal mixtures at harvest.

$(\min\{MC - 14, 0\})^2$.

Paddy purity is defined as the proportion of rice grains in the sample, excluding foreign matter like stones, soil, and straw. Sterile grains were excluded from the purity calculation. We dropped one sample with a paddy purity below 10%, as it no longer qualified as paddy. We also created a binary indicator for the presence of soil, which not only raises processing costs but also creates conditions favorable to fungal development.

Variety purity was assessed by examining a 1,000-grain subsample. If the sample contained a dominant single variety, the number of grains from that variety was counted. If multiple varieties were present in significant proportions, the sample was classified as "mixed variety." Such mixing often occurs when collectors purchase small quantities from many farmers.

The weight of 1000 grains reflects several factors, including MC, the ratio of immature grains, and grain size. After controlling for moisture, sterility, and variety, we interpret the remaining variation in weight as reflecting grain size and the share of immature grains. Note that while variation in grain size is identified as a key quality determinant in International Rice Research Institute (2013), we do not have a direct measure of it in our dataset.

Sterility was measured as the proportion of completely empty grains in the sample. Partially filled grains were not counted as sterile.

The presence of fungi was recorded as a binary indicator. Fungal growth can occur at various stages, such as harvesting, threshing, or storage. While high moisture levels promote fungal development, contamination shortly after harvest is often due to dirt and animal waste from inadequate threshing methods.

In addition, considering the possibility that price differences reflect variety differences, we also recorded some variety-specific characteristics, including (8) grain type (long vs. short), (9) grain color (white vs. red), and (10) grain translucency. Translucency was classified into three categories based on laboratory inspection: (Grade 1) less than 10% opacity (translucent, non-waxy, non-sticky), (Grade 2) 10-30% opacity, and (Grade 3) 50% or more opacity (waxy, glutinous, sticky rice). Translucent grains (Grade 1) are typically regarded as high-quality. Additionally, we recorded the variety name as reported by the seller.

Table 1 presents the reported rice varieties from the 2023 and 2022 surveys. Most observed trades involved Makalioka luxe (premium grade) and Makalioka ordinaire (standard grade). Tsemaka, an improved variety of Makalioka, was also frequently transacted. While traders can usually distinguish these varieties by the shape of paddy, most consumers regard them similarly. Together, these three varieties account for nearly 60% of all transactions.[3]

[3]Since Makalioka is a long-grain variety, if a seller reported a sample as Makalioka but the laboratory test showed short-grain variety, we reclassified it as "other".

Table 1: Variety

(A) Survey 2023

	Freq.	Percent
Bota	20	1.974
Dista	220	21.72
Makalioka luxe	96	9.477
Makalioka ordinaire	391	38.60
Mixed	100	9.872
Others	30	2.962
Rio	40	3.949
Tsemaka	116	11.45
Total	1013	100.00

(B) Survey 2022

	Freq.	Percent
Bota	13	4.676
Dista	65	23.38
Makalioka luxe	41	14.75
Makalioka ordinaire	78	28.06
Mixed	20	7.194
Others	9	3.237
Rio	21	7.554
Tsemaka	31	11.15
Total	278	100

Panel (A) and Panel (B) show the variety of the sampled grains for the 2023 and 2022 samples, respectively.

Another major variety is Dista, a local specialty. Dista is a pink-colored, sweet-flavored rice known for its high nutritional value and yields twice as high as other varieties. It accounts for about 25% of transactions. Due to its thin market, however, Dista prices are volatile and can be affected by the supply of a similar pink rice grown near Antananarivo.

Bota is a round, thick-grain rice similar to traditional Malagasy rice and is generally considered lower quality. Red-colored Bota is often mislabeled as Dista, which may contribute to observed price dispersion in transaction price.

Mixed variety refers to a blend of multiple varieties, typically different types of Makalioka. A special case is Rio, a mixture of Makalioka and Dista rice, recognizable by its blend of white and pink grains. However, Rio is sometimes simply labeled as "mixed rice," so some mixed samples may actually include Rio.

2.3 Observed trade patterns

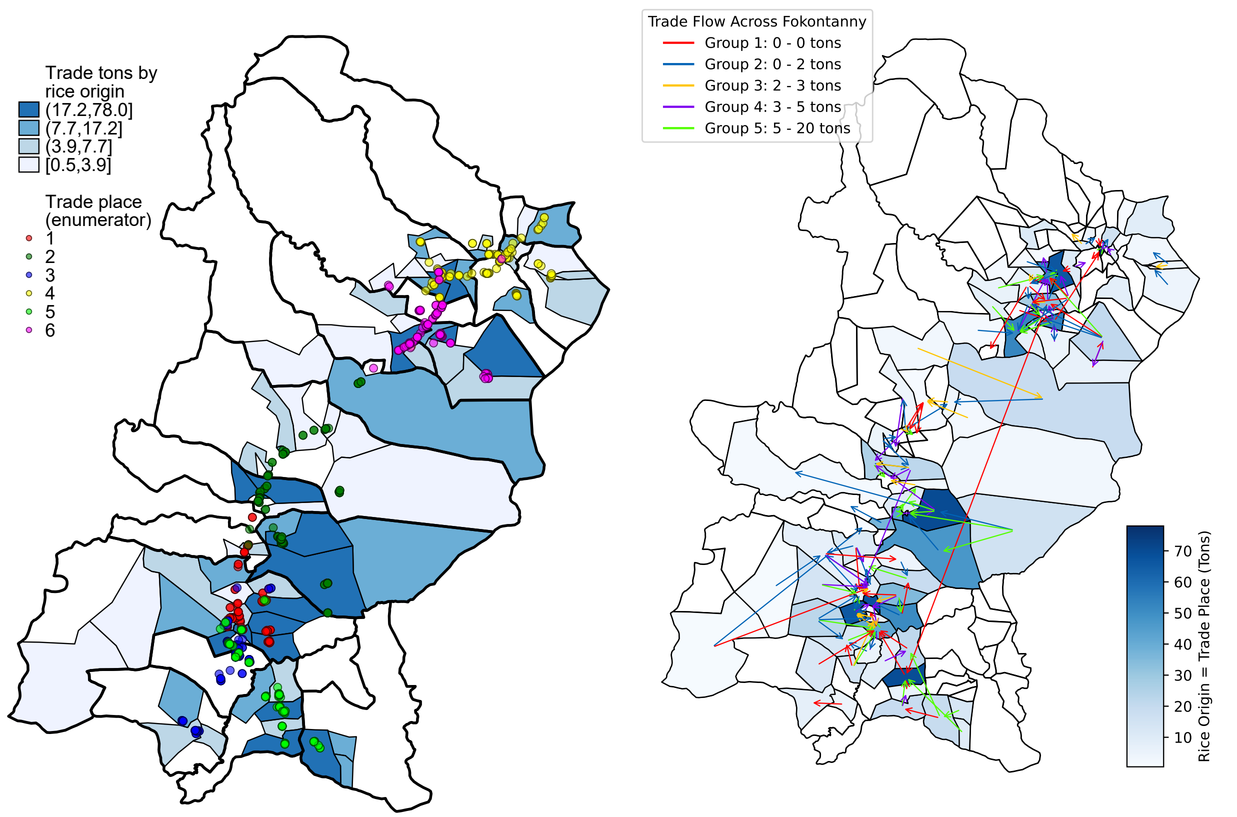
Figure 2 illustrates the trade patterns observed in the 2023 survey, during which enumerators accompanied brokers to transaction sites. Panel (A) shows the locations of these transactions as points, with different colors representing different enumerators. It also displays the aggregated volume of traded rice by origin village using a color gradient, where darker shades indicate larger trade volumes.

The figure reveals that the origins of traded rice were highly concentrated in a small number of villages. In particular, villages shaded in the darkest color (those with trade volumes exceeding 17.2 tons) account for over 70% of the total trade volume. Furthermore, the top 10% of villages by trade volume contributed more than 40% of the total volume traded.

Panel (B) illustrates trade flows (from the origin villages where rice was produced to the transaction locations) using arrows. It also displays the aggregated trade volume by production village through color shading, with darker shades indicating larger volumes. The trade flows suggest that

Figure 2: Trade place and trade flow (2023)

(A) Trade place and volume (B) Trade flow (from origin to trade place)



Panel (A) shows the trade place as points and represents the aggregated traded volume of rice for origin by color shading. The bold solid lines indicate the borders of a commune. Panel (B) indicates the trade flows by arrows, along with the amount of rice traded at the same village as the production site by color shading.

most rice was transported over relatively short distances, typically from production sites to nearby trade points along the main road.

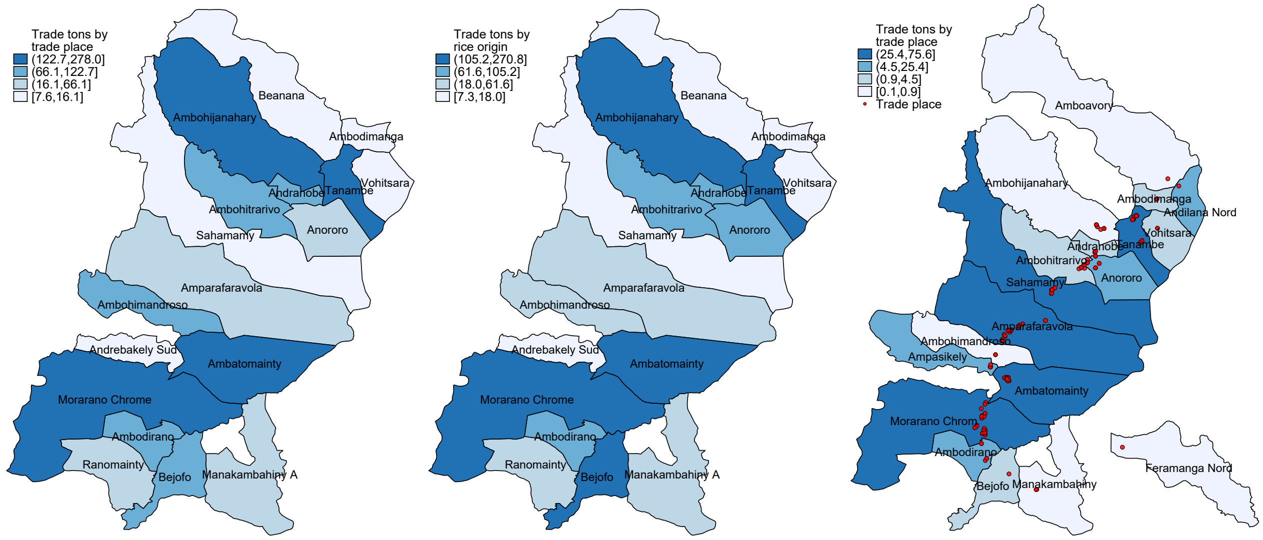
Figure 3 presents the geographic distribution of rice trade volumes, aggregated at the commune level.[4] Panels (A) and (B) display the 2023 data, showing trade volumes by transaction location and production origin, respectively. These maps reveal clear clustering patterns, with one prominent cluster in the north and another in the south, indicating spatial concentration in both trade and production.

Panel (C) shows the geographic distribution of trade volumes by transaction location in the 2022 survey. Unlike in 2023, the distribution was more evenly spread across communes, reflecting differences in survey design. In 2023, we followed active brokers, who connect buyers to local farmers, resulting in trade locations concentrated near production areas. In contrast, the 2022 survey targeted active traders operating along or near major roads, leading to a more even geographic

[4] A commune is a higher-level administrative unit encompassing multiple villages.

Figure 3: Geographical distribution of trade volume (aggregated at commune level)

(A) By trade place (2023) (B) By origin (2023) (C) By trade place (2022)



Panel (A) and Panel (B) show the aggregated trade volume (in tons) by commune of trade and commune of harvest, respectively, for the 2023 sample. Panel (C) shows the aggregated trade volume (in tons) by commune of trade for the 2022 sample, with exact trade locations marked by red dots.

spread of trade locations across the region.

3 Variations in price and quality

3.1 Price

We begin by documenting the degree of price dispersion and variation in paddy quality.

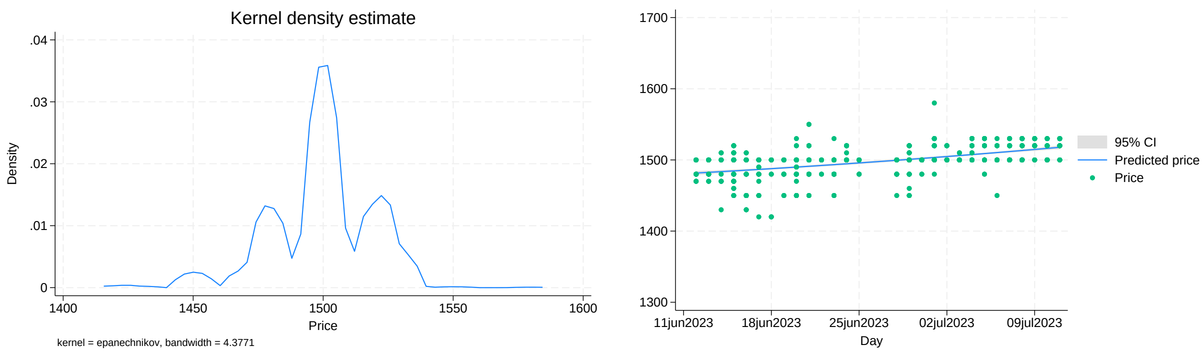
The left panels of Figure 4 display the distribution of paddy prices during the study periods, while the right panels plot transaction prices over time. In both years, the average price was approximately 1,500 MGA per kg (roughly 0.33–0.36 USD).[5] In 2023, prices remained relatively stable, as the survey took place shortly after the harvest, when local supply was still abundant. In contrast, 2022 prices gradually increased toward the end of the year due to tightening supply. Notably, in both years, we observe substantial price dispersion even on the same date, suggesting that factors beyond timing contributed to price variation.

Price dispersion was noticeably greater in the 2022 survey than in 2023. This may partly reflect differences in sampling strategy. In 2023, all recorded transactions were mediated by brokers who matched local sellers with outside truck buyers, ensuring that all trades occurred at a consistent stage in the supply chain. By contrast, the 2022 survey targeted a broader range of traders operating

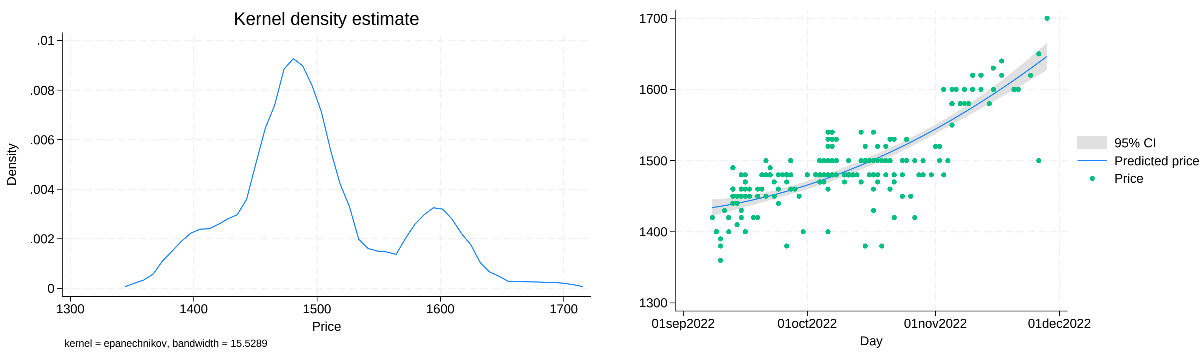
[5]The local currency is the Malagasy Ariary (MGA). In October 2022, 1,000 MGA equaled 0.24 USD; in June 2023, it was about 0.22 USD.

Figure 4: Rice Price Distribution

(A) Survey 2023



(B) Survey 2022



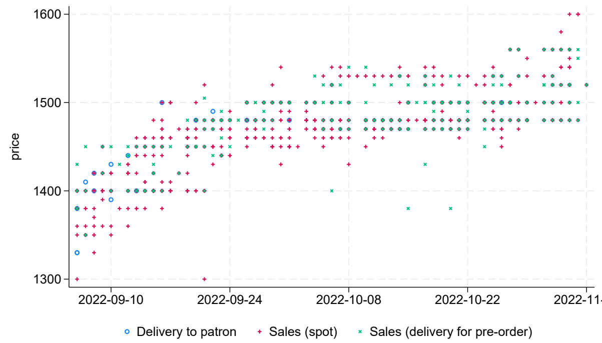
Panel (A) and Panel (B) show the trade price distribution as a kernel density plot on the left and trade prices with their respective trade dates on the right for the 2023 and 2022 samples, respectively.

at various stages. Some sold directly to trucks, while others sold to larger collectors or wholesalers, often through pre-orders or under patron-client relationships involving advance payments. These modes could plausibly be associated with lower prices due to financing arrangements.

However, our data indicate that buyer identity had no statistically significant effect on transaction prices. In regression analyses (not reported), dummy variables for pre-order sales and sales to patrons yielded coefficients close to zero and were statistically insignificant. To further explore this, we analyzed all recorded transactions, including those not sampled for quality testing. Figure 5 presents transaction prices by date, distinguishing between buyer types: blue markers for sales to patrons, red for spot sales, and green for pre-order sales. The figure shows no systematic differences in prices across transaction types.

The 2022 survey also recorded a few transactions with unusually low prices, around 1,400 MGA/kg. These cases involved farmers who had borrowed from a local rice bank, using their harvested rice as collateral. The rice bank valued the paddy at 1,400 MGA/kg and charged 2% monthly interest. Farmers often sold their rice quickly to nearby collectors or wholesalers to minimize interest accumulation, typically at the same 1,400 MGA/kg. While waiting for trucks from

Figure 5: Sales Price by Transaction Type 2022



The figure presents trade prices by trade dates, categorized by transaction type, for the 2022 sample.

Antananarivo (Tana), which offered higher prices and could have been more profitable, the urgency of loan repayment drove these lower-priced transactions.

3.2 Quality variables

Table 2 presents summary statistics for transaction prices, grain quality variables, and road access. Panel (A) shows results from the 2023 survey, while Panel (B) corresponds to 2022.[6] To capture variability, we report the mean, standard deviation, and the 10th, 25th, 50th (median), 75th, 90th, and 95th percentiles.

MC levels differed significantly between the two survey years. In 2023, only about 10% of samples exceeded the recommended 14% threshold, indicating relatively good paddy conditions. In contrast, nearly 60% of samples in 2022 exceeded this threshold. This contrast reflects differences in survey timing: the 2023 survey was conducted one month after harvest, when rice was still fresh and not stored long, while the 2022 survey took place four months after harvest, after prolonged storage.

Consistent with this, we found a higher likelihood of fungal presence in 2022. While fungi were detected in about one-seventh of 2023 samples, fungi was detected in one-third of 2022 samples. [7] These results indicate the inadequate post-harvest handling technology in this region.

The sterility rate was low in most samples across both years, though a few samples contained substantial proportions of empty grains. Paddy purity was 100% in both surveys, with no detectable contaminants such as stones, soil, or straw. Accordingly, no soil was detected in any sample.

The average variety purity index was 991 in 2023 and 964 in 2022. Most 2023 samples consisted

[6] In 2022, price data are missing for some grain samples due to the sampling strategy: grain samples were collected from traders' stores, and prices were recorded only if the corresponding paddy bags were sold during the survey period. Unsold paddy has no associated price.

[7] A regression of fungal presence on MC and $InappMC$ (defined as $(\min \mathrm{MC} - 14, 0)^2$) showed that $InappMC$ is positively associated with fungal detection, with an average partial effect of 0.05.

Table 2: Summary Statistics

(A) Survey 2023

	N	Mean	SD	p10	p25	p50	p75	p90	p95
Price	1013	1499.33	19.41	1480.0	1490.0	1500.0	1520.0	1520.0	1530.0
Moisture (MC)	1013	13.30	0.72	12.5	12.8	13.3	13.6	14.2	14.6
InappMC	1013	0.09	0.50	0.0	0.0	0.0	0.0	0.0	0.4
Sterility	1013	1.47	1.21	0.5	0.9	1.1	1.9	2.5	3.4
Paddy purity (%)	1013	100.00	0.00	100.0	100.0	100.0	100.0	100.0	100.0
Fungi detected	1013	0.14	0.35	0.0	0.0	0.0	0.0	1.0	1.0
Soil detected	1013	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.0
Weight for 1000 grains	1013	26.19	2.55	23.2	24.3	25.7	28.1	30.1	30.8
Variety purity	765	990.80	7.87	980.0	988.0	992.0	996.0	998.0	999.0
Varieties mixed	1013	0.24	0.43	0.0	0.0	0.0	0.0	1.0	1.0
Long grain	1013	0.81	0.39	0.0	1.0	1.0	1.0	1.0	1.0
Red grain	1013	0.38	0.49	0.0	0.0	0.0	1.0	1.0	1.0
Translucency	1013	0.81	0.39	0.0	1.0	1.0	1.0	1.0	1.0
Accessible by a 10-ton truck	1013	0.97	0.16	1.0	1.0	1.0	1.0	1.0	1.0
Accessible by a small vehicle	1008	1.00	0.00	1.0	1.0	1.0	1.0	1.0	1.0
Minutes from the main road by a small vehicle	1013	10.48	18.76	0.0	0.0	3.0	14.0	25.0	33.5

(B) Survey 2022

	N	Mean	SD	p10	p25	p50	p75	p90	p95
Price	246	1498.66	64.73	1420.0	1460.0	1480.0	1530.0	1600.0	1620.0
Moisture (MC)	278	14.47	0.92	13.3	13.9	14.4	15.0	15.7	16.2
InappMC	278	0.93	1.98	0.0	0.0	0.2	1.0	2.9	4.8
Sterility	278	3.20	4.84	1.0	1.1	2.1	3.4	5.4	7.9
Paddy purity (%)	278	100.00	0.00	100.0	100.0	100.0	100.0	100.0	100.0
Fungi detected	278	0.35	0.48	0.0	0.0	0.0	1.0	1.0	1.0
Soil detected	278	0.00	0.00	0.0	0.0	0.0	0.0	0.0	0.0
Weight for 1000 grains	278	26.36	2.68	23.0	24.4	26.4	27.8	29.8	30.6
Variety purity	238	964.02	39.38	917.7	956.3	974.5	990.0	995.0	997.0
Varieties mixed	270	0.15	0.36	0.0	0.0	0.0	0.0	1.0	1.0
Long grain	278	0.70	0.46	0.0	0.0	1.0	1.0	1.0	1.0
Red grain	278	0.40	0.49	0.0	0.0	0.0	1.0	1.0	1.0
Translucency	278	0.68	0.47	0.0	0.0	1.0	1.0	1.0	1.0
Accessible by a 10-ton truck	258	1.00	0.00	1.0	1.0	1.0	1.0	1.0	1.0
Accessible by a small vehicle	258	1.00	0.00	1.0	1.0	1.0	1.0	1.0	1.0
Minutes from the main road by a small vehicle	258	3.65	7.47	0.0	0.0	0.0	3.0	14.0	15.0

of a single variety, reflecting that traders could source full bags of rice directly from individual farmers soon after harvest. In contrast, in 2022, traders often had to combine smaller lots from multiple farmers, leading to more mixed varieties.[8]

Most rice traded in this region consists of long, white grains. The majority of samples had a Grade 1 translucency level, indicating high quality. Based on this, we created a dummy variable for translucency, equal to 1 if the grain was entirely Grade 1, and 0 otherwise. Because translucency does not apply to red or pink rice, we also created a separate indicator for the presence of red or pink grains.

Overall, these summary statistics indicate considerable variation in rice quality, with greater heterogeneity in 2022, reflecting longer post-harvest periods and more storage-related deterioration. In the next section, we examine the extent to which these quality differences explain price variation and affect measures of spatial market integration.

[8]The reported proportion of mixed varieties was somewhat higher in 2023. However, in most cases, traders explicitly labeled the rice as mixed, often referring to deliberate blends like "Rio."

Note that road access conditions also differed across the two surveys. In 2023, about 3% of transactions occurred in locations inaccessible to 10-ton trucks, though all sites were reachable by small vehicles. In contrast, in 2022, all trade locations were accessible by 10-ton trucks, consistent with the sampling focus on traders located along main roads.

4 Price differences by quality

4.1 Specifications

We now examine the extent to which quality differences explain observed price dispersion by estimating the following regression model:

$$price_{ijt} = F(\mathbf{X}_{ijt}) + \mathbf{A}_j\delta + \mu_t^{week} + \epsilon_{ijt}, \tag{1}$$

where $price_{ijt}$ is the transaction price of paddy sold by trader i in fokontany j during week t, and $\mathbf{X}_{ijt}$ is a vector of quality attributes. For mixed-variety paddy, we assign the variety purity variable a value of 1000, so that the coefficient on the mixed-variety dummy captures its effect relative to fully pure variety paddy. The vector $\mathbf{A}_j$ controls for accessibility of fokontany j, including a dummy for 10-ton truck access and travel time from the main road by hand tractor.[9] Week fixed effects (FEs) μ_t^{week} capture seasonal price fluctuations.

From this model, we construct an adjusted price net of accessibility and week FEs:

$$P_{ijt} \equiv price_{ijt} - \mathbf{A}_j\hat{\delta} - (\hat{\mu}_t^{week} - \hat{\mu}_1^{week}), \tag{2}$$

and a price-relevant quality index predicted from the covariates:

$$q_{ijt} = \hat{F}(\mathbf{X}_{ijt}),$$

where $\hat{F}(\cdot)$ is the estimated functional form of $F(\cdot)$. We then compute the quality-adjusted price residuals as

$$QP_{ijt} \equiv P_{ijt} - q_{ijt}. \tag{3}$$

This residual captures price variation not explained by quality, accessibility, or seasonality, thereby offering a more accurate measure to evaluate spatial market integration.

To quantify the explanatory power of the quality variables, we compute the net R^2, controlling for accessibility and week FEs:

$$R^{2,net} = \frac{Var(\hat{F}(\mathbf{X}_{ijt}))}{Var(price_{ijt} - \mathbf{A}_j\hat{\delta} - \hat{\mu}_t^{week})} = \frac{Var(q_{ijt})}{Var(P_{ijt})}.$$

[9] All fokontany were accessible by hand tractor.

We use three approaches to estimate $F(\cdot)$. Our baseline specification is OLS, assuming $F(\mathbf{X})$ is linear in $\mathbf{X}$. However, OLS may suffer from overfitting, and the predicted value of $F(\cdot)$ may be overly influenced by variables with statistically insignificant coefficients.

To address these concerns, we employ regularized regression methods, specifically LASSO (least absolute shrinkage and selection operator) and elastic net (EN, hereafter), which improve prediction by performing variable selection and shrinking coefficients toward zero.[10] Since shrinking coefficients may lead to underestimation of the impact of quality variables, we adopt a double-selection strategy similar to post-double selection LASSO. We first apply LASSO or EN for variable selection, then estimate $F(\cdot)$ using OLS with the selected variables. To enhance robustness in small samples, we implement a model averaging approach (MAA). Since variable selection can be sensitive to how the data is split into training and testing data, we run LASSO and EN with different splits, and then compute a score defined below for each covariate to be used for variable selection. To reduce correlation among splits due to strong predictors, we apply feature bagging. With only 12 features in total, we implement a leave-one-covariate-out (LOCO) strategy, wherein we exclude one feature in each split. We term these procedures post-LASSO OLS and post-EN OLS, which proceed in three steps:

1. Run LASSO or EN across 600 different sample splits (50 different sample splits for each LOCO). Let R_j^2 be the out-of-sample R^2 for the j-th split, and let s_{kj} be an indicator for whether variable x_k was selected in j-th split.

2. Compute each variable's selection score as $S_k = \frac{\sum_{j=1}^{600} L_j(x_k) R_j^2 s_{kj}}{\sum_{j=1}^{600} L_j(x_k) R_j^2}$, where $L_j(x_k)$ is an indicator if variable x_k is not excluded in LOCO.

3. Run OLS using variables with S_k above a threshold α.

Table 3 reports the selection score S_k for each quality variable. Variety grade variables are always selected across all sample splits in both years, indicating the importance of variety in paddy price. The weight for 1,000 grains is frequently selected in 2023, but less so in 2022. Given that weight will capture many attributes, these results suggest that weight itself may not be a fundamental quality predictor. Fungi presence and red grain indicators are frequently selected in both years, while variety purity and grain type emerge as key predictors only in 2023.

[10] EN improves predictive performance by addressing both variable selection and multicollinearity. It estimates the coefficients by solving

$$\arg\min_{\beta\in\mathbb{R}^p} \left\{ \frac{1}{2n} \sum_{i=1}^{n} (y_i - \beta_0 - \sum_{j=1}^{p} x_{ij}\beta_j)^2 + \lambda \left[\frac{(1-\alpha)}{2} \sum_{j=1}^{p} \beta_j^2 + \alpha \sum_{j=1}^{p} |\beta_j| \right] \right\},$$

which reduces to LASSO if $\alpha = 1$, and to Ridge regression if $\alpha = 0$. The parameter α is selected by cross-validation. Smaller α values lead EN to select more variables than LASSO.

Table 3: Score weighted by out-of-sample goodness of fit

(A) Survey 2023

	LASSO	Elastic Net
Continuous variables	.	.
Moisture (MC)	0.46	0.41
InappMC	0.33	0.39
Sterility	0.85	0.86
Weight for 1000 grains	1.00	0.99
Variety purity	0.84	0.85
Categorical variables	.	.
Varieties mixed	0.81	0.79
Red grain	0.34	0.44
Fungi detected	0.85	0.81
Soil detected	0.00	0.00
Variety grade	1.00	1.01
Long grain	0.33	0.37
Translucency	0.49	0.55

(B) Survey 2022

	LASSO	Elastic Net
Continuous variables	.	.
Moisture (MC)	0.09	0.07
InappMC	0.10	0.10
Sterility	0.20	0.15
Weight for 1000 grains	0.10	0.10
Variety purity	0.08	0.08
Categorical variables	.	.
Varieties mixed	0.09	0.11
Red grain	0.36	0.33
Fungi detected	0.33	0.32
Soil detected	0.00	0.00
Variety grade	1.01	1.00
Long grain	0.04	0.02
Translucency	0.19	0.13

Panels (A) and (B) report the weighted scores S_k for each quality variable using MAA for LASSO CV (column (1)), adaptive LASSO (column (2)), and EN (column (3)) in the 2023 and 2022 samples, respectively. The score ranges from 0 to 1.

We report results for post-LASSO OLS with $\alpha = 0.5$ (more selective) and post-EN OLS with $\alpha = 0.25$ (more inclusive). Varing the threhold value α has a minimal impact on out-of-sample R^2. If anything, higher α values slightly improve predictive performance, but the differences are quantitatively minor.

Our third approach applies MAA directly to prediction. We compute $\hat{F}(\mathbf{X}_{ijt})$ by averaging predicted prices across multiple sample splits. Specifically, we first obtain the predicted value $\hat{price}^b_{ijt}$ using EN based on equation (1) for each of the 600 LOCO-based splits. We then compute the predictor

$$\hat{price}_{ijt} = \frac{1}{B}\sum_{b=1}^{B} \hat{price}^b_{ijt}.$$

The corresponding quality-adjusted price residual is computed as

$$QP_{ijt} \equiv price_{ijt} - \hat{price}_{ijt}, \tag{4}$$

which is conceptually equivalent to the price residual defined in equation (3). This approach does not tell us which quality variables are important determinants of the transaction price, but it is a conventional MAA for prediction. We call this EN-MAA.

Note that although LASSO and EN can accommodate flexible functional forms of $F(\mathbf{X})$, our results indicate that linear models outperform more complex ones. Including second-order polynomial terms in $F(\mathbf{X})$ substantially deteriorates out-of-sample predictive accuracy. Appendix Figure 1 shows the cumulative distribution function (CDF) of out-of-sample R^2 based on 100 sample splits, revealing that the linear models first-order stochastically dominate the polynomial models. Therefore, we report results based solely on linear specifications.

4.2 Magnitudes of quality attributes

Table 4 reports the estimation results from the 2023 survey using the three approaches explained above: baseline OLS including all quality variables (Column (1)), post-LASSO OLS (Column (2)), and post-EN OLS (Column (3)), with p-values reported in parentheses. Table 5 shows the corresponding results from the 2022 survey.

In both years, rice variety was the most influential determinant of transaction price, with notably larger estimated effects in the 2022 survey. Compared to Bota (the reference category), the varieties Makalioka luxe, Makalioka ordinaire, and Tsemaka are associated with statistically significant price premiums of 16.7-17.6 MGA, 13.6-14.6 MGA, and 16.3-16.6 MGA, respectively, in 2023. In 2022, these premiums were substantially larger: 79.2-82.4 MGA for Makalioka luxe, 74.1-56.6 MGA for Makalioka ordinaire, and 87.8-89.6 MGA for Tsemaka. Mixed-variety rice was also transacted at higher prices, likely reflecting its composition–typically a blend of Makalioka varieties.

This discrepancy in Makalioka premiums between the two surveys may reflect differences in market conditions related to the timing of the surveys. Bota, being similar to traditional Malagasy rice, is more easily substituted with rice from other regions. In contrast, Makalioka is a high-quality variety produced only in this region. Immediately after harvest, Makalioka is relatively abundant, resulting in lower premiums, as observed in 2023. As time passes and its supply diminishes while traditional variety rice from other regions continuously enter the market, Makalioka premium increases, as observed in the 2022 survey. These results suggest that controlling for variety is essential when analyzing market integration, and its impact can be time-variant.

Most other quality attributes had negligible and statistically insignificant effects on transaction prices. For instance, the point estimates suggest that a one percentage point increase in sterility (the share of empty grains) reduced the price by only 0.5 MGA in 2023, while the average transaction price was approximately 1,500 MGA. Even a substantial reduction in sterility from the 90th percentile (25.0) to the 10th percentile (0.5) translates into a modest price increase of 12.25 MGA, or 0.8% of the average price. Similarly, the presence of fungi reduced price by only 1.3 MGA, while translucent rice was associated with a modest 2.7 MGA increase. None of these estimates was statistically significant.

In the 2022 survey, where the sample size is relatively small, post-LASSO ($\alpha = 0.5$) does not select quality attributes other than rice variety. post-EN ($\alpha = 0.25$) selected the presence of fungi as an additional predictor, but its estimated effect is small and statistically insignificant. In the 2023 survey, the only physical quality attribute with a statistically significant coefficient was the weight per 1,000 grains. However, as this variable had no significant effect in 2022, it is unlikely to be a robust or fundamental quality signal.[11]

[11] The 2022 survey also recorded whether the transaction was pre-ordered. When this variable was included in the

Table 4: Coefficients of the linear models: Survey 2023

	OLS - All Coefficients	post-LASSO OLS	post-EN OLS
Moisture (MC)	1.214		1.214
	(0.249)		(0.249)
InappMC	-0.219		-0.219
	(0.803)		(0.803)
Sterility	-0.503	-0.501	-0.503
	(0.201)	(0.206)	(0.201)
Weight for 1000 grains	-1.417	-1.233	-1.417
	(0.000)	(0.000)	(0.000)
Variety purity	-0.0772	-0.0717	-0.0772
	(0.239)	(0.272)	(0.239)
Varieties mixed=1	-1.851	-0.837	-1.851
	(0.308)	(0.535)	(0.308)
Red grain=1	1.471		1.471
	(0.487)		(0.487)
Fungi detected=1	-1.282	-1.268	-1.282
	(0.316)	(0.314)	(0.316)
Long grain=1	-1.019		-1.019
	(0.828)		(0.828)
Translucent	2.694		2.694
	(0.578)		(0.578)
Variety grade (ref=Bota)			
Dista	10.94	12.44	10.94
	(0.004)	(0.000)	(0.004)
Makalioka luxe	16.67	17.56	16.67
	(0.000)	(0.000)	(0.000)
Makalioka ordinaire	13.59	14.55	13.59
	(0.000)	(0.000)	(0.000)
Mixed	15.64	16.40	15.64
	(0.000)	(0.000)	(0.000)
Others	3.902	5.054	3.902
	(0.382)	(0.244)	(0.382)
Rio	2.031	3.628	2.031
	(0.660)	(0.394)	(0.660)
Tsemaka	16.32	16.95	16.32
	(0.000)	(0.000)	(0.000)
Constant	1488.7	1500.9	1488.7
	(0.000)	(0.000)	(0.000)
R^2	0.40	0.40	0.40
$R^{2,net}$	0.09	0.09	0.09
F	36.34	45.14	36.34
N	1013	1013	1013
Week FE	YES	YES	YES
Access FE	YES	YES	YES

Column (1) includes all available quality variables, while columns (2) and (3) include quality variables selected by post-LASSO ($\alpha = 0.5$) and post-EN ($\alpha = 0.25$), respectively. R^2 is the standard R^2, and $R^{2,net}$ is the R^2 net of week FEs and accesibility.

Table 5: Coefficients of the linear models: Survey 2022

	OLS - All Coefficients	post-LASSO OLS	post-EN OLS
Moisture (MC)	-1.857		
	(0.528)		
InappMC	0.919		
	(0.421)		
Sterility	0.195		
	(0.489)		
Weight for 1000 grains	0.531		
	(0.493)		
Variety purity	-0.0109		
	(0.784)		
Varieties mixed=1	-5.903		
	(0.684)		
Red rice	11.69		1.075
	(0.349)		(0.915)
Fungi detected	-2.265		-3.130
	(0.558)		(0.415)
Contain long grain	26.39		
	(0.197)		
Translucent	-23.93		
	(0.131)		
Dista	85.61	91.01	89.08
	(0.000)	(0.000)	(0.000)
Makalioka luxe	82.46	79.78	79.18
	(0.000)	(0.000)	(0.000)
Makalioka ordinaire	75.58	74.13	73.41
	(0.000)	(0.000)	(0.000)
Mixed	72.65	71.45	70.65
	(0.000)	(0.000)	(0.000)
Others	90.31	86.50	84.20
	(0.000)	(0.000)	(0.000)
Rio	76.19	79.33	77.18
	(0.000)	(0.000)	(0.000)
Tsemaka	89.19	89.55	87.76
	(0.000)	(0.000)	(0.000)
Constant	1328.5	1321.4	1323.9
	(0.000)	(0.000)	(0.000)
R^2	0.86	0.86	0.86
$R^{2,net}$	0.33	0.31	0.31
F	132.48	221.29	190.29
N	217	222	222
Week FE	YES	YES	YES
Access FE	YES	YES	YES

Column (1) includes all available quality variables, while columns (2) and (3) include quality variables selected by post-LASSO ($\alpha = 0.5$) and post-EN ($\alpha = 0.25$), respectively. R^2 is the standard R^2, and $R^{2,net}$ is the R^2 net of week FEs and accesibility.

Across all specifications, the explanatory power of quality attributes remains limited. In the 2023 survey, the value of $R^{2,net}$ was 0.09, indicating that only 9% of the variation in transaction prices (after controlling for accessibility and week FEs) is explained by observed quality variables. In the 2022 survey, the corresponding $R^{2,net}$ values were higher, ranging from 0.31 to 0.33, but this was primarily driven by large price differentials across rice varieties, rather than differences in physical quality.

These findings suggest that, at least in our study area, variation in physical quality is not well reflected in transaction prices, pointing to potential inefficiencies in the paddy transaction market. However, controlling for rice variety remains important, particularly when testing the law of one price. In the next section, we examine how adjusting for quality differences affects standard measures of market integration.

4.3 Magnitude of price differences by quality

In this subsection, we assess the relationship between the adjusted price net of accessibility and week FEs, P_{ijt} (defined in Equation (2)), and the quality-adjusted price residuals, QP_{ijt} (defined in equation (3) or (4)). Since P_{ijt} controls for the week FE with the first week as the reference category, it uses the first week's price as the benchmark, as reflected in equation 2. For consistency and ease of interpretation, we apply the same principle when computing QP_{ijt}. Specifically, we estimate equation 1 using the first week as the reference category, and then add the estimated intercept to the residual from equation (3) or (4) to construct QP_{ijt}.[12]

Figure 6 presents scatter plots of QP_{ijt} against P_{ijt}. The horizontal axis shows P_{ijt} estimated from the baseline OLS. The vertical axis displays three versions of QP_{ijt}, estimated using different regularized regression methods:

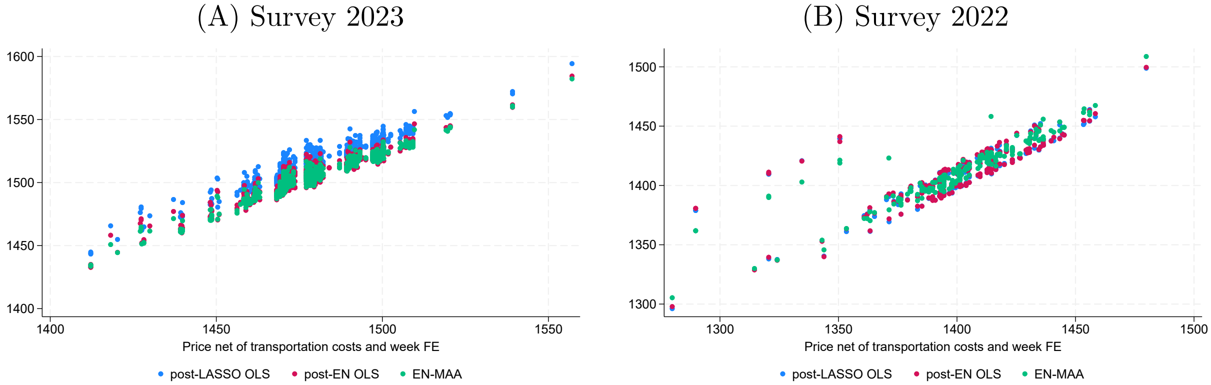
1. post-LASSO OLS with $\alpha = 0.5$ (blue dots),
2. post-EN OLS with $\alpha = 0.25$ (red dots), and
3. EN-MAA, based on equation (4) (green dots).

Panel (A) shows results from the 2023 survey, where we observe a tight relationship between QP_{ijt} and P_{ijt} across all estimation methods. This pattern is consistent with the small coefficient estimates reported in Table 4, suggesting that measured quality attributes explain only a small portion of the variation in transaction prices. The figure also illustrates the robustness of QP_{ijt} to the choice of estimation methods.

regression, it showed no significant association with price.

[12] Summary statistics for P_{ijt} and QP_{ijt} are reported in Appendix Table 1.

Figure 6: Relationship between P and QP



Panel (A) and Panel (B) plot the quality-adjusted price residual (QP_{ijt}) against the adjusted price net of accessibility and week FEs (P_{ijt}) for the 2023 and 2022 surveys, respectively. Blue, red, and green dots correspond to estimates from post-LASSO OLS ($\alpha = 0.5$), post-EN OLS ($\alpha = 0.25$), and EN-MAA, respectively.

Panel (B) presents the corresponding plot for the 2022 survey. Here, we observe larger discrepancies between QP_{ijt} and P_{ijt}, particularly for observations involving the Bota variety, which is consistently priced lower than other types. This finding supports the importance of controlling for variety when analyzing spatial price dispersion and market integration. Nevertheless, excluding Bota, the relationship between QP_{ijt} and P_{ijt} remains relatively tight, indicating that for most observations, price variation is not substantially driven by quality differences.

4.4 Reputation premium

Our earlier analysis showed that measured quality attributes had limited explanatory power for transaction prices. This finding contrasts insights from our field interviews, where traders and brokers frequently emphasized the importance of quality when sourcing rice (Ralandison et al., 2025). One potential explanation for this discrepancy is that many quality attributes were not readily observable to buyers at the time of the transaction. In practice, buyers assess rice quality superficially, typically by visually inspecting the husks and fungal contamination. During our field interviews, we observed no use of specialized tools to measure less visible attributes such as grain translucency or moisture content. Without such tools, subtle quality differences are unlikely to be detected or priced. Furthermore, detailed quality inspections are costly, making them impractical for the small-scale, high-frequency transactions typical of these markets.

Instead, buyers may depend on the general reputation of certain production areas. Indeed, brokers frequently cited specific areas as sources of particularly high- or low-quality rice. This suggests buyers may substitute costly individual assessments with coarse but informative signals based on origin, leading to a "reputation premium."

To examine this possibility, we assess the relationship between the estimated price-relevant quality index q_{ijt} and the quality-adjusted price residual QP_{ijt}, aggregated at the commune level.[13] In the destination market (typically Antananarivo, the capital of Madagascar), rice is not marketed or priced based on the commune's name or small geographic area of origin. Therefore, any systematic price residual differences conditional on quality can be interpreted as capturing a reputation premium. In this analysis, we exclude variety effects from the construction of the quality index, as variety differences would reflect product type rather than intrinsic quality. Since few, if any, quality attributes beyond variety are selected in the 2022 survey based on Lasso or elastic net, we focus exclusively on the 2023 survey.[14]

Figure 7 presents a scatter plot of commune-level averages of the quality index q_{ijt} and the quality-adjusted price residual QP_{ijt}, with marker sizes proportional to transaction volume. The figure reveals a positive association between the two variables: communes with higher average quality also tend to have prices that exceed what would be predicted based on the measured quality alone.

This pattern is consistent with theory under incomplete information: buyers may rely on regional reputation as a proxy when they cannot fully observe all relevant quality attributes. As a result, communes known for producing high-quality rice may attract more buyers and command a reputation-based price premium.

Importantly, the dispersion in price residuals across communes is substantially greater than the variation in the estimated quality index. This suggests that reputation may be a more important driver of price differences than objectively measured quality attributes.

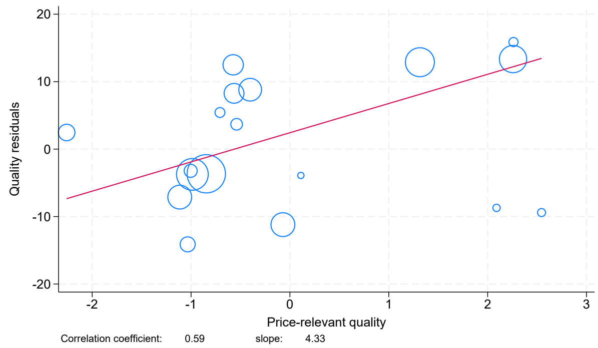
One may be concerned that this pattern was driven by unobserved quality attributes observed or perceived by traders but not captured in our data. However, our quality measures, developed and validated by IRRI, have been widely used across countries and are considered robust proxies for key grain characteristics (International Rice Research Institute, 2013). Moreover, in downstream consumer markets, rice from different communes is not priced differently, and there is no evidence of a commune-level price premium. This implies that the contribution of unmeasured quality attributes to the variation in price residuals is likely limited.

Another potential explanation for this positive relationship is a "fast-collection premium." Com-

[13] While ideally this aggregation should be done at the village level, our sample size was insufficient to support reliable aggregation at that scale.

[14] Appendix Figure 3 presents the spatial distributions of the quality index (q_{ijt}) and quality-adjusted price residuals (QP_{ijt}). Both measures are higher in northern communes, including areas frequently mentioned by brokers as known for high-quality rice, such as Ambatomainty, Ambohijanahary, Ambohitrarivo, Beanana, Morarano Chrome, and Tanambe. These areas tend to be flat and well-irrigated, with large farm plots and higher adoption of improved seeds and modern agricultural practices. They also benefit from better storage infrastructure and post-harvest handling. These characteristics are consistent with the interpretation that reputational advantages are associated with observable area-level features, even if not fully captured by our measured quality.

Figure 7: Quality and QP aggregated by commune (weighted by transaction volume): 2023



The figure plots commune-level averages of the estimated price-relevant quality (excluding variety) against the quality-adjusted price residuals (QP_{ijt}), for the 2023 (left) and 2022 (right) surveys. Marker size reflects the total trade volume of each commune.

munes known for high-quality rice tend to have larger farms and better storage capacity, allowing buyers to source rice more quickly. Buyers who prioritize rapid procurement may be willing to pay a price premium in these areas. To assess this possibility, we regress the commune-level average price residuals on the average estimated quality index, controlling for average collection time and average inventory size.

Table 6 presents the results from the 2023 survey. Column (1) reports the bivariate regression, which corresponds to the regression line shown in Figure 7. The point estimates indicate that a 1 MGA increase in the average price-relevant quality index is associated with a 2.1 MGA increase in the average price residual, suggesting that reputation premium has a stronger influence on transaction prices than the measured quality of individual paddy. In Column (2), we add the average collection time (measured as the number of days buyers needed to collect rice from their primary trading partners in each commune), which little affects the coefficient on the quality. In Column (3), we include the average inventory size (log-transformed) of each commune, which leads to a larger and statistically significant coefficient on the quality index. These results suggest that neither collection time nor inventory size explains the positive relationship between the average quality and the average price residuals, reinforcing the interpretation that reputation plays a significant role in shaping price differentials.

5 Impacts on measures of market integration

We now examine how product quality differences influence standard market integration measures. To quantify the degree of spatial integration, we adopt two complementary approaches, both grounded in the concept of the Law of One Price (LOP). According to the LOP, if markets are

Table 6: QP Quality Regression

	(1) QP	(2) QP	(3) QP
Quality (excluding variety)	2.095	2.040	3.286
	(0.258)	(0.287)	(0.007)
Truck time		1.323	
		(0.872)	
inventory			4.965
			(0.181)
_cons	1520.7	1518.7	1477.5
	(0.000)	(0.000)	(0.000)
N	16	16	16

The table shows the regression results, with p vales in parentheses.

fully integrated and function efficiently, identical products should trade at the same price across locations once transportation costs are accounted for. Persistent price gaps beyond these costs suggest the presence of market frictions or segmentation.

In the trade literature, market integration is commonly assessed through cross-sectional price dispersion, which is used to infer trade costs (Kano et al., 2013; Donaldson, 2018). In contrast, the LOP literature typically evaluates integration by examining price co-movements over time between regions (Ravallion, 1986; Baulch, 1997; Van Campenhout, 2007; Moser et al., 2009). Following these approaches, we conduct two analyses. First, we examine the extent to which differences in quality account for observed price variation across communes. Second, we investigate how controlling for quality differences affects the degree of price co-movement between the northern and southern regions.

5.1 Cross-sectional price dispersion

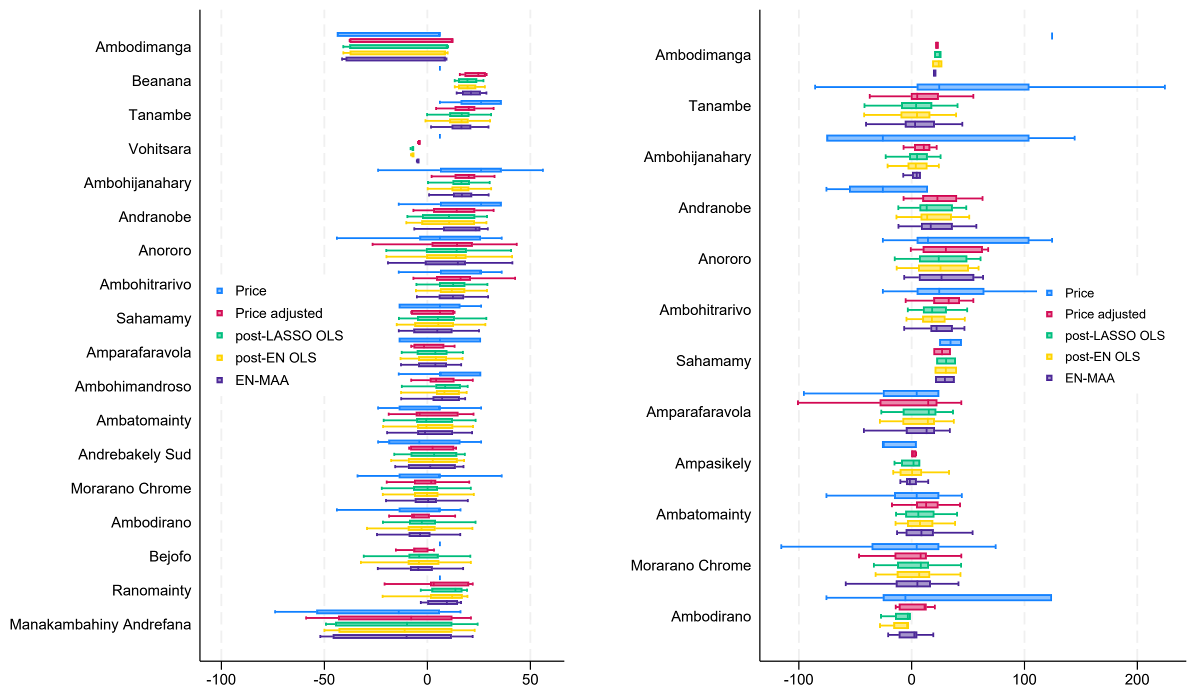
We begin by examining how adjusting for quality differences affects cross-sectional price variation. Figure 8 presents box plots for each commune, ordered from north to south. The blue box plots show the distribution of the observed prices. The red plots display the adjusted price net of accessibility and week FEs, P_{ijt}. Other colors indicate the quality-adjusted prices QP_{ijt}, estimated using three alternative MAA. For comparability, all price variables are demeaned, allowing them to be presented on a common scale.

Compared with the observed price $price_{ijt}$, adjusted price net of accessibility and week FEs, P_{ijt}, exhibits fewer extreme values, particularly in 2022. Additional adjustment for quality yields only a modest reduction in dispersion, consistent with the small coefficients on quality attributes reported earlier.

The box plots indicate the presence of price variations across and within communes. To assess

Figure 8: Differences in the observed price, week FE and accessibility adjusted price (P) and quality adjusted price (QP) across communes

(A) Survey 2023 (B) Survey 2022



Panel (A) and Panel (B) show the observed price, adjusted price net of accessibility and week FEs, (P_{ijt}), and quality-adjusted prices QP_{ijt} using the the post-LASSO OLS ($\alpha = 0.5$), post-EN OLS ($\alpha = 0.25$), and EN-MAA, across communes for the 2023 and 2022 samples, respectively.

how quality adjustment affects observed price variation across and within communes, we conduct an analysis of variance (ANOVA) that decomposes total price variation into between- and within-commune components (Table 7). The results show that most observed price variation arises from between-commune differences, especially in 2023. Adjusting for accessibility and week FEs reduced both components of variation, particularly in the 2022 data. Further adjusting for quality attributes only modestly reduced the between-commune variation and had much less impact on within-commune variation. Similar patterns emerge when we decompose price variation into between-trader and within-trader components using the 2022 survey data (Appendix Table 2).

Table 7: Results from ANOVA

(A) 2023 survey

	F-statistic	Between-groups MS	Within-groups MS
Price	21.188	5789.578	273.254
Price adjusted	27.850	4604.480	165.330
QP post-LASSO OLS	21.875	3572.212	163.305
QP post-EN OLS	21.572	3528.488	163.567
QP EN-MAA	23.713	3822.719	161.209
Quality post-LASSO OLS	5.458	111.775	20.480
Quality post-EN OLS	5.745	116.249	20.234

(B) 2022 survey

	F-statistic	Between-groups MS	Within-groups MS
Price	3.621	12778.306	3529.025
Price adjusted	3.861	2670.216	691.542
QP post-LASSO OLS	2.072	1104.678	533.045
QP post-EN OLS	2.053	1092.193	531.989
QP EN-MAA	2.494	1322.167	530.054
Quality post-LASSO OLS	4.618	949.072	205.521
Quality post-EN OLS	4.559	948.772	208.131

Each panel shows the F-statistic and decomposition of total variation into between- and within-commune components for the observed prices, P_{ijt}, and QP_{ijt} estimated via MAA-OLS with LASSO ($\alpha = 0.75$), EN ($\alpha = 0.25$), and EN-MAA.

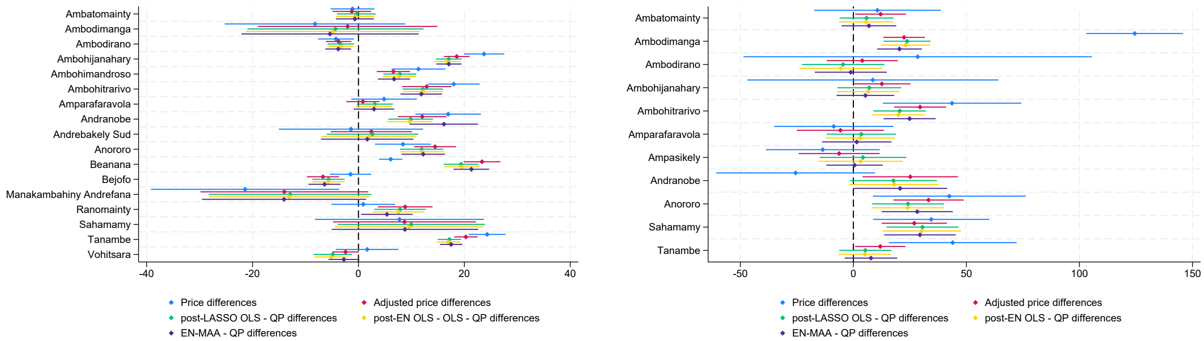
We also report the ANOVA results for the quality index, which show that quality varies more across communes than within communes. This pattern supports the interpretation of a reputation premium based on origin: the limited quality variation within communes suggests that visiting a commune with a higher average quality increases the likelihood of purchasing high-quality rice, leading to higher prices.

Finally, Figure 9 presents the estimated coefficients and 95% confidence intervals from regressions of various price measures on commune fixed effects, with Morarano Chrome used as the reference category. Since trade costs are often inferred from average price differentials across locations (Donaldson, 2018), these coefficients offer a direct measure of how quality adjustments affect estimated trade costs.

As in Figure 8, blue markers represent results based on the observed price, red markers correspond to the adjusted price net of accessibility and week FEs, P_{ijt}, and the remaining colors

Figure 9: Coefficients on the commune dummies in the price regressions

(A) Survey 2023 (B) Survey 2022



Panel (A) and Panel (B) plot the coefficients and their 95% confidence intervals obtained from the regression of price variables on commune dummies, with setting Morarano Chrome as the reference category for the trade price differences, week FE- and accessibility-adjusted price (P_{ijt}) differences, and the quality-adjusted price (QP_{ijt}) differences using the coefficients chosen with post-LASSO OLS ($\alpha = 0.5$), post-EN OLS ($\alpha = 0.25$), and EN-MAA, across communes for the 2023 and 2022 samples, respectively.

indicate the quality-adjusted prices, QP_{ijt}, estimated using three MAA.

The results show that adjusting for accessibility and week FEs substantially reduces the magnitude of the commune fixed effects in both survey years. In contrast, further adjusting for quality attributes has a relatively modest impact. These findings reinforce the earlier results, suggesting that quality differences explain only a small share of price variation across communes.

5.2 Time-series comovement

Next, we examine the time-series comovement of prices between regions, a common approach in the market integration literature. While this method requires a long and balanced panel, our data are limited to traded rice, characterized by infrequent transactions in specific communes. To create a panel suitable for time-series analysis, we aggregate the data into two broad regions–North and South–based on the geographic clustering of trade locations shown in Figure 3.

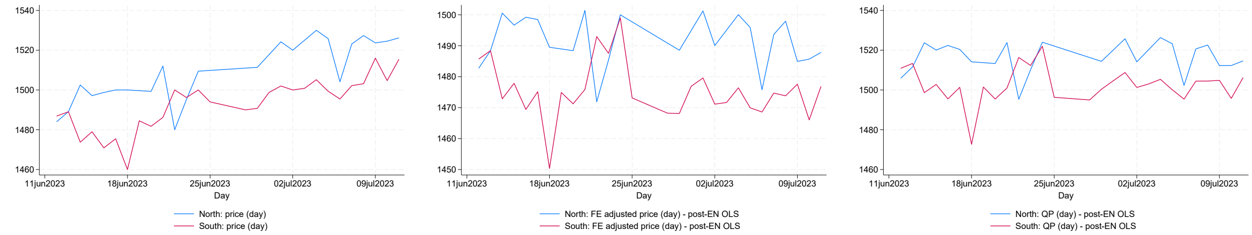
Figure 10 displays the time-series evolution of rice prices in both regions. Panel (A) presents the observed prices; Panel (B) shows prices adjusted for accessibility and week FEs (P_{ijt}); and Panel (C) further adjusts for quality differences (QP_{ijt}). Across all three panels, the price trends in the two regions are highly similar, suggesting that quality adjustments have limited influence on regional price dynamics and do not alter conclusions regarding market integration.

To formally assess spatial integration, we estimate a Threshold Autoregression (TAR) model, a widely used tool in the literature (Obstfeld and Taylor, 1997; Van Campenhout, 2007), which allows price differentials to evolve differently depending on whether they exceed a transportation

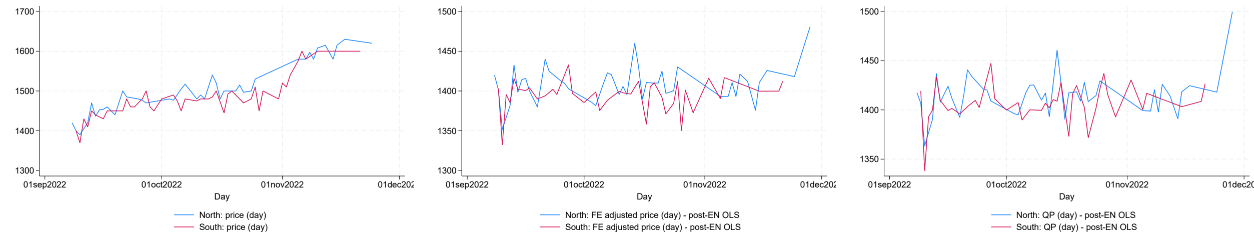
Figure 10: North-south price differences by date

(A) Observed price (B) Access & Week FEs adjusted (P) (C) Quality adjusted (QP)

(a) 2023 survey



(b) 2022 survey



Panel (A), Panel (B), and Panel (C) plot the trade price, price net of transportation costs and time fixed effects, P_{ijt}, and the quality-adjusted price QP_{ijt} for the north and south regions of the 2023 sample.

cost threshold:

$$\Delta m_t = \begin{cases} \rho_{\text{out}} m_{t-1} + \varepsilon_t & \text{if } |m_{t-1}| > \theta \\ \rho_{\text{in}} m_{t-1} + \varepsilon_t & \text{if } |m_{t-1}| \leq \theta, \end{cases}$$

where $m_t \equiv p_{n,t} - p_{s,t}$ is the price differential between the northern and southern regions in period t, and $\Delta m_t \equiv m_t - m_{t-1}$ is its change. The error term ϵ_t is assumed to be normally distributed. The model estimates three parameters: the threshold θ, and adjustment coefficients ρ_{out} and ρ_{in}.

Spatial integration predicts $\rho_{\text{out}} < 0$, indicating that price gaps above the threshold ($|m_{t-1}| > \theta$) trigger arbitrage and converge over time. Larger negative values of $\rho_{\text{out}} < 0$ indicate a faster rate of convergence. In contrast, when price gaps fall within the threshold (i.e., $|m_{t-1}| \leq \theta$), arbitrage is not profitable, and no adjustment is expected; thus, $\rho_{\text{in}} = 0$. This reflects the notion that small price differences within the bounds of transport costs do not necessarily signal market inefficiencies.

Because TAR models require relatively large sample sizes to reliably estimate the threshold, we also estimate a more parsimonious error correction model (ECM) as a robustness check:

$$\Delta m_t = \rho m_{t-1} + \varepsilon_t. \tag{5}$$

Tables 8 and 9 present the estimation results from the TAR and ECM, respectively. Each row represents a different price measure. The first row reports results using the observed (unadjusted) price ($price_{ijt}$). The second row uses the price adjusted for accessibility and week FEs (P_{ijt}). The remaining rows report results using the quality-adjusted price QP_{ijt}.[15]

Table 8: TAR analyses

(A) 2023 survey

	θ	ρ_{out}	SE(ρ_{out})	p-value	ρ_{in}	SE(ρ_{in})	p-value	N
Price	18.16321	-.8894003	.5094688	.0808561	-.2707509	.5340519	.6121725	22
Price adjusted	23.66553	-.5664092	.808313	.4834715	-.5649561	.2909288	.0521486	22
QP post-LASSO OLS	18.98767	-.7561503	.5503617	.169468	-.4614185	.3840723	.229602	22
QP post-EN OLS	21.01013	-.5757629	.54494	.2907116	-.407517	.3454817	.2381745	22
QP EN-MAA	21.39587	-.4405907	.6453322	.4947745	-.4816209	.3327017	.1477273	21

(B) 2022 survey

	θ	ρ_{out}	SE(ρ_{out})	p-value	ρ_{in}	SE(ρ_{in})	p-value	N
Price	15.55554	-1.329648	.9564104	.1644533	-.558633	1.407312	.6914041	12
Price adjusted	13.45947	-1.44156	.5598311	.0100243	.6886092	1.145356	.5476944	12
QP post-LASSO OLS	11.76819	-2.897315	1.159125	.0124344	-1.186862	.8172641	.1464356	12
QP post-EN OLS	11.79871	-3.129089	1.17945	.0079778	-1.35235	.781997	.0837456	12
QP EN-MAA	14.45496	-.8540851	1.465757	.5601004	.3708012	.8017747	.6437403	12

Panel (A) and Panel (B) present the results of a Threshold Autoregression (TAR) for differences between the north and south region for the 2023 and 2022 samples, respectively. Row 1 reports results for the raw price differentials, while rows 2-7 show results for the price net of accessibility and week FEs, P_{ijt}, and the quality-adjusted price, QP_{ijt}. The values of P_{ijt} and QP_{ijt} are calculated using the coefficients chosen with the post-LASSO OLS ($\alpha = 0.5$), post-EN OLS ($\alpha = 0.25$), and EN-MAA. Column (1) reports the estimated threshold. Columns (2) and (5) present the coefficients for the regions above and below the threshold, respectively, with their corresponding standard errors (SE) and p-values reported in columns (3) and (4) for the upper regime, and columns (6) and (7) for the lower regime.

Table 9: Error correction models

(A) 2023 survey

	Coefficient	SE	p-value
Price	-0.870	0.213	0.001
Price adjusted	-0.847	0.214	0.001
QP post-LASSO OLS	-0.835	0.175	0.000
QP post-EN OLS	-0.825	0.173	0.000
QP EN-MAA	-0.884	0.187	0.000

(B) 2022 survey

	Coefficient	SE	p-value
Price	-1.208	0.411	0.015
Price adjusted	-1.170	0.406	0.016
QP post-LASSO OLS	-1.209	0.260	0.001
QP post-EN OLS	-1.194	0.263	0.001
QP EN-MAA	-1.073	0.331	0.009

Panel (A) and Panel (B) report the results from the error correction model (5) for the 2023 and 2022 samples, respectively.

The estimated thresholds θ in Table 8 are small, suggesting limited transportation barriers between the northern and southern regions. The ρ_{out} coefficients are negative and often less than -1, indicating rapid adjustment and a high degree of spatial integration. Similarly, the ECM estimates in Table 9 yield negative coefficients close to or below -1, corroborating these findings.

More importantly, the estimates are fairly stable across different price measures, again indicating that quality differences have little effect on standard indicators of market integration. This reinforces our earlier conclusion that quality contributes only modestly to spatial price variation in our context.

[15] Standard errors of θ cannot be obtained because its asymptotic distribution is neither normal nor nuisance parameter free (Chan, 1993).

6 Conclusion

This study examines the extent to which spatial price dispersion in rice markets reflects underlying quality differences, drawing on detailed transaction-level data from rural Madagascar with laboratory-based quality assessments. We find that while rice quality varies across transactions, these differences explain only a modest share of price variation. In particular, controlling for measured quality—including moisture, sterility, and fungal presence—has a limited effect on both cross-sectional price dispersion and time-series indicators of market integration.

These findings point to significant information frictions in the market. While product variety emerges as a consistent and strong predictor of transaction prices, other quality attributes are not priced accurately, likely because they are difficult for buyers to verify at the point of sale. As a result, many buyers rely on village-level reputations to infer quality, which can lead to reputation-based premiums that reward known production areas but penalize others regardless of the actual characteristics of their output.

We also find that controlling for variety may be required when estimating trade costs. Variety accounts for a significant portion of observed price differences, and failure to adjust for it may exaggerate the estimated trade costs. However, further controlling for other quality attributes—beyond variety—has only a limited influence on estimated price gaps and convergence patterns.

The results carry two key policy implications. First, there is a need for accessible and affordable quality verification tools, such as portable moisture meters or community-based testing services, to reduce information asymmetries at the point of sale. Such tools would support more accurate pricing and reduce reliance on reputational heuristics, increasing individual farmers' incentive to improve quality. Second, policies that allow producers to credibly signal quality, such as collective branding or third-party certification, could improve price accuracy and help newer or smaller producers access better market opportunities. By improving the visibility and verifiability of quality, such interventions could align price signals more closely with product attributes, reduce mispricing, and enhance both the efficiency and inclusiveness of agricultural markets.

References

Baker, George, Robert Gibbons, and Kevin J Murphy, "Relational Contracts and the Theory of the Firm," *The Quarterly Journal of Economics*, 2002, *117* (1), 39–84.

Barrett, Christopher B and Jau Rong Li, "Distinguishing between equilibrium and integration in spatial price analysis," *American Journal of Agricultural Economics*, 2002, *84* (2), 292–307.

Baulch, Bob, "Transfer costs, spatial arbitrage, and testing for food market integration," *American Journal of Agricultural Economics*, 1997, *79* (2), 477–487.

Bold, Tessa, Kayuki C Kaizzi, Jakob Svensson, and David Yanagizawa-Drott, "Lemon technologies and adoption: measurement, theory and evidence from agricultural markets in Uganda," *The Quarterly Journal of Economics*, 2017, *132* (3), 1055–1100.

_ , Selene Ghisolfi, Frances Nsonzi, and Jakob Svensson, "Market Access and Quality Upgrading: Evidence from Four Field Experiments," *American Economic Review*, August 2022, *112* (8), 2518–52.

Campenhout, Bjorn Van, "Modelling trends in food market integration: Method and an application to Tanzanian maize markets," *Food policy*, 2007, *32* (1), 112–127.

Chan, Kung-Sik, "Consistency and limiting distribution of the least squares estimator of a threshold autoregressive model," *The annals of statistics*, 1993, pp. 520–533.

Demont, Matty, Pieter Rutsaert, Maimouna Ndour, and Wim Verbeke, "Reversing urban bias in African rice markets: evidence from Senegal," *World Development*, 2013, *45*, 63–74.

Donaldson, Dave, "Railroads of the Raj: Estimating the impact of transportation infrastructure," *American Economic Review*, 2018, *108* (4-5), 899–934.

Fackler, Paul L and Barry K Goodwin, "Spatial price analysis," *Handbook of agricultural economics*, 2001, *1*, 971–1024.

Hoffmann, Vivian and Ken Mwithirwa Gatobu, "Growing their own: Unobservable quality and the value of self-provisioning," *Journal of Development Economics*, 2014, *106*, 168–178.

International Rice Research Institute, *Quality Kit Manual* International Rice Research Institute 2013.

Kano, Kazuko, Takashi Kano, and Kazutaka Takechi, "Exaggerated death of distance: revisiting distance effects on regional price dispersions," *Journal of International Economics*, 2013, *90* (2), 403–413.

Macchiavello, Rocco, "Relational contracts and development," *Annual Review of Economics*, 2022, *14* (1), 337–362.

Minten, Bart, Thomas Assefa, and Kalle Hirvonen, "Can agricultural traders be trusted? Evidence from coffee in Ethiopia," *World Development*, 2017, *90*, 77–88.

Moser, Christine, Christopher Barrett, and Bart Minten, "Spatial integration at multiple scales: rice markets in Madagascar," *Agricultural Economics*, 2009, *40* (3), 281–294.

Negassa, Asfaw and Robert J Myers, "Estimating policy effects on spatial market efficiency: An extension to the parity bounds model," *American Journal of Agricultural Economics*, 2007, *89* (2), 338–352.

Obstfeld, Maurice and Alan M Taylor, "Nonlinear aspects of goods-market arbitrage and adjustment: Heckscher's commodity points revisited," *Journal of the Japanese and international economies*, 1997, *11* (4), 441–479.

Peterson-Wilhelm, Bailey, Lawton Nalley, Alvaro Durand-Morat, Aaron Shew, Francis Tsiboe, and Willy Mulimbi, "Quality determinates of rice price in open bag markets in Sub-Saharan Africa," *Journal of Agribusiness in Developing and Emerging Economies*, 2023, *13* (3), 361–378.

Ralandison, Tsilavo, Yutaka Arimoto, and Hisaki Kono, "Local search frictions in agricultural markets: A case of rice markets in Madagascar," 2025. Mimeo.

Ravallion, Martin, "Testing Market Integration," *American Journal of Agricultural Economics*, 1986, *68* (1), 102–109.

Tadelis, Steven, "What's in a Name? Reputation as a Tradeable Asset," *American Economic Review*, 1999, *89* (3), 548–563.

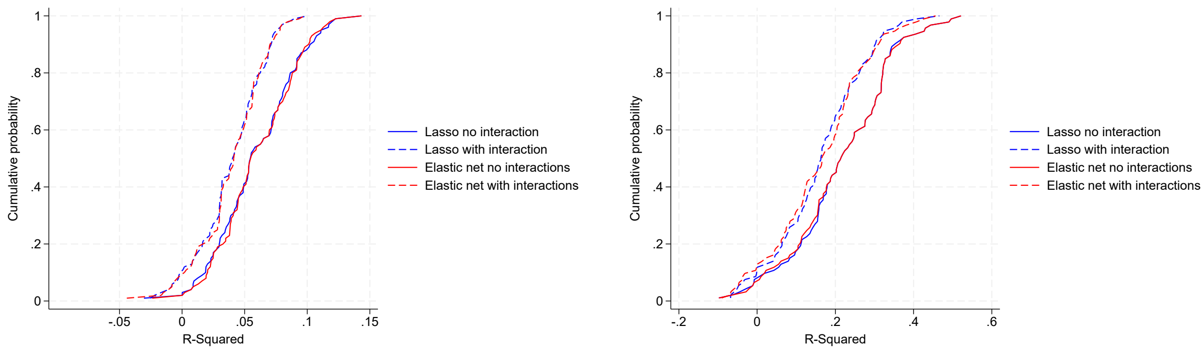
Twine, Edgar Edwin, Sali Atanga Ndindeng, Gaudiose Mujawamariya, Stella Everline Adur-Okello, and Celestine Kilongosi, "Consumer preferences for rice in East Africa," *British Food Journal*, 2023, *125* (13), 316–329.

Unnevehr, Laurian J, "Consumer demand for rice grain quality and returns to research for quality improvement in Southeast Asia," *American Journal of Agricultural Economics*, 1986, *68* (3), 634–641.

A Appendix Figures and Tables

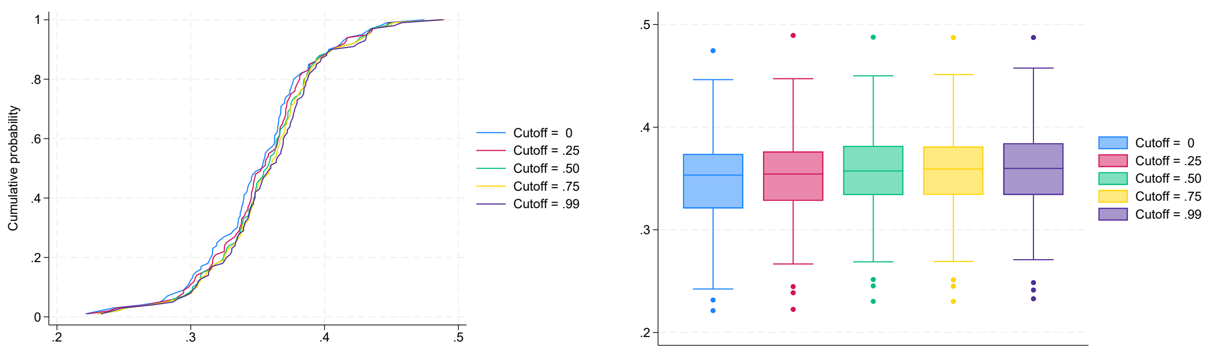
Appendix Figure 1: 100 iterations goodness of fit for week-FE adjusted price

(A) Survey 2023 (B) Survey 2022



Panel (A) and Panel (B) show the cumulative distribution functions of $R^{2,net}$ for out-of-sample predictions based on 100 iterations of CV LASSO and EN regression. The models regress time- and access-fixed-effect-adjusted prices on all quality variables (solid lines) and their interaction terms (dashed lines) for the 2023 and 2022 samples, respectively.

Appendix Figure 2: Out-of-sample goodness of fit for different values of α (LASSO)

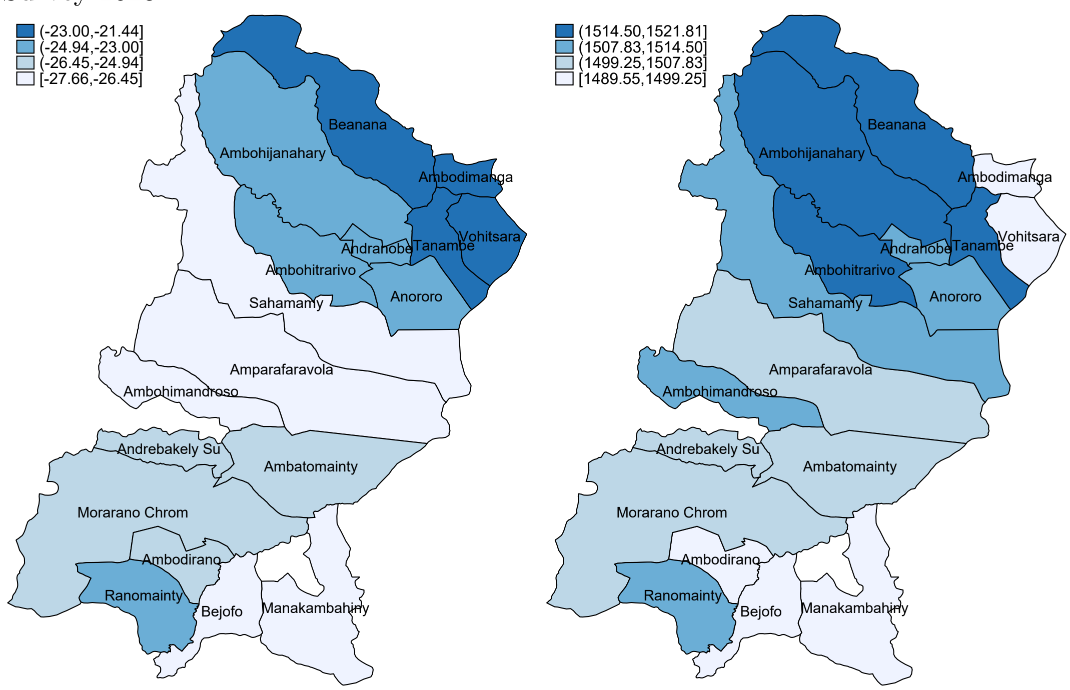


Panel (A) shows the CDF for the cross-validation LASSO $R^{2,net}$ for the out-of-sample prediction for different levels of α, while Panel (B) presents the corresponding box plots for the 2023 sample.

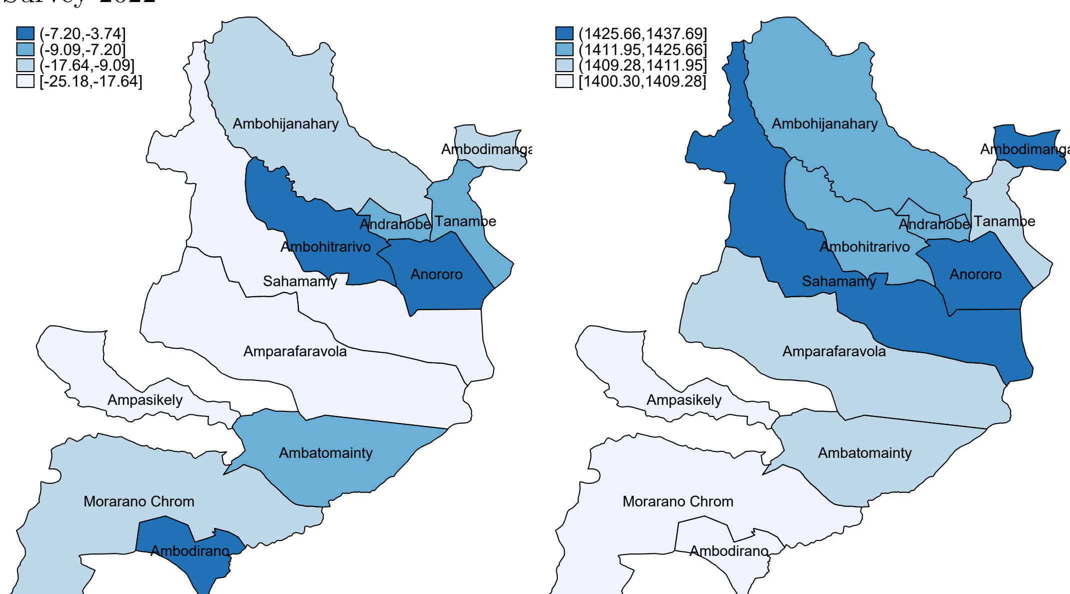
Appendix Figure 3: Geographical distribution of price-relevant quality (q) and quality adjusted price (QP) by commune: 2023

(A) Quality index (B) QP

Survey 2023



Survey 2022



Panel (A) and Panel (B) show the geographical distribution of the estimated price-relevant quality, q_{ijt}, on the left, and the quality-adjusted price, QP_{ijt}, on the right, by commune for the 2023 sample, respectively.

Appendix Table 1: Summary statistics of price measures

(a) 2023 survey

	Mean	SD	Min	Max	N
Price	1499.33	19.41	1420.00	1580.00	1013
Price adjusted	1481.26	15.77	1412.20	1557.11	1013
post-LASSO OLS	1517.89	15.08	1443.32	1594.34	1013
post-EN OLS	1507.75	15.06	1432.76	1584.47	1013
EN-MAA	1506.42	15.22	1433.65	1582.24	987

(b) 2022 survey

	Mean	SD	Min	Max	N
Price	1498.66	64.73	1360.00	1700.00	246
Price adjusted	1401.68	28.89	1279.86	1479.86	222
post-LASSO OLS	1410.98	23.97	1296.15	1498.84	222
post-EN OLS	1411.63	23.93	1297.89	1499.58	222
EN-MAA	1413.75	24.26	1305.36	1508.72	217

Appendix Table 2: ANOVA Trader 2022

	F-statistic	Between-groups MS	Within-groups MS
Price	2.991	9016.715	3014.492
Price adjusted	2.286	1498.664	655.474
QP post-LASSO OLS	2.739	1148.573	419.346
QP post-EN OLS	2.778	1154.072	415.469
QP EN-MAA	2.754	1173.376	426.135