Journal of Development Economics 171 (2024) 103336

Contents lists available at ScienceDirect

Journal of Development Economics

journal homepage: www.elsevier.com/locate/devec

Regular article

Returns to quality in rural agricultural markets: Evidence from wheat markets in Ethiopia ☆

Jérémy Do Nascimento Miguel

University of Bordeaux, CNRS, Bordeaux School of Economics, UMR 6060, F-33600 Pessac, France

ARTICLE INFO

Keywords:
Crops quality
Markets incentives
Agriculture
Africa

ABSTRACT

In many Sub-Saharan countries, farmers cannot meet the growing urban demand for higher quality products. While the literature has focused on production-side constraints to enhance smallholder farmers' output quality, there is scarce evidence of market-side constraints. Using a sample of 60 wheat markets in Ethiopia, I assess whether farmers received a price premium for supplying higher quality outputs. I exploit a unique feature of the data which precisely measures observable and less or unobservable quality attributes, and relate them to transaction prices. Observable attributes cannot serve as proxies for less observable ones. Transaction prices further reflect this, indicating that markets only reward quality attributes that are observable at no cost. However, these results hide cross-market heterogeneity. Farmers engage in relational contracts receive a higher price but similar rewards for quality. Observable quality attributes are better rewarded in markets with more traders per farmer, while unobservable attributes are rewarded in the presence of other value chain actors (i.e., grain millers and farmer cooperatives). Both regression and machine learning approaches support these findings.

1. Introduction

In many Sub-Saharan countries, national production of staple crops fails to meet the needs of local demand (OECD-FAO, 2016). In particular, local smallholder farmers cannot often supply higher quality products that are increasingly demanded by a growing urban population, causing further dependency on imports and gradual exclusion of smallholder farmers from these value chains. Improving smallholder farmers' output quality can be hampered by production-side constraints, through various combinations of market imperfections (e.g., credit, risk, or labor), weak extension systems, and attitudinal factors, as supported by a large literature (e.g., Benyishay and Mobarak, 2019; Karlan et al., 2014; Bold et al., 2017; Suri, 2011; Magnan et al., 2021; Kadjo et al., 2016; Duflo et al., 2011; Carter et al., 2013). Fewer studies have investigated quality issue from the perspective of output markets: the extent to which producers' uptake of quality-improving technologies depends on their expected market returns from it (Suri, 2011; Bernard et al., 2017; Hoffmann et al., 2013; Kadjo et al., 2016; Hoffmann and Moser, 2017; Bold et al., 2022).

Market rewards for higher quality output depend on the extent to which quality is easily and unambiguously observable. Many attributes define an agricultural product's quality. Some are readily observable to the naked eyes, such as size, purity, or color (hereafter *observable quality*), and can therefore be assessed at low cost. Others are only observable at the cost of a dedicated test, such as aflatoxin for maize and groundnuts, or flour-extraction rate for wheat (hereafter *unobservable quality*). Where observable and unobservable quality attributes are strongly correlated, farmers may rely on observable quality to obtain rewards for their investment in enhancing the unobserved quality of their product. When the correlation is weak, further investment is needed to assess unobservable quality (Hoffmann et al.,

☆ I would like to thank Gashaw Tadesse Abate and Tanguy Bernard for their tireless support and contribution to developing this line of research. I am also grateful to Nicholas Magnan, Erwin Bulte, Alain De Janvry, and Elisabeth Sadoulet, for their valuable comments and feedback on this work. I also thank the editor, Patrick Francois, and two anonymous referees for their valuable comments, Timothée Demont, Sarah Jenzen, Jonathan Malacarne, Jeremy Magruder, Louis Olié, Jed Silver, Quentin Stoeffler, as well as conference participants at 2021 NEUDC, UGA Development Workshop, 2022 MIEDC, and 2022 ICDE for their suggestions and comments. I additionally thank Delelegne Abera, Temesegen Keno Belissa, and Dejene Tefera for their work on implementing this project. This research was supported by the French National Research Agency (ANR-19-CE21-0001), the IPORA large research program, the CGIAR Research Program on Policies, Markets, and Institutions (PIM), as part of their flagship program on Inclusive and Efficient Value Chains, and by J-PAL and CEGA's Agricultural Technology Adoption Initiative (ATAI). The study received IRB approval from the University of California Berkeley, United States (study number 20192938). Any errors are the sole responsibility of the author.

E-mail address: jdnmiguel@u-bordeaux.fr.

https://doi.org/10.1016/j.jdeveco.2024.103336
Received 1 December 2022; Received in revised form 4 June 2024; Accepted 2 July 2024
Available online 9 July 2024

2021). Fafchamps et al. (2008) show that costly measures of unobservable quality attributes result in a lower price premium for these attributes and lower investment by farmers to enhance these product characteristics.

This paper provides some of the first empirical evidence of the relationship between quality attributes with different observability degree and related market prices in rural output markets. Assessing the relationship in this context is crucial, where quality certification bodies are mostly unavailable to smallholder farmers (Abate et al., 2021). Leveraging a unique data set on rural wheat farmer transactions, quality attributes, and market characteristics, I estimate whether farmers received a price premium for supplying higher quality output. I find that farmers receive a price premium only for observable quality attributes at no cost. My analysis also contributes to our broader understanding of the effects of local market conditions and their role in enhancing or inhibiting market efficiency, which is substantial.

The starting point is a new microdatabase covering 3485 farmers in 60 rural wheat markets in Ethiopia, collected during the 2019–2020 marketing season. Before selling her wheat to the trader, each farmer gave a subjective measure of her wheat's overall quality level (i.e., high, medium, and low grade), alongside the price obtained after the transaction was completed. Enumerators also collected a 1 kg sample from each farmer and used appropriate equipment to establish independent and precise measures of observable (i.e., purity content), less observable (i.e., moisture content), and unobservable (i.e., flour-extraction rate) quality attributes.[1] I use these measures to compute both an overall objective quality classification (i.e., high, medium, and low grade) and to measure each quality attribute independently. While farmers can easily improve purity content through sorting of their grain before visiting the market, agricultural practices (e.g., seed selection, use of fertilizers and pesticides, harvest technologies, storage conditions) are the main determinants of unobservable attributes.

I study the conditions under which traders will reward quality with a price premium. Intuitively, farmers will receive a price premium only for the observable attribute. However, when observable and unobservable attributes are strongly correlated, farmers can use the former to signal the supply of high-quality unobservable attributes. The data shows a clear positive relationship between the price obtained and overall quality classification (inclusive of all quality attributes). The results hold whether overall objective or subjective quality measures are used, suggesting that buyers recognize wheat quality in markets. Using an econometric approach, I find that a 1% to 8% price premium is observed for higher overall wheat quality. Turning to each quality attribute separately, I find no evidence of a correlation between them, suggesting that farmers and traders cannot use observable attributes as a reliable proxy of less observable ones. Further, while there is a clear positive relationship between price and observable quality measures (1% purer wheat gets a 14% higher price), there is no relationship with less and unobservable attributes (i.e., moisture and flour-extraction rate) despite significant heterogeneity across farmers and the crucial importance of these factors to millers downstream the value chain.

Anecdotal evidence suggests that traders receive a price premium for supplying high-quality unobservable attributes to processors and millers (Abate and Bernard, 2017). Therefore, traders should incur additional screening costs to assess unobservable attributes. These costs include fixed (sunk) costs (e.g., trucks) and variable costs (e.g., hired labor). Hence, I consider different market conditions that favor or inhibit quality recognition (e.g., Casaburi et al., 2013; Casaburi and Reed, 2022; Bergquist and Dinerstein, 2020). In particular, empirical evidence suggests that agricultural markets in Sub-Saharan Africa remain poorly integrated (Moser et al., 2009), face high transaction costs (Aker, 2010; Casaburi et al., 2013), experience unequal levels of competition (Bergquist and Dinerstein, 2020; Macchiavello and Morjaria, 2021) and limited access to infrastructure (De Janvry and Sadoulet, 2020). This implies that favorable market conditions such as market size, competition level, market infrastructure, and institutional arrangements affect buyer's (trader) incentives to measuring quality attributes, thereby increasing the price premium for unobservable quality.

I test these predictions together with three market-level conditions: market-based conditions that include market type (i.e., district market versus secondary market) and market day size (i.e., number of traders per farmer), institutional arrangements (i.e., informal trader-farmer relationship), and other value chain actors presence (i.e., wheat farmers cooperatives and milling plants).[2] I first provide suggestive evidence that traders are more likely to assess unobservable quality (i.e., moisture and test-weight) as market size increases and in markets with cooperatives or millers. Then, the results show a positive price premium on observable quality in central district markets. While higher market size is associated with a higher premium for attributes that can be easily (i.e., purity) and approximately (i.e., moisture) measured, the relationship disappears for the (non-observable) flour extraction rate in a two-stage least square estimate where the daily market size level is partly determined by market-day and weekly variations in rainfall, and religious days celebration.

I further analyze the role of relational contracting in quality recognition. Relational contracting is a relationship between a farmer and a trader involving the promise to trade facilitated by input or credit provision (Bulte et al., 2024). Given that quality supply is imperfectly visible and partly correlated with input usage, local actors may use relational contracting to limit information asymmetry regarding quality level. I show that these farmers receive a higher price but a similar premium for supplying high-quality wheat. This result suggests that farmers receive higher prices to avoid reneging on contracts and price premiums, compensating for their efforts to provide higher-quality output. I further show that differences in quality supply and bargaining power are unlikely to explain this result, yet further research is needed to identify these potential mechanisms exogenously.

Focusing on market infrastructure, the presence of a milling plant in the village's market is positively related to price premiums for flour extraction rate—the attribute millers value the most.[3] In contrast, a cooperative in the village's market is associated with higher prices for both observable and unobservable attributes. Then, I investigate whether traders sorting is a potential mechanism underlying these results. Intuitively, if preferences over quality or local amenities are misaligned between trader groups, then they would sort into different markets. This sorting increases quality premiums and farmers' awareness regarding quality. I find some evidence that differences in local quality supply may yield traders with specific characteristics (i.e., itinerant traders) to sort across markets. This result suggests that sorting along this line may explain my baseline results. In addition to quality, other value-chain actors (i.e., farmers' cooperatives and millers) are related to traders sorting along several dimensions (i.e., experience and social network size). Therefore, sorting can explain heterogeneity in quality premium along these market characteristics.

These findings derived from conventional econometric methods are largely confirmed using a machine learning approach that tests which market conditions and farmer characteristics best predict the price obtained. At market-level, this data-driven approach identifies market size, presence of cooperatives, and milling plants as key characteristics explaining overall price differences. At the same time, grain purity remains the strongest farmer-level predictor of price differences within a market.

[1] Moisture content can be partially – though imprecisely – assessed by breaking wheat kernels.

[2] Cooperatives have an active purchasing role in Ethiopian wheat market, see Minot et al. (2019) and Abate and Bernard (2017) for further details.

[3] These infrastructures are located in the market-location neighborhood.

Together, these results make three main contributions to the literature. First, they offer empirical evidence regarding quality recognition in rural agricultural markets in low-income countries. Existing work suggests that high transaction costs prevent price premiums for unobservable attributes on local markets (Fafchamps et al., 2008; Hoffmann et al., 2013; Abate and Bernard, 2017; Magnan et al., 2021). As a result, traders are willing to pay a price premium only for perfectly observable attributes such as color, visible damage, or grain size (Fafchamps et al., 2008; Minten et al., 2013; Kadjo et al., 2016). I show that the observable quality attribute weakly correlates with the unobservable attributes, preventing farmers and traders from relying on observed purity to signal flour extraction rate or moisture level. In line with previous work, the study also provides evidence that farmers are somewhat, but only partially, informed about the quality of their supply (Kadjo et al., 2016; Anissa et al., 2021). I find additional evidence consistent with the idea that local traders reward only observable quality attributes. This paper differs from Bold et al. (2022), who find that returns to selling high-quality output are null for Ugandan maize farmers in at least two aspects. First, while I assess returns to quality in an observational setting, Bold et al. (2022) estimate the treatment effect of a market access intervention on price returns and do not focus on pre-intervention price premiums. Second, our studies differ in how we each measure quality attributes. Although we both rely on an unobservable attribute (aflatoxins for them, test-weight here), they focus only on an aggregated bundle of observable attributes (e.g., purity, insects) without investigating the potential premium for each one. The findings from this paper provide new evidence suggesting that local traders reward specific attributes rather than overall quality and that focusing on average crop quality could yield underestimated returns to supplying high-quality crops.

Second, this paper contributes to an emerging body of literature on the role of local market conditions in transactions. Limited access to information, insufficient infrastructure, and local institutional arrangements restrict farmers' ability to exploit market opportunities (Aker, 2010; Bergquist and Dinerstein, 2020; Casaburi and Reed, 2022; Deutschmann et al., 2020). Small market size, particularly a lack of outside options for farmers to sell their produce, can reduce market prices and returns for quality. Previous work on quality recognition has failed to consider market conditions (Fafchamps et al., 2008; Kadjo et al., 2016; Magnan et al., 2021). The present paper adds to the literature by studying the interaction between market conditions and price premiums for unobservable and observable attributes. In particular, price premium vary across market size levels for observable attributes only at no or small cost. In addition, I show that while traders do pay a higher price, farmers engage in relational contracting, they do not pay a specific quality premium on top of it.

Third, I provide evidence of the demand-side constraints to agricultural quality upgrading. Public policies tend to concentrate on alleviating supply-side constraints to quality enhancement, through access to extension services, credit, inputs, and risk management devices (Carter et al., 2013; Duflo et al., 2011; Harou et al., 2022; Magnan et al., 2021). However, without explicit recognition of quality in local markets, such policies may fail to generate the kind of sustainable shift towards improving the supply of high-quality crops (Bernard et al., 2017; Bold et al., 2022; De Janvry and Sadoulet, 2020). Recent studies have adopted a demand-side approach and assume that improving local traders' capacity to recognize quality will encourage farmers' supply of higher-quality products (Bernard et al., 2017; Bold et al., 2022; Deutschmann et al., 2020; Abate and Bernard, 2017; Magnan et al., 2021). In a recent randomized controlled study in the Senegalese onion value chain, Bernard et al. (2017) highlight the importance of farmers' expectations regarding market conditions on investments in quality-enhancing inputs. More precisely, they show that while supply-side constraints are unlikely to explain low-quality supply, it can be explained by uncertainty about market rewards for high quality onions. They provide evidence that raising farmers' awareness of changes in local market conditions results in significant and rapid responses by farmers, leading to higher quality crops production. The findings from the present study add to this literature by further describing the role of market conditions in quality returns, distinguishing between observable and unobservable quality attributes.

The remainder of the paper is organized as follows. Section 2 provides additional background information on the Ethiopian wheat market. Section 3 presents the research design and the data used. Section 4 describes the main characteristics of the markets and farmers, and provides an overview of the key variables used in the analysis. Section 5 presents the empirical strategy, followed by the results in Section 6. Section 7 concludes.

2. Ethiopian wheat market

Wheat is one of the most important crops cultivated in Ethiopia, both as a source of food for consumers and as income for farmers. Wheat is grown mainly in the Central and Southern highlands by 5 million smallholder farmers, and it covers over 20% of the cereal production area (Shiferaw et al., 2014; Minot et al., 2019).[4] National demand for processed wheat is growing, driven by urban growth and changes in food habits (Worku et al., 2017). Imports increasingly satisfy this demand, and now represent almost one-third of domestic consumption. Despite significant investment and policies to increase local agricultural output over the last two decades, smallholder farmers remain unable to respond to the growing national demand for higher quality wheat Dercon et al. (2019).

High transaction costs and low quality of smallholders' output are key factors inhibiting development of the Ethiopian wheat value chain (Gebreselassie et al., 2017). Smallholder farmers have limited access to modern inputs such as fertilizer and improved seeds due to incomplete credit markets, an ineffective agricultural extension service, and climate shocks (Dercon and Christiaensen, 2011). Less than 1% of the wheat area is irrigated, making it vulnerable to drought (Seyoum Taffesse et al., 2012).[5] Inadequate infrastructure (e.g., few road networks, poor market information, restricted access to internet and phone networks) increases transaction costs and price volatility and reduces market integration, further contributing to limited market participation of these farmers (Minot et al., 2019). More recently, Ethiopia's agricultural strategy, led by the Federal Government of Ethiopia, has focused on transitioning towards smallholder farmers' inclusion and value chain development (Dercon et al., 2019; Tadesse et al., 2018). A key objective is to promote high-quality wheat production in order to achieve self-sufficiency.

Ethiopia's wheat value chain relies on a large and mostly uncoordinated network of rural middlemen (i.e., traders, wholesalers, brokers) whose influence has increased since the fall of the *Derg* Regime in 1991 (Dercon, 1995; Gabre-Madhin and Goggin, 2005; Gebreselassie et al., 2017).[6] Today, middlemen represent the main wheat buyers in local markets and ensure transportation from production areas to downstream actors such as millers in major urban demand centers (most importantly Addis Ababa).[7] It is often argued that middlemen use their dominant position and informational advantage over farmers to gain market power (Osborne, 2005).

[4] In this study, we refer to smallholders as those farm households cultivating less than 2 hectares.

[5] There are two rainy seasons: (i) the short rainy season (*Belg*) occurs between March and May, while (ii) the long rainy season (*Meher*) is between June and September.

[6] In 1980, the *Derg* government adopted a bundle of measures, called the quota systems, which taxed both farmers and traders, restricted trading licenses, and fixed grain prices. The collapse of the *Derg* regime led to the abolition of these quota systems.

[7] See Figure A.1 for a detailed map of production and market flows.

Formal grading systems and standards exist for many crops in Ethiopia, particularly wheat. Quality assessment and certification, however, are limited to large (often imported) consignments and are of limited use to smallholder farmers given their small transaction sizes (typically 200 kg) and the comparatively large fixed costs of quality assessment (Abate and Bernard, 2017; Anissa et al., 2021; Abate et al., 2021). Hence, spot market bargaining is based on weight and observable attributes (i.e., color, kernel size, presence of foreign matter, varietal mix). Abate and Bernard (2017) note that traders' bargaining is not based on unobservable quality attributes (i.e., flour-extraction rate). As a result, farmers can only increase their income by supplying larger volumes and investing in increase observable quality. Traders aggregate and mix individual farmers' produce and sell the aggregate output to downstream actors (e.g., millers, pasta factories, larger traders).

3. Research design and data sources

3.1. Sample selection and survey

The study was conducted in open-air markets, where smallholder farmers sell their produce mainly to traders. These markets are usually held on a predetermined day of the week throughout the wheat marketing season (Figure A.2A). When they are held on more than one day per week, there is typically a primary market day and a secondary market day. The marketing season starts between October and January according to local agro-ecological conditions, and ends with the long rainy season in June or July. Based on the market sample, Figure A.2B presents marketing season's length distribution (i.e., marketing season), the period during which the market is open regularly. Most spot-market wheat transactions happens during the marketing season, even though small transactions (a few kilograms) can occur out of this period, essentially on retail markets. On average, markets are open 18 weeks a year, ranging from 12 weeks to all year round for wholesale markets. Unique agroecological conditions and topographic-climate combination determines wheat production suitability and season length variation.

The paper uses data collected as part of a broader project conducted in Ethiopia's main wheat-producing areas: Amhara, Oromia, Southern Nations Nationalities and Peoples' Region (SNNPR), and Tigray (Figure A.3).[8] In the 2018–2019 marketing season, a census of all wheat markets in the regions was conducted to collect market-level information such as the estimated number of buyers and sellers, the volume traded, season length, and market facilities. From this census, the main wheat market and a secondary market were selected within each *woreda* (i.e., district). The main wheat market corresponds to the principal market in the *woreda* in terms of volume traded and number of participants. The secondary market was selected within 30 km of the district market. It operates during the same months of the year, but usually on a different day of the week.

In each market, and for two survey rounds, enumerators collected information from 30 selected wheat farmers who came to sell wheat during that day. Before the market day, two enumerators identified the two main market access roads. Then, these two enumerators were posted at the two main market access roads. They randomly surveyed one wheat farmer every 5 to 10 min from among those entering the market, surveying one over five farmers on average. This procedure allows the construction of a representative sample of the farmers commercializing wheat through that entrance at that time.

The first round of the survey was conducted in December 2019 and January 2020 and the second round in March 2020, early in the wheat marketing season and at peak supply time, respectively (Figure A.4). The final sample includes 3584 farmers, 1790 for the first survey round and 1694 for the second.[9]

On any given day, farmers were interviewed twice: once upon entering the market and once upon leaving it (Figure C.1). In the first interview, enumerators collected personal information about the farmers (e.g., age, gender, travel time to market), their overall wheat production (e.g., wheat plot area, volume produced), quantities and expected price for their sales on that particular market day, and self-assessed quality of their wheat (only in the March 2020 survey). The enumerators then purchased a 1 kg sample of wheat from each farmer to be analyzed later. They informed the farmers they would receive 25 Birr (i.e., 0.65 U.S. dollar) if they returned to answer another set of questions upon leaving the market. In the second part, the enumerators collected information on the wheat transactions the farmers had conducted that day, including price per kg and quantity sold.[10]

In each survey round, the enumerators collected market-level information regarding the specific market day as well as other market characteristics (Figure C.3).

3.2. Quality measures

The survey collected two aggregate quality measures: (i) subjective and (ii) objective. *Subjective* quality is based on farmers' perception of the quality of their product and is mainly based on visual inspection and experience. Subjective measures are usually considered inaccurate, while *objective* rely on formal grades and standards established by national or international authorities, assessed with appropriate equipment that is generally unavailable in local markets (Abate et al., 2021). Previous studies have relied on either objective (Hoffmann and Gatobu, 2014; Kadjo et al., 2016; Magnan et al., 2021; Deutschmann et al., 2020) or subjective (Fafchamps et al., 2008) measures of observable and unobservable quality attributes. I combine both approaches. First, the subjective measure is obtained from farmers' self-assessment of the quality of their wheat supply on that particular day. Farmers were asked to classify their wheat on a three-grade scale (i.e., low, medium, high). Second, three quality attributes were objectively measured using the 1 kg wheat sample purchased from the farmers:

1. **Moisture rate** assesses the water content in wheat kernels. This affects seed quality and storage life. Weather conditions during the growing season and storage conditions after harvest affect moisture content. High moisture content decreases the grain's protein content, while low moisture content results in a hard grain with low flour yield.
2. **Test-weight** measures grain density and gives the potential flour yield. It is the most important attribute for the majority of millers producing flour for bakeries.[11] Soil characteristics, weather conditions, agricultural practices and technology adoption affect test-weight. Increasing test-weight is costly for farmers. For instance, they need to apply nitrogen when nitrogen is deficient in their plot, apply it at good timing, which involves assessing soil quality, and use the adequate wheat variety according to their agroecological condition, which is often unavailable in the local market. In addition, harvesting at the "right" time also

[8] The data collection is part of a randomized controlled trial interrupted in March 2020 due to the COVID-19 pandemic. More information on the project summary can be found at Agricultural Technology Adoption Initiative and Agence Nationale de la Recherche.

[9] Note that while the same markets were surveyed twice, different farmers were interviewed across the two survey rounds. Only 58 markets were surveyed in the second survey round due to the COVID-19 pandemic.

[10] All farmers answered both interview parts, even if 1% did not sell their wheat. This high re-interview rate is unsurprising for at least two reasons. First, as the enumerators were posted at the main market entrance, the likelihood that a farmer used the same entrance twice is high. Second, farmers were paid for answering the second set of questions.

[11] Pasta industries are more concerned with protein content and generally seek to purchase durum wheat instead of white (or "bread") wheat. Farmers supplying on the markets of the current study essentially produce bread wheat.

affects test weight, which involves having information on wheat moisture content. Accurate measures are based on the weight of a standard volume of wheat, converted into kilograms per hectoliter —so-called test-weight. A high test weight indicates the grain is well filled, resulting in higher flour yield.
3. **Purity rate** is the share of wheat free of foreign matter such as stone or other cereals in the sample. High purity means that the grain sample is free of foreign elements. A grain sieve is used to separate foreign matter from a 100 g wheat sample. The residues are then weighed to give the rate of purity in the sample.

Enumerators brought wheat samples to the nearest quality-testing booth implemented as part of the broader project mentioned in Section 3.1. Well-trained operators with access to adequate equipment (e.g., hectoliter weight, grain moisture tester, sewing machine, diaphanoscope) were running the testing facility and tested each sample. On average, testing a 1 kg wheat sample takes 15 min and costs 0.4 US dollars to cover shop variable costs or 4.5 US dollars to cover fixed and variable costs, which represent 1 kg and 13 kg of wheat valued at the market price. Each of these dimensions was graded on a three-point scale based on the government's official grading system. An aggregate grade (i.e., low, medium, high) was then computed using the lowest factor approach.[12] This resulted in a minimum quality process, adopted for simplicity, and usable in a real market context.

It is costly and time-consuming for farmers to improve moisture content and test-weight, requiring investment in agricultural practices and technologies at planting and harvesting time. However, farmers can use traditional drying, sorting, and cleaning methods to increase purity levels before going to the market. The distinction between these attributes follows a continuous observability scale from observable to fully unobservable to the naked eye. The extraction rate is defined as an unobservable attribute as it is not readily observable to the naked eye. The lack of access to the required tools impends traders to measure extraction rate (Anissa et al., 2021). Moisture content is easier to observe than test-weight but harder than purity. Some experienced traders chew grains to get a rough idea about moisture content. Purity is fully observable and traders assess it easily at the transaction time by looking in the wheat bag. Although, sieve and scale are the necessary tools to obtain accurate tests, traders rely on visual inspection to assess purity (as informal interviews with traders confirm).

3.3. Traders survey

I mobilize an additional data source collected in a subset of *woreda* to explore potential mechanisms explaining my main results. From the 30 initial *woreda*, 15 markets were randomly drawn from this list. Then, using a list of active traders from the *woreda* trade office, a sample of 12 traders per *woreda* was selected. If a trader was unable to participate, she was randomly replaced. The survey was conducted in April 2022 and consists of 178 traders. More information about this dataset is available in Abate et al. (2023).

3.4. Additional data sources

3.4.1. Precipitation data

I combine this data with daily rainfall estimates obtained from the Climate Hazards Group Infrared Precipitation with Stations (CHIRPS) to identify the causal effects of market size on market price (see Section 5 for more details). Sparse or non-existent ground weather stations in low-income countries have led to increased adoption of satellite rainfall estimates. CHIRPS is a daily precipitation data set developed by the Climate Hazards Group (Funk et al., 2015), which provides information at a 0.5 arc-degree resolution. Dinku et al. (2018) demonstrate that CHIRPS estimates are the most accurate data in Ethiopia (and East Africa), despite lower accuracy in mountainous or coastal areas.

I use market-level precipitation data for the study period (December 2019 to March 2020) to construct instrument variables capturing (i) whether the market day when data was collected was a rainy day and (ii) whether heavy rain (i.e., higher than 10 mm) was recorded in the seven days before the survey date.

3.4.2. Population density data

I relate my market price data to population density using remotely sensed data at *kebele* level.[13] I rely on buildings recorded in Facebook's Data for Good program (Facebook, 2021) to construct population density measure at the *kebele* level. Since each market is localized in a distinct *kebele*, a specific population density measure it provided for each market.

The main advantage of this data over other high-resolution datasets, such as Open Street Maps, is that it consistently covers the whole study region. Maps are built by training a neural network algorithm over house satellite images. The primary output provides a 30-m spatial resolution map showing whether at least one house is found (example in Figure A.5). The map obtained is then combined with available census data and other population datasets to provide population estimates within the selected area. Tiecke et al. (2017) tested this approach to identify building and found it accurate in 18 low-income countries (including ten from Africa). Table 1 presents summary statistics from this data.

4. Descriptive evidence

The following section describes the wheat markets and smallholder farmers in greater detail, as well as descriptive evidence of the quality supply, the relationship between unobservable and observable attributes, and the farmers' perception of their supply quality.

4.1. Open air rural wheat markets

Table 1 presents summary statistics on market characteristics and market day conditions. The top panel displays time-invariant market characteristics such as the presence of price information board, the presence of millers or cooperatives, the length of the season, and market location at national and *woreda* level. Market-day specificities are displayed in the bottom panel, including enumerators' estimates of the number of sellers and buyers on a given day.

Market conditions are heterogeneous. As in Bernard et al. (2013), there is unequal distribution of cooperatives across markets: 60% of farmers have access to a market with a cooperative, and while millers are major wheat value chain actors, only 54% of farmers sell wheat at a market with or close to a mill. Cooperatives role expands beyond input sourcing, production, and marketing aggregation, they also act as buyers —though their share is limited (Abate and Bernard, 2017). Only one market has a price information board. On average, 40 traders and 560 farmers from nearby localities gather on a given market day. I use the ratio of the number of traders per farmer as the main indicator of standardized market size, similar to Krishna and Sheveleva (2017).[14, 15] On average, there are 13 traders per 100 farmers on any given market

[12] Grade and Standard institutions usually rely on the lowest factor approach to aggregate compliance with various standards into a single grade dimension. Following the lowest factor approach, a product is given the quality grade corresponding to the lowest standard satisfaction in any considered dimension.

[13] A *kebele* is the smallest administrative unit in Ethiopia.

[14] To facilitate interpretation, the variable was multiplied by 100 to re-scale.

[15] Ideally, I would have exogenously identify traders' markups, pass-through, entry barriers, or the shape of consumer demand to measure market competition (e.g., Dillon and Dambro (2017), Bergquist and Dinerstein (2020) and Ghani and Reed (2022)). Without data on traders' cost functions, I rely on the number of traders actively buying wheat on market day standardized

Table 1
Market characteristics.
Source: Author's computation based on 2019/2020 wheat markets survey.

	Mean	SD	N
Panel A: time-invariant market characteristics			
Length of the season (weeks)	24.2	14.16	60
Number of supply villages to the market	11.6	14.92	60
Price information board (0/1)	.017	.13	60
Miller (0/1)	.54	.5	60
Cooperative (0/1)	.61	.48	60
Distance to Addis Ababa (kms)	352.05	200.38	60
Distance to district town (kms)	8.05	9.18	60
Kebele Population	16,310	2443	60
Kebele population density (people/km^2)	1876	2442.75	60
Panel B: market-day specifities			
Religious day (0/1)	.07	.26	118
Market day rainfall (0/1)	.25	.44	118
Pre-market week rainfall (0/1)	.14	.351	118
Number of traders	39.94	58.13	118
Number of farmers	560.29	611.69	118
Number of traders per farmer	.13	.15	118

Notes. The table reports time-invariant market characteristics in panel A: market opening length in weeks, the number of villages supplying wheat to a market, the presence of a price information board at the market, milling plant or cooperative in the village's market, the distance to Addis Ababa and to the district capital in kms, the *kebele* population and density (people per square km). Panel B reports information gathered on market-day when surveys were recorded: whether it was a religious day, a rainy day, intense rainfall occurred the week before, the estimated number of traders this day, the estimated number of farmers this day, and the estimated number of traders per farmer this day.

day, albeit with significant heterogeneity. Fig. 1(a) presents the distribution of market size per market-day, distinguishing between main and secondary markets. The distribution is skewed to the right with a lower number of traders per farmer. I find no clear difference in market size across main and secondary markets, despite significant differences in the number of farmers and traders across market types (Fig. 2). This is confirmed by the formal tests presented in Table B.3.

Regarding trader quality assessment practices, purity testing is almost universal, as 99% of the traders surveyed do it. Moisture measurement is less common, with 85% doing it somehow, essentially bite testing. For test-weight, roughly 45% of traders are aware of it but do not test it on spot market. These practices vary across markets with different characteristics. Table B.4 and B.5 show the results. I find a positive association between local market size and measurement practices for unobservable quality. Traders are more likely to assess moisture and test weight when they are in an environment with more peers per farmer. In addition, while the presence of a mill plant in the village's market is uncorrelated with traders' practices, traders are 10 and 26 percentage points more likely to measure moisture and test-weight when a miller plant is present. These results suggest that traders adapt their quality practices to the market environment.

4.2. Smallholder farmers

The sample comprises mostly small-scale wheat producers (Table 2) with an average of 0.98 Ha of cultivated wheat and an average production of 2.7 tons. These figures are similar to those observed by Minot et al. (2019) in their detailed Ethiopian wheat supply chain analysis. Yields per hectare are low compared to the most productive countries at continental and global levels.[16] Smallholder farmers are mainly located in isolated areas and take about one hour to reach the marketplace. Transactions are small: half of the farmers supply less than 50 kg of wheat per transaction, corresponding to one standardized bag.[17] Last, with no formal contracts related to a lack of formal institutions, over half the farmers are involved in relational contracts with traders. Typically, these contracts involve credit provisions and pre-agreed prices. Relational contracts can have several purposes, such as minimizing the risk of contract breach when formal contract enforcement is lacking (Fafchamps, 2001), ensuring access to inputs (Ghani and Reed, 2022), or quality supply (Bulte et al., 2024). In addition, Bulte et al. (2024) find that side-selling threat makes relational contracts more difficult to sustain in higher competitive markets, depleting quality supply. They assume that quality is better rewarded in more competitive spot markets, lowering the incentive for farmers to engage in relationships to comply with their sale commitment, yielding traders who anticipate farmers' opportunistic behavior to reduce their supply of quality-enhancing inputs or credit.

by the number of farmers actively selling wheat as a market size proxy. While Macchiavello and Morjaria (2021) uses the number of mills within a 10 km radius as a competition measure, I prefer to call it market size because I do not observe actors behavior (e.g., collusion, trade cost).

Table 2
Farmers characteristics.
Source: Author's computations based on 2019/2020 wheat growers' survey.

	Mean	SD	N
Farmer characteristics			
Age	36.37	13.58	3484
Female (0/1)	.46	.49	3484
Travel time (min)	58.01	46.14	3483
Agricultural variables			
Wheat acreage	.98	.90	3484
Wheat production (kg)	2723.26	3431.44	3484
Quantity sold (kg)	83.08	129.95	3484
Trader relationship (0/1)	.54	.49	3484
Sold to usual trader (0/1)	.56	.49	3444
Transaction price in birr/kg	13.73	2.21	3444
Objective quality			
Purity (%)	93.40	4.75	2758
Moisture (%)	12.67	2.37	2895
Test-weight (%)	75.33	6.29	2764

Notes. The table reports farmers characteristics: farmer age, gender, travel time to the market in min, wheat area cultivated this season in hectares, total wheat production this season in kgs, the quantity sold on survey day in kgs, whether she has a durable relationship with a trader, if she sold to her usual trader that day, the price per kg obtained from selling wheat that day, the purity content in percent, the moisture content in percent, and the extraction rate in percent.

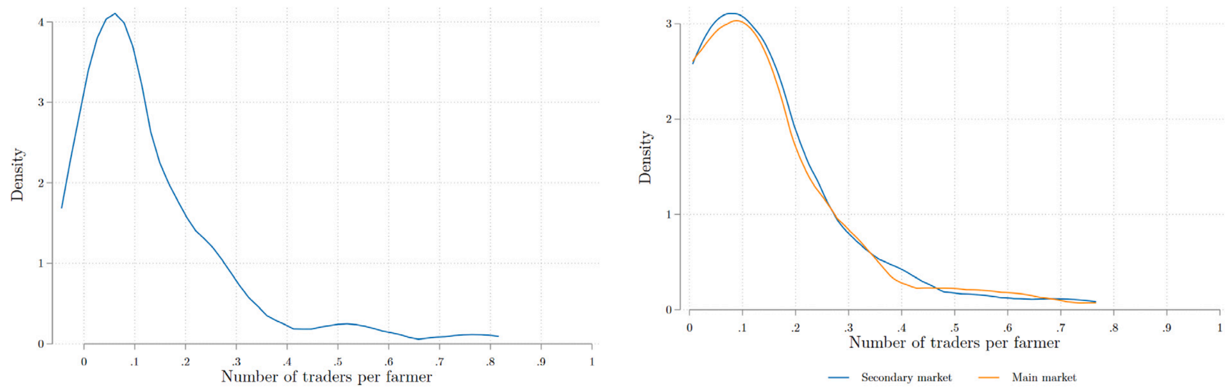
4.3. Quality supply

As explained above in Section 3, enumerators collected samples from farmers on market days and tested them for flour extraction rate (test-weight), moisture content, and purity content to obtain objective quality measures. Based on the overall grade, Fig. 3 shows that 43% of the wheat sample is of high quality, while almost 40% is of low quality at most (with low quality and no-grade grouped together). Quality distribution is consistent across the two survey periods (peak supply time and end of marketing season, see Figure A.6), indicative of no quality-related time-arbitrage (Kadjo et al., 2016).

Turning to each quality attribute separately, Fig. 4 displays their distributions in the sample. As discussed in Section 3, test-weight is

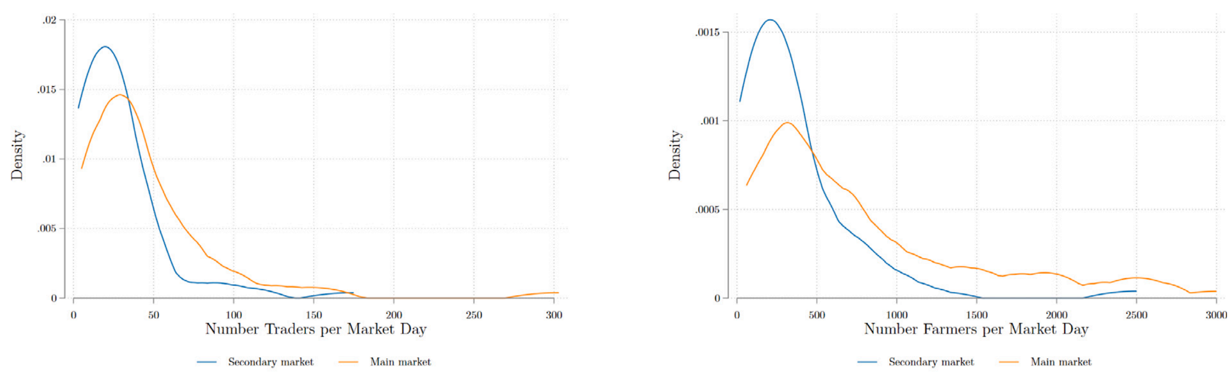
[16] Ethiopia's average yields is equal to 2.9 tons per hectare in 2020, 2.2 and 2.5 times lower than the two continental leaders Egypt and Zambia, respectively, and almost 3 times lower than global leaders such as Belgium and Netherlands (FAO, 2020).

[17] 1% of the farmers do not sell their wheat. They do not sell wheat essentially because price offered was too low or their usual buyer not present. The only observable difference between these farmers and those who sold their wheat is that they bring 12% less quantity on market day.



A. Number of traders per farmer in all markets B. Number of traders per farmer by market type

Fig. 1. Number of traders per farmer in all markets and by market type.
Notes. This figure shows the number of traders per farmer distribution. Number of traders per farmer is the ratio of the number of traders per farmer on a given market day. Panel A. displays the number of traders per farmer in all markets. Panel B. shows the number of traders per farmer in main (district) markets in blue line and in secondary markets in orange line. (For interpretation of the references to color in this figure legend, the reader is referred to the webversion of this article.)
Source: Author's computation based on 2019/2020 wheat markets survey.



A. Number of traders by market type B. Number of farmers by market type

Fig. 2. Number of traders and farmers on market day by market type.
Notes. This figure shows the distribution of the number of market actors on market day. Blue lines represent distribution on secondary markets. Orange lines represent distribution on main (district) markets. Panel A. displays the number of wheat traders across market types. Panel B. shows the number of wheat farmers across market types. (For interpretation of the references to color in this figure legend, the reader is referred to the webversion of this article.)
Source: Author's computation based on 2019/2020 wheat markets survey.

an unobservable attribute, moisture is partly unobservable, and purity content is observable.[18] While less than 1% of the wheat is not graded (i.e., below the lowest quality standard) for purity, the proportion of non-graded wheat reach almost 20% for test-weight and moisture. These differences may reflect the costs associated with producing higher quality for these attributes. While increasing purity is inexpensive (e.g., cleaning and sorting), enhancing test-weight and moisture require additional investment in inputs and practices. The differences may also reflect the absence of a price premium for these unobservable dimensions, reducing farmers' incentive to upgrade quality in these areas. It can also come from farmers' unawareness about unobservable attributes.

Fig. 5 investigates the correlation between observable (i.e., purity) and less observable (i.e., test-weight, moisture) attributes. A high correlation would imply that farmers or traders can rely on observable attributes to (partly) infer the level of unobserved ones (Barzel, 1982). However, no strong relationship can be observed in Fig. 5, such that farmers and traders cannot rely neither on purity to estimate test-weight or moisture level nor on moisture to estimate test-weight.[19]

Next, I investigate the relationship between quality and productivity, as one may suspect a trade-off, at farmer-level, between quality and quantity. For instance, are farmers more likely to supply larger volumes as opposed to higher quality if traders do not pay a premium for high-quality wheat. While I find a weak 15% correlation (but significant at 1%) between productivity and moisture content (Figure A.7B), this not the case for test-weight (3% and non significant correlation). The findings suggest that farmers tend not to specialize in either high-quality or high-volume production.[20]

[18] See Table B.1 for quality attribute thresholds.

[19] The correlation coefficients between purity and unobservable attributes are 0.18 and 0.22 and significant at 5% level for moisture content and test-weight, respectively. The correlation coefficient between test-weight and moisture content is 0.04 and not significant. The literature seems to consistently suggest that correlations below 0.2 and 0.25 are at most very weak and weak, respectively (Evans, 1996).

[20] However, these results should be taken with caution due to potential non-classical measurement errors in farmers' plot size estimation. See for instance Carletto et al. (2013) and Abay et al. (2019) for recent studies

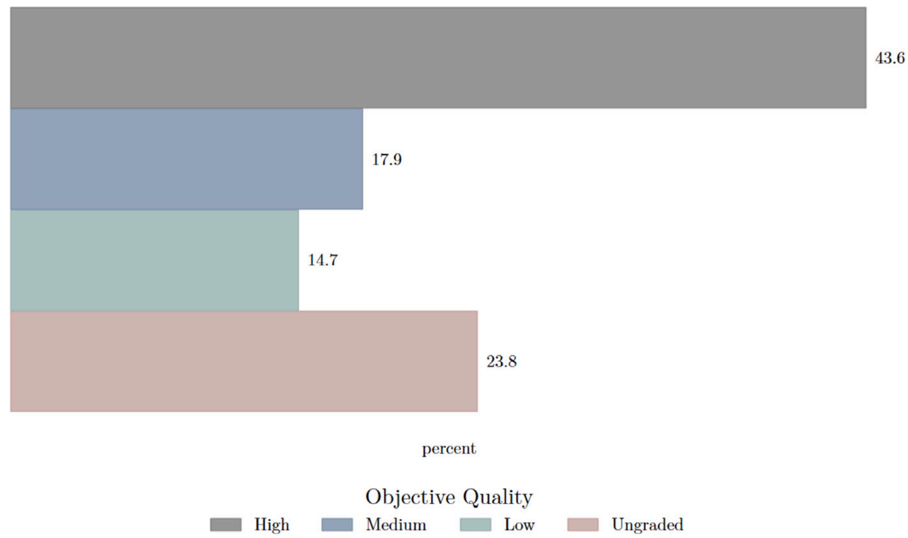
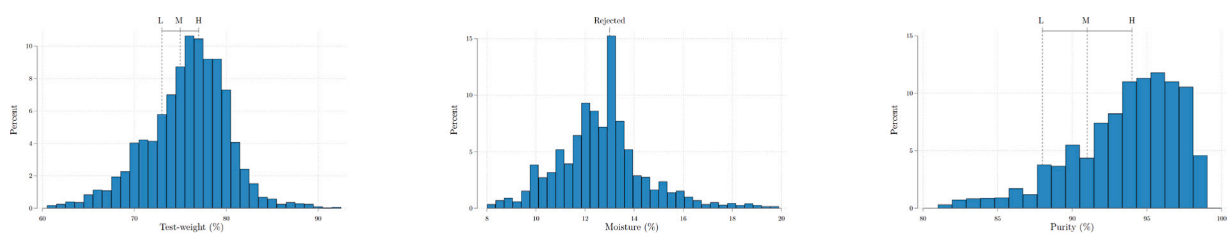
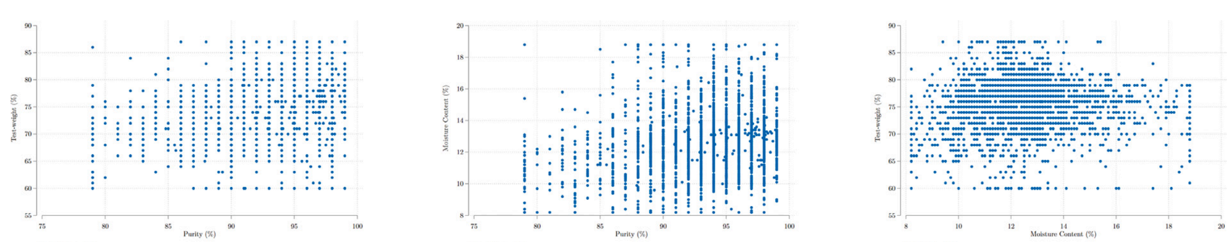


Fig. 3. Objective quality distribution.
Notes. The figure shows the distribution of wheat samples across quality grades based on objective assessment (i.e., laboratory test). The classification relies on three criteria: flour extraction rate (test-weight), moisture content, and purity content to obtain objective quality measure for each. Each of these dimensions was graded on a three-point scale based on the government's official grading system. Then, the aggregate grade (i.e., low, medium, high) relies on the lowest factor approach. Ungraded wheat corresponds to sample with moisture content higher than 13 percent.
Source: Author's computations based on 2019/2020 wheat growers' survey.



A. Test-weight (%) B. Moisture content (%) C. Purity content (%)

Fig. 4. Quality distribution by criteria.
Notes. The figures represents quality distribution for each quality criteria with vertical lines representing the threshold for different quality grades. Purity is the share of wheat free of foreign matter. Moisture is the share of water content in wheat kernels. Test-weight is the potential flour yield. For test-weight: low grade is for values between the two left vertical lines; medium grade is for values between the two right vertical lines; high grade is for values higher than the rightmost line. For purity content: low grade is for values smaller than the leftmost line; medium grade is for values between the two right lines; high grade is for values higher than the rightmost line. For moisture content: wheat is considered as no grade if the result is on the right of the vertical line.
Source: Author's computations based on 2019/2020 wheat growers' survey.



A. Test-weight and Purity content B. Moisture content and Purity content C. Test-weight and Moisture content

Fig. 5. Relationship between quality attributes.
Notes. The figure represents the relationship between quality attributes in each wheat sample. Purity is the share of wheat free of foreign matter. Moisture is the share of water content in wheat kernels. Test-weight is the potential flour yield. Panel A. shows the relationship between purity content and test-weight. Panel B. shows the relationship between purity content and moisture content. Panel C. shows the relationship between test-weight and moisture content. Correlation is the Pearson's correlation coefficient between the two variables considered.
Source: Author's computations based on 2019/2020 wheat growers' survey.

Table 3
Farmers' quality prediction by subjective quality.
Source: Author's computations based on 2019/2020 wheat growers' survey.

Prediction	Subjective quality			
	High	Medium	Low	Total
Accurate estimation %	48.1	16.7	42.6	28.3
Under estimation %	0.0	36.6	51.5	25.8
Over estimation %	51.9	46.7	5.9	45.9
Total %	100.0	100.0	100.0	100.0

Notes. This table shows farmers' quality prediction accuracy according to their subjective quality assessment. Subjective quality is individual perception about the wheat quality sold on the interview day. Prediction is a categorical variable capturing farmers' prediction accuracy: it is equal to accurate if farmer's subjective measure is equal to the objective quality measure; equals to under estimation if a farmer underestimates its quality (e.g., says low quality while true quality is medium); equals to over estimation if a farmer overestimates its quality (e.g., says medium quality while true quality is low).

I then examined farmers' own assessment of the quality of their produce and compare it with the objective estimates.[21] As seen in Table 3, only 28% of farmers accurately estimated the quality of their output: 26% underestimated it, and 46% overestimated it. I observe substantial differences when looking at predictions of farmers operating on spot markets or in relational contracting. Table B.6 shows that farmers engage in relational contract are 12 percentage points less likely to accurately estimate their output and 18 percentage points more likely to overestimate it. This accuracy gap may reflect farmers' beliefs that input or credit access through relational contracts yield to higher quality. Moreover, Figure A.8 displays prediction accuracy across market characteristics (i.e., market type, cooperative presence, and miller presence). Farmers predictions are varying when millers or cooperatives are present in the village's market, which may suggests some knowledge spillovers due to their presence. Thus, in line with Anissa et al. (2021), farmers are somewhat, but only imperfectly, aware of the quality of their supply. At least three reasons may explain this gap. First, farmers rely on an incomplete vector of mainly observable quality attributes for their assessment. Second, farmers might have perceived enumerators as government agents and so overrated their products to satisfy them.[22] Lastly, easier access to input through traders may bias farmers beliefs about the quality supply.

Last, I compare the effective market price farmers obtained by overall objective and subjective grade. As Fig. 6 shows, market prices are positively correlated with both objective and subjective aggregate quality assessment. The figure shows greater price dispersion for objectively higher quality wheat than for lower quality.

5. Empirical strategy

Following the above analytical framework, I describe the empirical strategy to estimate price returns to observable and unobservable quality attributes in rural Ethiopian wheat markets.

5.1. Econometric approaches

5.1.1. Baseline estimation

The price-quality relationship is estimated using the following equation based on ordinary least squares estimates:

$$lnY_{ijkt} = \beta_0 + \beta_1 Quality_{ijkt} + \beta_2 X_{ijkt} + \beta_3 X'_{jkt} + \gamma_j + \mu_t + \epsilon_{ijk} \tag{1}$$

where Y_{ijkt} is the wheat price per kg obtained by farmer i in market j located in *woreda* k at time t. Quality$_{ijkt}$ represents the overall wheat quality measure of farmer i in market j in *woreda* k at time t.[23] Vector X_{ijkt} includes farmer-level variables (i.e., age, gender, yearly wheat production, wheat plot area, travel time to market, wheat type, quantity sold on market day), and the vector X'_{jkt} includes market-level characteristics (e.g., the overall volume traded) at time t. The terms γ_j and μ_t are market and time (i.e., survey week) fixed effects, respectively. Standard errors ϵ_{ijk} are clustered at the *woreda* level.[24] The primary null hypothesis to be tested is whether $\beta_1 = 0$: price do not vary with wheat quality.

Next, I measure the price-quality relationship for every quality attributes using the following equation:

$$lnY_{ijkt} = \lambda_0 + \lambda_1 Q_{ijkt} + \lambda_2 X_{ijkt} + \lambda_3 X'_{jkt} + \gamma_j + \mu_t + \epsilon_{ijk} \tag{2}$$

where vector Q_{ijkt} includes quality attributes (i.e., purity level, moisture content, and test-weight) of farmer i wheat in market j in *woreda* k at time t. The primary null hypothesis to be tested is whether $\lambda_1 = 0$: price do not differ by wheat quality attributes.

Given the importance of farmer-trader relationship on quality supply in imperfect market settings (Bulte et al., 2024), I estimate Eq. (2) separately for farmers trading through spot markets and relational contracting. Comparing price-quality relationship under these two trading channels would be a way to rule out farmer endogenous choice in enhancing unobservable characteristics. If unobservable farmers characteristics are impacting simultaneously agricultural investments and bargaining power on spot markets, those with higher quality should receive a higher premium than those under relational contract. However, the lack of quality testing tools available make this situation unlikely, the most credible strategy to obtain a price premium for unobservable attribute would be to engage in relational contracting. I test whether $\lambda_1 Q^{\text{market}}_{ijkt} = \lambda_1 Q^{\text{contract}}_{ijkt}$: price for quality attributes do not differ on spot market and relational contracting.

5.1.2. Heterogeneity along market conditions

I then examine whether quality recognition varies with market conditions. Two categories of market conditions are considered: (i) market place, and (ii) alternatives to standard market transactions. Market-place conditions are defined as market characteristics directly related to spot market transactions between farmers and traders. I use two measures of market-place conditions: the market type (i.e., district or secondary market) and the market day size (i.e., number of traders per farmer). Alternatives to standard market transactions correspond to the different ways transactions are organized other than through traditional spot market exchanges. These alternatives are measured using two variables: (i) existence of a mill near the market site, and (ii) existence of a wheat producer cooperative on the market site. Quality price premium heterogeneity is estimated by market conditions using the following equation:

$$\begin{aligned} lnY_{ijkt} = \beta_0 &+ \beta_1 \text{Attribute}_{ijkt} + \beta_2 C_{jkt} + \beta_3 (\text{Attribute}_{ijkt} \times C_{jkt}) \\ &+ \beta_4 X_{ijkt} + \beta_5 X'_{jkt} + \gamma_j + \mu_t + \epsilon_{ijk} \end{aligned} \tag{3}$$

where C_{jkt} corresponds to market j condition in *woreda* k at period t. Attribute$_{ijkt}$ represent a given quality attribute (i.e., purity level, moisture content, and test-weight). I include market depth, the product of market day size and daily volume traded to control for potential differences across markets with similar attendance level. The primary

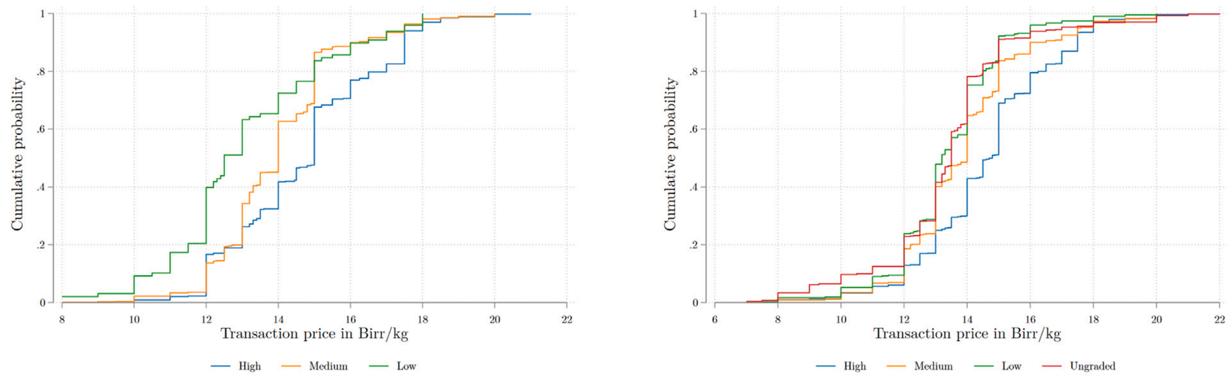
on measurement errors about the inverse size-productivity relationship in low-income countries agriculture.

[21] For comparability purposes with subjective quality, I group low-quality and no-grade.

[22] In line with a social desirability effect.

[23] I further test the relationship using a continuous aggregate quality measure, a normalized inverse-covariance weighted summary index (Anderson, 2008). While this approach can be closer to current traders' quality measurement behavior, its interpretation is less evident from a formal grades and standards system point of view.

[24] Following Abadie et al. (2017), standard errors are clustered at woreda level, which corresponds to the sampling process level.



A. Subjective quality B. Objective quality

Fig. 6. Price (in Birr/kg) by objective and subjective quality.
Notes. This figure shows the transaction price (in Birr/kg) distribution across quality levels. Panel A. relies on farmers' subjective assessment. Panel B. is based on laboratory test measurement. The objective classification relies on three criteria: flour extraction rate (test-weight), moisture content, and purity content to obtain objective quality measure for each. Each of these dimensions was graded on a three-point scale based on the government's official grading system. Then, the aggregate grade (i.e., low, medium, high) relies on the lowest factor approach. Ungraded wheat corresponds to sample with moisture content higher than 13 percent.
Source: Author's computations based on 2019/2020 wheat growers' survey.

null hypothesis to be tested is whether $\beta_3 = 0$: the relationship between price and quality does not depend on market conditions.

However, most of the market conditions are quite plausibly endogenous. For instance, the presence of cooperatives or the market type are likely to be an outcome of past agricultural policies; a farmer's decision to use an alternative to the standard market transaction process depends on unobserved factors and can also affect the return to quality. Related biases in the estimated parameters cannot be eliminated for market alternatives for at least two reasons. First, they can have long-term effects and spillover on farmers' marketing and agricultural performance, and on market transactions. Second, no administrative data or data on the previous marketing season is available to control for non random choices in infrastructure provision. Hence, the interpretation of the corresponding parameter estimates is limited to that of correlations.

Market day size is also (quite plausibly) endogenous for at least two reasons. First, unobservable factors can affect both traders' and farmers' behavior and consequently their market participation. Second, the relationship between market size and price may suffer from reverse causality bias. Indeed, markets within a *woreda* are close, 8 km on average, and this may result in spatial arbitrage by actors in their decision to participate in a given market. For instance, high-quality produce farmers can decide to sell their output in central markets to get a better price. Thus, the exogeneity assumption $E\left[\epsilon_{ijk}|C_{jkt}\right] = 0$ may be violated.

To identify the causal effects of market size on market price, I rely on the occurrence of holy days on market day and pre-week and market day rainfall as instruments for market-day size in Two-Stage Least Square framework. Religious days in Ethiopia are frequent and widely attended (Prunier, 2015). While there are 9 religious days officially recognized, it is widely accepted to take days off around the most important ones such as *Fasika* or *Eid al-Fitr*.[25] Market sales are a source of cash for farmers, thus religious days may increase their participation in markets to finance these celebrations (e.g., to buy specific food items). As market occurs only during morning, religious celebrations are unlikely to prevent farmers participation.

The recent literature has also investigated the relationship between rainfall and agricultural market performance. Rainfall has several implications on farmers' participation in markets and on volume traded due to poor road access (Salazar et al., 2019). Limited access to modern storage is another factor that makes farmers dependent on weather conditions (Hoffmann et al., 2021). For instance, rainfall may lead farmers to sell their wheat earlier than expected to avoid the risk of rot and future losses. Precipitation may also affect traders' participation in the market. If rainfall occurs either during market-day or within a few days before a market day, traders may expect farmers to be more likely to sell wet wheat and thereby increase traders' rot prevention storage costs. Search costs may also be increased as traders need to find a buyer quickly. In such weather conditions, expected net returns could be negative for some traders who may decide not to participate in the market.

I employ a simultaneous two-stage least squares approach, where market size is instrumented by whether the market day occurred on a holy day or on a rainy day, and whether heavy rainfall (i.e., over 10 mm) fell in the previous 7 days. Wheat price heterogeneity is then regressed on the predicted value of market size and the interaction of quality and predicted size as:

$$\begin{aligned}\text{1st Stage}: M_{jkt} &= \theta_0 + \theta_1 Z_{jkt} + \theta_2 (Z_{jkt} \times \text{Attribute}_{ijkt}) \\ &\quad + \theta_3 X'_{jkt} + \gamma_j + \mu_t + \phi_{ijk} \qquad (4a)\end{aligned}$$

$$\begin{aligned}\text{2nd Stage}: lnY_{ijkt} &= \beta_0 + \beta_1 Attribute_{ijkt} + \beta_2 \hat{M}_{jkt} \\ &\quad + \beta_3 (\text{Attribute}_{ijkt} \times \hat{M}_{jkt}) \qquad (4b) \\ &\quad + \beta_4 X_{ijkt} + \beta_5 X'_{jkt} + \gamma_j + \mu_t + \epsilon_{ijk}\end{aligned}$$

With Z_{jkt} indicating the vector of instruments. In the second stage, the wheat price per kg, (Y_{ijkt}), is regressed on the predicted value of market size ($\hat{M}_{jkt}$) obtained from the first stage. The interaction term gives the price premium heterogeneity by market size level.

5.2. Machine learning approaches

I extend the analysis of the quality-price relationship using a predictive model based on machine learning (ML) methods.[26] ML methods

[25] Five of them are Orthodox holidays: *Genna* on January 7th, *Timkat* on January 19th, *Siklet* and *Fasika* in spring, and *Meskel* on September 27th and 28th. Four of them are Islamic holidays and are moveable: *Ramadan*, *Mawlid*, *Eid al-Fitr*, and *Eid al-Adha*.

[26] ML literature uses specific terminology. The sample used to estimate the parameters is the *training* sample. Instead of estimating a model, it is *trained*. Covariates or predictors are called *features*. The dependent variable is referred to as *response* in the context of a regression model.

Table 4
Market price premium by objective and subjective quality.

	Objective quality		Subjective quality	
	(1)	(2)	(3)	(4)
High	0.08*** (0.03)	0.01** (0.01)		
Medium	0.03 (0.02)	0.01 (0.01)		
Low	0.01 (0.02)	−0.00 (0.01)		
High			0.12*** (0.03)	0.08*** (0.01)
Medium			0.08*** (0.02)	0.07*** (0.01)
Controls	Yes	Yes	Yes	Yes
Time FE	No	Yes	No	Yes
Market FE	No	Yes	No	Yes
Mean dep. var.	2.57	2.57	2.56	2.56
N	2901	2901	1676	1676
F-test (High = Medium)				
p-value	0.02	0.21	0.02	0.01

Notes. This table displays price premium per kg for objective (columns 1–2) and subjective quality measures (columns 3–4). The outcome is the price per kg in Birr (in log). Ungraded (low) quality grade is the omitted grade in columns 1–2 (3–4). *P*-value corresponds to the joint hypothesis test *p*-value that the coefficient on high-quality equals medium-quality. Mean dep. var. is the (log) average price per kg for ungraded wheat in columns (1–2) and low quality in columns (3–4). Controls include age of farmer i, gender of farmer i, yearly wheat production of farmer i, plot size of farmer i, travel time of farmer i to market j, type of wheat produced by farmer i, quantity sold by farmer i, and market day volume traded on market j. Standard errors (in parentheses) are clustered at *woreda* level. ***p < 0.01, **p < 0.05, *p < 0.1.

are typically better suited than econometric models when dealing with unconventional data or for the test of economic predictions in low-dimensional settings (Mullainathan and Spiess, 2017). On the other hand, they are more limited with respect to causal identification of parameters (Athey and Imbens, 2019). ML data-driven approaches do not rely on pre-specified parametric approaches resulting in functional form misspecification, but instead learn the relationship between variables directly from the data and optimally choose the parameter estimates over a broad set that is specific to the data.

I apply random forests (RF) and eXtreme Gradient Boosting (XGB) to predict wheat price in Birr per kg and to select the most accurate predictors.[27] I select these algorithms as they are more interpretable than Neural Networks, more versatile than Support Vector Machines, and repeated sampling makes them more accurate (Athey and Imbens, 2019).[28]

The main challenge in ML algorithms relates to their ease of interpretation. To overcome this issue, I present a measure of the importance of each feature, corresponding to the increase in the mean squared error of prediction when a given variable is randomly excluded from the model. A high feature importance increases the mean squared error due to the predictor's omission. However, it does not indicate the sign of the association between the feature and the response (i.e., predicted wheat price). Hence, I compute Shapley values (SHAP) to facilitate interpretation of the XGB results.

SHAP values correspond to the unexplained part of the model for each observation, and the sign of predictors are the association with the response.[29] A positive (negative) SHAP value indicates an increase (decrease) in the overall average predicted response due to the inclusion of a specific feature. A null SHAP value means no deviation from the average mean prediction. In other words, it corresponds to the feature's contribution to the difference between the current and the average prediction. Thus, the higher an absolute SHAP value, the more important the corresponding feature is for the model.

6. Results

In this section, I consider four different cases. First, I test whether the quality measures described in Section 4 are recognized in the market by a premium price. Second, I estimate the heterogeneous effect of quality attributes on price when interacted with market-based conditions. Third, I estimate whether alternatives to standard market transactions can help to enhance quality recognition. Finally, I use machine learning methods to identify the most important predictors of price.

6.1. Quality price premium

6.1.1. Overall grade

I first present results related to quality recognition using objective and subjective quality measures (Eq. (1)). Table 4 shows the presence of a price premium for high-quality wheat using overall quality measures.

Columns (1) to (4) show consistently positive and significant associations between quality and market price, although the introduction of market and time-fixed effects in columns (2) and (4) significantly reduce the point estimates. In the most conservative estimates, I find a 1% price premium per kilogram for objective high-grade compared to ungraded wheat (column 2), and an 8% premium for subjective high-grade wheat compared to low-grade in column (4). Prices for ungraded and low-quality are similar, indicating that traders do not perceive a significant difference in quality between these samples. Overall, these results suggest that farmers supplying higher quality output do receive a higher price.[30] These findings contrast with recent experimental ones in the Ugandan maize markets by Bold et al. (2022). They show that there is a lack of demand for high quality maize in the local markets. More precisely, while they provide evidence that providing services packages raised maize quality, traders did not pay higher prices for better quality products. However, the results from Table 4 do suggest minor differences in price premium between high and medium quality wheat. This may be the result of the aggregation of different quality attributes, which could hide the actual price returns of each of them individually. I examine these in further details below.

6.1.2. Quality attributes

The extent to which an attribute can be observed may play an important role in its recognition in the market (Fafchamps et al., 2008; Hoffmann and Gatobu, 2014; Abate and Bernard, 2017). While it is possible to assess the quality of different crop attributes, testing requires lab equipment. Since a homogeneous volume of grain is needed for the test, the per kg cost of testing decreases with the overall volume of grain to be assessed. Thus, objective quality testing is rarely performed in local markets (Abate et al., 2021), though they are

[27] See Hastie et al. (2009) and Chen et al. (2015) for more details on RF and XGB, respectively.

[28] To estimate the ML model, the features were standardized to ensure that their scale did not influence the feature's importance. The data were then randomly split into training (70%) and test samples (30%) using five-fold cross-validation during training. Next, the wheat price for farmers in the 30% test sample was predicted and the relevant statistics computed (e.g., out-of-sample mean squared error and R-squared). Finally, a grid search was conducted over a range of parameter values during model training, selected to minimize errors.

[29] See Amin et al. (2021) for more details on SHAP values.

[30] These results are robust to alternative objective quality measure (i.e., quality index). Table B.7 shows a 1% premium for a 1 standard deviation increases in quality index.

Table 5
Market price premium for different quality attributes.

	(1)	(2)
Purity	0.47***	0.14***
	(0.14)	(0.04)
Moisture	−0.02	0.01
	(0.07)	(0.02)
Test-weight	0.05	0.00
	(0.06)	(0.03)
Controls	Yes	Yes
Time FE	No	Yes
Market FE	No	Yes
Mean dep. var.	2.61	2.61
N	2712	2712

Notes. This table displays price premium per kg for quality attributes. The outcome is the price per kg in Birr (in log). Purity is the share of wheat free of foreign matter (in log). Moisture is the share of water content in wheat kernels (in log). Test-weight is the potential flour yield (in log). Mean dep. var. is the (log) average price per kg. Controls include age of farmer i, gender of farmer i, yearly (log) wheat production of farmer i, plot size of farmer i, travel time of farmer i to market j, type of wheat produced by farmer i, quantity sold by farmer i, and market day volume traded on market j. Standard errors (in parentheses) are clustered at the *woreda* level. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 6
Price premium for different quality attributes, with heterogeneity by market type.

	Purity (1)	Moisture (2)	Test-weight (3)
Quality	0.12**	0.03	−0.01
	(0.05)	(0.03)	(0.02)
District market × Quality	0.09	−0.02	0.11**
	(0.11)	(0.03)	(0.05)
Controls	Yes	Yes	Yes
Time FE	Yes	Yes	Yes
Market FE	Yes	Yes	Yes
Mean dep. var.	2.59	2.59	2.59
Joint significance	0.00	0.48	0.12
N	2726	2856	2731

Notes. This table displays price premium per kg for quality attributes, with heterogeneity by market type. The outcome is the price per kg in Birr (in log). Quality corresponds to the quality attribute specified at the top of the column. Purity is the share of wheat free of foreign matter (in log). Moisture is the share of water content in wheat kernels (in log). Test-weight is the potential flour yield (in log). District market equals 1 if market j is the district market in the *woreda*. Joint significance is the p-value associated to joint test that linear or interaction coefficients are equal to 0. Mean dep. var. is the (log) average price per kg in secondary markets. Controls include: age of farmer i, gender of farmer i, yearly wheat production of farmer i, plot size of farmer i, travel time of farmer i to market j, type of wheat produced by farmer i, quantity sold by farmer i, and market day volume traded on market j. Standard errors (in parentheses) are clustered at the *woreda* level. ***p < 0.01, **p < 0.05, *p < 0.1.

routinely performed at millers' levels, usually per unit of 5t of wheat (corresponding to the standard "*Isuzu*" truck load in rural Ethiopia). Using tested samples that we obtained from farmers, I can assess the extent to which unobservable quality is accurately perceived by farmers and rewarded by the market.

Using Eq. (2) I estimate the relationship between market price and objectively measured quality attributes. I present the results in Table 5. Of the three attributes, only purity (the easier to observe) is valued by traders (column 1). The estimated coefficients are smaller but remain significant after introducing market and time-fixed effects in column (2). On average, a 1% increase in purity is associated with a 14% price premium—equivalent to 0.84 Birr/kg. In comparison, there is no reward for quality attributes that are harder to observe, whether moisture content or flour extraction rate. Thus, results from Table 5 show that only the observable attribute is rewarded in markets by a price premium.

These results are well aligned with those of other studies in Sub-Saharan Africa. In Benin, Kadjo et al. (2016) find a 3% lower price for insect-damaged maize. In Kenya, Hoffmann et al. (2013) measure an observable quality attribute, discoloration, and an unobservable quality attribute, aflatoxin content. They find that maize prices are strongly correlated with maize discoloration (1% higher discoloration yields a 10% discount), but not with aflatoxin concentration. In Senegal, Bernard et al. (2017) find a 10% premium for higher quality onions. In Ethiopia, Abate and Bernard (2017) used test-weight as an indicator of wheat quality. They find that the average price Ethiopian wheat farmers receive is independent of test-weight level. More broadly, the findings in the present study contribute new evidence to the recent literature on demand-side constraints for quality-upgrading. In line with Fafchamps et al. (2008), attributes measurable with cost are neither valued on markets, nor by farmers themselves. These results speak also at a wider scale than output markets. In Tanzania, Michelson et al. (2021) focus on local input market and find that market prices are orthogonal to observable and unobservable input quality.

6.2. Market-based transactions

Next, I estimate Eq. (3) to examine whether price premiums vary with market-based conditions. I consider two market-based conditions in particular: the type of market (i.e., central and secondary markets) within the *woreda* and the market size on the given market day. The results are presented in Table 6 and show a significant and positive interaction between market type and test-weight on the price farmers obtained. Accordingly, a 1% increase in test-weight is associated with an 11% higher price, but only in district markets. In comparison, while there is a positive price premium for wheat purity, there are no apparent differences across market types. Last, I find no evidence of a relationship between moisture content and prices, on either types of markets.

Existing work on quality recognition in crop markets typically finds no price premium for unobservable attributes (Fafchamps et al., 2008; Hoffmann et al., 2013; Hoffmann and Gatobu, 2014; Abate and Bernard, 2017). Similarly, existing randomized controlled trials find that promoting information about unobservable attributes has a positive impact on price premiums (Bernard et al., 2017; Abate and Bernard, 2017). However, these past studies only consider a single market type. My results show a difference in quality recognition for test-weight between central and secondary markets, suggesting greater buyer interest of this attribute in district markets. It does not however necessary imply easier recognition of this attribute in district markets, a point I return to below.

Next, I consider the relationship between market size level (number of traders per farmer) and quality recognition. This is important as traders' market power can lead to major constraints in investment decisions and quality upgrading (Swinnen and Vandeplas, 2010, 2015). Where traders' market power is high, traders have no incentive to reward quality as farmers have limited outside options. In turn, a larger number of traders per farmer may result in a broader diversity of traders, including those with a higher valuation of higher-quality wheat. However, the existing literature in low-income countries on the topic mainly refers to global and export-oriented supply chains (Reardon and Hopkins, 2006; Swinnen and Vandeplas, 2010). Competition in local markets and market size have also seen a recent rise in academic interest (Dillon and Dambro, 2017; Bergquist and Dinerstein, 2020). While Dillon and Dambro (2017) do not find lack of competition in agricultural markets in SSA, Bergquist and Dinerstein (2020) provide new experimental evidence on imperfect competition among intermediaries from maize markets in Kenya. Given that local markets remain the principal option for farmers to sell their output, it is helpful to measure the extent to which market competition plays a role in quality recognition. For instance, Abate and Bernard (2017) find that Ethiopian

Table 7
Price premium for different quality attributes, with heterogeneity by the number of traders per farmer.

Quality variable:	OLS				2SLS			
	None (1)	Purity (2)	Moisture (3)	Test-weight (4)	None (5)	Purity (6)	Moisture (7)	Test-weight (8)
Number of traders per farmer	−0.14	−3.83**	−0.91**	−1.37***	−0.15	−18.40*	−3.06*	3.49
	(0.09)	(1.46)	(0.35)	(0.27)	(0.15)	(9.84)	(1.55)	(4.25)
Quality		0.06	−0.01	−0.01		−0.29	−0.09	0.08
		(0.04)	(0.01)	(0.02)		(0.21)	(0.06)	(0.10)
Number of traders per farmer × Quality		0.81**	0.30**	0.28***		3.96*	1.09*	−0.90
		(0.32)	(0.12)	(0.07)		(2.17)	(0.61)	(1.00)
Controls	No	Yes	Yes	Yes	Yes	Yes	No	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Market FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Joint significance		0.00	0.04	0.00		0.17	0.17	0.67
Mean dep. var.	2.61	2.61	2.61	2.61	2.61	2.61	2.61	2.61
N	3444	2726	2731	2856	3444	2726	2856	2731
F statistics (First stage)								
Number of traders per farmer					11.99	5.99	13.95	12.5
Interaction term					6.00	13.37	12.59	19.02
Over-identification *p*-value					0.14	0.55	0.22	0.44

Notes. This table displays price premium per kg for quality attributes, with heterogeneity by the number of traders per farmer. The outcome is the price per kg in Birr (in log). Quality corresponds to the quality attribute specified at the top of the column. Purity is the share of wheat free of foreign matter (in log). Moisture is the share of water content in wheat kernels (in log). Test-weight is the potential flour yield (in log). The number of traders per farmer is on a market day. The Table shows the OLS coefficients (columns 1–4) and 2SLS estimates (columns 5–8). In the 2SLS estimations the number of traders per farmer is instrumented based on the occurrence of a religious day, pre-week market day and market day rainfall. In addition, quality attribute is interacted with the previous instruments. Joint significance is the *p*-value associated to joint test that linear or interaction quality coefficients are equal to 0. *F* statistics is the Kleibergen–Paap *F*-statistic for weak identification for the number of traders per farmer and the interaction term. Over-identification *p*-value is the *p*-value for over-identification test of all instruments. Mean dep. var. is the (log) average price per kg. Controls include age of farmer i, gender of farmer i, yearly wheat production of farmer i, plot size of farmer i, travel time of farmer i to market j, type of wheat produced by farmer i, quantity sold by farmer i, market day volume traded on market j, and its interaction with market day number of traders per farmer. Standard errors (in parentheses) are clustered at the *woreda* level. ***p < 0.01, **p < 0.05, *p < 0.1.

wheat growers usually sell and buy food in their local *kebele* market and may therefore be captive to traders.

Table 7 shows how the relationship between price and quality differs with market size. Columns (2), (3), and (4) show that higher number of traders per farmer is positively correlated with market price premium, albeit average prices are lower in the larger markets.

However, as discussed in Section 5, there is concern with respect to the validity of the exogeneity assumption between market size and price. Thus, I rely on a 2SLS strategy to establish identification based on three instruments: occurrence of religious days, whether it rained in the pre-market week, and whether it rained on the market day. The interaction term which captures the heterogeneous effect of market size on quality price premium is also endogenous. Hence, I include interaction terms between instruments and quality attributes as additional instruments (Wooldridge, 2010). I first assess whether the instruments used are good predictors of market size. The results in Table B.8 show that rainfall and occurrence of a religious day have a significant and negative effect on market day size. The F-statistic of the first-stage regression associated with a test of the null hypothesis that all coefficients are zero is reported in Table 7. The F-statistic is satisfactory given the instrument available in the primary estimation in Column (5), indicative of non-weak instruments (Staiger and Stock, 1997). Apart from purity in Column (6), the F-statistics indicate that the instruments are good predictors of market size.[31] The results in Table B.9 show that religious days and rainfall have a positive and significant relationship with farmers' participation, possibly suggesting that farmers may sell more on a religious day to finance religious expenditure. The number of farmers in the market is higher when rainfall occurs pre-week and on market days, in all likelihood in a bid to sell wet wheat to prevent loss from rot. However, only rainfall during the week before market day has a significant and negative relationship on traders' participation. This supports the idea that some traders do not go on the following market day to avoid any additional costs related to the purchase of wet wheat (e.g., storage or screening costs).

Columns (1) and (5) of Table 7 show that a higher number of traders per farmer is negatively correlated with market price, even though not significantly so after accounting for endogeneity. While this result would be pretty surprising in a perfectly competitive market with homogeneous goods, this certainly is not the case in that context. Recent experimental studies document that imperfect competition and collusive agreements prevail in Kenyan and Sierra Leonean local crop markets (Bergquist and Dinerstein, 2020; Casaburi et al., 2013), this could also be the case here. In line with Table 7, Table B.10 provides descriptive evidence that market prices are negatively correlated with the number of traders, though imprecisely estimated. This additional result may suggest that other microeconomic models, such as Cournot, could better describe the Ethiopian wheat market. Data limitations do not allow me to control traders' behavior, size, and market composition within each market, and further data would be needed to test market structure formally (e.g., trader's marginal cost and shape of consumer demand). Nevertheless, standardizing the number of traders by farmers' attendance can capture market heterogeneity prevailing in this environment and partly correct for omitted variable bias. For instance, a larger number of traders per farmer can reflect a lower cost of entry.

The farmer-trader transaction is only the first in the value chain. More traders per farmer at this stage could imply a thinner supply-side market when traders sell their output to the next value chain actor (i.e., miller, broker). Traders may obtain a lower selling price in this context and, as a result, pay a lower price to farmers. Using two measures of market size on the supply-side (i.e., the number of millers and the millers per trader ratio), I further investigate this relationship focusing on the main upstream buyers, millers (Minot et al., 2019). Results in figures A.9 and A.10 provide evidence supporting the relationship between demand and supply-side constraints. Traders facing smaller markets on the demand side are more likely to face thinner markets when selling to millers up the value chain.

Accounting for endogeneity in the number of traders per farmer considerably affects the results. As reported in Table 7, 2SLS estimates

[31] As I define market day size as the number of traders per farmer, the negative relationship may be due to higher farmer participation or lower trader participation.

Table 8
Market price premium for different quality attributes, by trading channel.

	Spot market			Relational contract		
	(1)	(2)	(3)	(4)	(5)	(6)
Purity	0.64*	0.48*	0.16**	0.44***	0.41***	0.11***
	(0.34)	(0.24)	(0.06)	(0.15)	(0.14)	(0.04)
Moisture	−0.03	0.04	0.01	−0.09*	−0.07*	0.01
	(0.14)	(0.12)	(0.02)	(0.05)	(0.04)	(0.01)
Test-weight	0.03	0.01	0.01	0.17	0.17	0.04*
	(0.06)	(0.04)	(0.03)	(0.12)	(0.11)	(0.02)
Controls	No	Yes	Yes	No	Yes	Yes
Time FE	No	No	Yes	No	No	Yes
Market FE	No	No	Yes	No	No	Yes
Mean dep. var.	2.59	2.59	2.59	2.61	2.61	2.61
N	1311	1310	1309	1402	1402	1401

Notes. This table displays price premium per kg for quality attributes, with heterogeneity by trading channel. The outcome is the price per kg in Birr (in log). Purity is the share of wheat free of foreign matter (in log). Moisture is the share of water content in wheat kernels (in log). Test-weight is the potential flour yield (in log). In columns (1–3) the sample is restricted to farmers not engage in relational contract. In columns (4–6) the sample is restricted to farmers engage in relational contract. Mean dep. var. is the (log) average price per kg under relational contract (columns 1–3) and spot markets (columns 4–6). Controls include age of farmer i, gender of farmer i, yearly wheat production of farmer i, plot size of farmer i, travel time of farmer i to market j, type of wheat produced by farmer i, quantity sold by farmer i, and market day volume traded on market j. Standard errors (in parentheses) are clustered at the *woreda* level. ***p < 0.01, **p < 0.05, *p < 0.1.

point to larger price premium for purity and moisture content, as compared to OLS estimates, whereas the effect of test-weight becomes insignificant. In addition, the size of the interaction terms coefficient more than triples in 2SLS estimates in columns (6) and (7) compared to OLS estimates in columns (2) and (3).

These results indicate that price is more sensitive to purity and moisture as the number of traders per farmer increases. Hence, the incentive to supply high-quality wheat is higher in markets with more traders per farmer as traders offer higher price premiums for purer wheat. Since assessment of impurities does not entail additional cost for traders, rewarding purer wheat can be a differentiation strategy for them to secure the best wheat supply in such environment and to subsequently obtain higher prices. While moisture is harder to observe to the untrained naked eye, field observations suggest that experienced traders can approximate it by chewing grain. Thus, some traders can measure moisture free of cost, even if it is an imperfect estimation. Given that higher number of traders per farmer increases demand and alternative trading options for farmers, more traders may be interested in higher-quality wheat to preserve their margins and market share. No such approximation is available for test-weight (flour-extraction rate), in line with the lack of reward in all markets.

Given that relational contracting is widespread and such farmers are less likely to compete on spot markets, I also compute the number of traders per farmer without relational farmers. I explore the same relationship as in Table 7 but restricted my sample to farmers out of relational contracting using this alternative measure. Table B.11 shows the results. Reassuring, I still find a significant negative relationship between the number of traders per farmer and spot market price. In addition, while results are the same for OLS estimates, they differ slightly for 2SLS: even though price is more sensitive to purity and moisture as the number of traders per farmer increases, it is imprecisely estimated for purity. It is worth mentioning that data limitations make these last pieces of results more sensitive. In the same vein as excluding relational farmers from my measure, the same should be done on the traders' side. However, I do not have any information about the share of traders involved in relational contracting. One-side correction for relational contracting is likely to underestimate the number of traders per farmer, suggesting that results in Table B.11 are a lower bound.

These findings align with Bold et al. (2022), who find that the entry of buyers rewarding high quality increases the equilibrium price. In contrast, Bergquist and Dinerstein (2020) find that new entrants will not modify the market environment where pre-existing traders have significant market power as they will join collusive agreements with the incumbents. More broadly, the findings show that market conditions are key determinants of reward for quality. However, they also highlight the limits of market forces in rewarding unobservable crop attributes. Thus, alternatives to traditional market mechanisms can emerge as a second-best solution. These are discussed below.

6.3. Alternatives to standard market transactions and mechanisms

Results thus far suggest that, as a decentralized allocation mechanism, the market fails to reward unobservable attributes with a price premium. Fafchamps (2003) states that formal institutions in SSA are inefficient due to small transaction size. As a result, market actors may use alternative mechanisms to ensure quality provision. These mechanisms can be formal, such as providing agricultural inputs through cooperatives (Bernard et al., 2013; Deutschmann et al., 2020), certification services (Bernard et al., 2017), vertical integration (Deutschmann et al., 2020), or informal arrangements, such as farmer-trader relationships based on trust and repeated interactions (Fafchamps and Minten, 1999; Casaburi and Reed, 2022). I examine these issues focusing on the role of relational contracting and the presence of other value chain actors (i.e., mill and wheat producer cooperative in the village's market). Each one captures a slightly different aspect of farmers' alternatives to spot markets. In addition, I explore the potential role of traders sorting across markets.

6.3.1. Relational contracting

Farmer-trader relationships emerge as a credible alternative to minimize contract breach risk (Fafchamps, 2001). Without protection against opportunistic behavior, constructing personal trust through repeated interactions is often a reliable substitute to market allocations. Traders may use relational contracting to build a personalized relationship reducing uncertainty about quality supply.

I investigate whether farmers under relational contracts receive a specific premium for quality relatively to those selling on the spot market. Using Eq. (2), I estimate the relationship between market price and objectively measured quality attributes separately for farmers under spot market and relational contract. Table 8 shows how the relationship between price and quality differs with marketing channels. Columns (3) and (6) show that farmers receive similar premium under both marketing channels for supplying higher quality wheat. For instance, farmers in relational contracts (column 6) obtain a premium for a higher test-weight (the unobservable attribute) but it is not statistically different (p-value = 0.56) than those supplying on spot markets (column 3).

Homogeneity in quality supply under relational contracts and spot markets may explain the absence of differential premium across these

Table 9
Price premium for different quality attributes, with heterogeneity by the presence of other value chain actors.

	Purity (1)	Moisture (2)	Test-weight (3)	Purity (4)	Moisture (5)	Test-weight (6)
Quality	0.10	0.05	−0.02	0.03	−0.03	−0.00
	(0.06)	(0.03)	(0.02)	(0.07)	(0.02)	(0.03)
Millers × Quality	0.08	−0.04	0.09**			
	(0.08)	(0.04)	(0.04)			
Cooperative × Quality				0.17**	0.08**	0.07*
				(0.08)	(0.03)	(0.04)
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Market FE	Yes	Yes	Yes	Yes	Yes	Yes
Mean dep. var.	2.62	2.62	2.62	2.63	2.63	2.63
Joint significance	0.00	0.24	0.09	0.01	0.02	0.13
N	2726	2856	2731	2726	2856	2731

Notes. This table displays price premium per kg for quality attributes, with heterogeneity by the presence of other value chain actors. The outcome is the price per kg in Birr (in log). Miller equals 1 if a miller is present on market j. Cooperative equals 1 if a cooperative is present on market j. Quality corresponds to the quality attribute specified at the top of the column. Purity is the share of wheat free of foreign matter (in log). Moisture is the share of water content in wheat kernels (in log). Test-weight is the potential flour yield (in log). Joint significance is the p-value associated to joint test that linear or interaction quality coefficients are equal to 0. Mean dep. var. is the (log) average price per kg in markets without miller (columns 1–3) and without cooperative (columns 4–6). Controls include age of farmer i, gender of farmer i, yearly wheat production of farmer i, plot size of farmer i, travel time of farmer i to market j, type of wheat produced by farmer i, quantity sold by farmer i, and market day volume traded on market j. Standard errors (in parentheses) are clustered at the *woreda* level. ***p < 0.01, **p < 0.05, *p < 0.1.

channels. These contracts may involve inputs and loans provision before harvest conditional on future transaction at harvest time.[32] Relational contacts may rise crop value incentivizing farmers to renege on the contract and side-sell their output on the spot market. In other words, farmers engage in relational contracts who are producing higher quality output may be more likely to side sell on the spot markets. There is mixed evidence regarding relational contracts contribution in enhancing crop quality. While Fafchamps and Minten (1999) find that quality provision is not central in relational contracts for Malgassy crop traders, Macchiavello and Morjaria (2021) show in the Rwandan Coffee value chain that relational contracting is used to sustain quality supply through services provision. Closer to my context, Bulte et al. (2024) find that Ethiopian wheat farmers involved in relational contracting supply higher quality output. Table B.12 shows that farmers in relational contracts supply wheat with higher test-weight than those on spot markets. In addition, purity and moisture content are similar for relational contracts and spot markets farmers. This suggests that homogeneity in quality supply may explain price premium similarity.

A second possible explanation for the absence of a relationship between marketing channels and price premiums is that the average price is higher under relational contracts. Table B.13 shows that the average price for farmers under relational contract is two percentage points higher per kilogram than those supplied on the spot market. Farmers engage in relational contracts receive higher prices, avoiding reneging contracts and price premiums, compensating their efforts to provide higher quality output—though similar to premiums observed on the spot market. Then, I investigate whether bargaining power differences may explain this price difference. If farmers decide *ex-ante* what quality they want to produce, some unobservable characteristics affecting this choice might also impact the price one can obtain.[33] Given that prices under relational contracts are negotiated at the contracting time, only farmers supplying on the spot market can bargain for higher prices. To test for this possibility, Fig. 7 plots price residuals dispersion under relational contracts and spot markets. Unexplained price dispersion is higher under relational contracts than on spot markets.[34] This result suggests that bargaining power is unlikely to explain average price differences and similar returns to quality.

My results do not rule out that other farmers' and traders' motivations may explain the absence of premium differentials under both channels. For instance, repeated interactions, social norm conformity, and easier access to credit or input may incentivize farmers to supply higher-quality output to sustain this relationship. In addition, relational contracts involve repeated interaction before transaction time (e.g., input provisions), potentially reducing production costs. Therefore, narrowing returns to quality measurement to transaction price may represent a lower bound estimate for relational contract farmers when exact contract terms information are not available. This result suggests several avenues for further research to understand the mechanisms incentivizing farmers and traders to sustain such relationships while preserving crop quality.

6.3.2. *Presence of other value chain actors*

Value chains, like the wheat value chain in Ethiopia, can be long and involve many intermediaries (Osborne, 2005). Intermediaries increase final costs as each agent expects to make a profit. However, intermediaries are not the final buyers of the goods, and their demand for quality only depends on that of downstream value chain actors. For example, millers are the main end-buyers of wheat before its transformation into flour. Their demand is driven mainly by quality, as purity, moisture, and flour extraction rate (test weight) significantly affect the volume and quality of the flour. Thus, the presence of a mill near local markets is expected to reduce the length of the value chain and result in higher prices for higher quality wheat.

Then, I investigate the relationship between returns to quality and the presence of cooperatives near the market site. From field observations, cooperatives are often interested in higher-quality wheat that they aggregate under the cooperative's brand name. Several assess the quality of an individual farmer's wheat before aggregating it with others. However, from a farmer's perspective, selling to a cooperative has drawbacks in that payment is often made with a month's delay.

[32] I lack data about exact contract terms. This information relies only on fieldwork discussion with traders and farmers.

[33] While it could be the case for observable attributes, it is unlikely to be the case for less observable characteristics because they are uncorrelated to each other, and traders do not measure them. Results in Table 5 are robust to observable farmers' characteristics and the inclusion of time and market unobservable characteristics. Thus, unobservable farmers' characteristics are unlikely to drive these effects.

[34] I measure how much variation in price is left on a given market day after controlling for market and time fixed effects and farmers' characteristics and quality supplied. Relying on the specifications estimated in columns (3) and (6) of Table 8, I estimate their residuals and compare their distributions in Fig. 7.

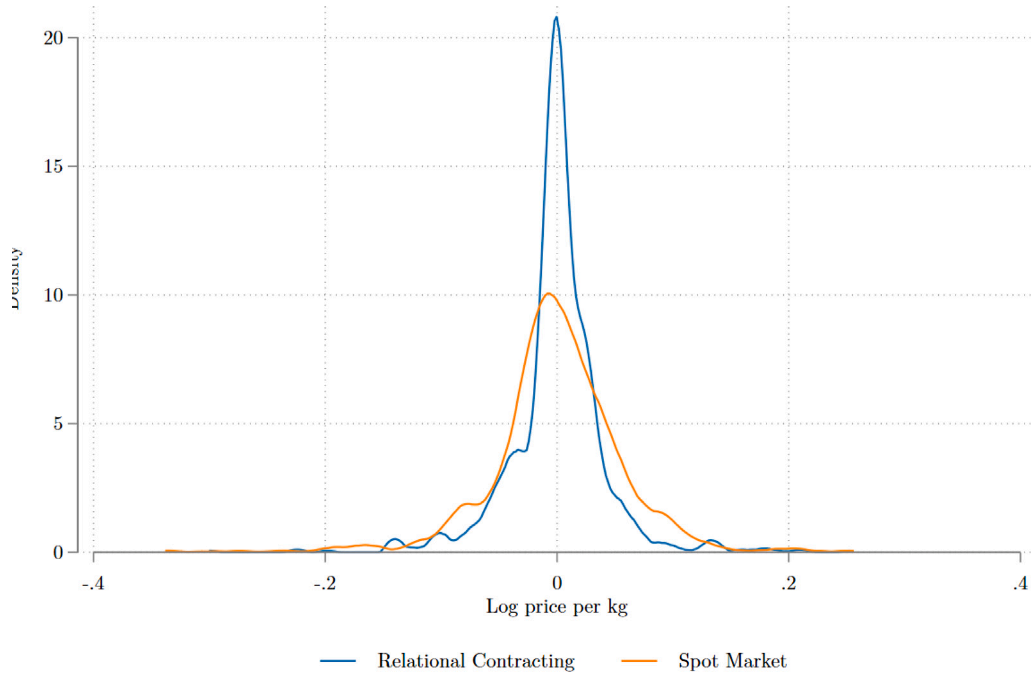


Fig. 7. Dispersion of (log) wheat price per kg, by marketing channel.
Notes.: This Figure shows the price dispersion of (log) wheat price by marketing channel after controlling for farmers characteristics, quality supplied, and market and time fixed effects. The blue curve shows price dispersion for farmers under relational contracts. The orange curve shows price dispersion for farmers selling on spot markets. (For interpretation of the references to color in this figure legend, the reader is referred to the webversion of this article.)
Source: Author's computations based on 2019/2020 wheat survey.

Table 10
Share of itinerant traders across markets with different level of quality attributes.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Test-weight	−0.00	0.02					0.00	0.02
	(0.02)	(0.02)					(0.02)	(0.02)
Moisture			−0.00	−0.05			0.00	−0.07**
			(0.04)	(0.03)			(0.05)	(0.03)
Purity					−0.04	0.07	−0.05	0.10
					(0.09)	(0.05)	(0.10)	(0.06)
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Zone FE	Yes	No	Yes	No	Yes	No	Yes	No
Woreda FE	No	Yes	No	Yes	No	Yes	No	Yes
Mean dep. var.	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
N	2701	2701	2696	2696	2826	2826	2682	2682

Notes. This table shows role of quality supply in itinerant traders decision to locate in market j. The outcome is the share of itinerant traders at the market level. Purity is the share of wheat free of foreign matter (in log). Moisture is the share of water content in wheat kernels (in log). Test-weight is the potential flour yield (in log). Mean dep. var. is the share of farmers in relational contract in a market. Controls include age of farmer i, gender of farmer i, yearly wheat production of farmer i, plot size of farmer i, travel time of farmer i to market j, type of wheat produced by farmer i, quantity sold by farmer i, and market day volume traded on market j. Standard errors (in parentheses) are clustered at the *woreda* level. ***p < 0.01, **p < 0.05, *p < 0.1.

Table 9 shows the heterogeneous price premium for each quality attribute by the presence of other value chain actors. The effect of the presence of millers in the village's market is first investigated. The findings are presented in columns (1) to (3). No additional rewards are paid upon the presence of a miller for moisture and purity attributes. In contrast, results in column (3) point to a positive reward for unobservable quality (test-weight) with the presence of a nearby mill. In these markets, a 1% increase in test-weight score leads to a 9% price premium. In Ethiopia, millers pay significant attention to flour-extraction rate. Two bundles of wheat, identical in observable attributes (e.g., purity), may exhibit substantial differences in flour-extraction rate, thereby affecting millers' final profit (Abate and Bernard, 2017). The presence of an on-site mill may affect rewards to such attributes through informational effects and reductions in the length of the value chain that otherwise dilute the incentive to procure higher quality wheat.

Lastly, the results in columns (4) to (6) of Table 9 show a positive effect of the presence of a cooperative on price rewards for all the quality attributes, whether observable or unobservable. On average, when there is a cooperative, a 1% increase in quality is associated with a price premium of 17% for purity, 8% for moisture, and 7% for test-weight. Cooperatives play a substantial role in rural markets by providing fertilizers and seeds on credit (Deutschmann et al., 2020; Bernard et al., 2008). Hence, farmers with access to cooperatives in the market may benefit from such agricultural technology and produce higher quality wheat. Indeed, 89% and 60% of Ethiopian farmers with access to cooperatives purchase fertilizers and seeds, respectively (Abate and Bernard, 2017). In Ethiopia, cooperatives usually provide quality assessments when they purchased output. Once aggregated, cooperatives may either

Table 11
Traders sorting across markets with different market characteristics.

Outcome variable:	% itinerant (1)	Trader's experience (2)	Social network size (3)
Cooperative	−0.07	−0.15*	0.08
	(0.05)	(0.08)	(0.16)
Millers	0.11**	0.33**	−0.36***
	(0.05)	(0.12)	(0.10)
Time FE	Yes	Yes	Yes
Zone FE	Yes	Yes	Yes
Mean dep. var.	0.25	2.29	2.52
N	3413	1739	1739

Notes. This table shows the role of other value chain actors in traders sorting in market j. The outcome in column (1) is the share of itinerant traders in market j, logarithm of traders' average experience in column (2), and logarithm of traders average social network size in column (3). Miller is a dummy equals 1 if a miller is present on market j. Cooperative is a dummy equals 1 if a cooperative is present on market j. In columns (2–3) sample is restricted to *woredas* included in Abate et al. (2023). Standard errors (in parentheses) are clustered at the *woreda* level. ***p < 0.01, **p < 0.05, *p < 0.1.

resell bulked wheat to millers or produce flour themselves. It is worth mentioning that selling output through a trader is the main marketing channel (Minot et al., 2019), and that if any selection had occurred, higher quality farmers would have chosen the spot market instead of the cooperative (Abate and Bernard, 2017). These results may inform cooperatives' role in upgrading quality in local markets.

6.3.3. Traders sorting

I am considering whether traders sorting across markets can be a potential mechanism through which quality is rewarded. Such sorting is essential to understand whether traders' ability to detect quality and interest in quality or general market forces can explain my results. While individual or firm sorting has been widely analyzed in developed countries and urban areas of developing countries, much less is known about sorting in the rural context of developing countries (e.g., Gáfaro and Pellegrina, 2022; Sayre, 2023).

Sorting would have happened if some market characteristics (e.g., local quality) determine the location decision of differential traders, that is, spatial sorting. The market-level data provides limited information about trader characteristics except for the number of resident and itinerant traders on a given market day. The main difference between itinerant and resident traders originates in their marketing process: itinerant traders buy wheat at the farm gate at a lower price, whereas resident ones are at the marketplace. Using itinerant traders' share in local markets gives a first insight into local market composition. For instance, a lower share could capture extremely well-connected markets where the lion's share of farmers travel to nearby markets, making itinerant trade unprofitable. However, such a measure provides only a limited, restrictive sense of sorting. Therefore, I rely on an additional dataset including more comprehensive information about trader characteristics, but that covers only 12 *woredas*.[35] Using this trader-level data, I look at potential sorting based on traders' trading years of experience and social network size outside the *woreda*. Traders obtain most of their human capital on the job and have more experience facilitating trading operations through, for instance, better information access or environment knowledge. Moreover, social networks constitute an element of social capital yielding efficiency gains on the job through transaction costs reduction (Fafchamps and Minten, 2002).

I first investigate if traders' location decisions are correlated with local quality supply. Table 10 shows the sorting results when I use the share of itinerant traders. Overall, I find some suggestive evidence supporting sorting in my preferred specification (column 8). There is a lower proportion of itinerant traders in districts with higher moisture content. This result can suggest that higher moisture increases costs for itinerant traders (e.g., sell quickly, reduce storage time), making this activity less profitable. Then, Table B.14 provides the sorting results relying on traders' experience in Panel A and social network size in Panel B. I do not find any evidence that traders with different experience or social network sizes are located in markets with different quality supplies. However, I show in Table B.15 that this potential sorting does not translate into a differential premium for quality.

Given that quality premiums vary across the presence of other value chain actors, such as cooperatives or millers (Table 9), I examine whether traders' characteristics vary across markets with cooperatives or millers. Table 11 displays the results: the share of itinerant traders is the outcome in column (1), the logarithm of traders' average experience in column (2), and traders' social network size in column (3). I find some evidence suggesting spatial sorting across markets with different value chain actors. Miller's presence matters in traders' decision to locate in a specific market: positively associated with the proportion of itinerant traders and traders' experience but negatively for traders with the broader social network. The role of a cooperative in traders' decisions is more nuanced, with less experienced traders operating in a market where a cooperative is present. These results identify sorting as one potential mechanism explaining heterogeneous price premiums across markets with different value chain actors.

6.4. Additional results and robustness checks

6.4.1. District specialization in wheat production

It is possible, of course, that some *woredas* drive the results observed in Tables 4 and 5. For instance, price premium might be lower (higher) for observable (unobservable) characteristics in wheat producing *woredas* because of the high volume supply. To address this, I separate *woredas* into two groups based on wheat specialization. A *woreda* is specialized in wheat production if wheat is the cereal with the highest share of total cultivated land area. Table B.16 shows the results for objective quality measures. Overall, doing this does not change the conclusion: farmers receive a price premium for higher quality wheat and attributes easier to observe. While the price premium paid for purest wheat is positive in both specialized and not specialized *woredas*, the coefficient is slightly lower and imprecisely estimated in the former—surely because of a smaller sample size. It is possible, for instance, that quality standards to obtain a premium are lower in unspecialized districts, as wheat production is smaller in these districts.

6.4.2. Market location characteristics

Several studies in SSA show that the geographic location of rural markets affects equilibrium prices (Vandercasteelen et al., 2018; Aker, 2010; Minot et al., 2019). Here, the question is examined using geographic and demographic variables related to market environment. Each variable captures a slightly different dimension. The first is based on the market's physical distance from Addis Ababa, the main demand center. The second captures the potential link between market price and population density (Bernard et al., 2008). In the most densely populated *kebele*, markets might be better integrated into the regional or national wheat market. These areas also derive substantial benefits from their positions in terms of economies of scale, which can reduce transaction costs. Areas with higher population density are also likely to be more urbanized and thus be subject to greater demand for quality (Vandercasteelen et al., 2018).

The results are presented in Table B.17. They point to an association between a market's geographical characteristics and a price premium for unobservable quality. In column (3), I find a positive interaction between distance to Addis Ababa and reward for unobservable quality. With distance to Addis Ababa possibly correlated with differences in soil quality across market locations (and therefore unobservable quality), caution should be taken in interpreting the result as market-driven. However, as only the interaction term is significant (and not test-weight alone), it confirms that the result is market-driven rather than due to differences in soil quality. According to population density, return on unobservable quality is higher in most populated areas in column (6).

[35] See Abate et al. (2023) for more details on the data.

6.4.3. Marketing time

Many smallholder farmers must deal with liquidity issues at harvest time to pay back agricultural loans or satisfy essential needs such as food or school fees (Stephens and Barrett, 2011; Dillon, 2020). Moreover, without access to affordable and efficient storage technology, stored outputs may suffer severe damages from fungi, rodents, mold, and insects. For these reasons, price premium on various quality attributes may differ across the dates of the survey rounds from which the data were obtained. The results are presented in Table B.18. Overall, I find only limited evidence that the transaction date is associated with differential rewards to quality. The results in column (1) suggest that traders pay a price premium for the purest wheat supplied. However, purity is not rewarded later in the commercialization season. This closely aligns with earlier work by Kadjo et al. (2016) on the rural maize sector in Benin.

6.4.4. Robustness check

In addition, I rely on the post-double selection (PDS) LASSO procedure presented in Belloni et al. (2013) to ensure that the choice of control variables did not bias the result. The main advantage of PDS is that it picks control variables consistently and avoids standard errors estimation issues. Table B.19 shows the association between quality and price, independent of market conditions as above in Table 5. As shown in Table 5, a price premium is only paid for purity. Table B.20 shows the association between price and quality by market type. The results are similar to those of Table 6. Table B.21 presents the results for the association between price and quality with heterogeneity by alternatives to market mechanisms. The results are identical to those observed in Table 5.

6.5. Identifying the most important price determinants, a machine learning approach

Previous work has suggested various farmer-level solutions to increase local agricultural prices (Karlan et al., 2014; Bergquist and Dinerstein, 2020; Casaburi and Reed, 2022). Often, this literature assumes that the main barriers to increasing price might be overcome at individual level. However, such interventions may have a limited impact if market conditions are the main price determinants. For instance, lack of infrastructure, limited information, and poor value chain integration may prevent farmers from obtaining higher prices, and farmer-level intervention will do little to overcome them. Here, I examine whether wheat price is more likely to be determined by market or by farmer characteristics.

Table B.22 presents the out-of-sample root mean squared error (RMSE) and square of the Pearson correlation coefficient for wheat price. There is little differences in performance between random forest (RF) and eXtreme Gradient Boosting (XGB). However, the XGB model appears more accurate as the confidence interval is smaller than for RF. All the features listed in Tables 1 and 2 were used.

My aim is to determine which features are the most predictive and the direction of the association with the response. Fig. 8 plots the Shapley values of the fifteen most predictive features using XGB. The SHAP values and the features are placed on the horizontal and vertical axis, respectively. Each dot represents a farmer. The average contribution of the corresponding variable in price prediction is on the vertical axis. A positive (negative) SHAP value represents an increase (decrease) in the predicted price across all possible combinations of the predictors. For instance, the "market volume" feature decreases the predicted values (the SHAP value is negative) for most observations when included in the model. Lighter colors imply smaller values of the feature: lower values of volume traded on the market are observed where SHAP is positive. However, it is not easy to fully understand the association between the feature and the predicted price from Fig. 8 alone.

Fig. 9 displays an astute way to visualize the association between features and predicted prices. Average SHAP values are plotted, then colored by the correlations between the feature and its SHAP values. Distance to Addis Ababa, the survey week, and the number of traders on the market day have high positive associations with wheat price, whereas the volume traded on market day, the number of farmers, and the absence of a cooperative or a miller have a negative relationship. Moreover, the quantity sold by farmers, the purity and the test-weight values are positively correlated with price. Most of the best wheat price predictors are market condition characteristics rather than farmer characteristics (i.e., quantity sold, purity, and test-weight). Otherwise, quality attributes (i.e., purity and test-weight) are among the most important price predictors. These results support previous ones, underscoring the importance of market conditions in the analyses of price premiums for quality and farmers' incentive to invest in improving the quality of their output. These warrant future research in better measuring and quantifying the potential effect of market conditions variation on quality price premium.

7. Summary and concluding remarks

Food crop quality is one of the main concerns that Sub-Saharan African countries must address to improve revenues for smallholder farming and thereby contribute to reduce poverty. A large number of empirical studies have considered supply-side approaches to alleviate farmers' constraints in quality-upgrading, such as liquidity, risk, information, and technology access (De Janvry and Sadoulet, 2020). Following recent empirical papers focusing on demand-site constraints (Bernard et al., 2017; Abate and Bernard, 2017; Bold et al., 2022), the present study presents evidence that imperfect market recognition of quality must be addressed to enhance quality supply. Using original survey data collected in 60 Ethiopian wheat markets, I examined the extent to which quality is rewarded in the Ethiopian wheat market. I found that farmers imperfectly interpret the quality they supply, and are imperfectly rewarded for their higher-quality wheat. While a significant price premium is paid to farmers for purer wheat (i.e., observable attribute), I find that low moisture content and test-weight (i.e., unobservable attributes) are not rewarded. This finding supports the idea that quality factors for unobservable attributes are not a current concern for traders.

Previous studies have implicitly assumed that farmers sell their output on homogeneous markets. However, market conditions are highly variable and location-specific, impacting transaction costs and affecting quality recognition. I present evidence of quality price premium variations across market conditions and identify various market features associated with the existence of a price premium to observable and/or non-observable quality attributes. Among other things, price premiums for observable quality attributes increase with number of traders per farmer. Moreover, farmers engage in relational contract do not receive higher premium rewarding higher-quality supply, though receiving a higher average price. In contrast, returns on non-observable quality attributes vary with the presence of millers and/or cooperatives near market sites. Lastly, I investigate further mechanisms showing that traders sorting across locations might be a mechanism explaining some of the results.

These findings are not based on a purposefully designed trial, and several of the highlighted relationships must be interpreted as exploratory. For instance, a potential intervention could consist of varying the existence of relational contracting and whether its formation is demand or supply-driven. Indeed, previous randomized experiments have only focused on varying traders' or farmers' marginal returns to relational contracts for those already using them, omitting that relational contracts' existence is endogenously determined (Casaburi and Reed, 2022; Deutschmann et al., 2020). The results suggest that current policies proposed to alleviate farmers' constraints (e.g., technology adoption subsidies, financial services, and extension services

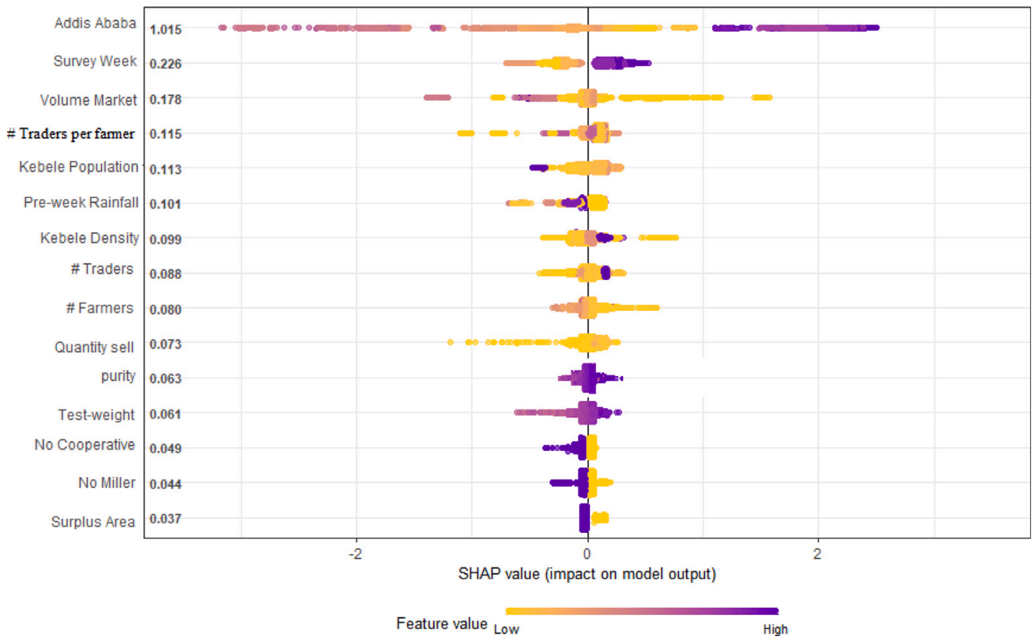


Fig. 8. Shapley values of the most predictive features of wheat price: eXtreme Gradient Boosting model.
Notes.: This Figure shows the Shapley (SHAP) values of the fifteen most predictive features using eXtreme Gradient Boosting. A positive (negative) SHAP value represents an increase (decrease) in the predicted variable (i.e., wheat price per kg) across all possible combinations of the features. The mean of SHAP values indicates the variable's average contribution in prediction on the vertical axis. Darker color corresponds to higher values of the predictor. (For interpretation of the references to color in this figure legend, the reader is referred to the webversion of this article.)
Source: Author's computations based on 2019/2020 wheat survey.

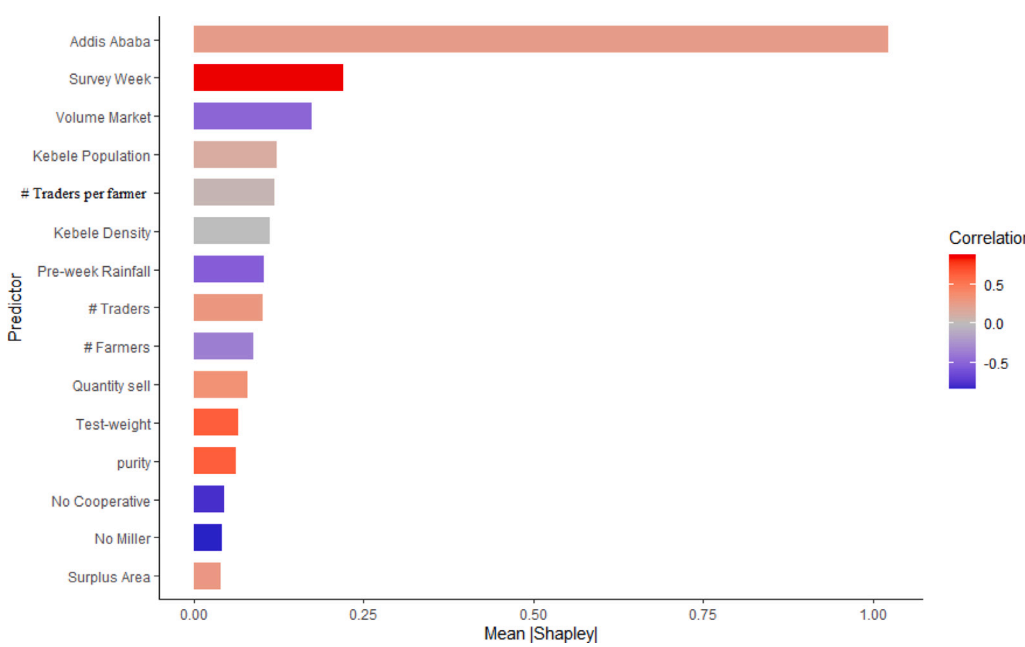


Fig. 9. Correlation between predictive features and predicted wheat price: eXtreme gradient boosting model.
Notes. This Figure shows the correlation between the fifteen most predictive features and SHAP values. It provides the direction of the association (red for positive and blue for negative), and the predictor's marginal contribution in prediction based on the mean SHAP values. (For interpretation of the references to color in this figure legend, the reader is referred to the webversion of this article.)
Source: Author's computations based on 2019/2020 wheat survey.

development) are limited in promoting quality-upgrading as long as quality is not fully rewarded in the market. Given the positive correlation between the number of traders per farmer and quality recognition, policymakers might be interested in promoting competition to enhance price premiums for quality, which may in turn increase farmers' returns from quality-upgrading.

However, implementing these policies in a weakly institutionalized and imperfect market context may worsen market functionality and have significant distributional effects. Market conditions are locally specific and organized around well-established rules and actors such as relational contracting. Radical shifts in such settings may negatively affect both farmers and traders (Macchiavello and Morjaria, 2021; Bulte et al., 2024). Hence, policy intervention must be evaluated on a case-by-case basis to address market issues experienced by local actors. For instance, some policies have promoted alternative marketing channels such as vertical coordination and cooperatives to enhance quality in local markets. However, they represent only a small share of local marketing channels. Other studies propose encouraging quality-upgrading through the promotion of third-party certification available to small-scale farmers to reveal unobservable attributes at low cost (Abate et al., 2021). On the farmer's side, recent evidence points to significant demand for such services (Anissa et al., 2021). The extent to which traders are willing to use such services, however, remains largely under-investigated (Abate et al., 2023).

CRediT authorship contribution statement

Jérémy Do Nascimento Miguel: Writing – review & editing, Writing – original draft, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data and codes will be made available in ATAI repository once paper is accepted for publication.

Appendix A. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.jdeveco.2024.103336.

References

Abadie, Alberto, Imbens, Guido W., Athey, Susan, 2017. When Should You Adjust Standard Errors for Clustering? National Bureau of Economic Research.
Abate, Gashaw Tadesse, Bernard, Tanguy, 2017. Farmers' Quality Assessment of Their Crops and Its Impact on Commercialization Behavior: a Field Experiment in Ethiopia. IFPRI - Discussion Papers 01624.
Abate, Gashaw Tadesse, Bernard, Tanguy, Bulte, Erwin, Do Nascimento Miguel, Jeremy, Sadoulet, Elisabeth, 2023. Buyers' Response to Third-Party Quality Certification. IFPRI Discussion Paper 02258.
Abate, Gashaw Tadesse, Bernard, Tanguy, De Janvry, Alain, Sadoulet, Elisabeth, Trachtman, Carly, 2021. Introducing quality certification in staple food markets in Sub-Saharan Africa: Four conditions for successful implementation. Food Policy 105, 102–173.
Abay, Kibrom A., Abate, Gashaw Tadesse, Barrett, Christopher B., Bernard, Tanguy, 2019. Correlated non-classical measurement errors, 'Second best' policy inference, and the inverse size-productivity relationship in agriculture. J. Dev. Econ. 139 (February), 171–184.
Aker, Jenny C., 2010. Information from markets near and far: Mobile phones and agricultural markets in niger. Am. Econ. J.: Appl. Econ. 2 (3), 46–59.
Amin, Modhurima Dey, Badruddoza, Syed, McCluskey, Jill J., 2021. Predicting access to healthful food retailers with machine learning. Food Policy 99 (October 2020), 101985.
Anderson, Michael L., 2008. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the Abecedarian, Perry Preschool, and Early Training Projects. J. Amer. Statist. Assoc. 103 (484), 1481–1495.
Anissa, Banawe Plambou, Abate, Gashaw, Bernard, Tanguy, Bulte, Erwin, 2021. Is the local wheat market a ' market for lemons '? Certifying the supply of individual wheat farmers in Ethiopia. Eur. Rev. Agric. Econ. 1–24.
Athey, Susan, Imbens, Guido W., 2019. Machine Learning Methods Economists Should Know About | Stanford Graduate School of Business. Working Paper.
Barzel, Yoram, 1982. Measurement cost and the organization of markets. J. Law Econ. 25 (1), 27–48.
Belloni, Alexandre, Chernozhukov, Victor, Hansen, Christian, 2013. Inference on treatment effects after selection among high-dimensional controls. Rev. Econ. Stud. 81 (2), 608–650.
Benyishay, Ariel, Mobarak, A. Mushfiq, 2019. Social learning and incentives for experimentation and communication. Rev. Econ. Stud. 86 (3), 976–1009.
Bergquist, Lauren Falcao, Dinerstein, Michael, 2020. Competition and entry in agricultural markets: Experimental evidence from Kenya. Amer. Econ. Rev. 110 (12), 3705–3747.
Bernard, Tanguy, Abate, Gashaw, Lemma, Solomon, 2013. Agricultural Cooperatives in Ethiopia : Results of the 2012 ATA Baseline Survey. International Food Policy Research Institute, Washington DC, pp. 1–23.
Bernard, Tanguy, De Janvry, Alain, Mbaye, Samba, Sadoulet, Elisabeth, 2017. Expected product market reforms and technology adoption by senegalese onion producers. Am. J. Agric. Econ. 99 (4), 1096–1115.
Bernard, Tanguy, Taffesse, Alemayehu, Gabre-Madhin, Eleni, 2008. Impact of cooperatives on smallholders ' commercialization behavior : evidence from Ethiopia. Agricult. Econ. 39, 147–161.
Bold, Tessa, Ghisolfi, Selene, Nsonzi, Frances, Svensson, Jakob, 2022. Market access and quality upgrading : Evidence from three field experiments. Amer. Econ. Rev. 112 (8), 1–57.
Bold, Tessa, Kaizzi, Kayuki, Svensson, Jakob, Yanagizawa-Drott, David, 2017. Lemon technologies and adoption: measurement, theory and evidence from agricultural markets in Uganda. Q. J. Econ. 132 (3), 1055–1100.
Bulte, Erwin, Do Nascimento Miguel, Jérémy, Anissa, Banawe P., 2024. Competition on agricultural markets and quality of smallholder supply: The role of relational contracting and input provision by traders. Econom. Dev. Cult. Chang. 72 (2), 603–632.
Carletto, Calogero, Savastano, Sara, Zezza, Alberto, 2013. Fact or artifact: The impact of measurement errors on the farm size-productivity relationship. J. Dev. Econ. 103 (1), 254–261.
Carter, Michael R., Laajaj, Rachid, Yang, Dean, 2013. The impact of voucher coupons on the uptake of fertilizer and improved seeds : Evidence from a randomized trial in Mozambique. Am. J. Agric. Econ. 95 (5), 1345–1351.
Casaburi, Lorenzo, Glennerster, Rachel, Suri, Tavneet, 2013. Rural roads and intermediated trade: Regression discontinuity evidence from Sierra Leone. SSRN Electron. J..
Casaburi, Lorenzo, Reed, Tristan, 2022. Using individual-level randomized treatment to learn about market structure. Am. Econ. J.: Appl. Econ. 14 (4), 58–90.
Chen, T, He, T, Benesty, M, Khotilovich, V, Tang, Y, Cho, H, Chen, K, 2015. Xgboost: extreme gradient boosting. In: R package version 0.4-2 1.4. pp. 1–4.
De Janvry, Alain, Sadoulet, Elisabeth, 2020. Using agriculture for development: Supply- and demand-side approaches. World Dev. 133, 105003.
Dercon, Stefan, 1995. On market integration and liberalisation: Method and application to Ethiopia. The Journal of Development Studies 32 (1), 112–143.
Dercon, Stefan, Christiaensen, Luc, 2011. Consumption risk, technology adoption and poverty traps: Evidence from Ethiopia. J. Dev. Econ. 96 (2), 159–173.
Dercon, Stefan, Gollin, Douglas, Dercon, Stefan, Gollin, Douglas, 2019. Agriculture's changing role in Ethiopia's economic transformation. In: The Oxford Handbook of the Ethiopian Economy. Oxford, pp. 448–467.
Deutschmann, Joshua W., Bernard, Tanguy, Yameogo, Ouambi, 2020. Contracting and quality upgrading: Evidence from an experiment in Senegal.
Dillon, Brian, 2020. Selling crops early to pay for school: A large-scale natural experiment in Malawi. J. Hum. Resour. 56 (4).
Dillon, Brian, Dambro, Chelsey, 2017. How competitive are crop markets in Sub-Saharan Africa? Am. J. Agric. Econ. 99 (5), 1344–1361.
Dinku, Tufa, Funk, Chris, Peterson, Pete, Maidment, Ross, Tadesse, Tsegaye, Gadain, Hussein, Ceccato, Pietro, 2018. Validation of the CHIRPS satellite rainfall estimates over eastern Africa. Q. J. R. Meteorol. Soc. 144, 292–312.
Duflo, Esther, Kremer, Michael, Robinson, Jonathan, 2011. Nudging farmers to use fertilizer: theory and experimental evidence from Kenya. Amer. Econ. Rev. 101 (6), 2350–2390.
Evans, James D., 1996. Straightforward Statistics for the Behavioral Sciences. Thomson Brooks/Cole Publishing Co.
Facebook, 2021. Facebook data for good population density maps.
Fafchamps, Marcel, 2001. Networks , communities and markets in Sub-Saharan Africa : Implications for firm growth and investment. J. Afr. Econ. 10 (suppl_2), 109–142.
Fafchamps, Marcel, 2003. Market Institutions in Sub-Saharan Africa: Theory and Evidence. MIT Press.
Fafchamps, Marcel, Hill, Ruth Vargas, Minten, Bart, 2008. Quality control in nonstaple food markets: Evidence from India. Agricult. Econ. 38 (3), 251–266.
Fafchamps, Marcel, Minten, Bart, 1999. Relationships and traders in Madagascar. J. Dev. Stud. 35 (6), 1–35.
Fafchamps, Marcel, Minten, Bart, 2002. Returns to social network capital among traders. Oxf. Econ. Pap. 54 (2), 173–206.
FAO, 2020. FAOSTAT Statistical Database. Tech. Rep., Food and Agriculture Organization of the United Nations.
Funk, Chris, Peterson, Pete, Landsfeld, Martin, Pedreros, Diego, Verdin, James, Shukla, Shraddhanand, Husak, Gregory, Rowland, James, Harrison, Laura, Hoell, Andrew, Michaelsen, Joel, 2015. The climate hazards infrared precipitation with stations - A new environmental record for monitoring extremes. Sci. Data 2, 1–21.
Gabre-Madhin, Eleni, Goggin, Ian, 2005. Does Ethiopia need a commodity exchange? An integrated approach to market development. EDRI-ESSP Policy Working Paper 4.

Gáfaro, Margarita, Pellegrina, Heitor S., 2022. Trade, farmers' heterogeneity, and agricultural productivity: Evidence from Colombia. J. Int. Econ. 137, 103598.

Gebreselassie, Samuel, Haile, Mekbib G., Kalkuhl, Matthias, 2017. The wheat sector in Ethiopia: Current status and key challenges for future value chain development.

Ghani, Tarek, Reed, Tristan, 2022. Relationships on the rocks: Contract evolution in a market for ice. Am. Econ. J. Microecon. 14 (1), 330–365.

Harou, Aurélie P, Madajewicz, Malgosia, Michelson, Hope, Palm, Cheryl A, Amuri, Nyambilila, Magomba, Christopher, Semoka, Johnson M, Tschirhart, Kevin, Weil, Ray, 2022. The joint effects of information and financing constraints on technology adoption: Evidence from a field experiment in rural Tanzania. J. Dev. Econ. 155, 102707.

Hastie, Trevor, Tibshirani, Robert, Friedman, Jerome, 2009. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer Science & Business Media.

Hoffmann, Vivian, Gatobu, Ken Mwithirwa, 2014. Growing their own: Unobservable quality and the value of self-provisioning. J. Dev. Econ. 106, 168–178.

Hoffmann, Vivian, Moser, Christine, 2017. You get what you pay for: the link between price and food safety in Kenya. Agric. Econ. (U. K.) 48 (4), 449–458.

Hoffmann, Vivian, Mutiga, Samuel, Harvey, Jagger, Nelson, Rebecca, Milgroom, Michael, 2013. Aflatoxin contamination of Maize in Kenya : Observability and mitigation behavior. In: Paper presented at the Agri- cultural and Applied Economics Association 2013 Annual Meeting.

Hoffmann, Vivian, Mutiga, Samuel K., Harvey, Jagger W., Nelson, Rebecca J., Milgroom, Michael G., 2021. Observability of food safety losses in maize: Evidence from Kenya. Food Policy 98 (April 2020), 101895.

Kadjo, Didier, Ricker-Gilbert, Jacob, Alexander, Corinne, 2016. Estimating price discounts for low-quality Maize in sub-Saharan Africa: Evidence from Benin. World Dev. 77, 115–128.

Karlan, Dean, Osei, Robert, Osei-Akoto, Isaac, Udry, Christopher, 2014. Agricultural decisions after relaxing credit and risk constraints. Q. J. Econ. 129 (2), 597–652.

Krishna, Kala, Sheveleva, Yelena, 2017. Wheat or strawberries? Intermediated trade with limited contracting. Am. Econ. J. Microecon. 9 (3), 28–62.

Macchiavello, Rocco, Morjaria, Ameet, 2021. Competition and relational contracts in the Rwanda coffee chain. Q. J. Econ. 1089–1143.

Magnan, Nicholas, Hoffmann, Vivian, Garrido, Gissele, Kanyam, Faniel, Opoku, Nelson, 2021. Information, technology, and market rewards: Incentivizing aflatoxin control in Ghana:. J. Dev. Econ..

Michelson, Hope, Fairbairn, Anna, Ellison, Brenna, Maertens, Annemie, Manyong, Victor, 2021. Misperceived quality: Fertilizer in tanzania. J. Dev. Econ. 148, 102579.

Minot, Nicholas, Warner, James, Lemma, Solomon, Kasa, Leulsegged, Tadesse, Gashaw, Rashid, Shahidur, 2019. The wheat supply chain in Ethiopia : Patterns , trends , and policy options. Gates Open Res. 3.

Minten, Bart, Murshid, K. A.S., Reardon, Thomas, 2013. Food quality changes and implications: Evidence from the rice value chain of Bangladesh. World Dev. 42 (1), 100–113.

Moser, Christine, Barrett, Christopher, Minten, Bart, 2009. Spatial integration at multiple scales: Rice markets in Madagascar. Agricult. Econ. 40 (3), 281–294.

Mullainathan, Sendhil, Spiess, Jann, 2017. Machine learning: An applied econometric approach. J. Econ. Perspect. 31 (2), 87–106.

OECD-FAO, 2016. Agriculture in Sub-Saharan Africa: Prospects and Challenges for the Next Decade. Tech. Rep..

Osborne, Theresa, 2005. Imperfect competition in agricultural markets: Evidence from Ethiopia. J. Dev. Econ. 76 (2), 405–428.

Prunier, Gérard, 2015. Understanding Contemporary Ethiopia: Monarchy, Revolution and the Legacy of Meles Zenawi. Press, Oxford University,

Reardon, Thomas, Hopkins, Rose, 2006. The supermarket revolution in developing countries : Policies to address emerging tensions among supermarkets, suppliers and traditional retailers, 18 (4) 522–545.

Salazar, César, Ayalew, Hailemariam, Fisker, Peter, 2019. Weather shocks and spatial market efficiency: Evidence from mozambique. J. Dev. Stud. 55 (9), 1967–1982.

Sayre, James, 2023. Farm to Firm: Clustering and Returns to Scale in Agricultural Supply Chains. Working Paper.

Seyoum Taffesse, Alemayehu, Dorosh, Paul, Gemessa, Sinafikeh Asrat, 2012. Crop production in Ethiopia: Regional patterns and trends. In: Food and Agriculture in Ethiopia: Progress and Policy Challenges. University of Pennsylvania Press Philadelphia, PA, pp. 53–83.

Shiferaw, Bekele, Kassie, Menale, Jaleta, Moti, Yirga, Chilot, 2014. Adoption of improved wheat varieties and impacts on household food security in Ethiopia. Food Policy 44, 272–284.

Staiger, B. Y. Douglas, Stock, James H., 1997. Instrumental variables regression with weak instruments. Econometrica 65 (3), 557–586.

Stephens, Emma C., Barrett, Christopher B., 2011. Incomplete credit markets and commodity marketing behaviour. J. Agric. Econ. 62 (1), 1–24.

Suri, Tavneet, 2011. Selection and comparative advantage in technology adoption. Econometrica 79 (1), 159–209.

Swinnen, Johan F.M., Vandeplas, Anneleen, 2010. Market power and rents in global supply chains. Agricult. Econ. 41, 109–120.

Swinnen, Johan, Vandeplas, Anneleen, 2015. Quality Standards, Value Chains, and International Development. Cambridge University Press, Cambridge.

Tadesse, Gashaw, Bernard, Tanguy, Brauw, Alan De, Minot, Nicholas, 2018. The impact of the use of new technologies on farmers ' wheat yield in Ethiopia : evidence from a randomized control trial. 49, 409–421.

Tiecke, Tobias G., Liu, Xianming, Zhang, Amy, Gros, Andreas, Li, Nan, Yetman, Gregory, Kilic, Talip, Murray, Siobhan, Blankespoor, Brian, Prydz, Espen B., Dang, Hai-Anh H., 2017. Mapping the world population one building at a time.

Vandercasteelen, Joachim, Beyene, Seneshaw Tambru, Minten, Bart, Swinnen, Johan, 2018. Big cities, small towns, and poor farmers: Evidence from Ethiopia. World Dev. 106, 393–406.

Wooldridge, Jeffrey M., 2010. Econometric Analysis of Cross Section and Panel Data. MIT Press.

Worku, Ibrahim Hassen, Dereje, Mekdim, Minten, Bart, Hirvonen, Kalle, 2017. Diet transformation in Africa: the case of Ethiopia. Agricult. Econ. 48, 73–86.