Relational Frictions Along the Supply Chain: Evidence from Senegalese Traders*

Edward Wiles ⓡ Deivy Houeix
MIT MIT

For the latest version, click here.

25 October 2024

Abstract

Search and trust frictions have historically made it hard for small firms in lower-income countries to buy inputs from foreign markets. The growth in smartphone ownership and social media usage has the potential to alleviate these barriers. We run a field experiment leveraging these technological tools to provide exogenous variation in both search and trust frictions in a large international import market. The design is informed by a dynamic relational contracting model featuring sequential search for suppliers and trust frictions in the form of adverse selection and moral hazard. In our search treatment, we connect a randomly selected 80% of 1,862 small garment firms in Senegal to new suppliers in Turkey. We then cross-randomize two trust treatments that provide additional information about the types and incentives of these new suppliers. Alleviating search frictions is sufficient to increase access to foreign markets: in all treated groups, firms are 25% more likely to have the varieties a mystery shopper requests and the goods sold are 32% more likely to be high quality. However, the trust treatments are necessary for longer-term impact: these groups are significantly more likely to develop the connections into relationships that persist beyond the study. These new relationships lead to increases in medium-run profit and sales, particularly among wholesalers in the upper tail. Finally, we use the treatment effects to estimate the model and evaluate counterfactuals where we set various combinations of the frictions to zero, finding that the largest gains come from eliminating adverse selection.

*Job Market Paper for Edward Wiles (edav@mit.edu). Author names are in ⓡ random order. We are immensely grateful to our advisors, David Atkin, Abhijit Banerjee, Dave Donaldson, Esther Duflo, Ben Olken, and Tavneet Suri, for their unwavering support throughout our work on this project. We also thank Miguel Fajardo-Steinhäuser, Ahmet Gulek, Isabela Manelici, Beatrice Montano, Jacob Moscona, Frank Schilbach, Advik Shreekumar, Jose Vasquez, Emma Wiles, and attendees of MIT's development lunch and trade lunch for numerous helpful discussions. We are grateful to Wave Mobile Money for providing some of the data used in this research. This project received funding from the George and Obie Shultz Fund at MIT, as well Private Enterprise Development in Low-Income Countries (PEDL). The Pre-Analysis Plan (PAP) for this study can be found on the AEA RCT Registry under ID 12522, `https://www.socialscienceregistry.org/trials/12522`.

1 Introduction

Search and trust frictions make it hard for small firms to access foreign inputs. Consider a clothing wholesaler in Senegal who wants to start selling high quality European-made jeans. First, they must find a supplier, most of whom are in Europe, as well as some way of observing when new varieties are introduced. Second, even if they manage to do this and make an order, they are typically required to pay before observing the quality. These issues are particularly severe in lower-income settings, whose vast informal sectors complicate information aggregation and lower state capacities makes contracts effectively unenforcable. Yet, while there is a long theoretical tradition in economics studying these frictions—dating back to Stigler (1961) on search and Shapiro (1983) on trust—it has proven much harder to learn how they work empirically as it is rare to observe exogenous variation that closely maps to the theory.

In recent years, there has been substantial growth in smartphone ownership and social media usage in lower-income countries, which has the potential to fundamentally change how small firms approach their buying and selling decisions. Not only are firms increasingly selling online—with e-commerce revenue in Africa estimated to have doubled between 2019 and 2024—but they are primarily doing so through social media rather than traditional platforms: in a survey of firms using e-commerce across six African countries, a large non-profit found that 99% do so using social media and 60% do so only using social media.[1] This growth in "social commerce" may reflect the fact that, in principle, social media could meaningfully alleviate both search and trust frictions in supply chains. For search, the supplier in Europe can send photos of their wares. For trust, domestic firms may be able to much more easily share information and coordinate action to discipline suppliers that cheat—a modern day version of the mechanism in Greif (1993). However, whether it actually does alleviate these frictions is ultimately an empirical question.

In this paper, we provide what is—to the best of our knowledge—the first experimental evidence on the role of search and trust frictions in international supplier relationships and the extent to which new technologies may be alleviating them. We designed a field experiment leveraging key features of social media to address various combinations of these frictions in the context of a large international import market. Specifically, we randomly allocated 1,862 small firms in the ready-to-wear garment industry in Dakar across treatment arms that connected them to new suppliers in Turkey and varied the information

[1] Statista estimate that total B2C e-commerce revenue across Africa was 18 billion USD in 2019, increasing to 34 billion USD in 2024 (Statista, 2023). The survey was conducted by the Global System for Mobile Communications Association, a large non-profit representing mobile network operators, in Egypt, Ethiopia, Ghana, Kenya, Nigeria, and South Africa (GSMA, 2023).

available about the types and incentives of these suppliers. We then measured how these interventions affected their access to foreign goods, supplier relationships, and profits and sales, using data from a mystery shopping exercise, real-time transactions from the largest mobile payments provider in Senegal, and a follow-up survey.

We first provide descriptive survey evidence that firms use social media extensively to learn about and interact with suppliers. 85% have bought directly from WhatsApp groups managed by suppliers in the past 12 months. The median firm is in 4 separate such groups, and for 27% of firms at least one of these groups is run by a supplier in a different country. When asked why they are in such groups, firms emphasise that the groups help them see more varieties and compare prices across suppliers. Other apps, such as TikTok, Instagram, and Facebook, are also popular, each used by around a third of firms.

In contrast, while most firms (88%) have heard of traditional B2B e-commerce platforms, such as Alibaba, they are seldom used. 85% of firms have never purchased from these platforms, and, of those who have, about half have done so only very rarely. When asked why, firms report that they find platforms too complicated and they don't trust the information on the platform. Social media offers substantial advantages on both fronts: most firm owners already use them regularly, and the social nature means that firms can interact personally with the suppliers with little need to trust WhatsApp or Facebook directly. The downside is that social media may be less effective at aggregating information and does not offer the enforcement and shipping benefits that some B2B platforms provide.

In order to make precise the role of search and trust frictions, we develop a model of relational contracting featuring sequential search and both adverse selection and moral hazard (which we refer to as trust frictions). A firm can either buy inputs from a local supplier without any frictions, or can pay a fixed search cost to match with a random foreign supplier. Foreign suppliers may be thought of as selling newer varieties, higher quality varieties, or the same varieties at a lower price. Foreign suppliers must take a costly but unobservable action to guarantee the goods are appropriate quality. Bad-type suppliers will never do this, while strategic-type suppliers will only do so if the future value of the relationship justifies the current-period cost. The fact that the contract is relational and the supplier has limited liability implies that it is optimal for the firm to offer one contract that pools both types of suppliers, which creates on-path learning and average quantity and value increasing over time. While the optimal contract is not available in closed form, we derive an equation showing that adverse selection and moral hazard create wedges between marginal revenue and marginal cost that distort quantity downwards.

The model highlights the types of variation that the ideal experiment aiming to isolate the three frictions would generate. To isolate search frictions, treatments should either create matches or lower the cost of finding a random new supplier. To isolate adverse selection, treatments should either directly give information about a particular supplier or improve the ability to learn this information over time, but should not affect suppliers' incentives. To isolate moral hazard, treatments should strengthen the incentives of the supplier or firms' perceptions of these incentives, but should not provide other information.

The experiment comprises three treatments, each targeting a specific friction: search, adverse selection, and moral hazard. In the search treatment, we add treated firms to the supplier WhatsApp groups of three different suppliers in Turkey. We inform firms that the suppliers were recruited by a local team in Turkey and export to Senegal, but do not provide any further details. We chose Turkey as our exporter country because it is the second largest source of ready-to-wear garments in Senegal (after China), and in this context Turkish-made garments command a large quality premium, which is well-suited for studying trust frictions because firms are worried about suppliers cheating on quality.[2]

We then cross-randomise the adverse selection and moral hazard treatments among the firms in the search treatment. In the adverse selection treatment, we add treated firms to a fourth WhatsApp group containing other firms matched with the same suppliers, the purpose of which is for them to share information about whether these suppliers are good or bad. Importantly, we seed these groups with initial information: treated firms receive a recommendation for one of the suppliers, based on real orders from the suppliers that we commissioned before the study started. Thus, the treatment both improves learning and directly provides information about suppliers' types.

In the moral hazard treatment, we inform firms that we will ask them to rate the study suppliers, and that any supplier receiving consistently negative feedback will be removed from the study, thereby losing access to 150-200 potential clients. Moreover, we inform firms that we have communicated this information to the suppliers and that they therefore have strong incentives to exert effort. Thus, the treatment shifts firms' perceptions about the incentives of the suppliers.

Altogether, we have five equally sized groups: Pure Control, Search Only, Search + Adverse Selection, Search + Moral Hazard, and Search + Adverse Selection + Moral Hazard.

Our primary outcome is a revealed preference measure of access to foreign goods. We designed a mystery shopping exercise in which trained surveyors, acting as real customers, attempt to purchase goods from all study firms. We then measured the type and quality

[2]This quality premium is large and salient: in a consumer survey in which we elicited willingness-to-pay for goods, randomising whether we said they made in Turkey or China, we found an average Turkey premium of 34% ($p < 0.01$).

of the goods that they sold to us. This has two advantages: it captures real behaviour (as opposed to self-reported outcomes), and it allows us to separately measure a horizontal dimension (access to a wide set of differentiated varieties) and a vertical dimension (access to high quality varieties). On the horizontal dimension, each good that we attempt to purchase is defined by 5 criteria, such as colour and sleeve style, and our outcome is an indicator for whether the firm has a good matching at least 3 criteria. On the vertical dimension, the outcome is an index that aggregates three measures: two based on a detailed quality scorecard that we designed together with hired experts, and one based on whether the good was made in Turkey (a well-known signal of quality in this setting). We pre-specified these outcomes and the regression specification that we use throughout the paper in our Pre-Analysis Plan (PAP).

We find that the treatments have a large and significant effect on access to foreign goods, on both the horizontal and vertical dimensions. First, we pool the four treated groups together (all of whom received the search treatment). On the horizontal dimension, treated firms are 9.0 percentage points ($p < 0.01$) more likely to have a suitable good (a 25.2% increase). On the vertical dimension, conditional on having a good, the index increases by 0.422 standard deviations ($p < 0.01$). The effect on the price is small, insignificant, and precisely estimated, so the horizontal and vertical gains do not come at the cost of a large price increase.

Second, disaggregating across the four treated groups, we find that the coefficients for the trust-treated groups are not significantly larger than the coefficients for Search Only. This does not necessarily mean that trust frictions do not exist: these are small orders (around 20-40 USD) and may be below the threshold at which trust frictions have bite. We thus conclude that (1) firms face constraints in accessing foreign goods, (2) alleviating search frictions improves this access, (3) social media can be an effective tool to do so.

However, while connections to new suppliers via social media improves access to foreign goods, the extent to which firms are able to realise this benefit depends on whether these connections develop into lasting relationships. We measure this using data from two sources: (1) a follow-up survey that we conducted after 3 months, and (2) real-time, transaction-level administrative data from the largest mobile money provider in Senegal, tracking most firm transactions for up to 9 months after the study started.

From the survey data, we find that (pooled) treatment increases the likelihood of having a regular supplier in Turkey by 3.7 percentage points ($p < 0.1$), a 22.2% increase relative to control. The effect is largest in the Search + Adverse Selection + Moral Hazard group at 7.5 percentage points. We find no effect on the total number of suppliers, suggesting that firms have substituted away from a local wholesaler and towards importing directly. From the mobile money data, we find that all treated firms order similar amounts from study suppliers in the first few months. However, this changes in the medium- to long-run: firms in the trust treatments order 286.2% ($p < 0.05$) more than Search Only in the 6 months after the study finishes. When we disaggregate, the coefficients on all three trust groups are positive and fairly similar. We thus conclude that that the trust treatments meaningfully increased the share of these connections that developed into relationships.

To understand whether alleviating these frictions ultimately flows through to producer surplus, we collect standard summary measures of monthly profit and sales in our follow-up survey. Pooling the treatments together, we find increases of 82.4 USD ($p < 0.05$) in profits and 237.4 USD ($p < 0.05$) in sales. These are 43.8% and 39.0% increases relative to control. The coefficient in the Search + Adverse Selection + Moral Hazard is, again, much larger than the others. The implied increases are very large, and the magnitudes reduce by around half (but remain significant) when we winsorize at the 1% level. When we look at distributional treatment effects, we find that these average results come primarily from the upper tail of the profit and sales distributions: we see large and significant increases starting at around the 75th percentile for the Search + Adverse Selection + Moral Hazard group. Overall, we conclude that firms are able to realise meaningful gains from accessing a new foreign supplier by using social media to overcome search and trust frictions.

Finally, we use the reduced-form treatment effects to estimate the parameters of the model and the distribution of match-specific productivity between firms and foreign suppliers. We find that the cost of finding a given foreign supplier is relatively low, but the distribution of match-specific productivity implies that firms would have to pay the search cost many times to find a good match. We consider a first counterfactual where we shut down the ability to ever purchase directly from foreign suppliers—for example, if these firms could not use technological tools to interact with foreign suppliers—and find that this would decrease profits by around one quarter. We then consider a second counterfactual where we progressively turn off the various frictions. The largest gains come from eliminating adverse selection, because when search costs are relatively low, the gains to improving the returns to searching are large.

This paper builds on several literatures. First, while there is a substantial theoretical tradition studying search and trust frictions, making empirical progress has proved more challenging. This is largely due to challenges in (1) observing buyer-seller relationships, and (2) obtaining exogenous variation closely related to theoretical predictions. Recently, a growing literature has begun to overcome some of these challenges, primarily on the data side (e.g., Antras and Foley (2015), Steinwender (2018), Macchiavello and Morjaria (2021, 2015), Startz (2024)). We bring two main contributions to this literature. First, this paper is the first experiment systematically testing theories of search and trust frictions in buyer-seller relationships. This allows us to overcome both challenges: we observe buyer-seller relationships through survey and mobile money data, and we create variation that is both exogenous and specifically designed to capture theoretical moments. Second, our descriptive evidence and experimental results highlight the role that new technologies can play in alleviating these frictions.

Second, we contribute to the literature on how e-commerce improves consumer and producer surplus by alleviating information frictions.[3] The bulk of this literature focuses on formal B2B or B2C e-commerce platforms, typically in higher-income settings. Our contribution is to bring both descriptive and experimental evidence that shows the large and distinct role that digital platforms—and social media in particular—are now playing in lower-income settings. In particular, we show that firms use social media extensively for buying and selling decisions, and that rather than alleviating the frictions by providing centralised repositories of information, social media enables and facilitates bilateral relational contracting.

Third, we contribute to a literature in international trade that emphasises how networks have historically played a large role in solving information frictions and enforcing contracts. For example, Greif (1993) highlights how 11th-century Maghribi traders were able to sustain a multilateral punishment system for overseas agents, and that informal information flows through social networks played a key role. The classic article testing these theories empirically is Rauch (1999), which shows that common language and proximity play a larger role in explaining trade in differentiated goods than homogeneous goods, consistent with a large role of shared ties for the types of goods subject to search and trust frictions. While our study is not about networks *per se*, one of the main channels through which social media may alleviate trust frictions is exactly the mechanism in Greif (1993). Our study thus highlights how social media facilitates a modern manifestation of this idea.

The rest of the paper is organised as follows. Section 2 describes the setting and presents descriptive evidence on the use of social media in supply chains. Section 3 describes the model. Section 4 describes the sample and experimental design. Section 5 describes the data and methods. Section 6 presents the results. Section 7 disaggregates the results by retailer versus wholesaler. Section 8 presents the model estimation. Section 9 concludes.

[3] See Ellison and Ellison (2009), Dinerstein, Einav, Levin, and Sundaresan (2018), Barach, Golden, and Horton (2020), Couture et al. (2021), Bai et al. (2023), or Goldfarb and Tucker (2019) for a recent review. The closest paper to ours is Alhorr (2024), who provides Facebook pages to female business owners in Jordan and finds positive effects on business outcomes, particularly for women who are less able to leave the house. Our paper differs as our focus is on how these technologies alleviate search and trust frictions in international trade.

2 Setting

2.1 Ready-to-Wear Garments in Dakar

Our study focuses on the ready-to-wear garment industry in Dakar, the capital city and economic hub of Senegal. The ready-to-wear garment industry exhibits substantial horizontal and vertical differentiation, making it ideal for our study: horizontal differentiation (a wide range of varieties) is well-suited for studying search frictions, and vertical differentiation (the presence of high and low qualities) is well-suited for studying trust frictions. It is also a large and important industry in its own right: in a consumer survey that we conducted with 400 households in Dakar, ready-to-wear garments represented an average of 6% of total household expenditure.

Within the ready-to-wear garment industry, our study places particular emphasis on goods made in Turkey. We chose Turkish-made goods for two reasons. First, Turkey is the second largest source of ready-to-wear garments in Dakar (after China). Second, Turkish-made goods have a reputation for being higher quality than goods made in China, which is ideal for studying trust frictions. In this setting, highlighting that a good is "Made in Turkey" is a very common way to signal quality. To quantify this, in our consumer survey we showed households an image of a product and randomised whether we said the good was made in Turkey or made in China. We then asked for their willingness to pay. We plot the CDF of willingness to pay in Figure 4, Panel (a). The Turkey CDF is shifted uniformly rightward relative to the China CDF, with an average premium of 34% ($p < 0.01$).

2.2 Sample

Firms in Senegal The main subjects of the study are 1,862 small firms in the ready-to-wear garments industry in Dakar. These firms are typical of small, informal, owner-operated businesses in many large cities in lower- and middle-income countries. 33% have a physical store in a market, while the remaining 67% operate exclusively online, primarily through social media. Firms with a physical store were recruited through a census in selected markets that sell both high and low quality goods; firms without a physical store were recruited through a combination of advertisements on Facebook and snowball sampling. The firms with a physical store are therefore broadly representative; recent years have seen substantial growth in the number of firms operating online-only businesses and those in our sample are typical of this phenomenon, but as there is no systematic database of such firms we cannot formally assess their representativeness.

At baseline, 91% of firms sell Turkish-made goods, with Turkish-made goods representing 40% of sales for the median firm. 33% of firms sell wholesale, which is an important margin of heterogeneity that we pre-specified in our PAP. 7% of retailers and 15% of wholesalers have travelled internationally for business at least once in the past 5 years, so the majority of these firms purchase their goods from other firms in Dakar or through e-commerce. 18% of retailers and 28% of wholesalers have at least one regular supplier based in Turkey. As we show in Section 2.3, these firms ubiquitously use social media, and WhatsApp in particular, to receive information from suppliers.

Firms have pessimistic beliefs about unknown foreign suppliers. 60% know multiple other firms that have had bad experiences ordering from a supplier online. To measure firms' priors, we asked them to consider a scenario in which they made an order from 10 unknown foreign suppliers, and to opine on what share of these orders would arrive with the anticipated quality. The median firm's response was that this would happen only 50% of the time.

Suppliers in Turkey The study involves connecting firms in Senegal with suppliers in Turkey. We worked with 30 suppliers, all of whom were based in Istanbul and were exporters of ready-to-wear garments to West Africa. We traveled to Turkey and conducted a census of two quarters in Istanbul that are well known for being textile wholesale and export hub for many parts of the world, including to several countries in West Africa. Among the suppliers that met our inclusion criteria, we then conducted a mystery shopping exercise to identify the most active. We focused on suppliers of Senegalese nationality, for a few reasons. First, while many Istanbul-based exporters to West Africa are of Turkish nationality, there is a sizeable Senegalese diaspora operating in this industry, reflecting the evidence in Rauch (2001) and Greif (1993) that cultural ties have historically played important roles in alleviating search and contracting frictions in long-distance trade. These shared ties certainly do not reduce trust frictions to zero, but they do likely facilitate reputation-based mechanisms. Second, they also solve some basic but important frictions such as language barriers and payment technologies. Since we expected our 1,862 firms in Senegal to be highly heterogeneous, introducing these barriers would have substantially complicated the experiment and distracted from the core issues of search and trust. Third, one of the main data sources for our outcomes is transaction-level data from the largest mobile money provider in Senegal. Senegalese suppliers in Turkey typically have accounts and use this system to accept payments, and so focusing on them offers substantial measurement advantages.

2.3 Social Media and e-Commerce in Supply Chains

In this section, we first describe the main channel that we use for the study and then we present statistics on social media usage and formal e-commerce platform usage.

Supplier Groups Many suppliers in our setting–both domestically and internationally–operate WhatsApp groups with their clients to advertise their goods, post prices, and highlight new items in stock. We will regularly refer to these as "supplier groups." We show examples in Figure 2. A typical group has one supplier and 50-100 clients, most of whom are regular or repeat customers. These are not discussion groups: the purpose is for the supplier to regularly post high quality photos and videos of their goods (typically only the supplier has permission to post). It is of course possible and common for buyers to negotiate with the supplier or to inquire about other goods–they can do so by simply sending a private message. These groups may be usefully thought of as virtual storefronts: clients can see what the supplier is selling, and can talk directly to the supplier about any queries.

These groups play a potentially important role in reducing search frictions and, to some extent, in reducing trust frictions. For search, firms can observe a very large number of goods from all over the world directly on their phone, and can easily negotiate and follow up as needed. Importantly, most firms also use social media extensively to sell to their own customers, and so these groups make it easy for them to forward relevant images to their own clients. For trust, a large group with many clients raises the cost of cheating because cheated buyers can message other members to share information. The group also makes it easier for the supplier to build a brand, further improving reputation-based mechanisms.

Social Media In Figure 3, we present statistics from our baseline survey with 1,862 firms in the ready-to-wear garment industry in Dakar. We focus here on sample-wide averages, but we also show in Appendix Figure A1 that the results are almost identical among firms with and without a physical store. In Panel (a), we plot the share of firms that reported using different types of social media to obtain information about suppliers, such as learning about new varieties or price information. WhatsApp is ubiquitous: 92% of firms use WhatsApp Status (a feature not often used in the United States in which content is broadcasted to all contacts for 24 hours), and 86% use supplier groups. TikTok, Instagram, and Facebook are also popular, each used by about a third of firms. Panel (b) shows the distribution of the number of unique supplier WhatsApp groups that firms belong to. Firms are in many supplier groups, with almost half of firms in 5 or more. Since these groups are very active, being in 5 such groups means that firms are observing a lot of information about different suppliers all the time. Importantly, these are not simply groups that they belong to but ignore: the distribution of the number of groups that they have bought inputs from in the past 12 months is almost identical.

To understand the reason such groups are so widely used, in Panel (c) we present the responses to a question asking what the main search-based advantages of supplier groups are. Firms highlight both how it allows them to see more varieties (both a wider set and higher quality) and how it allows them to compare prices across suppliers. Finally, in Panel (d), we show the location of the suppliers running these groups. The majority (81%) of firms are in a group with a supplier in Senegal, while a large minority are in at least one group with a supplier in a foreign country. 21% are in a group with a supplier in Turkey, 12% are in a group with a supplier in China, and 6% are in a group with a supplier in Dubai.[4] In total, 27% are in at least one supplier WhatsApp group where the supplier is based abroad. Since this focuses only on WhatsApp groups, this is a lower bound on the share of firms using social media more generally for international trade.

Firms are therefore familiar with the concept of using supplier WhatsApp group to transact with foreign suppliers, but, since only 21% are in a group with a supplier in Turkey, our experiment is still able to generate meaningful variation.

e-Commerce Platforms Traditional B2B e-commerce platforms, such as Alibaba, have also been shown to alleviate search and trust frictions (see Goldfarb and Tucker, 2019 for a review). Yet, in this setting, they are seldom used: 85% of firms have never purchased from these platforms, and, of those who have, about half have done so only very rarely. This is not because they have not heard of them (88% have). This reflects a broader trend in which large e-commerce companies have had limited success at penetrating African markets. On the surface, it may therefore seem puzzling why social media should be so heavily used in place of formal B2B platforms. In our follow-up survey, we asked firms why they don't use B2B platforms. Aside from the 39% firms who give no particular reason, the two most common answers are that firms find them too complicated to use (40%) and that firms do not trust them (33%).

These statistics confirm the anecdotal observations that led us to run this study: that small firms use social media extensively for their buying and selling activities, and that a sizable share of firms use it as a means of doing international trade.

[4] The share in China is likely a large underestimate of total social media interactions with China, as WhatsApp is blocked by China's firewall (it is useable with a VPN) and so other social media platforms–such as WeChat–are much more common. However, since WhatsApp is the main medium used to connect with suppliers in most countries, our survey focused on WhatsApp.

3 Theory

In this section, we describe a model of relational contracting featuring adverse selection and moral hazard with on-path learning, embedded within a sequential search framework. The goal is to make precise the role of the frictions and to highlight the types of variation necessary to identify them. In Section 8, we will use the reduced form treatment effects to estimate the parameters governing these frictions and consider counterfactuals in which we set them to zero to estimate the total potential gains. While the forces in the model are all canonical, we are not aware of existing literature that combines them in this way.[5]

3.1 Firms and Suppliers

The model is an infinite horizon repeated game with discrete time, all players have common discount factor δ, and all players are risk neutral. There are two sets of players: firms (the principal) and foreign suppliers (the agent). In the stage game, the firm can purchase input q and re-sell to consumers for revenue $r(q)$, where $r(q)$ is strictly increasing, concave, and has $\lim_{q\to\infty} r'(q) = 0$. The firm has outside option $\bar{U}$. The firm is not matched with a foreign supplier by default, and matching will be governed by a search process. We first discuss the principal-agent problem that arises conditional on matching, and then describe the search process and the determination of $\bar{U}$.

3.2 Relational Contracting with Foreign Suppliers

Goods sold by foreign suppliers can be high or low quality. For simplicity, we normalise the value of low quality goods to zero, and let the value of q_t high quality goods to the firm be $r(q_t)$, as before. Faced with an order for q_t, foreign suppliers can choose to take a costly but unobservable action $a_t \in \{0, 1\}$ that influences the probability that the goods are high quality. If they choose $a_t = 1$, then the goods are high quality with probability 1 and the supplier pays cost cq_t, where $c > 0$ is a constant marginal cost. If they choose $a_t = 0$, they instead pay marginal cost $c_0 < c$ but the goods are only high quality with probability $\lambda \in [0, 1)$. Avoiding the action may be interpreted as cutting costs by purchasing the goods from a cheaper manufacturer that only delivers with probability λ, or purchasing

[5] Our model differs from much of the relational contracts literature as the solution is non-stationary, due to both on-path learning and limited liability. Among prior work in this literature that studies non-stationary equilibria (Hörner (2002), Halac (2012), Yang (2013), Fong and Li (2017)), the principal-agent component of our model combines the on-path learning from Yang (2013) with the explicit treatment of limited liability from Fong and Li (2017). However, unlike both of these papers, we endogenise quantity, which complicates incentive design as the principal can choose not only the terms of the contract but also the stakes of the contract in each period. Martimort et al. (2017) take a mechanism design approach to analyse the endogenous quantity issue, but they focus on separating equilibria (ruled out under our version of the enforcement constraint and limited liability) and thus do not feature on-path learning.

the goods from a manufacturer that uses a risky technology.

If the firm orders from the supplier, they pay a transfer τ_t, determined endogenously. This transfer must be paid before receiving the good.[6] Foreign suppliers have outside option 0, reflecting the idea that relationships are separable for the supplier.

Adverse Selection: There are two types of foreign suppliers: good and bad. A supplier's type is fixed over time, known to the supplier, and unobservable to firms. The only difference between the two types is that bad types will never choose $a_t = 1$ (for example, because they have marginal cost $c_{\text{bad}} = \infty$), while good types will choose $a_t = 1$ if it is in their best interests to do so. We denote the firm's beliefs about the share of bad types after observing t realisations in which the goods were high quality as μ_t, with initial (correct) beliefs $\mu_0 \in (0, 1)$. The firm updates this belief each period using Bayes' Rule. We assume that λ is sufficiently low that, if the firm knew that the supplier was a bad type, they would prefer to exit the relationship and take their outside option, $\bar{U}$.

Moral Hazard: Good types will choose $a_t = 1$ if it is in their best interests to do so. Since this is a repeated game, it may be possible to induce them to do this by the promise of future rewards tied to repeated high quality realisations. To parameterise the extent of moral hazard, we assume that if a good type chooses $a_t = 0$, the firm will find out and find a way to enforce the contract with probability $1 - \xi$.[7] This implies that any equilibrium in which good type suppliers choose $a_t = 1$ must satisfy a standard Dynamic Incentive Compatibility Constraint (DICC), as follows

$$\delta(1 - \lambda)V_{t+1} \geq \xi(c - c_0)q_t, \tag{DICC}$$

where $V_{t+1} \equiv \sum_{n=1}^{\infty} \delta^n \left(\tau_{t+1+n} - cq_{t+1+n}\right)$ is the discounted sum of future profits from the relationship. Intuitively, the supplier has already received τ_t. If they choose $a_t = 1$, they earn $\delta V_{t+1} - cq_t$; if they choose $a_t = 0$, they earn $\delta\lambda V_{t+1} - \xi c_0 q_t - (1 - \xi)cq_t$. This constraint imposes a ceiling on the contractible quantity, governed by the current value of the relationship, V_{t+1}, and the extent of moral hazard, ξ.

Contracts: In period 0, the firm makes a take-it-or-leave-it offer of a long-term contract to

[6] One could extend the model to allow for an endogenous share α_t to be paid upfront, and the remaining $(1 - \alpha_t)$ to be paid after quality is observed. In our empirical setting, almost 100% of contracts involve full payment upfront, so for simplicity we simply set $\alpha_t = 1$ for all t. Including endogenous α_t creates a two-sided moral hazard problem, as the buyer has the option to refuse payment even if the good is high quality. Since there are two instruments, $\{\tau_t\}$ and $\{\alpha_t\}$, the equilibrium involves using one instrument to discipline the supplier and the other to discipline the firm, and $\alpha_t = 1$ would arise endogenously if the moral hazard problem on the buyer's side is very severe.

[7] There are other equivalent ways to microfound this moral hazard parameter, ξ. For example, one could interpret it as a psychic dislike of dishonesty. All that matters for our purposes is that it is a parameter in $(0, 1)$ where the DICC holds trivially when it is near 0 and we recover the standard DICC when it is near 1.

the supplier, which consists of a sequence $\{q_t, \tau_t\}_0^\infty$. The contract is relational, meaning that is is not enforceable in court: both parties can agree on a long-term plan, but are unable to commit to it. This implies that both parties face an enforcement constraint that requires that in every period t it must be optimal for them to take the contracted action. For the foreign supplier, this is already implied by their limited liability constraint (see below), so this is only relevant for the firm, for whom it implies the additional constraint $U_t \geq \bar{U}$ for all t. The enforcement constraint prevents the firm from (for example) promising large bonus payments to the supplier in the distant future, because when the distant future arrives the firm will prefer to renege on these contracted payments and instead simply take their outside option. We focus on trigger strategies in which both parties take their outside option if any player has ever deviated from the contract, because this provides maximal incentives.

Limited Liability: The contract must satisfy $\tau_t \geq cq_t$ for all t; that is, the supplier must make weakly positive profits period-by-period. We impose this restriction to prevent large period-0 rent extractions that we rarely see in our empirical setting. Instead, in our discussions with suppliers, we often heard of selling at cost early in the relationship, which this restriction permits (and which will occur in equilibrium). Note that this constraint means we can ignore the supplier's participation constraint and enforcement constraint.

Equilibrium: We focus on Perfect Public Equilibria in which the good type supplier chooses $a_t = 1$. We also restrict attention to pooling equilibria. This is without loss, as the inability to commit combined with the assumption that the bad type is inefficient ensures that the firm will immediately terminate the contract if they ever learn that the supplier is a bad type. The bad type thus always earns 0 upon revealing their type, and limited liability plus the DICC ensures they can always earn a strictly positive expected payoff by mimicking the good type. Similarly, limited liability prevents the good types from signalling their type by making period-0 losses that the bad types would not be willing to incur.

Before stating the full dynamic program, it is convenient to denote the firm's expected stage payoff as $y(q_t, \tau_t, \mu_t) \equiv (1 - \mu_t(1 - \lambda))r(q_t) - \tau_t$, and the supplier's stage payoff as $\pi(q_t, \tau_t) \equiv \tau_t - cq_t$.[8] We can then write the continuation values of the firm, U_t, and the supplier, V_t, recursively as

$$
\begin{aligned}
U_t &\equiv y(q_t, \tau_t, \mu_t) + \delta\left[(1 - \mu_t(1 - \lambda))U_{t+1} + \mu_t(1 - \lambda)\bar{U}\right], \\
V_t &\equiv \pi(q_t, \tau_t) + \delta V_{t+1}.
\end{aligned}
$$

[8]The bad type suppliers behave entirely mechanically, so there is no need to write out their payoff, and thus in general whenever we refer to the payoff or decision of a foreign supplier it is implied that it is a good type.

Standard arguments in the dynamic moral hazard and endogenously incomplete markets literatures establish that it is equivalent–and substantially easier–to solve this problem by explicitly letting the firm choose these continuation values and encoding them as state variables.[9] This implies that the firm solves the following recursive dynamic program in each t:

$$W_t(U_t, V_t, \mu_t) = \max_{q_t, \tau_t, U_{t+1}, V_{t+1}} y(q_t, \tau_t, \mu_t) + \delta(1 - \mu_t(1 - \lambda))W_{t+1}(U_{t+1}, V_{t+1}, \mu_{t+1})$$

subject to the following constraints:

$$\delta(1 - \lambda)V_{t+1} \geq \xi(c - c_0)q_t \qquad (DICC)$$

$$U_{t+1} \geq \bar{U} \qquad (DEC)$$

$$\tau_t \geq cq_t \qquad (LL)$$

$$y(q_t, \tau_t, \mu_t) + \delta\left[(1 - \mu_t(1 - \lambda))U_{t+1} + \mu_t(1 - \lambda)\bar{U}\right] \geq U_t \qquad (PK_f)$$

$$\tau_t - cq_t + \delta V_{t+1} \geq V_t, \qquad (PK_s)$$

with $\mu_{t+1} = \mu_t\lambda/(1 - \mu_t + \mu_t\lambda) < \mu_t$ if the good is high quality in t and $\mu_{t+1} = 1$ if the good is low quality in t, and with $U_0 = \bar{U}, V_0 = 0$, μ_0 given. We have already introduced the first three constraints. The final two constraints are known as promise-keeping constraints as they ensure that the continuation utilities promised in the previous period, U_t and V_t, are actually delivered through a combination of stage payoffs and future promises.

The solution to this program is generally not available in closed form. In Appendix B, we provide a detailed derivation of several properties of the solution, which we summarise briefly here. The optimal contract looks similar to a dynamic version of the typical "sell the firm to the agent" solution in static models without risk aversion. In particular, there exists some finite T^* such that the supplier will earn zero stage profits for periods $t = 0, 1, 2, ..., T^* - 1$ and then earn the entire surplus (net of the firm's outside option) for all $t > T^*$. Intuitively, the firm makes the supplier the residual claimant for most of the relationship—which is the most efficient way to provide incentives—and extracts surplus in the early periods as these minimise incentive distortions. This is stark, but not unreasonable: the supplier sells at cost at the beginning of the relationship while its reputation is being established, and reaps the benefits of its reputation later on. Moreover, backloading of incentives of this form is a very general prediction of models of reputation and dynamic moral hazard (Shapiro (1983), Banerjee and Duflo (2000), Ray (2002)).

[9]This approach was originally developed somewhat independently in different theoretical contexts by Spear and Srivastava (1987), Abreu, Pearce, and Stacchetti (1990), and Thomas and Worrall (1988). Golosov, Tsyvinski, and Werquin (2016) provide an excellent review in the context of incomplete markets models.

We now make two remarks relevant to intuition and the interpretation of the experiment.

Remark 1 (Intensive Margin)**.** *Quantity and value ordered from the foreign supplier are constrained by both adverse selection and moral hazard.*

The FOCs of this program yield the following equation relating marginal revenue and marginal cost,

$$r'(q_t^*) = \underbrace{\frac{1}{1-\mu_t+\mu_t\lambda}}_{\text{Adverse Selection}} \underbrace{(1+\xi\rho_t^*)}_{\text{Moral Hazard}} c,$$

where ρ_t^* is a weakly positive function of Lagrange multipliers. Both wedge terms are weakly greater than 1, and so distort q_t^* downwards relative to the level that equates marginal revenue and marginal cost. Learning over time will alleviate this as μ_t decreases, although the exact effects are somewhat complicated as ρ_t^* can evolve non-monotonically. A similar equation can be derived for value, τ_t. Treatments that alleviate these frictions should therefore increase quantity and value ordered, at least among firms that would have imported directly anyway.

Remark 2 (Extensive Margin)**.** *The period-0 value of the relationship is decreasing in μ_0 and ξ.*

This follows from a straightforward application of the Envelope Theorem, and highlights that the extent of adverse selection and moral hazard will limit the possible gains from trade, perhaps to the point where no trade occurs. Treatments that alleviate these frictions should therefore increase the propensity of firms to import directly.

3.3 Search

Matching with a foreign supplier is costly. A firm can pay a one-time search cost $s > 0$ to match with a random foreign supplier. Upon matching, the firm immediately observes a realisation of a match-specific productivity term, $\psi \sim G$, that is fixed over time, implemented by decomposing $c = \tilde{c}/\psi$, where $\tilde{c}$ is common across all supplier-firm pairs. We will denote the period-0 value of a relationship with a foreign supplier with match-specific productivity ψ as $U_0(\mu_0, \xi, \psi)$. Search is sequential, meaning that if a firm's current best option delivers discounted utility $\tilde{U}$, then the firm will search if $E_\psi[\max\{U_0(\mu_0, \xi, \psi), \tilde{U}\}] - s \geq \tilde{U}$. Standard arguments then imply that there exists a cutoff value $\bar{U}$ such that the firm will search if and only if their current best option is less than $\bar{U}$, and that their expected return to doing so is exactly $\bar{U}$.[10] This reservation value thus defines their outside option.

[10]The cutoff value is defined implicitly by equating marginal benefit of searching with marginal cost, that is, $\int_{\bar{\psi}(\bar{U})}^{\infty}(U_0(\mu_0, \xi, \psi) - \bar{U})f(\psi)d\psi = s$, where $\bar{\psi}(\bar{U})$ is the value of ψ for which $U_0(\mu_0, \xi, \psi) = \bar{U}$. This cutoff $\bar{U}$ is decreasing in s.

Finally, all firms can purchase instead from a local supplier, the cost of which is not heterogeneous and does not involve any frictions. If the value of purchasing from the local supplier is lower than $\bar{U}$, then they will search; otherwise, they will remain with their local supplier.

3.4 Implications for the Experiment

The model highlights the three frictions that we will study in the experiment: search, adverse selection, and moral hazard. Both adverse selection and moral hazard reduce the intensive and extensive margins of transacting with foreign suppliers, while the search friction reduces the extensive margin only. To relax the search friction, we need a treatment that either lowers the cost of matching with foreign suppliers or improves the likelihood of finding a suitable match. To relax adverse selection, we need a treatment that improves beliefs (or improves the ability to learn) about a supplier that a firm has been matched with. To relax moral hazard, we need a treatment that either improves the supplier's incentives directly or changes firms' perceptions about the cost to the supplier of not honouring the contract.

Importantly, the ideal experiment targeting adverse selection should have no effect if $\mu_0 \approx 0$; they should not affect the relationship if there is only one type of supplier. Similarly, the ideal experiment targeting moral hazard should have no effect if $\xi \approx 0$; they should not affect the relationship if the strategic type always chooses $a_t = 1$. We describe our experimental design that aims to achieve this in the next section.

4 Experimental Design

4.1 Treatment Conditions

The goal of the experiment is to generate variation that identifies the three frictions: search, adverse selection, and moral hazard (where we refer to the latter two jointly as trust frictions).

Search 80% of firms receive the Search treatment. The purpose of this treatment is to generate exogenous variation in the cost of finding a supplier of Turkish-made goods. We add treated firms to the supplier WhatsApp groups of 3 different suppliers.[11] The

[11]While this treatment bundles the role of connecting them to a supplier and the role of doing so specifically using social media, our view–and our experience from various pilots–is that these are effectively inseparable. If we had simply connected them by, say, giving their phone numbers, the suppliers would have simply added the firms to the WhatsApp groups anyway. This is simply how business takes place in this market.

suppliers to match with are selected at random, subject to being a match to the merchant's chosen sector. We do not give firms any information about the suppliers, except to say that they were recruited by a team in Turkey in a similar manner to how the firm itself was recruited. We communicate to the control group that unfortunately we cannot add them to any supplier groups at this time, but that we might do so at the conclusion of the study.

Adverse Selection 50% of firms in the Search treatment condition are treated with the Adverse Selection treatment. As the model highlights, identifying adverse selection requires either improving the ability to learn or providing information directly. This treatment consists of doing both. We add firms in this treatment condition to a fourth WhatsApp group. This group does not contain any suppliers, but instead contains other firms in the study that were matched with the same suppliers. We explain that all members of this fourth group have been matched with the same three suppliers and that the purpose of the group is to share information about them. A member of the study team moderates discussion and encourages firms to share information from time to time. We do not tell suppliers about the existence of these groups, so only the information of the firms–and not the incentives of the suppliers–are affected.[12]

Since no firms have experience with the supplier at this point, we seed the groups with initial information. Treated firms receive a phone call 2-3 days after recruitment from a recommender.[13] The recommender is part of a team of firms–who are not subjects in the study–that we hired prior to the study to make mystery orders from all of the suppliers. The recommender describes their experience ordering from one of the suppliers that the firm was matched with and sends a photo of the item that they ordered. They also explain that they, too, are in the information-sharing WhatsApp group, and post a similar message there.

Despite the fact that treated firms did not know the recommender personally, they generally took this information seriously for two reasons. First, at the end of the baseline survey, we ask all firms if they would be willing to call a few other firms to discuss their experiences working with the study suppliers. They are therefore not surprised when they receive this call. Second, one of the reasons that social media is so ubiquitously used for commerce is precisely the social nature: even if they don't know the recommender, they can ask questions and assess the preferences and knowledge of the recommender.

[12]We do not think that firms would expect to be able to credibly communicate the role of these groups to suppliers to try to improve incentives as the firm would have no easy way to convince the supplier that this is not cheap talk, especially since we make clear that we have not told the suppliers about these groups.

[13]All firms not in this treatment condition–both the other 50% of firms treated with Search and all firms not treated with Search–instead receive a "placebo" phone call from a surveyor, asking them for their opinion about supplier WhatsApp groups in general.

Moral Hazard 50% of firms in the Search treatment are treated with the Moral Hazard treatment, cross-randomised with the Adverse Selection treatment. As the model highlights, identifying moral hazard requires either shifting the suppliers' incentives or shifting firms' perceptions of suppliers' incentives. To do this, we read the following information to treated firms at the end of the baseline survey:

> I have one last piece of information to give you. As you know, you have been added to WhatsApp groups of Senegalese suppliers in Turkey.
>
> We work with many suppliers in our study. We want to assure you that they are motivated.
>
> We would like to collect feedback on these suppliers so that we can recommend the best ones in the future. To do this, we will ask the merchants in the study [such as yourself] to rate your experience with the suppliers we have presented to you on a scale of 1 to 5 on product arrival and quality. These reviews help identify the best suppliers, which is beneficial to them and allows us to continue recommending them to others. They are therefore motivated.
>
> If a supplier gets bad ratings, we will investigate and remove them from the study if they did not do a good job. They will therefore lose access to around 150 merchants if they do not do a good job.
>
> I will give you a phone number that you can use to give your rating or report a problem.
>
> Lastly, I want to emphasise that the suppliers are aware that they are being rated and that, if they receive bad ratings, they will be removed from the study. We can thus assure you that they are motivated.

After delivering this message, the surveyor provides a business card to the firm. The business card has a phone number to call, and prominently highlights that this number should be used to rate the suppliers and/or to signal any problems. Untreated firms receive a similar card, but without any mention of ratings or suppliers—instead saying that the phone number is for questions about the study. Both cards can be seen in Appendix Figure A2. All suppliers are told a similar message about how the ratings will work.

Note that the experiment does not randomise the incentives provided to suppliers. Instead, it provides high-powered incentives to all suppliers and randomises whether we tell this fact to firms. We make clear to firms not in this treatment condition that we in no way vouch for or provide guarantees about the suppliers in the study–our only role is to make connections. This treatment should not have an impact in a model of adverse selection only because it does not provide any information to the firms about the supplier–crucially, we do not share the ratings with the firms.[14]

Sub-Treatments We cross-randomised two additional sub-treatments within the pure control group. The first sub-treatment aims to test whether the binding constraint behind the lack of traditional B2B e-commerce platform usage is that firm owners do not understand how to use them. We thus provide a short training on Alibaba that covers how to install the app, how to search for products, how to contact suppliers, and how to make purchases and arrange delivery. The second sub-treatment is a placebo check for the fact that, in the Adverse Selection treatment condition, we have connected firms to each other. To ensure that results are not driven by grouping firms per se, we thus also create similar groups here, where none of the firms have been connected with any suppliers.

4.2 Randomisation and Balance Check

Overall, there are five equally likely groups: Pure Control, Search Only, Search + Adverse Selection, Search + Moral Hazard, and Search + Adverse Selection + Moral Hazard. We randomly assigned firms to one of these five groups, stratifying on product group (men's clothing, women's clothing, or shoes & bags), an indicator whether the firm has a physical store, and an indicator for whether the firm had prior direct importing experience. Any misfits, due to integer indivisibility or other issues, were unconditionally randomised across the five cells.

Since this is an RCT, treatment is independent of pre-randomisation covariates by construction, absent errors in the randomisation protocol. To check that the randomisation protocol operated as expected, we report a balance check in Appendix Table A1. The differences in means across treatment groups are all small and insignificant, and a joint test across variables has p-value 0.724.

5 Outcomes, Data, and Empirical Methodology

5.1 Data and Outcomes

Consumer Survey In March 2024, we conducted a 15-20 minute survey with 400 households. We use this to calculate two sets of summary statistics. First, we use it to measure

[14] In principle, since bad types will eventually draw a low quality realisation, firms could wait and attempt to infer types by observing whether the supplier is still around after a given number of periods. We do not think this happens for two reasons. First, our observation from pilots was that firms typically decide relatively quickly whether to pursue a relationship with the study suppliers. Second, the script does not imply that the enforcement process is particularly fast.

the relationship between consumer willingness to pay and various important variables in our analysis. Second, we use it to calculate statistics on household clothing expenditures.

Baseline Survey Upon recruiting a firm into the study, between November 2023 and January 2024, we conducted a 30 minute baseline survey. The survey contained questions on their supplier relationships, social media usage, e-commerce usage, and profit and sales. For profit, we use the summary survey question from De Mel, McKenzie, and Woodruff (2009), and a similar question for sales.

Access to Foreign Goods To test whether reducing search and trust frictions via social media improves firms' access to foreign goods, we need a measure of access to foreign goods. We constructed a novel measure of this using a detailed mystery shopping exercise. The goal was to measure access to foreign goods on a horizontal dimension (access to more differentiated varieties) and on a vertical dimension (access to higher quality varieties), as well as any effects on price. Around two weeks after recruitment, firms are contacted over WhatsApp by a mystery shopper, played by a trained surveyor. Firms are not aware that the customer is employed by the survey team, but are expecting to be contacted by customers, as we explain to them at the end of the baseline survey that we will put them in touch with customers who often buy high-quality goods.

The mystery shopper explains that they would like to purchase a certain high quality product for an event. Each product is defined by five horizontal criteria that are largely unrelated to quality, such as colour, sleeve style, and presence of a graphic (see Appendix Figure A3 for two examples). The mystery shopper proceeds with the purchase–including asking about price and delivery–if the firm has a good with at least three of the five criteria. The primary outcome for this horizontal component, pre-specified in our PAP, is an indicator for whether the firm had a good with at least three criteria.

If the firm has such a good, the mystery shopper buys it in a random 80% of cases.[15] Then, once the good arrives in our office, two tailors and an expert shoemaker assess its quality according to a 50-point scorecard that we developed.[16] To validate the quality measure, we also gave the surveyors conducting the consumer survey a subset of these goods to present and elicit willingness to pay (WTP). We show a binscatter of the relationship be-

[15]If the random draw indicates to not buy the good, the mystery shopper explains that they have had a change of plans. They offer a nominal payment of 2.5 USD as a gesture of gratitude for the firm's time. We piloted different ways of doing this, and found that this procedure was natural and largely avoided upset.

[16]We designed this scorecard together with the these hired experts specifically for this study. Vitali (2024), who studies the relationship between consumer search costs and firm location choices in Kampala, takes a similar approach to measure the quality of garments. Although the details of the scorecards are quite different, we benefited greatly from showing her scorecard to our hired experts as an example of what we had in mind.

tween the quality score and WTP in Panel (a) of Figure 5. There is a clear positive relationship, although it becomes flat in the left tail, reflecting the fact that beyond a certain point consumers simply view goods as "low quality". In Panel (b), we classify goods as "high quality" or "low quality" (defined as whether a good is above the median quality score of its product type), and plot the CDF of consumer WTP separately for these two sets of goods. The high quality CDF is shifted rightwards of the low quality CDF, with an average premium of 35%.

The outcomes are these two measures: the high quality indicator and the raw 50-point quality score. The rationale for the binary outcome is that it is not vulnerable to a long left tail of quality scores that, as we saw in Figure 5, are not meaningful in terms of WTP.

Finally, we also attempt to infer whether the good was manufactured in Turkey. As we showed in Figure 4, there is a large premium for Turkish-made goods since it is a strong signal of quality. Thus, while the other two vertical outcomes measure quality directly, in practice quality is not fully observable to consumers and so product origin plays an important role in consumer WTP.[17] For most goods, we record this information from the label, and the outcome is 1 if the label says "Made in Turkey" and 0 if it says it was made elsewhere.[18]

Followup Survey We conducted a 30-minute followup survey with similar questions to the baseline survey between February and April 2024, around 3 months after a firm is recruited to the study. We successfully surveyed 1671 firms, or 90% of the sample. The followup rate is very similar and not significantly different across the four treated groups, but is 5 percentage points higher and statistically significantly different in the pure control group. The main outcomes are questions about the number and location of the firms' suppliers, their profits and sales, and their e-commerce use.

Mobile Money To go beyond survey data, we also use real-time, transaction-level data from the largest mobile money provider in Senegal, Wave Mobile Money, made available for this study. This data contains the universe of transactions between the phone numbers of firms in the study and the phone numbers of study suppliers. This data complements survey-based measures and has several advantages: (1) it is dynamic, so we can see

[17]We pre-specified this outcome, but did not attach it to either the horizontal or vertical dimensions. Since the consumer preference for Turkish-made goods reflects a preference for quality, it seems more fitting to include it under the vertical dimension.

[18]For the small share of goods for which the label does not indicate the origin, we ask the hired experts to (independently) give their opinions as to whether the good was made in Turkey (based on sewing patterns, product style, and any other characteristics), and set the outcome to 1 if they both opine that it was made in Turkey and 0 otherwise. We report robustness to various alternative definitions in the appendix.

transaction profiles over time, (2) it continues long after the followup survey, (3) it is not self-reported.

While we cannot know the exact share of transactions taking place through this medium, we expect that it is relatively large, at least for retailers, for a few reasons. First, we asked the non-study firms that we hired to mystery order from all suppliers prior to the study (mentioned in Section 4.1) to record how the supplier asked them to pay, and in 100% of cases they were asked to pay with this particular mobile money provider. Second, in the baseline survey, 86% of firms reported that they often use this provider to pay suppliers when making payments at distance. We thus expect that we see most small-to-medium sized orders, but likely miss larger orders as—anecdotally—these are more likely to take place with more formal methods such as bank transfers or international transfer services (such as Western Union and Moneygram). Since wholesalers tend to make larger orders and have significantly more experience with formal methods, we expect that this dataset is more representative of retailers than wholesalers.

5.2 Empirical Methodology

Our primary empirical method, specified in our Pre-Analysis Plan (PAP), is to estimate the following OLS specification

$$y_i = \alpha + \sum_{j=1}^{4} \beta_j T_{ji} + \delta y_{0i} + \gamma_s + \rho' X_i + \varepsilon_i, \tag{1}$$

where y_i is the outcome for firm i and T_{ji} for $j = \{1, 2, 3, 4\}$ are indicators for treatment arms Search Only, Search + Adverse Selection, Search + Moral Hazard, and Search + Adverse Selection + Moral Hazard. y_{0i} is the outcome measured at baseline, if available. γ_s are stratum fixed effects. X_i are firm-level covariates, selected by Double Lasso, following the method of Belloni, Chernozhukov, and Hansen (2014).[19] We also report a version of the same regression where we pool the four treated groups.

Inference Our primary method of inference is randomisation inference, as recommended by Athey and Imbens (2017) and Young (2019). In particular, we compute two-sided p-values for the sharp null of no treatment effect using 5,000 permutations of the t-statistic. We report conventional formula-based heteroskedasticity-robust standard errors in parentheses, but we do not use these for inference directly. As each regression involves four

[19] This means that, prior to each regression, we run lasso to predict y_i and each T_i and include the union of selected covariates. In practice, and as Cilliers, Elashmawy, and McKenzie (2024) note is generally the case among published papers using PDS Lasso, this tends to select few covariates and thus makes little difference. We report versions of the main tables in Appendix C where exclude all covariates. The results are very similar.

main coefficients, we also calculate and report Romano-Wolf (RW) multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005).

Indexes To account for multiple hypothesis testing across outcomes, for any table that presents more than one outcome corresponding to the same family of outcomes, we also report the results on an index that aggregates the outcomes using the standardised inverse-variance weighted method of Anderson (2008). Since the disaggregated regressions may include different covariates y_{0i} and X_i, before indexing we first residualise each outcome using the covariates that were included in their respective regressions.

Quantile Regression As some of our outcomes, most notably profit and sales, may have thick tails and/or exhibit non-uniform distributional treatment effects (see, e.g., Meager, 2022), we also included in our PAP that we may use quantile regression to examine distributional treatment effects where relevant. For these, we follow the same specification as described above, except that we omit the stratum fixed effects (γ_s) and the vector of covariates (X_i) as quantile regressions are much more demanding and the covariate selection procedure in Belloni, Chernozhukov, and Hansen (2014) is designed for linear treatment effect models.

6 Results

In Section 6.1, we present results on access to foreign goods, as measured by our mystery shopping activity. In Section 6.2, we present results on supplier relationships. In Section 6.3, we present results on profit and sales. In almost all tables, we show the pooled regression in Panel A, and then disaggregate across the treated groups in Panel B, with standard errors in parentheses and Romano-Wolf multiple-testing adjusted p-values in square brackets.

6.1 Access to Foreign Goods

As described in Section 5, we designed a mystery shopping exercise to obtain a revealed preference of access to foreign goods. Table 1 reports the outcomes of this exercise.

Horizontal In Column 1 , we report the main horizontal outcome, which is an indicator for whether the firm had a product with at least three horizontal criteria.[20] Pooling the

[20]Our PAP specified this indicator variable as the main outcome. Nonetheless, as a robustness check, we use the raw number of criteria as an outcome in Columns 3 and 4 of Appendix Table A2. The pattern of the results is the same.

treatments together, treated firms are 9.0 percentage points more likely to find a suitable good, significant at the 1% level. This is a 25.2% increase from the control mean of 35.7%. In Panel B, we see that the effect is broadly similar in all four treatment groups.[21] The results thus show that relaxing the Search friction by connecting firms with foreign suppliers via social media led to a sizeable increase in the set of varieties that they can provide to real customers.

Vertical The outcome in Column 2 is an indicator for whether the product is "High Quality", defined as whether the product's quality score is greater than its product-group median. Pooling the treatments together, treated firms are 14.0 percentage points more likely to be high quality, significant at the 5% level. This is a 32.5% increase from the control mean of 43.1%. In Panel B, we see that the coefficient is positive and similar in all four treatment groups.

The outcome in Column 3 is the raw quality score out of 50. Here, the effect is both small and insignificant. In fact, the coefficients are negative in both groups with the Moral Hazard treatment. As we noted in Section 5, this outcome is vulnerable to a long left tail having an outsize influence that is not particularly meaningful. This is indeed what happens: in Appendix Figure A4 we plot the CDF of quality score by treatment status, and we see a handful of very low quality goods in exactly this group.

The outcome in Column 4 is an indicator for whether the product was made in Turkey. Pooling the treatments together, treated firms are 17.6 percentage points more likely to supply a good saying "Made in Turkey", significant at the 1% level, a 36.9% increase from the control mean of 47.7%.[22] In Panel B, we see that the effect is positive for all four treatment groups.

Finally, to account for multiple hypothesis testing across outcomes, we aggregate these three outcomes into a vertical index using the standardised inverse-variance weighting method recommended in Anderson (2008). The pooled coefficient is 0.422 standard deviations, significant at the 1% level, and is similar across the four separate treatments. The results thus show that relaxing the Search friction by connecting firms with foreign suppliers via WhatsApp groups led to a sizeable increase in their access to higher quality goods.

[21] In our PAP, we noted that we would also report the same outcome separated into extensive margin (agreeing to sell a Turkish-made product at all) versus intensive margin (how suitable was the product provided). We do this in Appendix Table A2. We see that, while the extensive margin effect is positive, the effect comes mostly through the intensive margin.

[22] As a robustness check, we report various alternative outcome definitions in Appendix Table A3, including specifications where we only use the information on the label and where we only use the expert tailors' judgements. The pattern of results is the same.

Price In Column 6 of Table 1, the outcome is the unit price charged by the firm. The effect is positive, but small, insignificant, and precisely estimated, so we can rule out modest price increases.

Summary Putting together these results, we find that all of the treated groups saw large and significant increases in access to foreign goods. In particular, treated firms are able to sell a wider set of varieties and higher quality varieties. The fact that we find no large effects on the price, while there are gains from variety and quality, suggests that consumer surplus has increased.

Across all of these outcomes, the pattern is fairly consistent: the results are largely driven by relaxing the Search friction. We therefore conclude that finding a supplier of Turkish-made goods is costly, and that WhatsApp can play an important role in alleviating this friction. This does not mean that the trust frictions do not exist: these are small orders, and so for many firms the risk is sufficiently low that relaxing the trust frictions is unlikely to have a large effect. Nonetheless, we can at least conclude that trust frictions cannot be so large as to prevent firms from experimenting with new suppliers.

6.2 Supplier Relationships

The mystery shopping exercise shows that relaxing the search friction (through social media) improves firms' access to foreign goods on both horizontal and vertical dimensions. However, to realise these gains in practice, firms need to overcome the trust frictions (if any) and develop these connections into relationships. This section examines whether the treatments caused new relationships to develop, as well as what happened to previous relationships.

6.2.1 Survey Data

In our follow-up survey, conducted after 3 months, we asked firms how many regular suppliers they had, and where those regular suppliers were based. We defined a regular supplier as any supplier from whom the firm had made at least two orders, and intended to continue the relationship. We analyse these outcomes in Table 2. In Columns 1, the outcome is an indicator for whether the firm has a regular in Turkey. In Column 2, the outcome is the number of regular suppliers in Turkey. In both cases, pooling all treatments together–including Search Only–shows that treatment caused firms to develop new relationships with suppliers in Turkey. The pooled coefficients are 3.7 percentage points (an increase of 22.2%), significant at 10% and 0.083 suppliers (an increase of 37.4%), significant at 5%.

Disaggregating by treatment arm, for both outcomes the coefficient for Search + Adverse Selection + Moral Hazard (the group where we relax both trust frictions) is substantially larger, around twice the pooled coefficient. This suggests that relaxing trust frictions, and in particular relaxing them together, increased the likelihood that these new connections developed into regular relationships.

Having seen evidence that new supplier relationships developed, it is therefore natural to ask whether these new relationships complement or substitute for existing relationships. We provide some evidence to address this in columns 4, 5, and 6. In column 4, the outcome is the total number of suppliers; in column 5, the outcome is the number of suppliers in Senegal. The general direction looks closer to a world of substitutes: the coefficients on the total number of suppliers are close to zero, and the coefficients on the number of suppliers in Senegal are of similar magnitude (but opposite sign) to the coefficients on the number of suppliers in Turkey. Finally, column 6 shows an indicator for whether the firm said that they have ended a relationship with a regular supplier in the past 3 months. The coefficients are generally negative, which is also suggestive of substitutes.

6.2.2 Mobile Money Data

As discussed in Section 5, we use data made available for this research from the largest mobile money provider in Senegal to directly observe transactions between firms in the study and study suppliers. Before turning to formal regression results, we first show broad patterns. In Figure 6, we plot cumulative order value from study suppliers by treatment group over the course of the study.[23] The dashed line shows when we finished our mystery shopping activities. The figure shows two striking patterns. First, over the mystery shopping period, the total value ordered is very similar across the four treatment groups. This suggests that trust treatments were not necessary to make small experiments with the study suppliers when there is no demand risk as a customer is already present. The total value ordered is much larger than the total value purchased by our mystery shoppers, so this not simply coming from buying and re-selling to us, but it is a bit challenging to interpret this as a secondary goal of the mystery shopping was to lower the cost for firms to experiment. Second, almost immediately after the mystery shopping ends, the Search Only line becomes flat, suggesting that these relationships were not lasting. In contrast, in all three of the trust treatments, the total value ordered continues to increase well beyond when the mystery shopping ended, suggestive of continuing relationships.

We formally test these patterns in Table 3. In this table, the omitted category is Search

[23]The pure control group is omitted from the figure because they were not connected to the study suppliers. Reassuringly, we find a negligible number of orders from such firms.

Only, as this outcome is not defined for pure control. Column 1 shows the effect of the trust treatments on the probability of making any order from the study suppliers. We find no significant effect, either in the pooled coefficient or the disaggregated coefficients. In Columns 2 and 3, we test the observation from Figure 6 that only the trust treatments appear to continue ordering after the mystery shopping ends. The outcome in both cases is the total value ordered after the mystery shopping ends, aggregated over weeks to the firm level. We did not pre-specify this outcome, but rather included it after observing the pattern in Figure 6. The coefficient pooling all three trust treatments together in the Poisson regression implies an $e^{1.351} - 1 \approx 286.1\%$ increase, significant at the 5% level. The pooled coefficient in the OLS regression is similar and also significant at the 5% level. When we disaggregate the treatments, we find that the coefficient is positive in all three trust treatment groups. While it is larger in the Moral Hazard groups, the standard errors are large. Finally, in columns 5 and 4, we report the effects on total value ordered over the entire course of the study. The coefficients are all positive, although they are not significant.

The OLS coefficients in Table 3 are very similar between total value after the mystery shopping and total value, which is in line with the first observation from Figure 6 that the differences only open up after the mystery shopping ends. Moreover, the fact that the trust treated groups order more after the mystery shopping ends (and thus more overall), but there is no effect on the probability of ever making an order, highlights that these are lasting relationships formed during the study period rather than new firms that start to order after the mystery shopping ends.

6.2.3 Supplier Relationships: Summary

Putting together the results from Sections 6.2.1 and 6.2.2, we find that trust treatments increased the probability that these new connections developed into relationships. From the survey data, these effects come primarily from the Search + Adverse Selection + Moral Hazard group, and, to some extent, the Search + Adverse Selection group. In the mobile money data, these effects come from all three groups.

6.3 Profit and Sales

6.3.1 Mean Results

In order to see whether improved access to foreign goods and new relationships flow through to profits, we report the results on profits and sales in Table 4. For profit, we use the summary survey question from De Mel, McKenzie, and Woodruff (2009). For sales, we use a similar summary question.

Columns 1-2 show the results on raw profit and sales. We find large and statistically significant effects. For profit, the pooled coefficient is 82.4 USD, significant at the 5% level, or a 43.8% increase from the control mean. For sales, the pooled coefficient is 237.4 USD, significant at the 5% level, or a 39.0% increase from the control mean. When we look at the effects of the four treatments separately in Panel B, we see a similar pattern for both outcomes. While the effect is positive in all four groups, it is substantially larger and highly significant in Search + Adverse Selection + Moral Hazard group (including after adjusting for multiple hypothesis testing), which is also reflected in the linear combinations in Panel C. The same pattern holds when we combine these two outcomes into an index.

To limit the influence of outliers, in Columns 4-5 we report the results when we winsorize the outcomes at the top 1%. The coefficients decrease in magnitude by around half on average, but the same pattern remains: the Search + Adverse Selection + Moral Hazard group has a very large and highly significant coefficient, including after adjusting for multiple hypothesis testing.

6.3.2 Distributional Results

It is well-known in the literature studying small firms in lower-income countries that profit and sales tend to be thick-tailed, and that these tails can have outsize effects on the coefficients in OLS regressions (Meager (2022)). Thus, as discussed in Section 5.2 and specified in our PAP, we use quantile regression to examine distributional effects.[24]

Quantile Treatment Effects In Figure 7 Panel (a), we show the quantile treatment effects for profit for percentiles 5-95. Across all four groups, the coefficients are small and generally insignificant for percentiles 5-65. However, starting from around the 75th percentile, the Search + Adverse Selection + Moral Hazard group coefficient becomes large and significant. The coefficients for Search Only and Search + Adverse Selection are also large at the 95th percentile, but are very noisy. We report the same analysis for sales in Panel (b). The results are similar: there is little evidence of an effect for percentiles 5-65, but it begins to increase at around the 75th percentile for the Search + Types + Actions group, with some positive but noisy effects for Search Only at the 95th percentile.

The increasing trend in both profit and sales from percentiles 75 to 95 also suggests there may be potentially very large effects in the top 10 percentiles. We thus report the same approach for percentiles 90-99 in Appendix Figure A6. With the caveat that these are very

[24] To verify that treatment effects are not driven by measurement error in the tails, we called back all firms whose profit was more than 5 times that at baseline and exceeded a high threshold at endline, and we asked them to confirm their previous survey response. Out of 13 such firms, 12 confirmed that their previous response was correct.

demanding specifications, the coefficients are generally large and positive, further suggestive of treatment effects at the top of the distribution.

Threshold Regression An alternative way to analyse distributional treatment effects is to construct a series of indicator variables that are 1 if the outcome is greater than t, for a range of t, and run OLS regressions where these indicators are the outcome variable (using the specification in Section 5.2). This has the advantage of being unaffected by high variance in the tails: all that matters is whether the outcome is above the threshold t. The results, reported in Appendix Figure A5, are similar to the quantile regressions: large and positive treatment effects near the top of the distribution for the Search + Adverse Selection + Moral Hazard group, and some suggestive evidence of positive effects for the other groups at the very top.

6.3.3 Profit and Sales: Summary

In Table 4, we saw large average effects on profit and sales. This suggests that search and trust frictions have quantitatively important effects on firm profits, and that alleviating using social media can unlock large gains. As in Section 6.2, the effects are concentrated in the group with trust interventions.

We do not think that these effects are simply the result of a few outliers that happen to be in the treatment group, for several reasons. First, the positive distributional effects are coming from at least the top 5% of the distribution, which is considerably more than a few outliers. Second, the threshold regressions use indicators as their outcomes and thus are immune to the risk of a few observations having outsize influence. Third, the p-values in Table 4 highlight that the patterns we observe are very unlikely under the null, including after adjusting for multiple hypothesis testing. Fourth, the results in Table 4 are robust to winsorising the top 1% of firms. Fifth, as a placebo check, we compute the same quantile figures using baseline profit as the outcome in Appendix Figure A7 (we do not have a measure of sales at baseline), and find no evidence of this pattern.

A positive effect driven by the upper tail of the profit distribution is not unusual in the literature studying firms in lower- and middle-income countries. For example, Meager (2022) aggregates the results of six RCTs on microcredit, and finds consistent evidence of little-to-no effects on profit throughout most of the distribution, with large but uncertain effects near the top. Another example is De Mel, McKenzie, and Woodruff (2013), who study the effect of formalisation among small firms in Sri Lanka, and find profit results driven by the upper tail. Moreover, an effect concentrated among a relatively small number of firms is consistent with a small subset of firms developing meaningful relationships with the study suppliers.

6.4 e-Commerce Platforms

Our final set of outcomes relate to the Alibaba training sub-treatment. The goal of this treatment was to test whether the binding constraint explaining the very limited use of formal e-commerce platforms is that firms find these platforms too complex. Thus, we regress outcomes relating to Alibaba use against an indicator for whether the firm was in the Alibaba training group. Since we only randomised this training among firms that received none of the main treatments (i.e., the pure control group), we exclude firms that received any of the main treatments from this regression for ease of interpretation (although the results turn out to be the same if we include them).

We report the results in Table 5. The training has a first-stage: treated firms are 6.5pp more likely to have heard of Alibaba, 11.3pp more likely to have searched for goods on Alibaba, and 8.7pp more likely to have compared prices on Alibaba with prices from their regular supplier. However, they are no more likely to have actually made a purchase from Alibaba. The coefficient is 1.4pp and the standard errors are small enough to at least rule out modest to large effects. These results provide strong evidence against the hypothesis that the binding constraint is that firms struggle to understand how to use the platform. While our experiment was not designed to directly test social media against formal platforms, we can speculate that the fact that firms clearly prefer social media as their main way of doing e-commerce likely reflects something deeper about how social media–in this context–relaxes frictions in a way that formal platforms do not.

7 Retailers vs Wholesalers

Our study contains both retailers and wholesalers. In particular, 33% of the sample have at least some wholesaler clients. These types of firms are quite different, and it was *ex ante* unclear we should expect direct connections to foreign suppliers via social media to be more relevant for retailers or wholesalers. Retailers are generally too small to travel, and thus their baseline technology to access foreign goods (in absence of social media) is typically to buy from local wholesalers. Retailers may benefit from direct connections if it allows them to access a greater set of varieties, or to get better prices by shortening the supply chain. Wholesalers, on the other hand, are significantly more likely to travel at baseline. However, travel is expensive, and if social media allows them to partially substitute for travel, then the implications for profit are large. We thus specified that we would examine this dimension of heterogeneity in our PAP, which is what we do in this section.

We report the indexes for the main outcomes in Table 6. We use the same specification as that used throughout the paper, except that we now interact each treatment indicator with an indicator for whether the firm has wholesale clients. Throughout the table, the standard errors on the interaction coefficients are very large, so at best we can speculate on these differences. For the Foreign Market Access outcomes in Columns 1 and 2, the results appear fairly similar across retailers and wholesalers. For the index on suppliers in Turkey, the interaction coefficients are roughly equal in magnitude to the retail coefficients, perhaps suggesting a larger effect for wholesalers. Finally, there are some differences in the profit and sales index: there are essentially no (detectable) effects for retailers, while for wholesalers the pattern replicates that from the overall sample.

We thus conclude that the treatment does improve the foreign market access of both retailers and wholesalers, as evidenced by the increase in both the horizontal and vertical outcomes. The increase in foreign suppliers thus suggests that retailers substitute away from their local suppliers, but, since their scale of operations is small and this process takes time, any increases in profits in the short to medium run are sufficiently small that we are unable to detect them. Wholesalers, on the other hand, see a large increase in profit. This may be because wholesalers buy and sell in large quantities, so a new supplier can have more immediate implications for profit, or it may be because they have saved on travel costs. While we do not know for sure which one is the case, in Appendix Table A6, we show results on indicators for whether the firm has travelled internationally for business in the past 3 months (collected in the followup survey). The effects for wholesalers are negative and very large, with a few significant coefficients, but the standard errors are also large so these effects should be thought of as suggestive at best.

8 Model Estimation

In this section, we use the results of the experiment to estimate the model from Section 3. The goal is to estimate the parameters governing the search and trust frictions and to evaluate the gains from trade available if the frictions were to be completely or near-completely removed.

Modifications to the Model In order to make the model estimable, we modify the basic model in a few ways. First, we need to implement a functional form for the revenue function, $r(q)$, which we do in the following way. The firm faces and internalises a constant elasticity residual demand curve for goods, $Q^{-1/\sigma} = \nu P$, where $\sigma > 1$ is the elasticity of demand and ν is a demand shifter. The firm can produce the aggregate good Q by purchasing inputs from either a given foreign supplier that they are matched with, q_f, or their existing supplier, q_e, which we let be perfectly substitutable.[25] The firm can purchase from their existing supplier in unlimited quantities at constant price p_e without frictions. Timing is such that the firm chooses (q_f, q_e), observes whether the foreign order q_f is high quality and therefore whether $Q = q_f + q_e$ or $Q = q_e$, and then sells to their downstream buyer. The firm's stage game payoff is therefore the following,

$$(1-\mu_t(1-\lambda))z(q_{ft}+zq_{et})^{\frac{\sigma-1}{\sigma}} + \mu_t(1-\lambda)q_{et}^{\frac{\sigma-1}{\sigma}} - \tau_t - p_e q_{et}, \tag{2}$$

where $z > 0$ is a productivity term. The firm will always choose $q_{et} > 0$, but may choose $q_{ft} = 0$ if either μ_t or the transfers $\{\tau_s\}_t^\infty$ required to incentivise the foreign supplier are sufficiently large. Firms only update their beliefs if they order positive quantity, and to prevent arbitrarily small orders to allow near-costless learning we impose a minimum order size of $\underline{q} > 0$, which we calibrate later and is in fact a feature of our empirical setting.

Second, we need sufficient heterogeneity to match real data. The model in Section 3 had one dimension of heterogeneity in the form of match-specific productivity ψ, implemented by defining the constant marginal cost of the foreign supplier as $c = \tilde{c}/\psi$. We set $\tilde{c} = 1$ and then let ψ be distributed lognormal with parameters (ψ_μ, ψ_σ), which we will estimate. The match-specific productivities are important theoretically as they define firms' beliefs about the value of searching, and empirically as they rationalise the fact that two otherwise-identical firms matched with study suppliers may order different amounts. We also allow the productivity, z, to be heterogeneous, drawn from the empirical distribution implied by baseline profits.[26] This is important to rationalise the firm size distribution at baseline.

Third, we need to take a stand on what firms' outside options are and what how to account for the fact that some firms already have foreign suppliers at baseline. The sequential search process described in Section 3 implies that there exists a cutoff value $\bar{z}$ such that, at baseline, all firms with $z > \bar{z}$ will search until they eventually find a foreign supplier with sufficiently high match-specific productivity. This cutoff is a function of parameters to be estimated. We interpret the baseline equilibrium as the very long-run of the model, meaning that all firms with $z > \bar{z}$ have found a good-type foreign supplier and fully

[25] In principle, these could be imperfect substitutes if foreign goods are local and foreign goods represent horizontally differentiated varieties. We focus on the case of perfect substitutes as it allows for analytical solutions, which dramatically speeds up the computation as this (albeit simple) problem must be solved within an inner loop in the program to solve the dynamic model.

[26] Specifically, assuming that the firm faces one input price, the model implies that profits for firm i are given by $\pi_i = z^\sigma \frac{(1-\sigma)^{1-\sigma}}{\sigma^\sigma} p_i^{1-\sigma}$, where p_i is the input price. Re-arranging allows us to write z_i as a function of π_i, p_i, σ. We use monthly profit from the baseline survey for π_i, average price of the most common input from the baseline survey for p_i, and we calibrate σ as described below. This gives an empirical distribution of z_i.

learned their type. For such firms, p_e represents the price offered by their foreign supplier (which is itself a draw of ψ). For all other firms, p_e represents the cost of buying from a local supplier in Dakar, which we will calibrate below. A firm's outside option, $\bar{U}$, is then defined by the value of only purchasing from their existing supplier forever,

$$\bar{U} = \left(\max_{q_e} q_e^{\frac{\sigma-1}{\sigma}} - p_e q_e \right) / (1-\delta).$$

Parameters Our goal is to estimate the parameters governing the three frictions: the search cost, s, the share of bad types, μ_0, and the moral hazard multiplier, ξ. We also need to estimate the parameters governing the lognormal distribution of match-specific productivity, (ψ_μ, ψ_σ), which may also be thought of as a characterising the search friction. We thus estimate these five parameters, and calibrate all others. We describe the values and origins of the calibrated parameters, (σ, δ, p_l), in Appendix Table A7.

Moments Since the ultimate goal is to use the model to extrapolate from the treatments to consider counterfactuals where the frictions are even lower (or removed altogether), we estimate the model by simulating the impact of treatment in the model and then matching the reduced form treatment effects. We select the four treatment effects on winsorized profit, the four treatment effects on likelihood of having a supplier in Turkey, and the three treatment effects on mobile money order value post mystery shopping.

We simulate the three treatments in the model as follows. For search, we implement this as the firm being matched (at zero cost) to a foreign supplier with match-specific productivity equal to the maximum of three draws from the distribution (to capture the idea that the firm is matched to three foreign suppliers in the experiment). For adverse selection, we implement this as the firm receiving one high quality signal realisation without having to purchase anything, meaning that they update as a Bayesian and thus beginning the relationship with $\mu_1 = \mu_0\lambda/(1-\mu_0(1-\lambda)) < \mu_0$. This intends to capture the part recommendation call from the adverse selection treatment that explicitly told the firm about one positive order experience. For moral hazard, we implement this as treated firms playing a joint punishment strategy. Specifically, they face a modified DICC of the form $(1-\lambda)\delta V_{t+1} \geq \xi c q_t - (1-\lambda)\delta N \tilde{V}_{t+1}$, where $\tilde{V}_{t+1}$ is the average relationship value among other firms and N is the empirical number of firms that ever ordered from the average supplier.

Solving and Estimating the Model In order to compute the simulated treatment effects, we need to solve the model. The model is dynamic with a non-stationary optimal contract, meaning that solving it numerically is non-trivial. We use the method of Marcet and Marimon (2019), which involves rewriting the original Lagrangean recursively and then defining a Saddle Point Functional Equation (SPFE), which is analogous to a standard Bellman Equation for saddle point problems. We can then use standard dynamic programming techniques, and in particular we iterate on the value function implied by the SPFE. For estimation, we use Simulated Method of Moments (SMM) with a weighting matrix equal to the inverse of the variance-covariance matrix of the empirical moments, and with 100 randomized initial points for the numerical minimization algorithm. We provide a more detailed description of the procedure to solve and estimate the model in Appendix D.

Results We obtain $[\mu_0, \xi, s, \psi_\mu, \psi_\sigma] = [0.68, 0.59, \$13.1, -3.05, 0.40]$. The value of $\mu_0 = 0.68$ suggests a fairly severe adverse selection problem: around two-thirds of suppliers are bad types. The value of $\xi = 0.59$ is in between a standard model of moral hazard and a model with perfect enforcement (at least among good types), although the overall extent of moral hazard depends on whether and how severely the DICCs bind. The search cost of $s = 13.1$ USD may not, on the surface, seem especially large. However, the estimated parameters governing the distribution of match-specific productivity imply that the probability of a random foreign supplier being a good enough match to transact with is in the order of 10%. An average firm would thus need to pay the search cost many times before finding a suitable match.

Counterfactuals We now use the estimated model to evaluate two types of counterfactuals. In the first type, we consider a counterfactual where no firms can import directly. This is equivalent to setting $s = \infty$. With the baseline estimated parameters, the discounted value of profits, U_0, is 7,953 USD. With the calibrated monthly discount factor of 0.96, this implies a long-run average monthly profit of around 300 USD, which is broadly in line with our baseline profit statistics. When we set $s = \infty$, average lifetime discounted profits decrease to 6,005 USD, a decrease of around one quarter. The ability to trade with foreign suppliers thus quantitatively matters. Note that, since the search cost is fixed, the firms that do trade with foreign suppliers at baseline are larger.

For the second type of counterfactuals, we set various combinations of the parameters governing the frictions to zero or near zero. We show the results in Table 7. We first cut the search cost in half–reducing it to zero is not well-defined because as $s \to 0$ firms will search indefinitely to obtain arbitrarily good matches. This increases lifetime discounted profits by relatively little, only raising it to 8,014 USD. This happens because the likelihood of finding a match that will be worth developing a relationship with, governed by the distribution of match-specific productivity and the trust frictions, is relatively low. So relatively few firms are induced to search when the costs are lowered. Setting ξ to zero has almost no effect, at least while μ_0 is so large. Intuitively, if the majority of suppliers are bad types, then improving incentives has little effect. In contrast, setting μ_0 to 0 has a much larger effect, raising lifetime discounted profits to 9,661 USD, or around a 20% increase. This magnitude is not unreasonable when considering the profit effects. The effect is large both because the high value of μ_0 prevents many relationships from forming altogether and necessitates significantly time spent learning even within formed relationships. Finally, note also that the impact of setting $\xi = 0$ increases substantially from almost nothing to around 150 USD when $\mu_0 = 0$. This happens because the treatments are complements: in the model, if μ_0 is very high, then the firm wants to choose low values of q_t for early t. But since moral hazard only has bite by constraining q_t, this endogenous belief-driven desire to lower q_t means that moral hazard imposes little pressure. Overall, the largest gains from reducing the frictions beyond the implemented treatments come from reducing adverse selection.

9 Conclusion

We study the nature of search and trust frictions in international trade, and the extent to which social media can alleviate them. Finding and developing relationships with suppliers is a first order issue for firms, but one rife with these information frictions. Providing rigorous evidence on these issues is therefore important, both to understand the descriptive fact that firms use social media ubiquitously to interact with suppliers and to guide policy aiming to improve business outcomes.

Our results suggest that social media can meaningfully reduce these frictions. First, we find that connecting firms with foreign suppliers improves their access to foreign inputs, as shown by the results of our mystery shopping exercise. This is not *ex ante* obvious: instead of using social media to interact directly with suppliers abroad, firms could buy from a local wholesaler or travel abroad themselves. Yet, the fact that we find relatively large effects suggests that substantial search frictions are still present, and that direct connections via social media can meaningfully reduce them.

Our mystery shopping results are largely driven by relaxing the search friction. Ultimately, though, whether firms are able to realise this better foreign market access depends on whether these connections develop into lasting relationships. Our findings, using both survey data and mobile money data, suggest that this is more likely to happen among firms in the trust treatments. This suggests that trust frictions are present in this setting, and that using social media to share information and coordinate action can reduce them. Ultimately, we find that our treatments cause some firms to develop meaningful new relationships and increase profit.

The widespread use of social media suggests that there is substantial demand for e-commerce, so it may therefore seem puzzling why firms don't use formal platforms that are explicitly designed to alleviate search and trust frictions. The fact that our Alibaba training has no effect on their propensity to buy inputs from it suggests that the reason is more fundamental than simply lack of knowledge. The exact reasons for limited use seem like fruitful topics for future research, but, the fact that firms clearly prefer social media as their main way of doing e-commerce likely reflects something deeper about how social media–in this context–relaxes frictions in a way that formal platforms do not.

Taken together, our results show that search and trust frictions are large, and that social media can meaningfully reduce them. Moreover, it has the potential to change the structure of supply chains for some firms: retailers can import directly, and wholesalers can save on travel costs. Policies that improve smartphone access and mobile connectivity are therefore likely to benefit firms through this channel and to potentially reshape supply chains. Similarly, social media companies are already introducing e-commerce features onto their apps (e.g., better search features or better virtual storefronts), and this is likely to further these changes. Our results thus suggest that the rapidly developing digital landscape in lower and middle income countries is likely to lead to substantial benefits for small businesses and require researchers, policymakers, and organisations to update how they think about how firms find, learn about, and develop relationships with suppliers in these contexts.

References

D. Abreu, D. Pearce, and E. Stacchetti. Toward a theory of discounted repeated games with imperfect monitoring. *Econometrica: Journal of the Econometric Society*, pages 1041–1063, 1990.

L. Alhorr. Virtual windows through glass walls? 2024.

M. L. Anderson. Multiple inference and gender differences in the effects of early intervention: A reevaluation of the abecedarian, perry preschool, and early training projects. *Journal of the American statistical Association*, 103(484):1481–1495, 2008.

P. Antras and C. F. Foley. Poultry in motion: a study of international trade finance practices. *Journal of Political Economy*, 123(4):853–901, 2015.

S. Athey and G. W. Imbens. The econometrics of randomized experiments. In *Handbook of economic field experiments*, volume 1, pages 73–140. Elsevier, 2017.

J. Bai, M. X. Chen, J. Liu, X. Mu, and D. Y. Xu. Stand out from the millions: Market congestion and information friction on global e-commerce platforms, 2023.

A. V. Banerjee and E. Duflo. Reputation effects and the limits of contracting: A study of the indian software industry. *The Quarterly Journal of Economics*, 115(3):989–1017, 2000.

M. A. Barach, J. M. Golden, and J. J. Horton. Steering in online markets: the role of platform incentives and credibility. *Management Science*, 66(9):4047–4070, 2020.

A. Belloni, V. Chernozhukov, and C. Hansen. Inference on treatment effects after selection among high-dimensional controls. *Review of Economic Studies*, 81(2):608–650, 2014.

J. Cilliers, N. Elashmawy, and D. McKenzie. Using post-double selection lasso in field experiments. *World Bank working paper*, 2024.

V. Couture, B. Faber, Y. Gu, and L. Liu. Connecting the countryside via e-commerce: evidence from china. *American Economic Review: Insights*, 3(1):35–50, 2021.

S. De Mel, D. J. McKenzie, and C. Woodruff. Measuring microenterprise profits: Must we ask how the sausage is made? *Journal of development Economics*, 88(1):19–31, 2009.

S. De Mel, D. McKenzie, and C. Woodruff. The demand for, and consequences of, formalization among informal firms in sri lanka. *American Economic Journal: Applied Economics*, 5(2):122–150, 2013.

M. Dinerstein, L. Einav, J. Levin, and N. Sundaresan. Consumer price search and platform design in internet commerce. *American Economic Review*, 108(7):1820–1859, 2018.

G. Ellison and S. F. Ellison. Search, obfuscation, and price elasticities on the internet. *Econometrica*, 77(2):427–452, 2009.

Y.-f. Fong and J. Li. Relational contracts, limited liability, and employment dynamics. *Journal of Economic Theory*, 169:270–293, 2017.

A. Goldfarb and C. Tucker. Digital economics. *Journal of economic literature*, 57(1):3–43, 2019.

M. Golosov, A. Tsyvinski, and N. Werquin. Recursive contracts and endogenously incomplete markets. In *Handbook of Macroeconomics*, volume 2, pages 725–841. Elsevier, 2016.

A. Greif. Contract enforceability and economic institutions in early trade: The maghribi traders' coalition. *The American economic review*, pages 525–548, 1993.

GSMA. E-commerce in africa: Unleashing the opportunity for msmes. `https://www.gsma.com/solutions-and-impact/connectivity-for-good/mobile-for-development/gsma_resources/e-commerce-in-africa-unleashing-the-opportunity-for-msmes/`, 2023. Accessed: 2024-07-10.

M. Halac. Relational contracts and the value of relationships. *American Economic Review*, 102(2):750–779, 2012.

J. Hörner. Reputation and competition. *American economic review*, 92(3):644–663, 2002.

R. Macchiavello and A. Morjaria. The value of relationships: evidence from a supply shock to kenyan rose exports. *American Economic Review*, 105(9):2911–2945, 2015.

R. Macchiavello and A. Morjaria. Competition and relational contracts in the rwanda coffee chain. *The Quarterly Journal of Economics*, 136(2):1089–1143, 2021.

A. Marcet and R. Marimon. Recursive contracts. *Econometrica*, 87(5):1589–1631, 2019.

R. Marimon and J. Werner. The envelope theorem, euler and bellman equations, without differentiability. *Journal of Economic Theory*, 196:105309, 2021.

D. Martimort, A. Semenov, and L. Stole. A theory of contracts with limited enforcement. *The Review of Economic Studies*, 84(2):816–852, 2017.

R. Meager. Aggregating distributional treatment effects: A bayesian hierarchical analysis of the microcredit literature. *American Economic Review*, 112(6):1818–1847, 2022.

J. E. Rauch. Networks versus markets in international trade. *Journal of international Economics*, 48(1):7–35, 1999.

J. E. Rauch. Business and social networks in international trade. *Journal of economic literature*, 39(4):1177–1203, 2001.

D. Ray. The time structure of self-enforcing agreements. *Econometrica*, 70(2):547–582, 2002.

J. P. Romano and M. Wolf. Exact and approximate stepdown methods for multiple hypothesis testing. *Journal of the American Statistical Association*, 100(469):94–108, 2005.

C. Shapiro. Premiums for high quality products as returns to reputations. *The quarterly journal of economics*, 98(4):659–679, 1983.

S. E. Spear and S. Srivastava. On repeated moral hazard with discounting. *The Review of Economic Studies*, 54(4):599–617, 1987.

M. Startz. The value of face-to-face: Search and contracting problems in nigerian trade. 2024.

Statista. E-commerce in africa. `https://www.statista.com/study/83935/e-commerce-in-africa/`, 2023. Accessed: 2024-07-10.

C. Steinwender. Real effects of information frictions: When the states and the kingdom became united. *American Economic Review*, 108(3):657–696, 2018.

G. J. Stigler. The economics of information. *Journal of political economy*, 69(3):213–225, 1961.

J. Thomas and T. Worrall. Self-enforcing wage contracts. *The Review of Economic Studies*, 55 (4):541–554, 1988.

A. Vitali. Consumer search and firm location: Theory and evidence from the garment sector in uganda. *Working Paper*, 2024.

H. Yang. Nonstationary relational contracts with adverse selection. *International Economic Review*, 54(2):525–547, 2013.

A. Young. Channeling fisher: Randomization tests and the statistical insignificance of seemingly significant experimental results. *The quarterly journal of economics*, 134(2):557–598, 2019.

A. Young. Asymptotically robust permutation-based randomization confidence intervals for parametric ols regression. *European Economic Review*, 163:104644, 2024.

Tables

Table 1: Access to Foreign Goods

	Horizontal	Vertical				Price
	Find Product ≥ 3 Criteria	High Quality Dummy	Quality Score (/50)	Made in Turkey	Index	Price (USD)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled						
Treatment	0.090***	0.140**	-0.372	0.176***	0.422***	0.715
	(0.030)	(0.067)	(0.580)	(0.066)	(0.138)	(0.698)
Panel B: Individual Treatments						
Search Only	0.134***	0.181**	0.350	0.143*	0.512***	1.396
	(0.038)	(0.082)	(0.675)	(0.080)	(0.167)	(0.887)
	[0.002]	[0.097]	[0.826]	[0.138]	[0.012]	[0.339]
Search + Adverse Selection	0.048	0.165*	-0.081	0.136	0.419**	0.317
	(0.038)	(0.088)	(0.871)	(0.083)	(0.180)	(0.843)
	[0.201]	[0.155]	[0.924]	[0.138]	[0.059]	[0.900]
Search + Moral Hazard	0.100***	0.104	-1.152	0.189**	0.370**	1.032
	(0.038)	(0.084)	(0.826)	(0.080)	(0.171)	(0.855)
	[0.022]	[0.346]	[0.431]	[0.046]	[0.059]	[0.493]
Search + AS + MH	0.077**	0.104	-0.662	0.235***	0.381**	0.063
	(0.038)	(0.083)	(0.836)	(0.078)	(0.178)	(0.832)
	[0.076]	[0.346]	[0.771]	[0.010]	[0.059]	[0.933]
Control Mean	0.357	0.431	43.064	0.477	0.000	19.990
% Increase (Pooled)	25.2%	32.5%	-0.9%	36.9%	N/A	3.6%
Adjusted R^2	0.09	0.05	0.35	0.10	0.02	0.39
N	1579	359	359	361	360	642

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 5000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all treated groups. Panel B shows the coefficients corresponding to treatment indicators for each of the four treatment arms. All regressions include covariates selected by Double Lasso (Belloni, Chernozhukov, and Hansen, 2014), as well as stratum fixed effects and the outcome measured at baseline (if available).

Column 1 is an indicator that is one if the merchant finds a good that matches at least 3 horizontal criteria, and is missing if the merchant never replied to the mystery shopper or was otherwise unreachable. Column 2 is an indicator for whether the good's quality score is above the median product-group quality score. Column 3 is the raw quality score. Column 4 is an indicator for whether the good is made in Turkey, primarily inferred based on whether the label says . See the text for full details of how this outcome is constructed. Column 5 is the Anderson (2008) index combining the vertical outcomes. Column 6 is the price in USD, which is only measured conditional on the firm finding a good matching at least three horizontal criteria.

Table 2: Supplier Relationships (Followup Survey)

	Regular Suppliers in Turkey			Previous Suppliers		
	Any Sup in Turkey	Num Sup in Turkey	Index	Num Sup Total	Num Sup in Senegal	Ended with Sup
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled						
Treatment	0.037*	0.083**	0.135**	-0.102	-0.147	0.060***
	(0.021)	(0.035)	(0.062)	(0.164)	(0.167)	(0.021)
Panel B: Individual Treatments						
Search Only	0.024	0.081	0.116	-0.101	-0.114	0.049*
	(0.027)	(0.051)	(0.081)	(0.216)	(0.216)	(0.028)
	[0.579]	[0.258]	[0.251]	[0.880]	[0.901]	[0.116]
Search + Adverse Selection	0.049*	0.063	0.164*	-0.171	-0.272	0.066**
	(0.029)	(0.047)	(0.085)	(0.218)	(0.220)	(0.028)
	[0.193]	[0.303]	[0.126]	[0.840]	[0.532]	[0.053]
Search + Moral Hazard	0.003	0.003	0.019	-0.118	-0.076	0.073***
	(0.026)	(0.041)	(0.078)	(0.201)	(0.207)	(0.029)
	[0.909]	[0.938]	[0.803]	[0.880]	[0.901]	[0.046]
Search + AS + MH	0.075***	0.188***	0.245***	-0.014	-0.124	0.053*
	(0.028)	(0.057)	(0.086)	(0.222)	(0.224)	(0.028)
	[0.031]	[0.004]	[0.016]	[0.948]	[0.901]	[0.116]
Control Mean	0.167	0.222	0.000	3.700	3.213	0.135
% Increase (Pooled)	22.2%	37.4%	N/A	-2.8%	-4.6%	44.4%
Adjusted R^2	0.14	0.12	0.01	0.32	0.29	0.05
N	1680	1680	1680	1681	1681	1671

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 5000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all treated groups. Panel B shows the coefficients corresponding to treatment indicators for each of the four treatment arms. All regressions include covariates selected by Double Lasso (Belloni, Chernozhukov, and Hansen, 2014), as well as stratum fixed effects and the outcome measured at baseline (if available).

Column 1 is 1 if the merchant says that they have a regular supplier in Turkey. Column 2 is the number of regular suppliers in Turkey. Column 3 is the Anderson (2008). Column 4 is the total number of regular suppliers. Column 5 is the number of regular suppliers in Senegal. Column 6 is 1 if the merchant has ended a relationship with a regular supplier in the past 3 months. A regular supplier is defined as a supplier from whom the merchant has made two or more orders with an intention of continuing the relationship.

Table 3: Order Value (Mobile Money Data)

	Any Order	Value Post Mystery Shopping		Total Value	
	Any Order	Order Value (OLS)	Order Value (Poisson)	Order Value (OLS)	Order Value (Poisson)
	(1)	(2)	(3)	(4)	(5)
Panel A: Pooled					
Any Trust Treatment	-0.004	4.325**	1.351**	3.490	0.396
	(0.020)	(1.885)	(0.494)	(2.638)	(0.288)
Panel B: Individual Treatments					
Search + Adverse Selection	0.001	2.121	0.830	2.223	0.255
	(0.024)	(1.602)	(0.544)	(2.779)	(0.306)
	[0.982]	[0.318]	[0.368]	[0.713]	[0.701]
Search + Moral Hazard	0.004	6.194***	1.638**	4.557	0.476
	(0.025)	(2.887)	(0.505)	(3.733)	(0.327)
	[0.982]	[0.023]	[0.035]	[0.501]	[0.449]
Search + AS + MH	-0.019	4.644	1.425	3.681	0.445
	(0.024)	(4.413)	(0.859)	(5.314)	(0.510)
	[0.778]	[0.353]	[0.368]	[0.713]	[0.701]
Control Mean	0.134	1.511	1.549	7.964	7.964
% Increase (Pooled)	-3.0%	286.2%	286.1%	43.8%	48.6%
Adjusted R^2	0.05	0.01	0.04	0.02	0.02
N	1500	1500	1459	1500	1499

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 5000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all trust treated groups, where Search Only is the omitted category. Panel B shows the coefficients corresponding to treatment indicators for each of the three treatment groups with trust treatments. All regressions include covariates selected by Double Lasso (Belloni, Chernozhukov, and Hansen, 2014), as well as stratum fixed effects and the outcome measured at baseline (if available).

Column 1 is an indicator for whether the merchant ever ordered from a study supplier. Column 2 is the total value of orders. Column 3 is the total value of orders, analysed with Poisson regression. Column 4 is the total value of orders. Column 5 is the total value of orders, analysed with Poisson regression. Mystery shopping took place during the first 13 weeks of the study. All values are in USD.

Table 4: Profit and Sales

	Raw			Winsorized (1%)		
	Profit (USD)	Sales (USD)	Index	Profit (USD)	Sales (USD)	Index
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled						
Treatment	82.4**	237.4**	0.205**	45.5**	115.0	0.109
	(31.6)	(107.5)	(0.081)	(21.2)	(79.0)	(0.070)
Panel B: Individual Treatments						
Search Only	31.7	333*	.226**	22.1	188	.129
	(28.8)	(177)	(.115)	(25.4)	(114)	(.096)
	[0.550]	[0.132]	[0.109]	[0.612]	[0.234]	[0.388]
Search + Adverse Selection	43.5	59.5	.056	30.8	52.7	.029
	(37.7)	(112)	(.096)	(30.7)	(98.9)	(.095)
	[0.550]	[0.843]	[0.591]	[0.612]	[0.820]	[0.892]
Search + Moral Hazard	15.8	7.83	.093	8.21	-31.2	.031
	(25.1)	(102)	(.079)	(23.2)	(85.2)	(.076)
	[0.568]	[0.947]	[0.410]	[0.728]	[0.820]	[0.892]
Search + AS + MH	254***	571**	.46**	128***	261**	.256**
	(89.3)	(261)	(.195)	(39.2)	(125)	(.112)
	[0.003]	[0.067]	[0.042]	[0.003]	[0.120]	[0.074]
Control Mean	188.3	609.5	0.000	188.3	609.5	0.000
% Increase (Pooled)	43.8%	39.0%	N/A	24.2%	18.9%	N/A
Adjusted R^2	0.13	0.17	-0.01	0.25	0.31	0.00
N	1351	1378	1431	1351	1378	1431

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 5000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all treated groups. Panel B shows the coefficients corresponding to treatment indicators for each of the four treatment arms. All regressions include covariates selected by Double Lasso (Belloni, Chernozhukov, and Hansen, 2014), as well as stratum fixed effects and the outcome measured at baseline (if available).

Column 1 is total profit from the past 30 days in USD. Column 2 is total sales from the past 30 days in USD. Column 3 is the Anderson (2008) index combining the previous two columns. Column 4 is total profit from the past 30 days in USD, winsorizing the top 1%. Column 5 is total sales from the past 30 days in USD, winsorizing the top 1%. Column 6 is the Anderson (2008) index combining the previous two columns. Profit is measured using the survey question from De Mel, McKenzie, and Woodruff (2009). Sales is measured using a similar survey question.

Table 5: Effect of Alibaba Training

	Heard of Alibaba (1)	Searched on Alibaba (2)	Compared Prices with Supplier (3)	Bought on Alibaba (4)
e-Commerce Treatment	0.065***	0.113**	0.087*	0.014
	(0.025)	(0.052)	(0.049)	(0.035)
Control Mean	0.908	0.423	0.319	0.135
% Increase	7.2%	26.8%	27.4%	10.7%
Adjusted R^2	0.07	0.09	0.14	0.22
N	340	340	340	340

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 5000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. Conventional robust standard errors are reported in parentheses. This table shows the effect of the Alibaba training treatment on Alibaba usage.

Table 6: Heterogeneity by Retailer vs Wholesaler

	Horiz Dummy (1)	Vert Index (2)	Sup Turk Index (3)	Prof Sales Index (1%) (4)
Search Only	0.194***	0.386**	0.059	0.038
	(0.047)	(0.192)	(0.094)	(0.087)
Search + AS	0.049	0.356*	0.183*	0.004
	(0.047)	(0.209)	(0.097)	(0.082)
Search + MH	0.131***	0.322	-0.017	0.050
	(0.047)	(0.199)	(0.091)	(0.065)
Search + AS + MH	0.036	0.423*	0.165	0.036
	(0.049)	(0.223)	(0.103)	(0.071)
S Only * Wholesaler	-0.158*	0.181	0.134	0.304
	(0.083)	(0.421)	(0.188)	(0.270)
S + AS * Wholesaler	0.048	-0.066	-0.027	0.120
	(0.083)	(0.424)	(0.184)	(0.256)
S + MH * Wholesaler	-0.024	-0.008	0.106	-0.071
	(0.084)	(0.418)	(0.179)	(0.219)
S + AS + MH * Wholesaler	0.114	-0.002	0.155	0.646**
	(0.083)	(0.417)	(0.189)	(0.300)
Control Mean Retail	0.374	0.010	0.003	0.079
Control Mean Wholesale	0.315	-0.030	-0.007	-0.195
Adjusted R^2	0.01	-0.00	0.00	0.01
N	1579	361	1680	1431

Note: This table shows the main results with treatment interacted with an indicator for whether a firm sells wholesale.

Table 7: Counterfactuals

Scenario	Discounted Profits (USD)
Baseline	7,952
$s = 6.5$	8,014
$s = 6.5, \mu_0 = 0$	9,661
$s = 6.5, \xi = 0$	8,016
$s = 6.5, \mu_0 = \xi = 0$	9,813

Figures

Figure 1: Design Tree

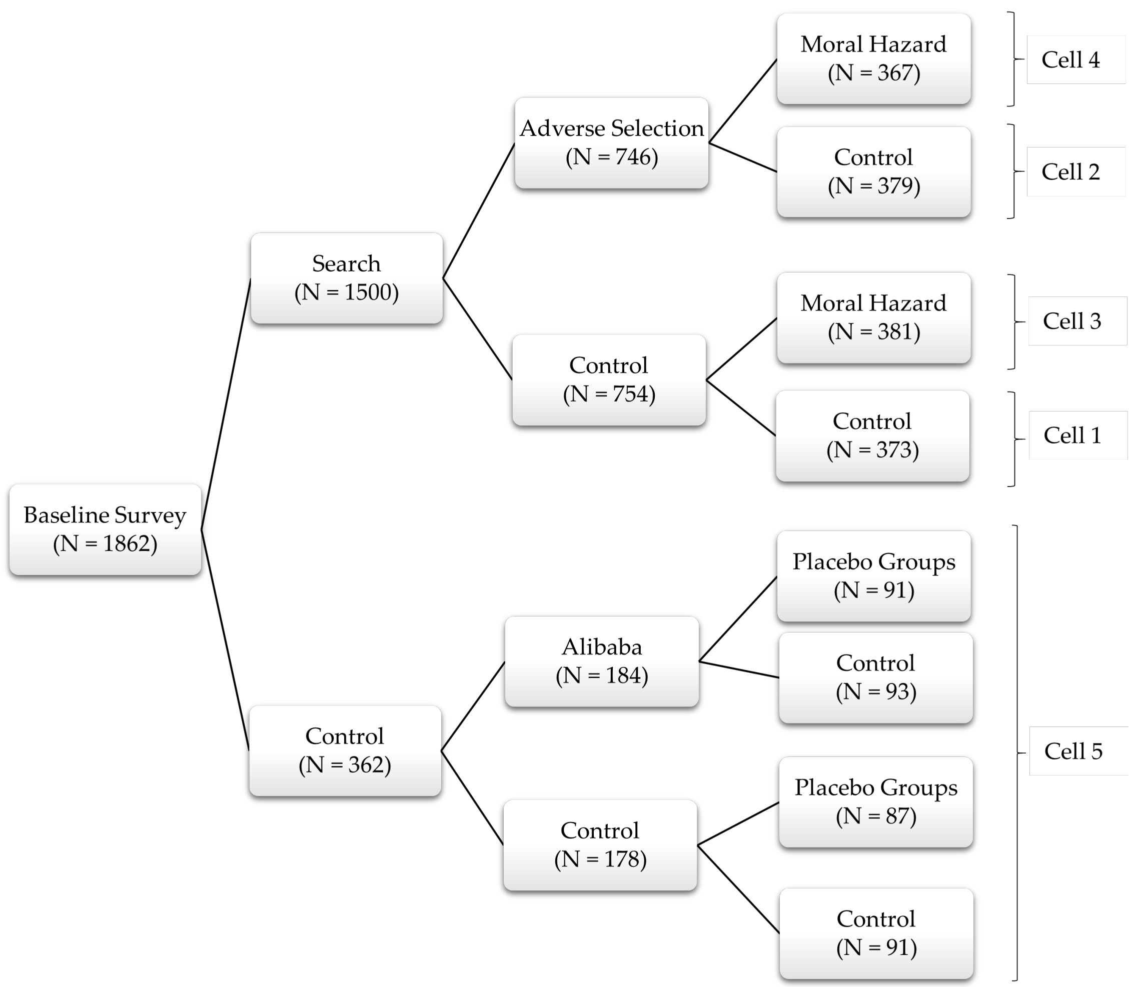


Figure 2: Supplier WhatsApp Groups

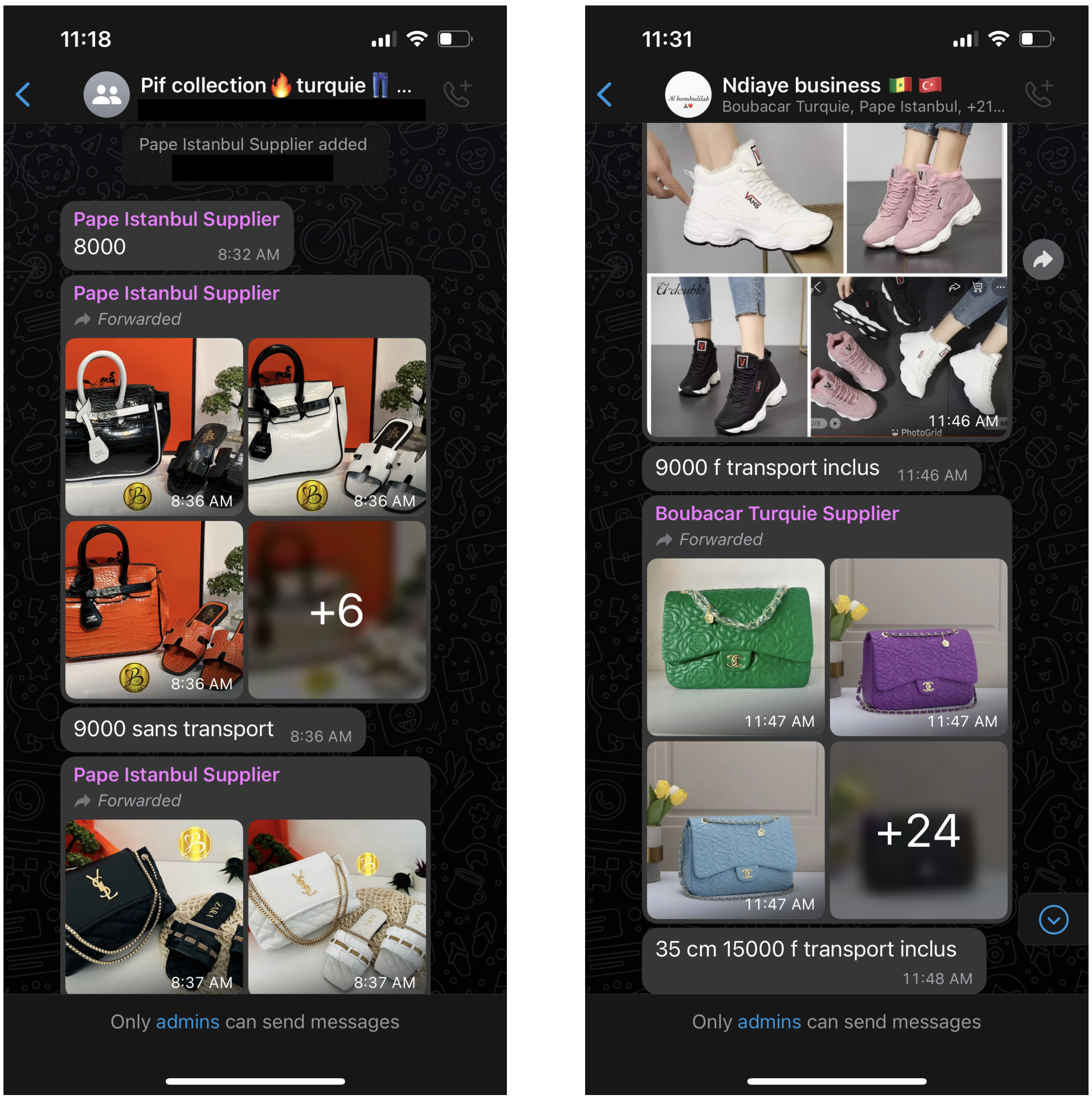
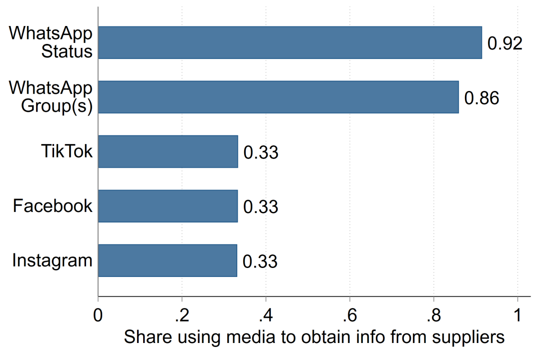
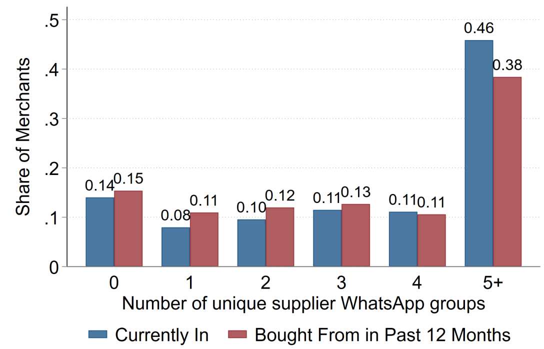


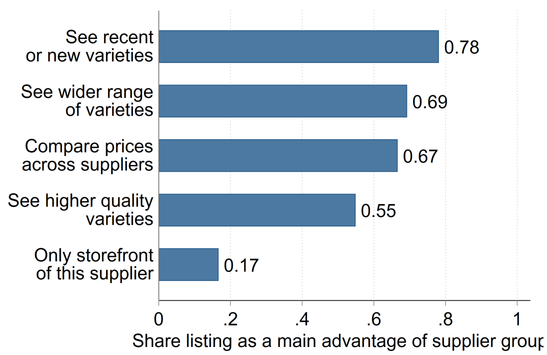
Figure 3: Firm Social Media Usage to obtain Information about Suppliers



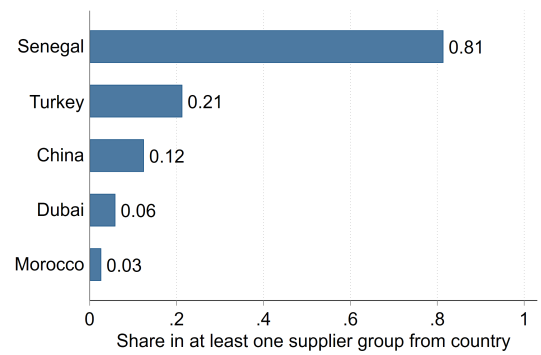
(a) Social Media Usage



(b) Supplier WhatsApp Groups



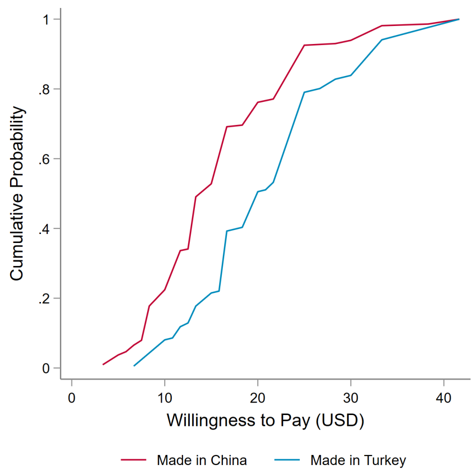
(c) Advantages of Supplier Groups



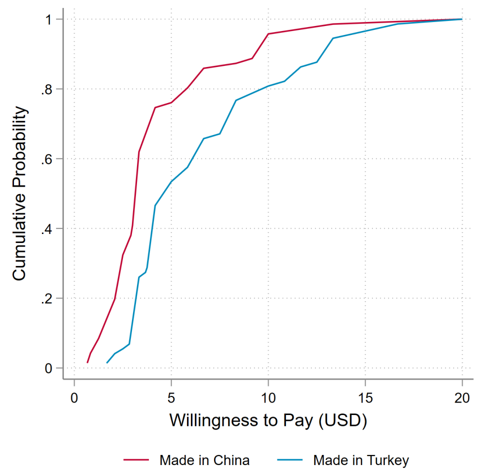
(d) Supplier Groups, by Supplier Country

Note: This figure shows a number of statistics about how firms in our sample use social media to obtain information about suppliers. All data is from our baseline survey with 1,862 firms. Panel (a) shows the results of a question asking firms to select all social media that they use to obtain information about suppliers for their business. Panel (b) shows the distribution of the number of supplier WhatsApp groups a firm is in at the time of the baseline survey, as well as the distribution of the number of such groups that the firm has directly made at least one purchase from in the past 12 months. Supplier WhatsApp groups are defined as WhatsApp groups in which the primary purpose is for suppliers to advertise their wares to downstream clients. Panel (c) shows the results of a question asking firms that use supplier WhatsApp groups to select all reasons why they find these groups useful. Panel (d) shows the share of firms who are in at least one supplier WhatsApp group where the supplier located and based in the country listed.

Figure 4: Willingness to Pay by Product Origin



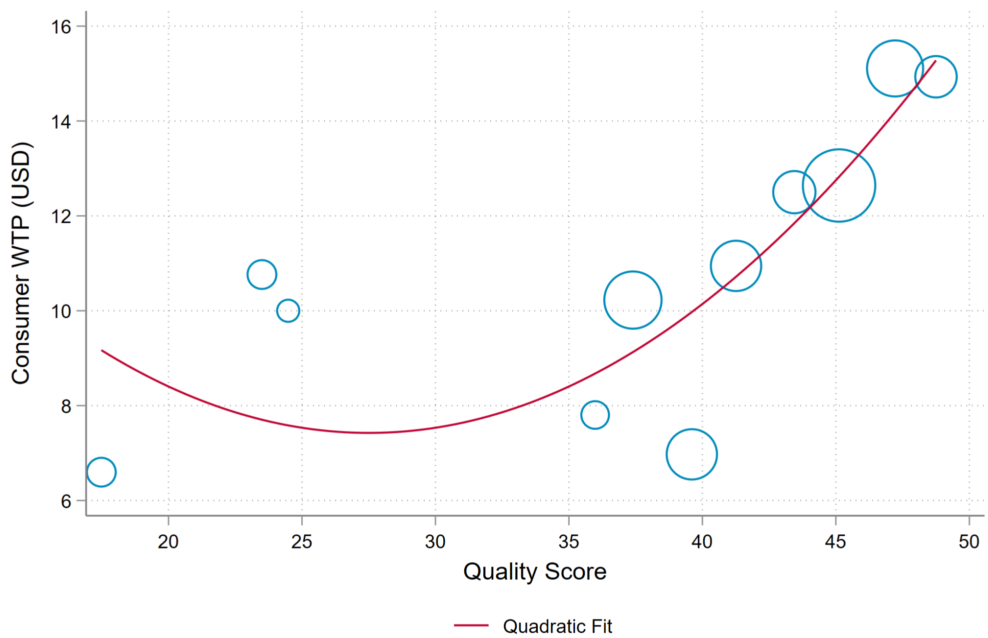
(a) Consumer Survey ($N = 400$)



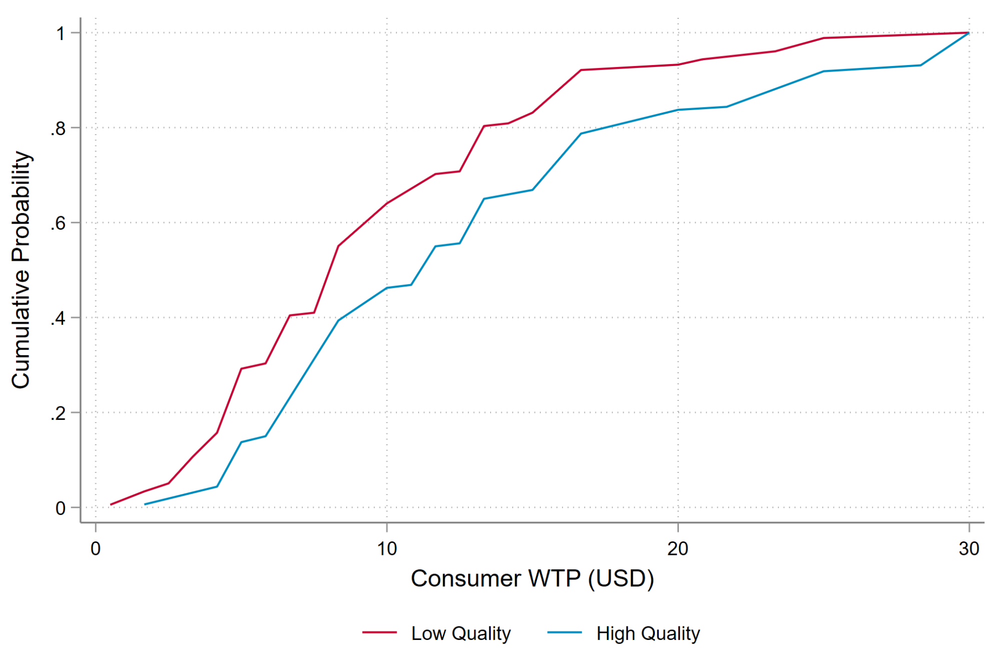
(b) Firm Survey ($N = 144$)

Note: This figure shows the results of two exercises in which we showed various images of garments to survey respondants, randomised whether we said the good was made in Turkey or made in China, and elicited willingness to pay for the garments. Panel (a) shows the CDF of WTP in the consumer survey, separately by whether we said the good was made in Turkey or China. The distribution is truncated at 40 USD for ease of readability. Panel (b) shows the CDF of WTP for a small, separate survey of firms (for a different set of goods), with distribution truncated at 20 USD for ease of readability.

Figure 5: Consumer Willingness to Pay for Quality



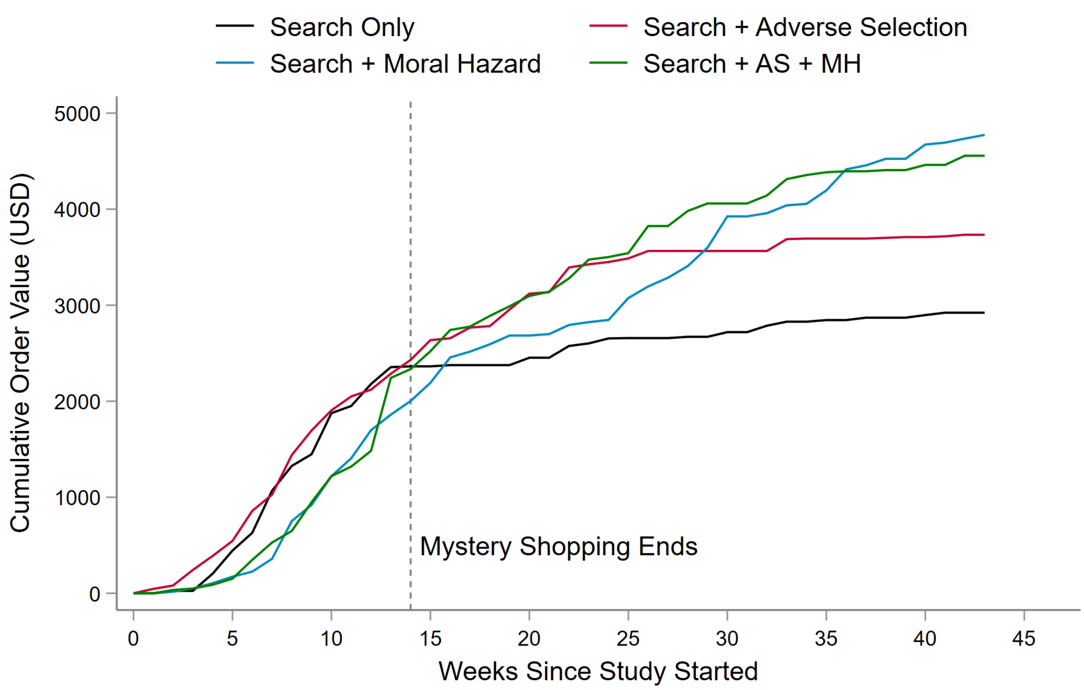
(a) Raw Score



(b) CDF of WTP by High or Low Quality

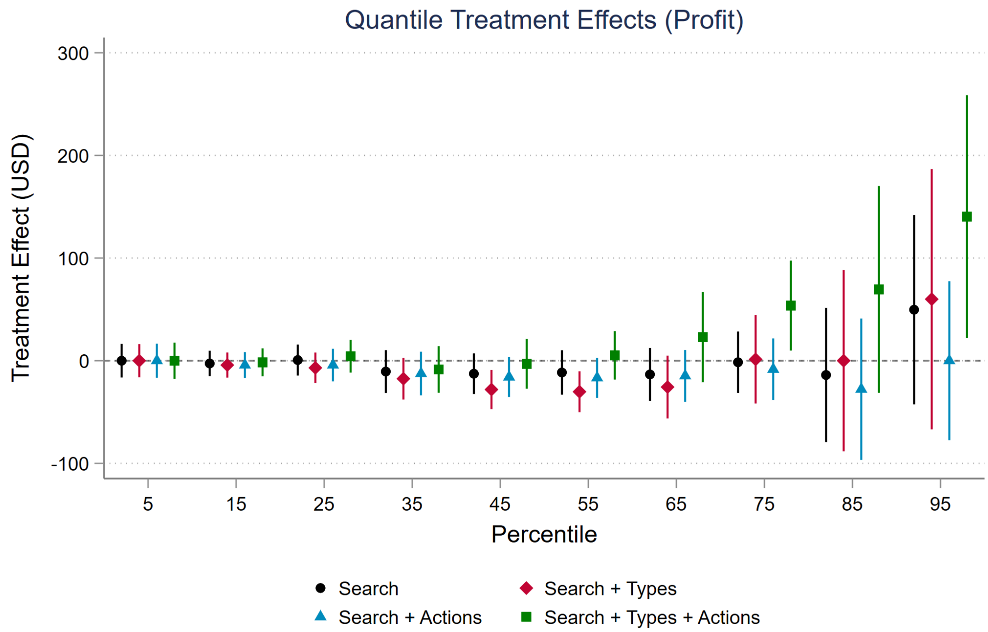
Note: Panel (a) shows a binscatter of consumer willingness to pay for garments (as measured by the consumer survey) against the quality score of the garments. The size of each bubble is proportional to the number of observations. Panel (b) shows the CDF of consumer willingness to pay separately based on whether the garment met our definition of high quality. We truncate willingness to pay at 30 USD to avoid unnecessarily stretching the x-axis. See the main text for full details on the consumer survey and variable construction.

Figure 6: Cumulative Order Value (Mobile Money Data)

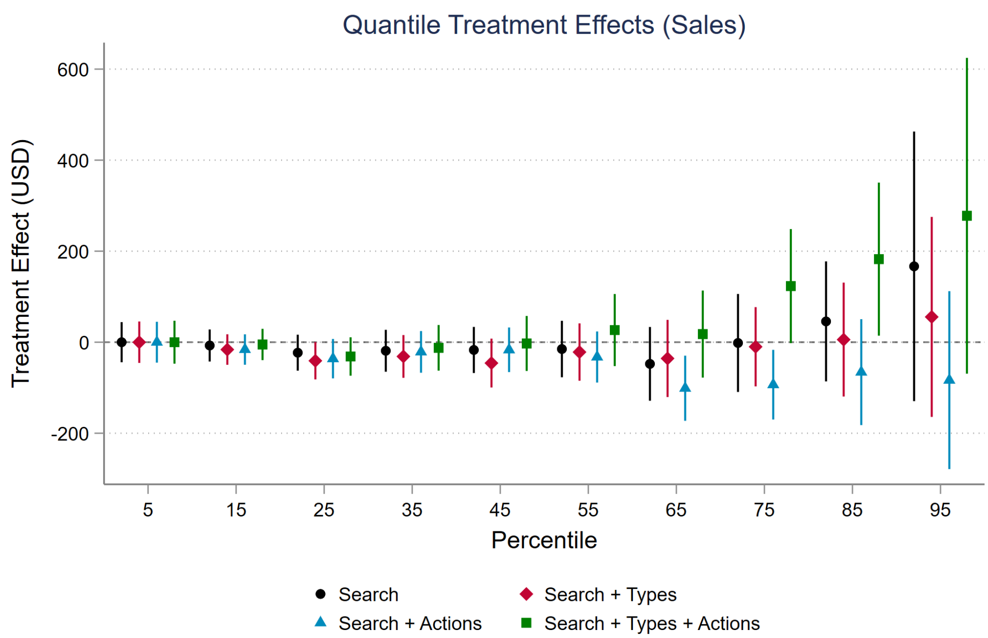


Note: This figure shows the total order value from study suppliers, according to the mobile money data, in each treatment group as a function of number of weeks since the study begun (16 November 2023). Pure control is omitted as they were not connected to any study suppliers.

Figure 7: Quantile Treatment Effects



(a) Profit



(b) Sales

Note: This figure shows the coefficients from quantile regressions of profit and sales on the four treatment groups. All quantile regressions include the outcome measured at baseline (if available), but otherwise do not include any covariates. We plot 95% confidence intervals constructed using randomisation inference, defined as the set of sharp nulls that do not reject at the 5% level.

Appendix A – Additional Tables and Figures

Tables

Table A1: Balance Table

Variable	Control (1)	Search (2)	Search AS (3)	Search MH (4)	Search Types MH (5)	Joint p-value (6)
Female	0.49	0.51	0.49	0.53	0.49	0.74
Online Only	0.67	0.66	0.68	0.7	0.65	0.56
Business Age	4.85	4.74	5.2	4.72	5.21	0.57
Share Cust Turkey	0.43	0.45	0.45	0.44	0.45	0.96
Any Reg Supp Turkey	0.22	0.21	0.24	0.21	0.19	0.67
Travelled Business (5y)	0.09	0.08	0.11	0.09	0.11	0.43
Profit USD (30 Days)	221.35	221.08	262.49	195.36	235.63	0.24
Bought Alibaba Ever	0.16	0.13	0.15	0.18	0.13	0.23
N	362	373	379	381	367	

Note: The table shows the mean for each variable in each of the five treatment cells. The final column shows the p-value from regressing the variable on indicators for each treatment (where the control group is omitted) and conducting a test that all coefficients are zero. Finally, we run a multinomial logit of treatment group against all of the variables in the table, for which a joint test that all coefficients are zero has a p-value of 0.724.

Table A2: Horizontal Outcomes (Detailed)

	Extensive vs Intensive Margin		Number of Criteria	
	Agree Search	Find Product Conditional	Num Criteria Unconditional	Num Criteria Conditional
	(1)	(2)	(3)	(4)
Panel A: Pooled				
Treatment	0.019	0.099***	0.347***	0.364**
	(0.025)	(0.035)	(0.125)	(0.141)
Panel B: Individual Treatments				
Search Only	0.019	0.141***	0.647***	0.706***
	(0.031)	(0.044)	(0.162)	(0.179)
	[0.858]	[0.004]	[0.001]	[0.000]
Search + Adverse Selection	0.003	0.064	0.164	0.191
	(0.032)	(0.045)	(0.160)	(0.185)
	[0.929]	[0.153]	[0.342]	[0.347]
Search + Moral Hazard	0.041	0.100**	0.374**	0.323*
	(0.030)	(0.044)	(0.158)	(0.176)
	[0.435]	[0.058]	[0.048]	[0.160]
Search + AS + MH	0.014	0.091**	0.202	0.229
	(0.032)	(0.045)	(0.159)	(0.181)
	[0.858]	[0.075]	[0.342]	[0.347]
Control Mean	0.781	0.457	1.650	2.111
% Increase (Pooled)	2.4%	21.7%	21.0%	17.2%
Adjusted R^2	0.06	0.03	0.04	0.02
N	1579	1269	1579	1269

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 5000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all treated groups. Panel B shows the coefficients corresponding to treatment indicators for each of the four treatment arms. All regressions include covariates selected by Double Lasso (Belloni, Chernozhukov, and Hansen, 2014), as well as stratum fixed effects and the outcome measured at baseline (if available).

Column 1 is 1 if the merchant agrees to sell or search for the product, and is missing if the merchant never replied to the mystery shopper or was otherwise unreachable. Column 2 is 1 if the merchant found a suitable product, conditional on agreeing to sell or search for the product. Column 3 is the number of horizontal criteria of the product, and is 0 if the merchant either did not agree to sell or search for a product or agreed but never sent any product. Column 4 is the number of horizontal criteria of the product, conditional on agreeing to sell or search for a product, and is 0 if the merchant agreed but never sent any product.

Table A3: Vertical Outcomes (Detailed)

	From Turkey		
	Made in Turkey (Label)	Made in Turkey (Tailor Judgement)	Made in Turkey (Label + Tailors)
	(1)	(2)	(3)
Panel A: Pooled			
Treatment	0.153*	0.142	0.168**
	(0.082)	(0.088)	(0.072)
Panel B: Individual Treatments			
Search Only	0.109	0.263**	0.178**
	(0.097)	(0.101)	(0.087)
	[0.270]	[0.037]	[0.120]
Search + Adverse Selection	0.162	0.058	0.136
	(0.099)	(0.106)	(0.088)
	[0.229]	[0.586]	[0.129]
Search + Moral Hazard	0.133	0.145	0.179**
	(0.094)	(0.104)	(0.087)
	[0.262]	[0.369]	[0.120]
Search + AS + MH	0.209**	0.102	0.183**
	(0.099)	(0.103)	(0.089)
	[0.116]	[0.510]	[0.120]
Control Mean	0.500	0.585	0.554
% Increase (Pooled)	30.6%	24.3%	30.3%
Adjusted R^2	0.17	0.05	0.09
N	252	284	328

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 5000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all treated groups. Panel B shows the coefficients corresponding to treatment indicators for each of the four treatment arms. All regressions include covariates selected by Double Lasso (Belloni, Chernozhukov, and Hansen, 2014), as well as stratum fixed effects and the outcome measured at baseline (if available).

Column 1 is 1 if the label says , 0 if the label says for X other than Turkey, and missing otherwise. Column 2 is 1 if both tailors independently determined that the product was made in Turkey, and is 0 if both tailors independently determined that the product was not made in Turkey. It is missing if the tailors disagreed. For shoes, as there was only one expert shoemaker, we take their opinion directly. Column 3 is an indicator that combines the label and tailor measures of whether the good was made in Turkey. It is equal to the label measure where available, and the tailor measure otherwise.

Table A4: Supplier Relationships (Further Results on Substitution)

	Reg Supp in China		Media for Suppliers			Forward Media	
	Any Supp in China	Num Supp in China	Uses Facebook	Uses TikTok	Uses Instagram	Fwd Photo for Search	Fwd Photo for Price
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Panel A: Pooled							
Treatment	-0.025	-0.014	-0.084***	-0.025	-0.027	-0.054**	-0.097***
	(0.016)	(0.035)	(0.027)	(0.025)	(0.026)	(0.027)	(0.030)
Panel B: Individual Treatments							
Search Only	-0.030	-0.023	-0.092***	-0.045	-0.039	-0.058	-0.107***
	(0.019)	(0.040)	(0.033)	(0.032)	(0.032)	(0.036)	(0.039)
	[0.326]	[0.923]	[0.018]	[0.409]	[0.542]	[0.207]	[0.016]
Search + Adverse Selection	-0.029	-0.028	-0.064*	-0.005	-0.040	-0.037	-0.091**
	(0.019)	(0.045)	(0.034)	(0.032)	(0.033)	(0.035)	(0.038)
	[0.326]	[0.923]	[0.061]	[0.931]	[0.542]	[0.287]	[0.030]
Search + Moral Hazard	-0.018	0.009	-0.080**	-0.039	-0.016	-0.063*	-0.081**
	(0.020)	(0.044)	(0.033)	(0.032)	(0.033)	(0.035)	(0.038)
	[0.423]	[0.939]	[0.034]	[0.465]	[0.841]	[0.207]	[0.033]
Search + AS + MH	-0.023	-0.013	-0.104***	-0.010	-0.012	-0.060*	-0.110***
	(0.020)	(0.045)	(0.033)	(0.033)	(0.033)	(0.035)	(0.039)
	[0.423]	[0.939]	[0.007]	[0.931]	[0.841]	[0.207]	[0.016]
Control Mean	0.099	0.167	0.328	0.290	0.279	0.716	0.659
% Increase (Pooled)	-25.3%	-8.4%	-25.6%	-8.6%	-9.7%	-7.5%	-14.7%
Adjusted R^2	0.12	0.36	0.09	0.12	0.12	0.04	0.03
N	1680	1680	1671	1671	1671	1671	1565

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 5000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all treated groups. Panel B shows the coefficients corresponding to treatment indicators for each of the four treatment arms. All regressions include covariates selected by Double Lasso (Belloni, Chernozhukov, and Hansen, 2014), as well as stratum fixed effects and the outcome measured at baseline (if available).

Column 1 is 1 if the merchant says that they have a regular supplier in China. Column 2 is 1 if the merchant says that they have a regular supplier in China. Column 3 is 1 if the merchant says that they use Facebook to learn about suppliers. Column 4 is 1 if the merchant says that they use TikTok to learn about suppliers. Column 5 is 1 if the merchant says that they use Instagram to learn about suppliers. Column 6 is 1 if the merchant says that they have forwarded a photo or video from a supplier group to a regular supplier to try to obtain a similar product in the past 3 months. Column 7 is 1 if the merchant says that they have forwarded a photo or video from a supplier group to a regular supplier to try to obtain a better price in the past 3 months. A regular supplier is defined as a supplier from whom the merchant has made two or more orders with an intention of continuing the relationship.

Table A5: Profit and Sales (Poisson Regression)

	Profit		Sales	
	Profit 30 Days (USD)	Profit 30 Days Winsorized 1% (USD)	Sales 30 Days (USD)	Sales 30 Days Winsorized 1% (USD)
	(1)	(2)	(3)	(4)
Panel A: Pooled				
Treatment	0.307*	0.188	0.184	0.085
	(0.132)	(0.112)	(0.138)	(0.123)
Panel B: Individual Treatments				
Search Only	0.088	0.074	0.315	0.198
	(0.147)	(0.133)	(0.210)	(0.158)
	[0.913]	[0.831]	[0.421]	[0.503]
Search + Adverse Selection	0.097	0.119	-0.041	-0.026
	(0.189)	(0.162)	(0.167)	(0.159)
	[0.913]	[0.831]	[0.830]	[0.887]
Search + Moral Hazard	-0.047	-0.009	-0.151	-0.147
	(0.138)	(0.143)	(0.174)	(0.143)
	[0.913]	[0.957]	[0.667]	[0.533]
Search + AS + MH	0.865***	0.480***	0.475**	0.260
	(0.214)	(0.142)	(0.212)	(0.162)
	[0.002]	[0.012]	[0.161]	[0.379]
Control Mean	188.3	188.3	609.5	609.5
% Increase (Pooled)	35.9%	20.7%	20.2%	8.9%
Adjusted R^2	0.34	0.38	0.46	0.48
N	1351	1351	1378	1378

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 5000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all treated groups. Panel B shows the coefficients corresponding to treatment indicators for each of the four treatment arms. All regressions include covariates selected by Double Lasso (Belloni, Chernozhukov, and Hansen, 2014), as well as stratum fixed effects and the outcome measured at baseline (if available).

Column 1 is total profit from the past 30 days in USD. Column 2 is total profit from the past 30 days in USD, winsorizing the top 1%. Column 3 is total sales from the past 30 days in USD. Column 4 is total sales from the past 30 days in USD, winsorizing the top 1%. Profit is measured using the survey question from De Mel, McKenzie, and Woodruff (2009). Sales is measured using a similar survey question.

Table A6: Travel

	Wholesalers			Retailers		
	Any Travel	Travel China	Travel Turkey	Any Travel	Travel China	Travel Turkey
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled						
Treatment	-0.033	-0.035	-0.020	-0.004	-0.021*	0.003
	(0.036)	(0.030)	(0.019)	(0.014)	(0.011)	(0.005)
Panel B: Individual Treatments						
Search Only	-0.009	-0.016	-0.025	-0.004	-0.022*	-0.004
	(0.049)	(0.039)	(0.026)	(0.018)	(0.013)	(0.004)
	[0.973]	[0.883]	[0.668]	[0.988]	[0.237]	[0.710]
Search + Adverse Selection	0.003	-0.014	-0.014	-0.001	-0.026**	0.004
	(0.042)	(0.035)	(0.022)	(0.018)	(0.012)	(0.008)
	[0.973]	[0.883]	[0.668]	[0.997]	[0.095]	[0.828]
Search + Moral Hazard	-0.069*	-0.061**	-0.024	-0.008	-0.015	0.010
	(0.039)	(0.031)	(0.020)	(0.017)	(0.014)	(0.008)
	[0.236]	[0.149]	[0.608]	[0.980]	[0.287]	[0.606]
Search + AS + MH	-0.054	-0.047	-0.019	-0.001	-0.022	0.001
	(0.038)	(0.032)	(0.020)	(0.020)	(0.014)	(0.006)
	[0.349]	[0.331]	[0.668]	[0.997]	[0.237]	[0.870]
Control Mean	0.130	0.090	0.040	0.041	0.033	0.004
% Increase (Pooled)	-25.4%	-38.9%	-50.0%	-9.8%	-63.6%	75.0%
Adjusted R^2	0.16	0.27	0.07	0.04	0.07	0.02
N	546	546	546	1125	1125	1125

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 5000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all treated groups. Panel B shows the coefficients corresponding to treatment indicators for each of the four treatment arms. All regressions include covariates selected by Double Lasso (Belloni, Chernozhukov, and Hansen, 2014), as well as stratum fixed effects and the outcome measured at baseline (if available).

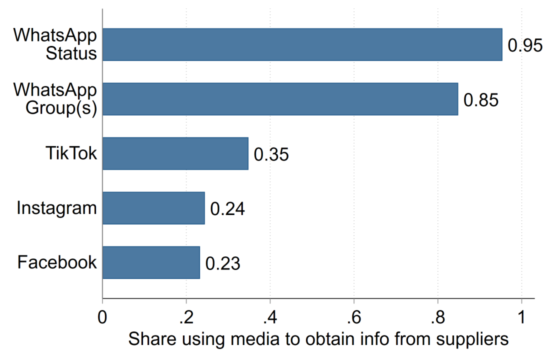
Column 1 is 1 if the firm travelled for business in the past 3 months. Column 2 is 1 if the firm travelled for business to China in the past 3 months. Column 3 is 1 if the firm travelled for business to Turkey in the past 3 months. Column 4 is 1 if the firm travelled for business in the past 3 months. Column 5 is 1 if the firm travelled for business to China in the past 3 months. Column 6 is 1 if the firm travelled for business to Turkey in the past 3 months. Travel is 1 if either the firm owner or someone closely involved with the firm travelled internationally for firm-specific business in the past 3 months.

Table A7: Calibrated Parameters

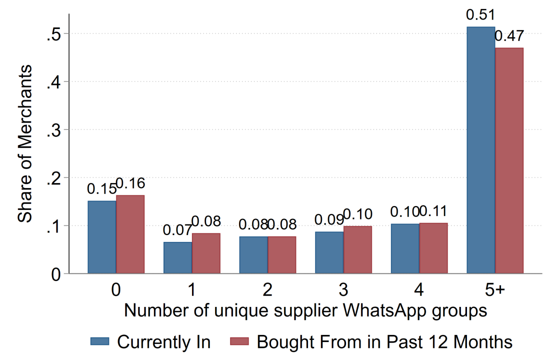
Parameter	Value	Origin
σ	3.02	Average markup from baseline survey.
δ	0.96	Empirical survival rate from baseline to endline survey.
p_l	12.29	Average input price from baseline survey among firms without a foreign supplier.
λ	Varies	Calibrating using baseline survey question.

Figures

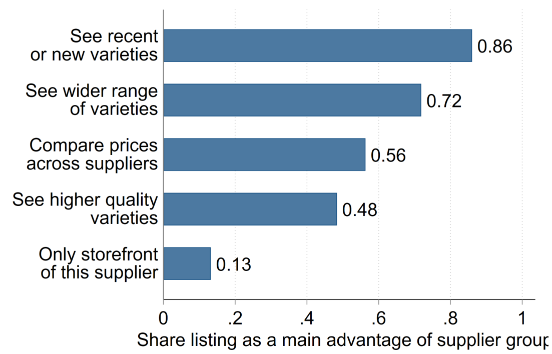
Figure A1: Social Media Usage (Physical Store Only)



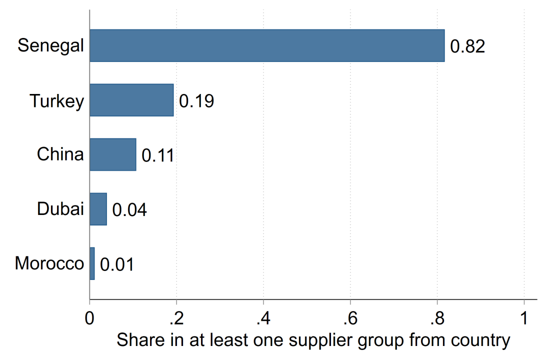
(a) Social Media Usage



(b) Supplier WhatsApp Groups



(c) Advantages of Supplier Groups



(d) Supplier Groups, by Supplier Origin

Note: This figure shows a number of statistics about how firms in our sample use social media to obtain information about suppliers. It is the same as Figure 3, but instead calculates statistics only for the 607 firms that have physical stores. Panel (a) shows the results of a question asking firms to select all social media that they use to obtain information about suppliers for their business. Panel (b) shows the distribution of the number of supplier WhatsApp groups a firm is in at the time of the baseline survey, as well as the distribution of the number of such groups that the firm has directly made at least one purchase from in the past 12 months. Supplier WhatsApp groups are defined as WhatsApp groups in which the primary purpose is for suppliers to advertise their wares to downstream clients. Panel (c) shows the results of a question asking firms that use supplier WhatsApp groups to select all reasons why they find these groups useful. Panel (d) shows the share of firms who are in at least one supplier WhatsApp group where the supplier located and based in the country listed.

Figure A2: Business Cards for Hidden Actions Treatment

(a) Treatment

(b) Control

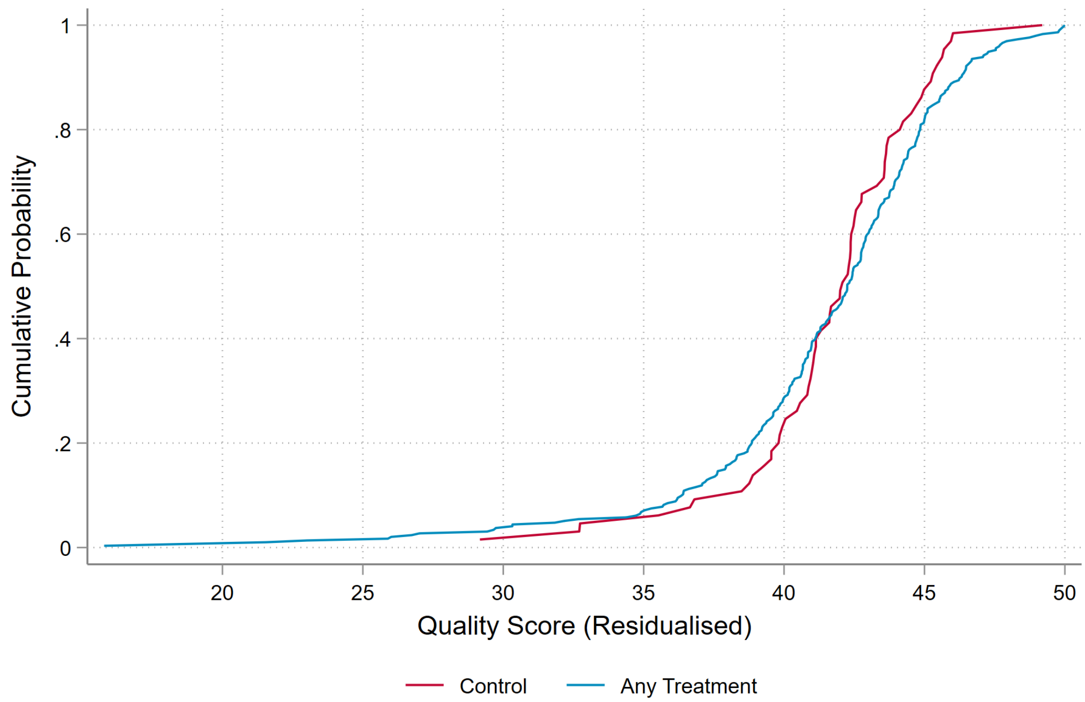
Figure A3: Mystery Shopping Goods (Examples)

1. *Red*
2. *One colour*
3. *No large logo/print*
4. *3 black buttons*
5. *Size L*

1. *Black*
2. *Long sleeves*
3. *Knee length*
4. *Turtleneck*
5. *Ribbed Cotton*

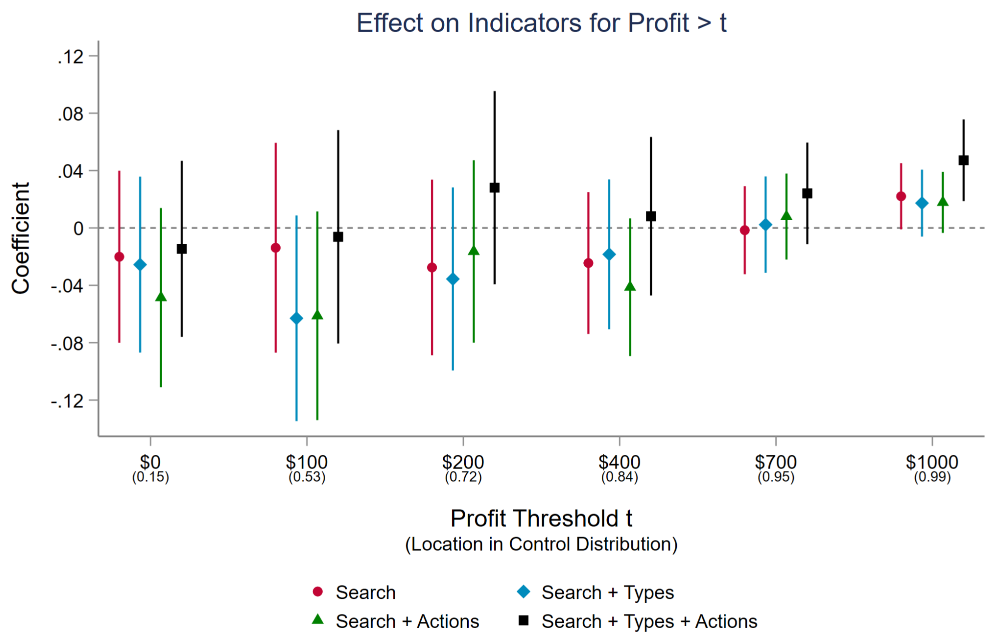
Examples of goods requested in the mystery shopping exercise. In total, there were 28 different goods.

Figure A4: Quality Score Distribution

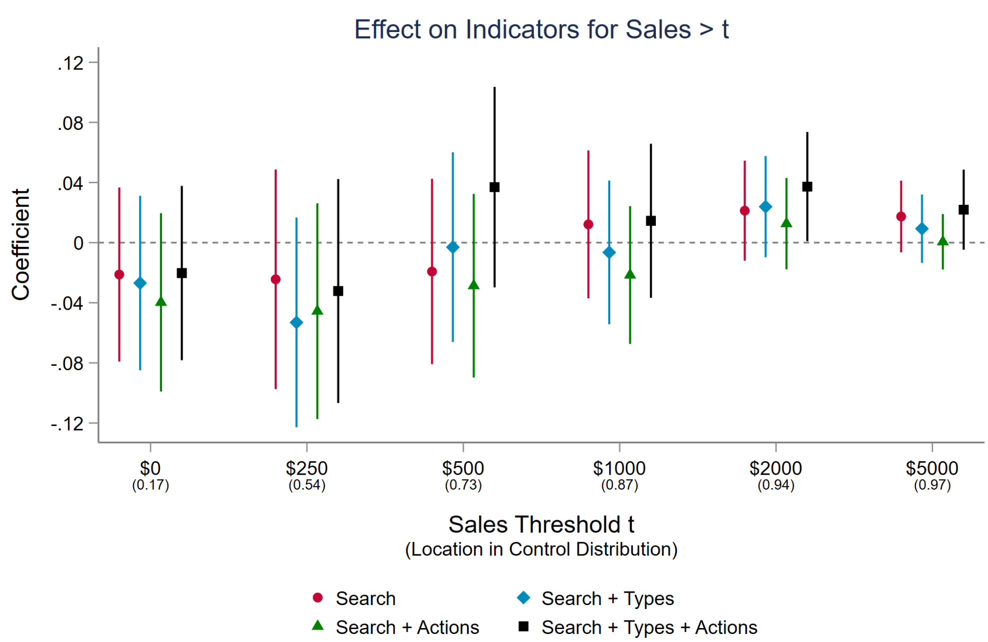


Note: This figure shows CDF of the quality score separately by treatment status, with all four treatment groups (Search Only, Search + Types, Search + Actions, Search + Types + Actions) pooled for ease of readability. To be consistent with the regression in the table, we first residualise quality using stratum fixed effects and the covariates selected in the regression.

Figure A5: Threshold Regressions for Profit and Sales



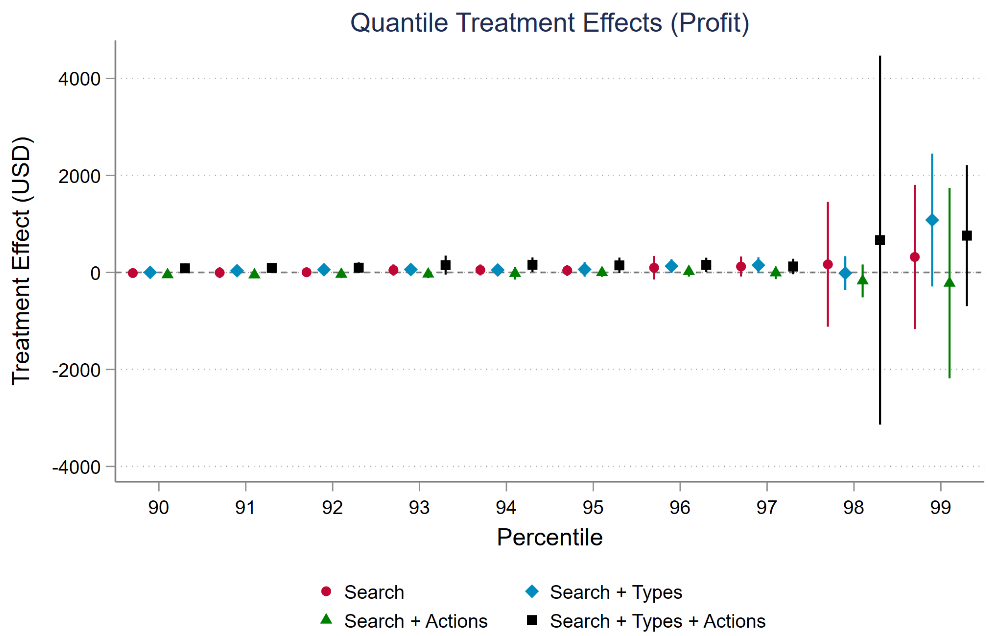
(a) Profit



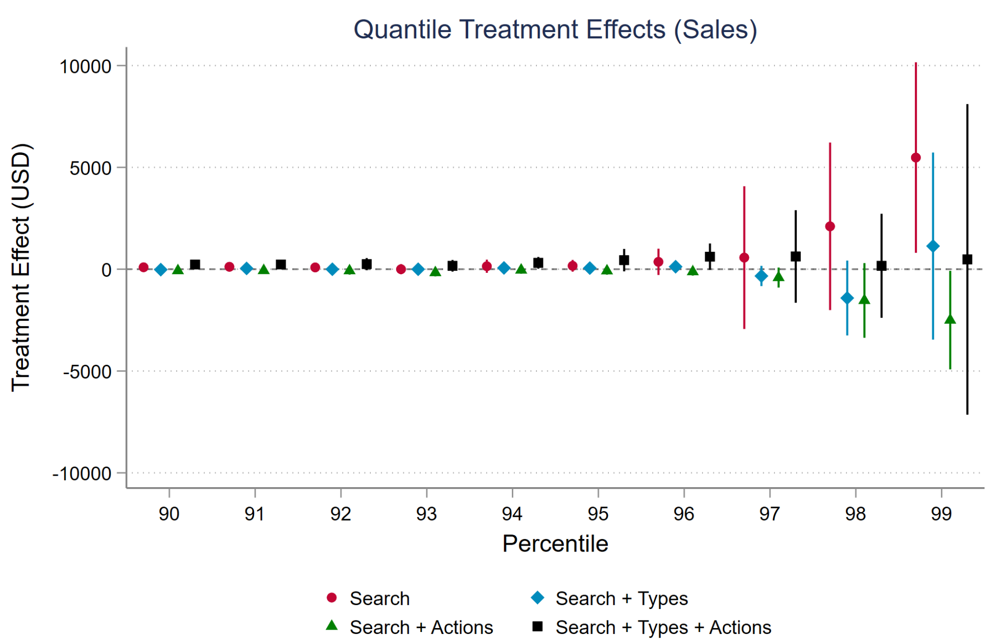
(b) Sales

Note: This figure shows the coefficients from regressions of indicators for whether profit and sales are above some threshold t, for a range of t. All regressions include covariates selected by Double Lasso (Belloni, Chernozhukov, and Hansen, 2014), as well as the outcome measured at baseline (if available). The numbers in parentheses show the percentiles at which t is located in the distribution of the pure control group. We plot 95% confidence intervals constructed using randomisation inference, defined as the set of sharp nulls that do not reject at the 5% level, using the procedure in Young (2024).

Figure A6: Quantile Treatment Effects (90-99)



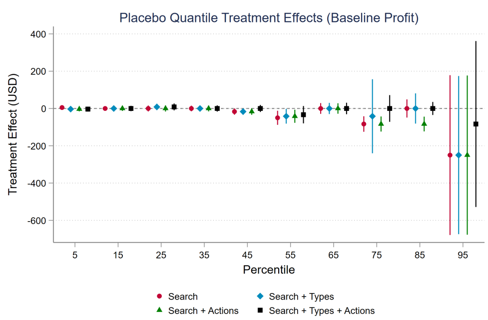
(a) Profit



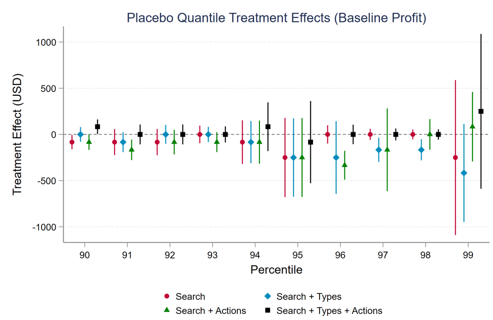
(b) Sales

Note: This figure shows the coefficients from quantile regressions of profit and sales on the four treatment groups, for quantiles 90-99. All quantile regressions include the outcome measured at baseline (if available), but otherwise do not include any covariates. We plot 95% confidence intervals constructed using randomisation inference, defined as the set of sharp nulls that do not reject at the 5% level.

Figure A7: Placebo Check: Quantile Treatment Effects on Baseline Profit



(a) Percentiles 5-95



(b) Percentiles 90-99

Note: This figure shows the coefficients from quantile regressions of baseline profit on the four treatment groups, intended as a placebo test. As the only covariate included in the main quantile regressions is the outcome measured at baseline, which is itself the outcome here, we do not include any covariates. We plot 95% confidence intervals constructed using randomisation inference, defined as the set of sharp nulls that do not reject at the 5% level.

Appendix B – Mathematical Appendix

In this Appendix, we prove further detail on the model. In particular, we formally state and prove various properties of the optimal contract.

Assumption 1. *(No trade with bad types)*

$$\max_q \lambda r(q) - c_0 q < \max_q r(q) - p_l q$$

This assumption states that if a firm knew that the supplier was a bad type, they would prefer to order from the local supplier.

Proposition 1. *It is not optimal to offer a menu of contracts that fully separates good and bad suppliers.*

Proof. Suppose the contrary, and consider the state of the world where the supplier is a bad type. Since the menu fully separates the types, the firm's posterior is then that the supplier is a bad type with probability 1. Because the supplier has limited liability, the maximum that the firm can earn under any such contract is the full surplus, i.e., $(\max_q \lambda r(q) - c_0 q)/(1-\delta)$. Assumption 1 implies that the firm can always do better than this, because at the very least they can buy from a local supplier in every period. Since the contract is relational, the firm would thus renege before sending the first transfer. Thus, the expected payoff to the bad type from accepting the revealing contract is 0. So long as the contract recommended to the good type involves positive quantity, the bad type can always earn a positive expected payoff by accepting the good type's contract, because limited liability ensures that $\tau_t \geq cq_t > c_0 q_t$ for all t. Therefore, the bad type would not accept the contract that reveals their type, which is a contradiction. □

The original program is as follows:

$$\begin{aligned}
L = & \min_{\{\rho_t\},\{\eta_t\},\{\gamma_t\}} \max_{\{q_t\},\{\tau_t\}} \sum_{t=0}^{\infty} \delta^t \left(1-\mu_0(1-\lambda^t)\right) \left((1-\mu_t(1-\lambda)r(q_t) - \tau_t + \delta\mu_t(1-\lambda)\bar{U}\right) \\
& + \sum_{t=0}^{\infty} \delta^t \rho_t \left[\sum_{\tau=t+1}^{\infty} \delta^{\tau-t}(R_\tau - cq_\tau) - \xi(c-c_0)q_t \right] \\
& + \sum_{t=0}^{\infty} \delta^t (1-\mu_0(1-\lambda^t))\eta_t \left[\sum_{\tau=t}^{\infty} \delta^{\tau-t} \left(1-\mu_\tau(1-\lambda^{\tau-t})\right) \left((1-\mu_\tau(1-\lambda)r(q_t) - \tau_t + \delta\mu_t(1-\lambda)\bar{U}\right) - \bar{U} \right] \\
& + \sum_{t=0}^{\infty} \delta^t \gamma_t \left[\tau_t - cq_t\right]
\end{aligned}$$

The modified program is as follows.

$$\begin{aligned}
W_t(U_t, V_t, \mu_t) = {} & y(q_t, \tau_t, \mu_t) + \rho_t \left(\delta(1-\lambda)V_{t+1} - \xi(c-c_0)q_{t+1}\right) \\
& + \eta_t \left(U_{t+1} - \bar{U}\right) \\
& + \gamma_t \left(\tau_t - cq_t\right) \\
& + \nu_t^b \left(y(q_t, \tau_t, \mu_t) + \delta(1-\mu_t(1-\lambda))U_t + \delta\mu_t(1-\lambda)\bar{U} - U_t\right) \\
& + \nu_t^s \left(\tau_t - cq_t + \delta V_{t+1} - V_t\right) \\
& + \delta(1-\mu_t(1-\lambda))W_{t+1}(U_{t+1}, V_{t+1}, \mu_{t+1})
\end{aligned}$$

The FOCs are as follows

$$(1-\mu_t(1-\lambda))r'(q_t)(1+\nu_t^b) = \rho_t\xi(c-c_0) + \left(\gamma_t + \nu_t^s\right)c = 0 \tag{q_t}$$

$$1 + \nu_t^b = \gamma_t + \nu_t^s \tag{τ_t}$$

$$\rho_t(1-\lambda) + \nu_t^s = -(1-\mu_t(1-\lambda))\frac{\partial W_{t+1}}{\partial V_{t+1}} \tag{V_{t+1}}$$

$$\eta_{t+1} + \delta(1-\mu_t(1-\lambda))\nu_t^b = -\delta(1-\mu_t(1-\lambda))\frac{\partial W_{t+1}}{\partial U_{t+1}} \tag{U_{t+1}}$$

Substituting the FOC for τ_t into the FOC for q_t and re-arranging gives

$$r'(q_t) = \frac{1}{1-\mu_t(1-\lambda)}\left(1 + \xi\frac{\rho_t}{1+\nu_t^b}\frac{c-c_0}{c}\right)c.$$

The Envelope Condition implies that $\frac{\partial W_{t+1}}{\partial V_{t+1}} = -\nu_{t+1}^s$ and $\frac{\partial W_{t+1}}{\partial U_{t+1}} = -\nu_{t+1}^b$. Combining FOCs 2-4 then gives the following equation relating γ_t and γ_{t+1} that we will make extensive use of in the following proofs.

$$\rho_t(1-\lambda) + 1 - \gamma_t + \mu_t(1-\lambda)\nu_t^b = (1-\gamma_{t+1})(1-\mu_t(1-\lambda)) + \frac{\eta_{t+1}}{\delta}. \tag{3}$$

Remark 3. $\nu_{t+1}^b > \nu_t^b \iff \eta_{t+1} > 0$

Proof. This follows immediately from the FOC for U_{t+1} after substituting in the Envelope Condition. □

The optimal contract is generally not available in closed form, but we state and prove some properties in the following proposition.

Proposition 2. *There exists finite T^* such that:*

1. *The agent earns zero stage profits for all $t < T^*$ (i.e., LL binds for all $t < T^*$).*

2. *The principal earns $(1-\delta)\bar{U}$, that is, zero stage profits net of their outside option, for all $t > T^*$ (i.e., DEC binds for all $t > T^*$).*

3. *q_t is strictly increasing for $t < T^*$.*

4. *ICC binds for at least some $t < T^*$.*

Proof. We prove this through a series of Lemmas, which we state and prove below. For parts 1 and 2, see Lemma 3. For part 3, see Lemma 1. For part 4, see Lemma 5. □

Before formally stating the Lemmas, we first provide an intuitive sketch of the approach. First, we show that q_t must be strictly increasing whenever LL binds. Intuitively, a (weakly) decreasing q_t despite beliefs improving would imply that ICC strongly becomes "more binding" over time. But when LL is binding, the supplier is earning zero stage profits, so the ICC must be getting less binding over time.

Second, we show that if LL in t binds, then DEC in t cannot bind. Intuitively, both parties cannot be earning their outside option at the same time, as belief improving and q_t growing would imply that in other periods one of them must be making a loss.

Third, we show that the problem can be divided into two phases: LL will bind for all early periods and be slack for all late periods. "Backloading" results of this kind are standard in the dynamic moral hazard literature. Backloading happens for two reasons. The first reason is that incentives must be given to the agent at some point, and backloading incentives is efficient because it improves both early and late ICCs (whereas frontloading or even-loading still improves early ICCs but improves late ICCs less). The second reason is that adverse selection means that the "good type" agent is more patient than the principal, as the good type knows their own type. This means that it is always cheaper for the principal to backload payments. As a corollary, the DEC must bind for all late periods, as the principal wants to backload as much as possible, and will continue to do so until the DEC binds.

Finally, we show that ICC must bind for some t in the early section. Intuitively, if it didn't, then the principal would just extend the early period–where they earn all the surplus–for longer. The only reason to ever end this early phase is precisely because an ICC eventually binds (the principal has to pay the agent eventually, and in absence of the ICC would always prefer not to).

Lemma 1. *If LL binds in $t+1$, then $q_{t+1} > q_t$.*

Proof. Suppose that LL binds in $t+1$ and $q_{t+1} \leq q_t$. The FOC for q_t is

$$r'(q_t) = \frac{1}{1-\mu_t(1-\lambda)}\left(1+\xi\frac{\rho_t}{1+\nu_t^b}\frac{c-c_0}{c}\right)c.$$

We already know that $\nu_{t+1}^b \geq \nu_t^b$, and that $\mu_{t+1} \leq \mu_t$. Then, the hypothesis $q_{t+1} \leq q_t$ implies that $\rho_{t+1} > \rho_t$. Then, we can write

$$\delta(1-\lambda)V_{t+1} < (1-\lambda)V_{t+1} = (1-\lambda)\left(R_{t+1} - cq_{t+1} + \delta V_{t+2}\right) = \delta(1-\lambda)V_{t+2} = \xi(c-c_0)q_{t+1} \leq \xi(c-c_0)q_t.$$

The first equality is a definition, the second equality follows from the fact that LL binds in $t+1$, the third equality follows from the fact that $\rho_{t+1} > \rho_t$ implies that $\rho_{t+1} > 0$, which implies that ICC binds in $t+1$. The final inequality follows from the hypothesis that $q_{t+1} \leq q_t$. Thus, we have shown that

$$\delta(1-\lambda)V_{t+1} < \xi(c-c_0)q_t.$$

But this implies that ICC in t is violated, which is a contradiction. □

Lemma 2. *If LL in* $t+1$ *binds, then DEC in* $t+1$ *is slack.*

Proof. Suppose that DEC in $t+1$ binds. Then, we can write

$$\begin{aligned}
U_t &= (1-\mu_t(1-\lambda))r(q_t) - \tau_t + \delta(1-\mu_t(1-\lambda))U_{t+1} + \delta\mu_t(1-\lambda)\bar{U} \\
&= (1-\mu_t(1-\lambda))r(q_t) - \tau_t + \delta\bar{U} \\
&\leq (1-\mu_t(1-\lambda))r(q_t) - cq_t + \delta\bar{U} \\
&< (1-\mu_{t+1}(1-\lambda))r(q_t) - cq_t + \delta\bar{U} \\
&\leq (1-\mu_{t+1}(1-\lambda))r(q_{t+1}) - cq_{t+1} + \delta\bar{U} \\
&\leq (1-\mu_{t+1}(1-\lambda))r(q_{t+1}) - cq_{t+1} + \delta(1-\mu_{t+1}(1-\lambda))U_{t+2} + \delta\mu_{t+1}(1-\lambda)\bar{U} \\
&= (1-\mu_{t+1}(1-\lambda))r(q_{t+1}) - R_{t+1} + \delta(1-\mu_{t+1}(1-\lambda))U_{t+2} + \delta\mu_{t+1}(1-\lambda)\bar{U} \\
&= U_{t+1} \\
&= \bar{U}.
\end{aligned}$$

The first line is the definition of U_t. The second line follows from DEC binding in $t+1$. The third line follows from LL in t. The fourth line follows from $\mu_{t+1} < \mu_t$. The fifth line follows from Lemma 1. The sixth line follows from DEC in period $t+2$, i.e., $U_{t+2} \geq \bar{U}$. The seventh line follows from the hypothesis that LL binds in $t+1$, i.e., $\tau_t = cq_t$. The eighth line is the definition of U_{t+1}. The ninth line follows from DEC binding in $t+1$.

The above thus establishes that $U_t < \bar{U}$. But this is a violation of DEC in period t, which is a contradiction. □

Corollary 1. *If LL is slack in t, then it is also slack in $t+1$.*

Proof. LL slack in t means $\gamma_t = 0$. Suppose it binds in $t+1$, which means $\gamma_{t+1} > 0$. Lemma 2 then implies that $\eta_{t+1} = 0$. But, Equation (3) gives

$$\gamma_{t+1}(1 - \mu_t(1-\lambda)) = \frac{\eta_{t+1}}{\delta} - \rho_t(1-\lambda) - \mu_t(1-\lambda)(1+\nu_t^b).$$

It must then be that $\eta_{t+1} > 0$, which is a contradiction. □

Corollary 2. *If LL is slack in t, then DEC binds in $t+1$.*

Proof. Corollary 1 implies that $\gamma_t = \gamma_{t+1} = 0$. Then, the FOC implies

$$\frac{\eta_{t+1}}{\delta} = \rho_t(1-\lambda) + \mu_t(1-\lambda)(1+\nu_t^b) \geq \mu_t(1-\lambda)(1+\nu_t^b).$$

The final term is strictly positive, which implies that $\eta_{t+1} > 0$. □

Lemma 3. *If any trade occurs, then there exists finite $T^* \geq 1$ such that (i) LL binds for all $t < T^*$ and is slack for all $t \geq T^*$, and (ii) DEC is slack for all $t < T^*$ and binds for all $t > T^*$.*

Proof. For (i): Corollary 1 shows that $\gamma_t = 0 \implies \gamma_{t+1} = 0$. Thus, if there exists T^* such that $\gamma_t = 0$, then $\gamma_s = 0$ for all $s \geq T^*$. We already know that $\gamma_0 = 1$, so this is not the case for $t = 0$. Suppose that $\gamma_t > 0$ for all t. Then, the supplier earns zero profit, which implies that all ICCs will fail unless $q_t = 0$ for all t, which violates the supposition that trade occurs at some point. Thus, there must be at least one $t \geq 1$ such that LL is slack in t. If there are multiple, define T^* as the earliest such t.

For (ii): Since LL is slack for all $t \geq T^*$, Corollary 2 implies that DEC binds for all $t > T^*$. □

Corollary 3. *$\nu_t^b = 0$ for all $t \leq T^*$, and $\nu_{t+1}^b > 0$ for all $t > T^*$.*

Proof. This follows from Remark 3 and Lemma 3. □

Lemma 4. *$\gamma_{t+1} \leq \gamma_t$, with inequality strict if $\gamma_t \in (0,1)$.*

Proof. If $\gamma_{t+1} = 0$, then this holds trivially. We thus need to establish the claim for $\gamma_{t+1} > 0$. Suppose then that $\gamma_{t+1} > \gamma_t$ with $\gamma_{t+1} > 0$. Then, Equation (3) implies

$$0 > \rho_t(1-\lambda) - \frac{\eta_{t+1}}{\delta} + \mu_t(1-\lambda)(1+\nu_t^b - \gamma_{t+1}).$$

Since $\gamma_{t+1} > 0$, Lemma 2 implies that $\eta_{t+1} = 0$. We are thus left with

$$\begin{aligned} 0 &> \rho_t(1-\lambda) + \mu_t(1-\lambda)(1+\nu_t^b - \gamma_{t+1}) \\ &= \rho_t(1-\lambda) + \mu_t(1-\lambda)(1+\nu_{t+1}^b - \gamma_{t+1}) \\ &= \rho_t(1-\lambda) + \mu_t(1-\lambda)\nu_{t+1}^s \end{aligned}$$

where the first equality follows from the fact that the FOC for U_{t+1} implies that $\eta_{t+1} = 0 \implies \nu_t^b = \nu_{t+1}^b$, and the second equality follows from the FOC for R_{t+1}. But the RHS is weakly positive, so this is a contradiction, which establishes that $\gamma_{t+1} \leq \gamma_t$.

Then, to establish the claim about strict inequality, suppose instead that $\gamma_{t+1} = \gamma_t$. We instead have

$$0 = \rho_t(1-\lambda) + \mu_t(1-\lambda)\nu_{t+1}^s,$$

which is only possible if $\rho_t = \nu_{t+1}^s = 0$. But this implies that $\gamma_{t+1} = 1 + \nu_t^b \geq 1$. If $\gamma_t \in (0,1)$, this implies $\gamma_{t+1} > \gamma_t$, which is a contradiction. $\square$

Lemma 5. *ICC binds for some $t < T^*$.*

Proof. We prove this by iterating forward Equation (3). Since $\eta_t = \nu_t^b = 0$ for all $t < T^*$, the equation can be written

$$\gamma_t = \mu_t(1-\lambda) + \rho_t(1-\lambda) + \gamma_{t+1}(1-\mu_t(1-\lambda)).$$

Starting with $\gamma_0 = 1$ and iterating this until $T^* - 1$, for which $\gamma_{t+1} = 0$, we get

$$1 + (1-\mu_0(1-\lambda^{T^*}))\frac{\eta_{T^*}}{\delta} = \sum_{t=0}^{T^*-1}(1-\mu_0(1-\lambda^t))[\mu_t(1-\lambda) + \rho_t(1-\lambda)].$$

Note that $LHS \geq 1$. The first term on the RHS simplies to

$$\sum_{t=0}^{T^*-1}(1-\mu_0(1-\lambda^t))\mu_t(1-\lambda) = (1-\lambda)\mu_0 \sum_{t=0}^{T^*-1}\lambda^t = \mu_0(1-\lambda^{T^*}) < 1.$$

Thus, it cannot be that $\rho_t = 0$ for all $t \leq T^* - 1$, as otherwise $LHS > RHS$. So there must be some $t \leq T^*$ for which the ICC binds. $\square$

Appendix C – Main Tables without Covariates

We pre-specified that we would use the specification in Equation (1). Nonetheless, in this section, we replicate all of the main tables in the analysis using the following simpler regression specification that does not include any covariates:

$$y_i = \alpha + \sum_{j=1}^{4} \beta_j T_{ji} + \varepsilon_i.$$

Table C1: Access to Foreign Goods (No Covariates)

	Horizontal	Vertical				Price
	Find Product $\geq$ 3 Criteria	High Quality Dummy	Quality Score (/50)	Made in Turkey	Index	Price (USD)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled						
Treatment	0.101***	0.107	-0.636	0.145**	0.456***	1.656**
	(0.031)	(0.068)	(0.591)	(0.068)	(0.139)	(0.840)
Panel B: Individual Treatments						
Search Only	0.141***	0.139*	-0.177	0.112	0.492***	2.444**
	(0.039)	(0.084)	(0.683)	(0.081)	(0.174)	(0.991)
	[0.003]	[0.231]	[0.951]	[0.315]	[0.025]	[0.051]
Search + Adverse Selection	0.062	0.154*	-0.026	0.108	0.360*	1.409
	(0.039)	(0.087)	(0.845)	(0.085)	(0.179)	(1.049)
	[0.116]	[0.229]	[0.966]	[0.315]	[0.058]	[0.287]
Search + Moral Hazard	0.121***	0.076	-1.284	0.154*	0.452**	1.746*
	(0.039)	(0.084)	(0.843)	(0.080)	(0.170)	(0.976)
	[0.005]	[0.577]	[0.357]	[0.139]	[0.025]	[0.171]
Search + AS + MH	0.076**	0.062	-0.934	0.203**	0.505***	0.932
	(0.039)	(0.085)	(0.846)	(0.081)	(0.178)	(0.994)
	[0.088]	[0.577]	[0.548]	[0.053]	[0.025]	[0.347]
Control Mean	0.357	0.431	43.064	0.477	0.000	19.990
% Increase (Pooled)	28.3%	24.8%	-1.5%	30.4%	N/A	8.3%
Adjusted R^2	0.01	0.00	0.32	0.09	0.01	0.23
N	1579	359	359	361	361	642

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 2000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all treated groups. Panel B shows the coefficients corresponding to treatment indicators for each of the four treatment arms.

Column 1 is an indicator that is one if the merchant finds a good that matches at least 3 horizontal criteria, and is missing if the merchant never replied to the mystery shopper or was otherwise unreachable. Column 2 is an indicator for whether the good's quality score is above the median product-group quality score. Column 3 is the raw quality score. Column 4 is an indicator for whether the good is made in Turkey, primarily inferred based on whether the label says . See the text for full details of how this outcome is constructed. Column 5 is the Anderson (2008) index combining the vertical outcomes. Column 6 is the price in USD, which is only measured conditional on the firm finding a good matching at least three horizontal criteria.

Table C2: Supplier Relationships (Followup Survey) (No Covariates)

	Regular Suppliers in Turkey			Previous Suppliers		
	Any Reg Sup in Turkey	Num Reg Sup in Turkey	Index	Num Reg Sup Total	Num Reg Sup in Senegal	Ended with Reg Sup
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled						
Treatment	0.035	0.079**	0.138**	-0.032	-0.092	0.070***
	(0.023)	(0.038)	(0.062)	(0.193)	(0.191)	(0.022)
Panel B: Individual Treatments						
Search Only	0.018	0.069	0.118	0.092	0.088	0.062**
	(0.029)	(0.055)	(0.081)	(0.269)	(0.266)	(0.029)
	[0.762]	[0.325]	[0.240]	[0.854]	[0.917]	[0.061]
Search + Adverse Selection	0.056**	0.077*	0.172**	-0.154	-0.236	0.067**
	(0.030)	(0.048)	(0.084)	(0.253)	(0.249)	(0.029)
	[0.127]	[0.240]	[0.101]	[0.854]	[0.736]	[0.050]
Search + Moral Hazard	-0.001	-0.009	0.023	-0.197	-0.168	0.091***
	(0.029)	(0.043)	(0.078)	(0.248)	(0.244)	(0.029)
	[0.974]	[0.835]	[0.789]	[0.818]	[0.817]	[0.008]
Search + AS + MH	0.069**	0.183***	0.244***	0.139	-0.049	0.057**
	(0.031)	(0.062)	(0.087)	(0.258)	(0.251)	(0.029)
	[0.076]	[0.007]	[0.018]	[0.854]	[0.917]	[0.061]
Control Mean	0.167	0.222	0.000	3.700	3.213	0.135
% Increase (Pooled)	21.0%	35.6%	N/A	-0.9%	-2.9%	51.9%
Adjusted R^2	0.00	0.01	0.00	-0.00	-0.00	0.00
N	1680	1680	1680	1681	1681	1671

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 2000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all treated groups. Panel B shows the coefficients corresponding to treatment indicators for each of the four treatment arms.

Column 1 is 1 if the merchant says that they have a regular supplier in Turkey. Column 2 is the number of regular suppliers in Turkey. Column 3 is the Anderson (2008). Column 4 is the total number of regular suppliers. Column 5 is the number of regular suppliers in Senegal. Column 6 is 1 if the merchant has ended a relationship with a regular supplier in the past 3 months. A regular supplier is defined as a supplier from whom the merchant has made two or more orders with an intention of continuing the relationship.

Table C3: Order Value (Mobile Money Data) (No Covariates)

	Any Order	Value Post Mystery Shopping		Total Value	
	Any Order	Order Value (OLS)	Order Value (Poisson)	Order Value (OLS)	Order Value (Poisson)
	(1)	(2)	(3)	(4)	(5)
Panel A: Pooled					
Trust Treatment	-0.003	4.292**	1.346**	3.803	0.390
	(0.020)	(1.904)	(0.485)	(2.813)	(0.286)
Panel B: Individual Treatments					
Search + Adverse Selection	0.001	1.944	0.827	2.261	0.250
	(0.025)	(1.483)	(0.542)	(2.825)	(0.307)
	[0.931]	[0.018]	[0.041]	[0.410]	[0.444]
Search + Moral Hazard	0.008	6.255***	1.637**	4.915	0.481
	(0.025)	(2.803)	(0.511)	(3.726)	(0.335)
	[0.812]	[0.368]	[0.343]	[0.700]	[0.696]
Search + AS + MH	-0.017	4.681	1.410	4.240	0.427
	(0.024)	(4.629)	(0.828)	(5.733)	(0.497)
	[.]	[.]	[.]	[.]	[.]
Control Mean	0.134	1.511	1.511	7.964	7.964
% Increase (Pooled)	-2.2%	284.1%	284.2%	47.8%	47.7%
Adjusted R^2	-0.00	0.00	0.03	-0.00	0.01
N	1500	1500	1500	1500	1500

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 2000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all trust treated groups, where Search Only is the omitted category. Panel B shows the coefficients corresponding to treatment indicators for each of the three treatment groups with trust treatments.

Column 1 is an indicator for whether the merchant ever ordered from a study supplier. Column 2 is the total value of orders. Column 3 is the total value of orders, analysed with Poisson regression. Column 4 is the total value of orders. Column 5 is the total value of orders, analysed with Poisson regression. Mystery shopping took place during the first 13 weeks of the study. All values are in USD.

Table C4: Profit and Sales (No Covariates)

	Raw			Winsorized (1%)		
	Profit (USD)	Sales (USD)	Index	Profit (USD)	Sales (USD)	Index
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled						
Treatment	57.3*	204.0*	0.213**	24.1	83.0	0.109
	(28.8)	(108.1)	(0.081)	(20.9)	(83.6)	(0.069)
Panel B: Individual Treatments						
Search Only	.566	254	.228**	-5.26	107	.126
	(29.1)	(187)	(.115)	(26.2)	(125)	(.096)
	[0.988]	[0.392]	[0.109]	[0.836]	[0.570]	[0.411]
Search + Adverse Selection	30.6	56	.069	18.4	38.1	.035
	(36.4)	(119)	(.092)	(31.5)	(111)	(.092)
	[0.630]	[0.662]	[0.482]	[0.746]	[0.726]	[0.880]
Search + Moral Hazard	-27.4	-103	.106	-27.4	-121	.033
	(24.8)	(103)	(.078)	(24.8)	(93.6)	(.076)
	[0.572]	[0.522]	[0.325]	[0.515]	[0.396]	[0.880]
Search + AS + MH	244***	640**	.463**	120***	325**	.253**
	(93.2)	(282)	(.196)	(45.2)	(152)	(.112)
	[0.012]	[0.047]	[0.043]	[0.022]	[0.081]	[0.081]
Control Mean	188.3	609.5	0.000	188.3	609.5	0.000
% Increase (Pooled)	30.4%	33.5%	N/A	12.8%	13.6%	N/A
Adjusted R^2	0.01	0.01	0.01	0.01	0.01	0.00
N	1351	1378	1431	1351	1378	1431

Note: p-values are computed using randomisation inference. Specifically, we compute the randomised randomisation-t p-value from Young (2019) using 2000 reps. * $p < 0.1$ ** $p < 0.05$ *** $p < 0.01$. We report conventional robust standard errors in parentheses, although we do not use these directly for inference. We also report Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Panel A shows the coefficient from a regression on an indicator that pools all treated groups. Panel B shows the coefficients corresponding to treatment indicators for each of the four treatment arms.

Column 1 is total profit from the past 30 days in USD. Column 2 is total sales from the past 30 days in USD. Column 3 is the Anderson (2008) index combining the previous two columns. Column 4 is total profit from the past 30 days in USD, winsorizing the top 1%. Column 5 is total sales from the past 30 days in USD, winsorizing the top 1%. Column 6 is the Anderson (2008) index combining the previous two columns. Profit is measured using the survey question from De Mel, McKenzie, and Woodruff (2009). Sales is measured using a similar survey question.

Appendix D – Further Details on Solving and Estimating the Model

In this Appendix, we provide more detail on our algorithm to numerically solve and estimate the model.

Solving the Model

As is often the case in dynamic optimisation, the original infinite horizon program is very difficult to work with directly. It is much more tractable to find a way to work with a recursive formulation. The challenge is that, unlike standard dynamic problems encountered in macro, we have two constraints (the DEC and DICC) that are forward looking, and, in particular, forward-looking to an infinite horizon.

The literatures on dynamic moral hazard and limited commitment typically deal with this in one of two ways. One way is to define continuation values as state variables, which completely summarise the future and thus allow the constraints to be written recursively. This "promised utility" approach was originally developed somewhat independently in different theoretical contexts by Spear and Srivastava (1987), Abreu, Pearce, and Stacchetti (1990), and Thomas and Worrall (1988), and in fact we take this approach in Appendix B when we derive qualitative properties of the optimal contract.

The other approach, pioneered by Marcet and Marimon (2019), follows the idea that the original Lagrangean can be rewritten recursively as a pseudo planner's problem, where the Pareto weights are state variables that evolve endogenously to perfectly summarise historical binding constraints. Intuitively, if the DICC is binding in period 0, which implies that the agent must be delivered a certain amount of utils at some point in the future, the Pareto weight on the agent increases over time to ensure that the planner delivers precisely the required amount of utils. This allows the problem to be written recursively because the principal can trade off the benefit of making a constraint "more binding" today against the cost of increasing next period's Pareto weight on the agent. The recursive formulation delivers a Saddle Point Functional Equation, which is analogous to the familiar Bellman Equation but for saddle point problems, which satisfies a number of familiar properties that permit the use of dynamic programming techniques.

The key advantage of the Marcet and Marimon (2019) approach over the promised utility approach is that the feasible set of Pareto weights is known. This is important, because one needs to know the feasible set in order to numerically solve the model. In contrast, in the promised utility approach, we would need to know the feasible set of continuation values, which are endogenous objects that likely depend in complicated ways upon the model parameters. This is not a problem for qualitatively analysing the model, which is why we use this approach in Appendix B, but is a problem for numerically solving it. It is also not an insurmountable obstacle: Abreu et al. (1990) provide an algorithm that can be used within an inner loop to compute the feasible set, and many papers in the literature fruitfully pursue this approach. Nonetheless, the Pareto weight approach completely sidesteps this issue, which is why we use it here.

We first rewrite our problem as a recursive Lagrangean and thus derive the Saddle Point Functional Equation. Define $y(q_t, \tau_t, \mu_t) \equiv (1 - \mu_t(1-\lambda))r(q_t) - \tau_t$. Then, the original dynamic program is as follows

$$\begin{aligned}\max_{\{q_t\},\{\tau_t\}} & y(q_0, \tau_0, \mu_0) + \delta\mu_0(1-\lambda)\bar{U} \\ & + \delta(1-\mu_0(1-\lambda))[y(q_1,\tau_1,\mu_1) + \delta\mu_1(1-\lambda)\bar{U} \\ & + \delta(1-\mu_1(1-\lambda))[y(q_2,\tau_2,\mu_2) + \delta\mu_2(1-\lambda)\bar{U} \\ & + \delta(1-\mu_2(1-\lambda))[...\end{aligned}$$

subject to

$$\begin{aligned}\sum_{n=1}^{\infty}\delta^n(\tau_{t+n} - cq_{t+n}) \geq \xi cq_t & \qquad \forall t \\ U_t \geq \bar{U} & \qquad \forall t \\ \tau_t \geq cq_t & \qquad \forall t,\end{aligned}$$

and with μ_t evolving according to Bayes' Rule. Rewriting the objective function as an infinite sum and including the constraints with Lagrange multipliers, the program can be expressed as follows

$$\begin{aligned}L = & \min_{\{\rho_t\},\{\eta_t\},\{\gamma_t\}} \max_{\{q_t\},\{\tau_t\}} \sum_{t=0}^{\infty}\delta^t\left(1-\mu_0(1-\lambda^t)\right)\left((1-\mu_t(1-\lambda)r(q_t) - \tau_t + \delta\mu_t(1-\lambda)\bar{U}\right) \\ & + \sum_{t=0}^{\infty}\delta^t\rho_t\left[\sum_{\tau=t+1}^{\infty}\delta^{\tau-t}(R_\tau - cq_\tau) - \xi cq_t\right] \\ & + \sum_{t=0}^{\infty}\delta^t(1-\mu_0(1-\lambda^t))\eta_t\left[\sum_{\tau=t}^{\infty}\delta^{\tau-t}\left(1-\mu_\tau(1-\lambda^{\tau-t})\right)\left((1-\mu_\tau(1-\lambda)r(q_t) - \tau_t + \delta\mu_t(1-\lambda)\bar{U}\right) - \bar{U}\right] \\ & + \sum_{t=0}^{\infty}\delta^t\gamma_t\left[\tau_t - cq_t.\right]\end{aligned}$$

Then, with some algebra, we can collect the Lagrange terms directly inside the first infinite sum to express the Lagrangean as a function of "Pareto weights", Lagrange multipliers, state variables, and time-invariant functions.

$$L = \min_{\{\rho_t\},\{\eta_t\},\{\gamma_t\}} \max_{\{q_t\},\{\tau_t\}} \sum_{t=0}^{\infty} \delta^t \big[(\zeta_t^b + \eta_t)\beta_t h_0^b(q_t, \tau_t, \mu_t) + \eta_t \beta_t h_1^b(q_t, \tau_t, \mu_t) \\ + (\zeta_t^s + \gamma_t) h_0^s(q_t, \tau_t) + \rho_t h_1^s(q_t, \tau_t)\big]$$

where $\zeta_t^b \equiv \zeta_0^b + \sum_{\tau=1}^{t-1} \eta_\tau$, $\zeta_t^s \equiv \zeta_0^s + \sum_{\tau=1}^{t-1} \rho_\tau$, $\beta_t = \prod_{s=0}^{t-1}(1 - \mu_s(1-\lambda))$, and

$$\begin{aligned}
h_0^b(q_t, \tau_t, \mu_t) &\equiv \big[(1 - \mu_t(1-\lambda))r(q_t) - \tau_t + \delta\mu_t(1-\lambda)\bar{U}\big] \\
h_1^b(q_t, \tau_t, \mu_t) &\equiv -\bar{U} \\
h_0^s(q_t, \tau_t) &\equiv \tau_t - cq_t \\
h_1^s(q_t, \tau_t) &\equiv -\xi c q_t
\end{aligned}$$

ζ_t^b and ζ_t^s behave like Pareto weights and are equal to the sum of all prior Lagrange multipliers from the forward-looking constraints for the principal (the buyer) and the agent (the seller), respectively. This is the sense is which they fully summarise the shadow cost of constraints from earlier periods of the problem. For example, if DICC is "very binding" in early periods, meaning ρ_t is large for early t, then this is reflected in a large value of ζ_t^s for later t, which causes the "planner" to endogenously choose a high value of τ_t and thus give the agent utility.

We can then write this recursively as a Saddle Point Functional Equation (SPFE) as follows,

$$W(\zeta_t^b, \zeta_t^s, \mu_t, \beta_t) = \min_{\eta_t, \rho_t, \gamma_t} \max_{q_t, \tau_t} \quad (\zeta_t^b + \eta_t)\beta h_0^b(q_t, \tau_t, \mu_t) + \eta_t \beta h_1^b(q_t, \tau_t, \mu_t) \\ + (\zeta_t^s + \gamma_t) h_0^s(q_t, \tau_t) + \rho_t h_1^s(q_t, \tau_t) + \delta W(\zeta_{t+1}^b, \zeta_{t+1}^s, \mu_{t+1}, \beta_{t+1}),$$

subject to

$$\begin{aligned}
\zeta_{t+1}^b &= \zeta_t^b + \eta_t \\
\zeta_{t+1}^s &= \zeta_t^s + \rho_t \\
\beta_{t+1} &= \beta_t(1 - \mu_t(1-\lambda)),
\end{aligned}$$

and $\mu_{t+1} = \mu_t\lambda/(1 - \mu_t(1-\lambda))$ if $q_t > 0$ and high quality is observed, $\mu_{t+1} = 1$ if $q_t > 0$ and low quality is observed, and $\mu_{t+1} = \mu_t$ if $q_t = 0$.

The Saddle Point Functional Equation is analogous to the Bellman Equation for saddle point problems, and Marcet and Marimon (2019) prove that–under some regularity conditions–it has the usual desirable properties associated with dynamic programming problems. In particular, this means that we can obtain solutions to the original problem by using dynamic programming techniques to solve for the value function and policy functions associated with the SPFE.

We solve the model using value function iteration. Since W is homogeneous of degree one in the Pareto weights, note that we can write $W(\zeta_t^b, \zeta_t^s, \mu_t, \beta_t) = \zeta_t^b W(1, \frac{\zeta_t^s}{\zeta_t^b}, \mu_t, \beta_t)$. This means that we can eliminate one state variable. We define a discrete grid over the three state variables and interpolate over the grid using a shape-preserving spline. Within each iteration, the FOCs for τ_t combined with our results in Appendix B allow us to obtain analytical solutions for η_t, γ_t, and τ_t as a function of the state variables. We can then use the FOCs for q_t and q_{ot} to obtain analytically solve for these. Unfortunately, the FOC for ρ_t involves a derivative of the value function so we cannot obtain analytical solutions. Since ρ_t is bounded below at zero, we first evaluate the derivative at zero. If it is positive, we set $\rho_t^* = 0$; otherwise, we use a numerical minimiser to solve for ρ_t^*.

Once we have obtained the value function, we iterate forward from the initial conditions at $t = 0$ to get the solution path. Since the agent's utility is linear in τ_t, the value function has a kink in the neighbourhood of $\beta_t \zeta_t^b = \zeta_t^s$, which Marimon and Werner (2021) show results in inconsistent promises. We resolve this by imposing their Envelope Selection Condition.

Estimation

The above sub-section describes how we solve the model for a given guess of the parameters. In order to estimate the parameters, we need to find the parameters that best match the empirical moments, and thus solve the model for many combinations of parameters. We do this in two steps. First, we solve the model for a grid of (μ_0, ξ, c), and interpolate over the grid using a shape-preserving spline. This gives *functions* for all of the relevant theoretical objects, such as $y_t(\mu_0, \xi, c)$. This means that we do not need to further solve the model as we can simply evaluate these functions at a given (μ_0, ξ, c). Second, we use the above functions to compute the theoretical moments and then use SMM to estimate the parameters.