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

 (\mathbf{r})

Edward Wiles MIT Deivy Houeix MIT

Please click here for the latest version.

03 July 2025

Abstract

Search and trust frictions have historically made it hard for small firms in lowerincome countries to buy inputs from foreign markets. The growth in smartphone ownership and social media usage has the potential to alleviate these barriers. Informed by a dynamic model of relational contracting, we run a field experiment leveraging these technological tools to provide exogenous variation in (1) search frictions and (2) trust frictions (adverse selection and moral hazard) in a large international import market. In the search treatment, we connect a randomly selected 80% of 1,862 small garment firms in Senegal to new suppliers in Türkiye. We then cross-randomize two trust treatments that provide additional information about the types (adverse selection) and incentives (moral hazard) of these new suppliers. Reducing search frictions is sufficient to increase access to foreign markets: in all treated groups, firms are 26% more likely to have the varieties a mystery shopper requests, and the goods sold are 30% more likely to be high quality. However, the trust treatments are necessary for longer-term impact: using both transaction-level mobile payments data and a follow-up survey, we show that these groups are significantly more likely to develop the connections into relationships that persist beyond the endline survey. These new relationships lead to increases in medium-run profit and sales. Finally, we use the treatment effects to estimate the model and evaluate counterfactuals where we further alleviate various combinations of the frictions.

^{*}Author names are in (f) random order. We are immensely grateful to our PhD 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 Isaiah Andrews, Aaron Berman, Miguel Fajardo-Steinhäuser, Ahmet Gulek, Isabela Manelici, Beatrice Montano, Jacob Moscona, Vishan Nigam, Ashesh Rambachan, Frank Schilbach, Advik Shreekumar, Jose Vasquez, Emma Wiles, Alwyn Young, 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. Author contacts: edward.m.wiles@gmail.com, deivy.houeix@gmail.com.

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 seeing what this supplier sells. Second, even if they manage to do this and decide to make an order, they must often pay before observing quality—exposing them to both adverse selection and moral hazard. These issues are particularly severe in lower-income settings, whose vast informal sectors complicate information aggregation and lower state capacities make contracts unenforceable. Yet, while there is a long theoretical tradition studying these frictions—dating back to Stigler (1961) on search and Shapiro (1983) on trust—it has proven harder to analyze them empirically as researchers rarely observe variation that is both exogenous and a close map onto the specific theoretical objects.

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 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.¹ This growth in "social commerce" may reflect the fact that 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 at minimal cost. 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 the first experimental evidence on the extent to which search and trust frictions limit access to foreign input markets and whether information transmitted via social media can alleviate them. We designed a field experiment leveraging key features of social media to alleviate these frictions in the context of a large international import market. Specifically, we randomly allocated 1,862 small garment firms in Dakar across treatment arms that connected them to new suppliers in Türkiye and varied the information 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 Senegal's largest mobile payments provider, and a follow-up survey.

¹In a survey across six African countries, a large non-profit found that among firms that use some ecommerce, 60% use social media exclusively (GSMA, 2023). Statista estimate that total B2C e-commerce revenue across Africa increased from 18 billion USD in 2019 to 34 billion USD in 2024 (Statista, 2023).

We first provide novel descriptive evidence from our baseline survey showing that firms use social media extensively in ways consistent with it alleviating search and trust frictions. For search, we document that 86% of firms are in WhatsApp groups managed by suppliers (both domestic and international), with the median firm in four. These groups help firms see what suppliers sell, but it is unclear how much of the friction it alleviates as firms must still find the suppliers—and thus join the groups—in the first place. For trust, one-quarter of firms are in WhatsApp groups with other firms for the purpose of sharing business information, and two-thirds of firms have recommended or warned against particular suppliers to other firms in the past year. However, these networks are often highly fragmented, so it is unclear how relevant the information that they provide is to the average firm. While we focus on WhatsApp, other apps, such as TikTok, Instagram, and Facebook, are also popular, each used by around a third of firms. Motivated by the widespread use of "social commerce" among businesses, we design an experiment that leverages social media to create exogenous variation in search and trust frictions in international trade.

To make precise what we mean by search and trust frictions, to understand how the frictions interact, and to discuss what types of variation might separate them, we develop a model of relational contracting featuring sequential search for suppliers 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 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 ensure that the goods are high quality. Bad-type suppliers will never do this, while good-type suppliers will only do so if the future value of the relationship exceeds the current-period cost. We characterize the optimal contract and derive an equation showing that adverse selection and moral hazard create wedges between marginal revenue and marginal cost that distort quantity downwards. Moreover, these trust frictions lower the initial value of a relationship and thus the return to searching.

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 treatment is designed to be of real-world

interest in its own right, while also mapping onto one of the three frictions (search, adverse selection, and moral hazard). In the search treatment, we add treated firms to the supplier WhatsApp groups of three different suppliers in Türkiye. We inform firms that the suppliers were recruited by a local team in Türkiye and export to Senegal, but do not provide any further details.² Thus, the treatment creates new matches between Senegalese firms and suppliers abroad.

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. This group aims for them to privately 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 mystery orders that we commissioned prior to the study. Thus, the treatment directly provides (positive) information about a supplier's type and improves learning.

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. We emphasize that we have made this clear to the suppliers and that they therefore have strong incentives to exert effort. Thus, the treatment aims to shift firms' perceptions about the suppliers' incentives.

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 buy goods from all firms. We then measured the type and quality of the goods that they sold to us. This has two advantages: it captures real behaviour, 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 buy 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 Türkiye (a strong signal of quality in this setting). We pre-specified these outcomes and the regression specification that we use

²We chose Türkiye as the exporter country because it is the second largest source of ready-to-wear garments in Senegal (after China), and Turkish-made garments command a large quality premium in this setting, which is well-suited for studying trust frictions as firms are typically worried about suppliers cheating on quality.

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 dimensions. First, we pool the four treated groups together (all of whom received the search treatment). On the horizontal dimension, treated firms are 9.3 percentage points (p < 0.01) more likely to have a suitable good (a 26.1% increase). On the vertical dimension, conditional on having a good, the index increases by 0.412 standard deviations (p < 0.01). The effect on the price is positive, but small, insignificant, and precisely estimated, so the horizontal and vertical gains do not come at the cost of a large price increase.

Second, when we disaggregate 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: rather, these mystery transactions involve small orders (around USD 20-40), which may fall below the threshold where trust becomes a binding constraint. Taken together, the results suggest 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.

While alleviating search frictions by connecting firms 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 a large share of transactions for up to 18 months after the study started.

From the survey data, we find that (pooled) treatment increases the likelihood of having a regular supplier in Türkiye by 3.7 percentage points (p = 0.069), a 22.2% increase relative to control. Here, disaggregating the treatments matters: the effect comes primarily from the groups with the trust treatments, and is largest in the Search + Adverse Selection + Moral Hazard group (the group with both trust treatments) at 7.5 percentage points (p < 0.01). 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 202.6% (p = 0.079) more than Search Only in the 15 months after the study finishes. When we disaggregate this effect into the separate trust treatments, the coefficients are all positive, with Search + Adverse Selection + Moral Hazard again the largest. We thus conclude that that the trust treatments 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 = 0.028) in profits and 245.2 USD (p = 0.042) in sales. These are 43.8% and 40.2% 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 experimental treatment effects as moments to estimate the model. We solve the model numerically using the recursive Lagrangian formulation of Marcet and Marimon (2019), and then calculate theoretical treatment effects, which we match to the experimental ones using Simulated Method of Moments. The estimates imply that firms behave in line with around half of suppliers in the marketplace being bad types, and that the moral hazard parameter is close to the unity assumed by many models of moral hazard. The estimated search costs are large but noisy. With the estimated parameters, we evaluate counterfactuals where we alleviate the frictions beyond the experiment. Unilaterally cutting search costs in half and removing half of the bad types from the market (substantially reducing adverse selection) have large and similar effects on lifetime discounted profit. Lowering moral hazard on its own has a much smaller effect: for most firms, this is not the binding constraint, as the adverse selection problem is sufficiently severe that they do not want to order large quantities anyway. However, as in the experiment, there is a strong complementarity between interventions that lower adverse selection and moral hazard, and the gains are very large to alleviating both simultaneously.

Overall, our results show that both search and trust frictions meaningfully limit the ability of small firms to buy inputs from foreign markets, and social media can reduce these barriers. This provides new evidence on the nature of information frictions in international trade, and also shows how the rapidly evolving digital landscape can change how firms find, learn about, and develop relationships with suppliers.

This paper builds on several literatures. First, while there is a substantial theoretical tradition studying search and trust frictions between firms, making empirical progress has proved challenging. This is largely due to challenges in (1) observing buyer-seller relationships, and (2) obtaining exogenous variation that isolates theoretical forces. A recent literature has begun to overcome some of these challenges, primarily on the data side (e.g., Allen (2014), Antras and Foley (2015), Steinwender (2018), Macchiavello and Morjaria (2015, 2021), Bergquist, McIntosh, and Startz (2024), Startz (2024)). We contribute in two ways. 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 how firms in lower-income countries can and do use new technologies to overcome these frictions.

Second, we contribute to the literature on digital trade, which has documented how the internet (Freund and Weinhold (2004), Fernandes et al. (2019), Akerman et al. (2022)) and online information aggregation platforms like Alibaba (Lendle et al. (2016), Chen and Wu (2021), Carballo et al. (2022)) increasingly facilitate international trade and reduce the impact of distance. We contribute by showing descriptively that many firms (in fact, the vast majority in our setting) in lower-income countries that buy and sell online do not do so through formal platforms designed for B2B trade and instead use social media—"social commerce" rather than "e-commerce". Moreover, we show experimentally that social media mitigates the same information frictions that online platforms aim to alleviate.

Third, we contribute to a literature in trade and development that emphasises the importance of networks when formal institutions to solve information frictions and enforce contracts are inadequate (Rauch (1999), Karlan et al. (2009), Fisman et al. (2017), Cai and Szeidl (2018), Boken et al. (2024)).³ In a classic article, Greif (1993) highlights how 11th-century Maghribi traders sustained a multilateral punishment system for overseas agents through informal information flows over social networks. While our study does not focus on the role of 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 presents the setting and descriptive evidence on how firms use social media in supply chains. Section 3 describes the model. Section 4 describes the experimental design. Section 5 describes the data and methods. Section 6 presents the results. Section 7 presents the model estimation. Section 8 concludes.

³There is also a wide literature outside of economics that analyzes how shared ties substitute for formal institutions among particular groups or geographies. For example, Cohen (1969) studies the Hausa trade diaspora in Yoruba towns in 1960s Nigeria, Weidenbaum and Hughes (1996) studies the "bamboo network" of Chinese entrepreneurs across Southeast Asia in the latter half of the 20th century, and Chin et al. (1996) studies the role of Korean immigrants in Los Angeles in 1968-1977 in facilitating trade in the wig industry.

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 Türkiye. We chose Turkish-made goods for two reasons. First, Türkiye 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 Türkiye" 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 Türkiye 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 Türkiye 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 readyto-wear garments industry in Dakar. These firms are typical of small, informal, owneroperated 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. 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 Türkiye. 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 orders from 10 unknown foreign suppliers, and to opine as to how many such orders would arrive with the anticipated quality. The median firm's view was that this would happen only 50% of the time.

Suppliers in Türkiye The study involves connecting firms in Senegal with suppliers in Türkiye. We work with 30 suppliers, all of whom are based in Istanbul and exporters of ready-to-wear garments to West Africa. We conducted a census of two neighborhoods of 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 three main reasons. First, a sizable Senegalese diaspora operates in the Türkiye–West Africa export industry, alongside Turkish exporters. This reflects general patterns documented by Greif (1993) and Rauch (2001), where shared ties have played a key role in reducing search and contracting frictions in long-distance trade. These shared ties may facilitate reputation-based mechanisms and reduce—but certainly do not eliminate—trust frictions. Second, Senegalese suppliers help reduce logistical barriers, such as language differences and incompatible payment technologies. Given the heterogeneity of the 1,862 firms in our sample, removing these barriers allows us to isolate the core frictions of interest: search and trust. Third, most Senegalese suppliers in Türkiye rely on the same mobile money system used by firms in Senegal. This provides a unique measurement advantage, as we can track transactions directly through the platform's administrative data.

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 high-

light 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). Buyers can negotiate with the supplier or inquire about other goods 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 a lesser 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 could privately message other members to share information. The group could also make it easier for the supplier to build a brand, 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 (in which content is broadcast to all contacts for 24 hours, a feature used less often in the United States), 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 why 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 Türkiye, 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 Türkiye, 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 we also document here in which large e-commerce companies have had limited success at penetrating African markets. At first glance, the widespread reliance on social media—rather than formal B2B platforms—may seem surprising. To understand this, we included follow-up survey questions asking firms why they do not use B2B platforms. Beyond the 39% firms who gave no specific reason, the two most common answers were that firms find them too complicated to use (40%)—often due to language barriers and bank access requirements—and that firms do not trust them (33%).

These statistics confirm the 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.

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 7, we will use the reduced form treatment effects to

⁴The 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 usable with a VPN) and so other social media platforms–such as WeChat–are much more common.

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 \overline{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 \overline{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 a lower cost $(1 - \xi)cq_t$ with $\xi \in (0, 1)$ but the goods are only high quality with probability $\lambda \in (0, 1)$. Avoiding the action and cutting a share ξ of the cost may be interpreted as purchasing the goods from a cheaper manufacturer that only delivers with probability λ , or as shipping the goods with a cheaper exporting service where the goods only arrive with probability λ .

If the firm orders from the supplier, they pay a transfer τ_t , determined endogenously. This

⁵Our 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.

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 are unskilled and thus unable to use this productive technology), 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, \overline{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. In particular, any equilibrium in grim trigger strategies in which good type suppliers choose $a_t = 1$ must satisfy the following standard Dynamic Incentive Compatibility Constraint (DICC),

$$\delta(1-\lambda)V_{t+1} \ge \xi cq_t \tag{DICC}$$

where $V_{t+1} \equiv \sum_{n=1}^{\infty} \delta^n (\tau_{t+1+n} - cq_{t+1+n})$ is the discounted sum of future profits from the relationship. Intuitively, the supplier has already received τ_t , they will only choose $a_t = 1$ if the increased probability $(1 - \lambda)$ of obtaining the future value of the relationship, δV_{t+1} , exceeds the additional cost of ξcq_t . This constraint imposes a (time-varying) ceiling on the contractible quantity. Note that the constraint will have more bite for larger ξ , and will hold trivially for ξ near 0.

Contracts: In period 0, the firm makes a take-it-or-leave-it offer of a long-term contract to the supplier, which consists of a sequence $\{q_t, \tau_t\}_0^\infty$. The contract is relational, meaning that it 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 $U_t \ge \overline{U}$ for all *t*. The enforcement constraint prevents the firm from (for example) promising large

⁶One 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* for simplicity.

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 these strategies provide maximal incentives.

Limited Liability: The contract must satisfy $\tau_t \ge cq_t$ 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, which we show in Proposition 1 in Appendix B is without loss. Intuitively, the inability to commit combined with the parameter restriction implying 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 combined with the DICC ensures they can always earn a strictly positive expected payoff by mimicking the good type.⁷

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

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

$$V_t \equiv \pi(q_t, \tau_t) + \delta V_{t+1}.$$

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

⁷Clearly, this result relies on the parameter restriction that ensures that the firm would never voluntarily trade with bad types. However, this is not important for either the intuition or the bulk of the results: the separating equilibrium that could be obtained in absence of this restriction is qualitatively very similar to the pooling equilibrium. This is because dynamic moral hazard is sufficient to generate an upward sloping quantity profile due to the well-known backloading intuition. Nonetheless, we impose the restriction here because it seems closer in spirit to our empirical setting, in which many firms take their outside option (purchasing from a local wholesaler).

⁸Bad type suppliers behave entirely mechanically, so we do not write out their payoff and, in general, whenever we refer to the payoff or decision of a foreign supplier, it is implied that it is a good type.

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:

$$\delta(1-\lambda)V_{t+1} \ge \xi cq_t \qquad (DICC)$$

$$U_{t+1} \ge \bar{U} \tag{DEC}$$

$$\tau_t \ge cq_t \tag{LL}$$

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

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

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 $U_0 = \overline{U}$, $V_0 = 0$, 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 Proposition 2 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 a 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 result 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 incentives is a very general prediction of models of reputation and dynamic moral hazard (Shapiro (1983), Banerjee and Duflo (2000), Martimort et al. (2017)).

We now state two results relevant to intuition and the interpretation of the experiment.

Result 1 (Intensive Margin). Quantity and value purchased from the foreign supplier are dis-

⁹This 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.

torted downwards by both adverse selection and moral hazard.

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

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

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. We prove in Proposition 2 in Appendix B that this distortion decreases over time, both because of learning (i.e., μ_t decreases) and because it is optimal to backload incentives for the agent (i.e., ρ_t decreases).¹⁰ A similar equation can be derived for value, τ_t . Treatments that alleviate these frictions should therefore increase quantity and value on the intensive margin.

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

The result follows from a straightforward application of the Envelope Theorem (although we provide a formal proof in Proposition 3 in Appendix B), 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 replacing $r(q_t)$ with $r(\psi q_t)$. 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} .¹¹ This reservation value thus defines their outside option.

¹⁰Strictly speaking, ρ_t evolves non-monotonically in general. However, we prove that it can never increase by enough to offset the learning effect. Indeed, if μ_0 is sufficiently small, then ρ_t monotonically decreases.

¹¹The 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. Result 2 above implies that the return to searching is decreasing in μ_0 and ξ .

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 (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. The 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 Türkiye 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 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 does both. We add treated firms 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 periodically encourages firms to share information. We do not tell suppliers about the existence of these groups, so only the information sets of the firms–and not the suppliers' incentives–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 explain that they are also in this fourth 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.

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 high-lights, 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 to emphasize that study suppliers have strong incentives:

¹²We do not expect firms to be able to credibly communicate the role of these groups to suppliers in an attempt to improve incentives. The firm would have no straightforward way to convince suppliers that such claims are not cheap talk—especially given that we explicitly informed firms that suppliers were not told about the groups. Consistent with this, we found no evidence that such behavior occurred in practice.

¹³All firms not in this treatment condition instead receive a "placebo" phone call from a surveyor, asking them for their opinion about supplier WhatsApp groups in general.

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.

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 to firms. We make clear to control firms that we do not vouch for or provide guarantees about the study suppliers—our only role is to make connections. This treatment should not have an impact in a model of adverse selection because it does not provide information to the firms about supplier type–crucially, we do not share the ratings with the firms.¹⁴

¹⁴In 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 quickly whether to initiate a relationship with the study suppliers (indeed, this is embedded within the model as the firm makes a one-time take-it-or-leave-it offer for the long-term plan upon matching). Second, the script does not imply that the enforcement process is particularly fast.

Sub-Treatments We cross-randomised two additional sub-treatments within the pure control group, although, as we specified in our Pre-Analysis Plan, they are not of primary interest. 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 connecting firms *per se* (as in Cai and Szeidl (2018)), we thus also create similar groups here, where none of the firms have been connected with any suppliers.¹⁵

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 (men's clothing, women's clothing, shoes & bags), whether the firm has a physical store, and 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 0.600. We therefore conclude that our analysis consistently estimates average treatment effects.

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 the relationship between consumer willingness to pay and various important variables in

¹⁵We find no meaningful effects of the placebo groups. This was our strong prior, as the nature of the treatment is very different to Cai and Szeidl (2018), where the firms met in-person every month for a year. In our study, we connected firms only via a WhatsApp group, which is well-suited for pooling information about a specific topic but much less likely to lead to broader networking discussions.

our analysis. Second, we use it to calculate statistics on household clothing expenditures.

Baseline and One Week Surveys Upon recruiting a firm (see Section 2.2) we conducted a 30 minute baseline survey between November 2023 and January 2024. The survey contained questions on their supplier relationships, social media usage, e-commerce usage, and 30-day profits and sales. In a short additional follow-up survey one week later, we asked about their beliefs about their access to Turkish-made goods, and their perceptions about the trustworthiness of one of the suppliers that they were matched with.¹⁶

Access to Foreign Goods To evaluate whether social media reduces search and trust frictions in a way that expands firms' access to foreign goods, we designed a novel mystery shopping exercise that directly induces trade. This approach allows us to measure access to foreign goods along both a horizontal dimension (access to more differentiated varieties) and a vertical dimension (access to higher quality varieties), as well as any changes in pricing. 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 part of 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. This exercise enables us to capture firms' sourcing capabilities and pricing in a natural, incentive-compatible way, shortly after treatment exposure.

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.¹⁷ Then, once the good arrives in our office, two tailors and a shoemaker assess its quality according

¹⁶As a measure of the effects of the search and trust treatments on these perceptions (albeit after only one week and self-reported), we report treatment effects on these outcomes in Appendix Table A2. All treatments significantly increase self-reported access to Turkish-made goods, and we reject the joint null that the trust treatments have no effect on self-reported trustworthiness (although only the Search + AS + MH group is individually significantly different).

¹⁷If 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.

to a 50-point scorecard that we developed.¹⁸ 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 between 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 or below the median quality score of its product type), and plot the CDF of consumer WTP separately by quality. 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.

We also attempt to infer whether the good was manufactured in Türkiye. 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.¹⁹ For most goods, we record this information from the label, and the outcome is 1 if the label says "Made in Türkiye" and 0 if it says it was made elsewhere.²⁰

Finally, the mystery shopping also had a secondary goal of providing treated firms with an opportunity to experiment with the study suppliers. Nothing in the procedure makes this explicit, but, if a firm was considering making an order, then the mystery shopper reduces the risk that they will be unable to find a buyer.

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 90% of the sample (1671 firms). 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

¹⁸We 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.

¹⁹We 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.

²⁰For 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 Türkiye (based on sewing patterns, product style, etc.), and set the outcome to 1 if they both opine that it was made in Türkiye and 0 otherwise.

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) we can see transaction profiles over time, (2) it continues 15 months 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

$$y_i = \alpha + \sum_{j=1}^4 \beta_j T_{ji} + \delta y_{0i} + \gamma_s + \rho' X_i + \varepsilon_i,$$
(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).²¹ We also report a version of

²¹This 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, this tends to select few covariates and thus makes little difference. Nonetheless, in Appendix D, we reproduce the main tables without covariates. The results are very similar.

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 pvalues for the sharp null of zero treatment effect using 2,000 permutations of the *t*-statistic. We report conventional heteroskedasticity-robust standard errors in parentheses, but we do not use these for inference directly. As each regression estimates four coefficients, we also calculate Romano-Wolf multiple-testing adjusted p-values in square brackets (Romano and Wolf, 2005). Since there are three combinations of trust treatments, we report the p-value of a joint test that all four coefficients are equal (computed by permuting the F-statistic), which will be the case if the trust treatments have no effect.

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 Because some of our outcomes, 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. For these, we follow the same specification as 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 most tables, we show the pooled regression in Panel A, and then disaggregate across the treated groups in Panel B, with standard errors in parentheses, Romano-Wolf multiple-testing adjusted *p*-values in square brackets, and the *p*-value of a joint test that all coefficients are equal at the bottom.

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.²² Pooling the treatments together, treated firms are 9.3 percentage points more likely to find a suitable good (p = 0.001). This is a 26.1% increase from the control mean of 35.7%. In Panel B, we see that the effect is broadly similar in all four treatment groups (and we cannot reject the joint null that they are all equal), which implies that the effect is not larger for the groups with the trust treatments. The results thus show that relaxing the Search friction by connecting firms with foreign suppliers via social media led to a sizable 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 13.1 percentage points more likely to be high quality (p = 0.044). This is a 30.4% 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 that, while the treatment CDF is to the right of the control CDF from the 30th percentile onwards, it also contains a handful of very low scoring goods.

The outcome in Column 4 is an indicator for whether the product was made in Türkiye. Pooling the treatments together, treated firms are 16.2 percentage points more likely to supply a good saying "Made in Türkiye", a 34.0% increase from the control mean of 47.7% (p = 0.023).²³ In Panel B, we see that the effect is positive and similar for all four treatments.

²²Our PAP specified this indicator variable as the main outcome. As a robustness check, we use the raw number of criteria as an outcome in Appendix Table A3. The pattern of the results is the same. Our PAP also noted that we planned to disaggregate this outcome into an extensive margin effect (agreeing to sell a Turkish-made product at all) versus an intensive margin effect (how suitable was the product provided). We thus also show this in Appendix Table A3, and find that the effect comes mostly through the intensive margin.

²³As a robustness check, we report various alternative ways of defining whether a good was made in Türkiye in Appendix Table A4. The pattern of results is the same.

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 0.412 standard deviations (p = 0.004), and we cannot reject the null that the coefficient is the same for all four treatment groups. 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.

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: Relaxing Search Frictions Expands Access to Foreign Varieties 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 consistent: the results are 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 necessarily mean that the trust frictions do not exist: these are small orders, and so for many firms the risk may be sufficiently low that relaxing the trust frictions does not 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, 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 supplier in Türkiye. In Column 2, the outcome is the number of regular suppliers in Türkiye. In both cases, pooling all four treatment groups together shows that treatment caused firms to develop new relationships with suppliers in Türkiye. The pooled coefficients are 3.7 percentage points (an increase of 22.2% relative to control mean of 16.7%) with p = 0.069, and 0.083 suppliers (an increase of 37.4% relative to the control mean of 0.222) with p = 0.024.

Disaggregating by treatment arm, for both outcomes, the coefficient for Search + Adverse Selection + Moral Hazard is substantially larger: this group sees a 7.5 percentage point increase in the likelihood of having a supplier in Türkiye, and a 0.188 increase in the number of suppliers in Türkiye, both of which are highly significant (including after adjusting for multiple hypothesis testing). The effect is also larger for the Search + Adverse Selection group. Formally, we can reject the null that all four coefficients are equal. This suggests that relaxing trust frictions, and in particular relaxing them together, increased the likelihood that these new connections developed into regular relationships.

In the second part of Table 2, we provide evidence on whether these new relationships complement or substitute for existing relationships. 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 Türkiye (although they are noisy). 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 study firms and study suppliers. Before turning to formal regression results, we first present broad patterns in the raw data. In Figure 6, we plot cumulative order value over time from study suppliers aggregated across the four treatment groups.²⁴ 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

²⁴We omit Pure Control from the figure because they were not connected to any study suppliers. Reassuringly, we find very few orders from such firms.

treatment groups. This suggests that trust treatments did not increase the likelihood of making small experiments with the study suppliers (at least when there is low demand risk, since a customer is already on hand). 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—as discussed in Section 5—interpreting this is a little challenging because 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 flattens, suggesting that most of 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, where the omitted category is Search Only. In Column 1, we test the observation from Figure 6 that only the trust treatments appear to continue ordering after the mystery shopping ends. The outcome is the total value ordered after the mystery shopping ends. 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 implies an 202.6% increase (p = 0.079). When we disaggregate the treatments, we find that the coefficient is positive in all three trust treatment groups, although the *p*-value for the joint *F*-test is 0.180. While the coefficient is largest in the group with both trust treatments, the standard errors are too large to meaningfully distinguish between the trust groups. In Columns 2 and 3, we decompose total value into the number of orders and average value per order. The coefficients on average order size are both large and significant.

We repeat the same three outcomes with the mystery shopping period included in Columns 4, 5, and 6. The broad pattern is similar, which is in line with the observation from Figure 6 that the differences only open up after the mystery shopping ends. In interpreting the magnitudes, note that these are unconditional averages and that most firms did not develop relationships, and that the largest transactions tend to take place off-platform.

6.2.3 Summary: Trust Treatments Convert Matches into Supplier Relationships

Putting together the results from Sections 6.2.1 and 6.2.2, we find that trust treatments made it more likely for these new connections to develop into lasting relationships. In both datasets, the largest effects come from the Search + Adverse Selection + Moral Hazard group.

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 (p = 0.028), or a 43.8% increase from the control mean. For sales, the pooled coefficient is 245.2 USD (p = 0.042), or a 40.2% increase from the control mean. When we disaggregate across the four treated groups separately in Panel B, we see that (for both outcomes), while the effect is positive in all four groups, it is substantially larger and highly significant in the Search + Adverse Selection + Moral Hazard group (including after adjusting for multiple hypothesis testing). Formally, we can reject the null that the four coefficients are equal. 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 coefficient on the Search + Adverse Selection + Moral Hazard group is very large and highly significant, 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 large 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.²⁵

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

²⁵To 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 threshold at endline, and we asked them to confirm their previous response. Out of 13 such firms, 12 confirmed that their previous response was correct.

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.²⁶

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 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 Summary: Substantial Profit Gains from Reducing Search and Trust Frictions

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 groups 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 sharp null, including after adjusting for multiple hypothesis testing. Fourth, as a placebo check,

²⁶While quantile treatment effects are only interpretable as the effect for firms at a given quantile under a rank preservation assumption, we think that it is likely here that these are indeed treatment effects for firms that are larger at baseline. As one piece of suggestive evidence, in Appendix Figure A7 we present treatment effects on the transition matrix between baseline and endline profit: the largest coefficients are in the northeast region, suggesting that treatment makes large firms more likely to remain large.

we compute the same quantile figures using baseline profit as the outcome in Appendix Figure A8 (we did not measure 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 concludes that the evidence suggests precise zero effects on profit throughout most of the distribution, and 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, which is what we observe in Section 6.2.

6.4 Firms Prefer Social Media to Formal 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 6.5 percentage points more likely to have heard of Alibaba, 11.3 percentage points more likely to have searched for goods on Alibaba, and 8.7 percentage points 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 1.4 percentage points and the standard errors are small enough to 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 B2B platforms currently do not.

6.5 Heterogeneous Treatment Effects: Higher Profit Gains for Wholesalers

We pre-specified a set of heterogeneity analyses to explore how treatment effects vary across firm characteristics, with detailed tables reported in Appendix C. Several patterns emerge.

First, we compare retailers and wholesalers. We define wholesalers as the 33% of the sample for whom some (or all) of their sales are wholesale to other firms. While we find no consistent heterogeneity across mystery shopping or relationship outcomes, the profit gains appear concentrated among wholesalers. This aligns with the quantile treatment effects presented in Section 6.3, which suggest that firms higher in the profit distribution benefit more from the intervention. One possible explanation is that wholesalers, who typically buy and sell in bulk, stand to gain more immediately from access to new suppliers. Alternatively, the effects may reflect reduced travel costs for sourcing, given that wholesalers are more likely to travel abroad for procurement. While we cannot definitively distinguish between these mechanisms, Appendix Table A8 provides suggestive evidence: wholesalers in the treatment group are less likely to have traveled internationally for business in the three months preceding the follow-up survey. The estimated effects are large and mostly negative, with a few statistically significant coefficients, though wide confidence intervals suggest they should be interpreted with caution.

Second, we examine whether treatment effects differ by firms' mode of operation—onlineonly versus those with a physical store. The horizontal mystery shopping effects (e.g., custom order responsiveness) are more pronounced among online-only firms, likely reflecting their greater flexibility in adapting to buyer requests. However, we find no systematic differences in vertical mystery shopping outcomes or relationship measures. Interestingly, the profit effects are instead concentrated among firms with physical storefronts, many (45%) of which are wholesalers. This suggests that while online-only firms may be more responsive to initial contact, firms with physical stores may be better positioned to convert new relationships into sustained profit gains. Given that online-only firms are disproportionately women-operated (63%), we observe a similar pattern when disaggregating by the gender of the firm owner.

Finally, we find no meaningful heterogeneity in treatment effects based on prior importing experience, initial exposure to Turkish-made goods, or existing membership in a Turkish supplier WhatsApp group. This suggests that the intervention may be useful for firms regardless of their baseline access to international sourcing networks.

7 Model Estimation

In this section, we use the results of the experiment to estimate the model from Section 3. We have two main goals. First, we want to estimate the parameters governing the search and trust frictions, which are of general interest to users of these types of models. Second, since our treatments are only one particular way of alleviating the frictions, we want to evaluate the gains from trade available if the frictions were to be counterfactually alleviated in different combinations, to different magnitudes, or through different interventions.

Modifications to the Model In order to make the model estimable, we need to make a few modifications. First, we need to implement a functional form for the revenue function, r(q). We assume that 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 foreign supplier that they are matched with, q_f , or their existing supplier, q_e , which we let be perfectly substitutable.²⁷ 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 buyers. The firm's stage game payoff is therefore the following,

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

where z > 0 is a general productivity term and ψ is a match-specific productivity term. We can see from inspection that 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.²⁸

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 . We set $\tilde{c} = 1$ and then let ψ be distributed lognormal with parameters ($\psi_{\mu}, \psi_{\sigma}$), which we will estimate. The match-specific productivities are important theoretically as they define firms'

²⁷In principle, these could be imperfect substitutes if foreign goods are local and foreign goods represent horizontally differentiated varieties. We choose perfect substitutes as it allows for analytical solutions, which substantially speeds up the computation as this problem must be solved within an inner loop.

²⁸To prevent a discontinuity in learning when going from $q_{ft} = 0$ to any positive value, while also ensuring the principal's choice set remains convex, we impose a technical assumption that for q_{ft} below some small order size q, posteriors are given by $\mu_{t+1} = \mu_t + (\mu_{t+1}^{\text{Bayes}} - \mu_t) \cdot q_{ft}/q$, where μ_{t+1}^{Bayes} is the posterior implied by Bayes' Rule. Making learning a continuous function of q_{ft} helps ensure the problem satisfies appropriate regularity conditions. Intuitively, it captures the idea that there is a minimum order size (which is indeed the case in our empirical setting).

beliefs about the value of searching, and empirically as they rationalise the fact that two otherwise-identical firms matched with the same suppliers may order different amounts. We also allow the productivity, z, to be heterogeneous, drawn from the empirical distribution implied by baseline profits.²⁹ This rationalises the firm size distribution at baseline.

Third, we need to take a stand on what firms' outside options are and 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, firms with $z > \bar{z}$ will import directly, meaning that they 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 satisfactory foreign supplier and are in the late phase of the relationship, where they earn $(1 - \delta)\bar{U}(z)$ every period. For such firms, we then set p_e to be the constant price that delivers exactly $(1 - \delta)\bar{U}(z)$ if the firm could buy frictionlessly. 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 . As the experiment has four treatment effects per outcome, we include two additional, related parameters to better match the combinations of treatments. First, we include the discount factor, δ , which is important for dynamic incentives and thus has been a target parameter in previous studies of dynamic moral hazard (e.g., Einav, Finkelstein, and Schrimpf (2015)). We interpret one period as one month. Second, we include the probability of low effort generating high quality goods, λ . This parameter governs the speed of learning and, indirectly, the strength of the interaction between adverse selection and moral hazard: when learning is slow (i.e., λ is high), moral hazard becomes less consequential. Finally, we also need to estimate the parameters governing the lognormal distribution of match-specific productivity, ($\psi_{\mu}, \psi_{\sigma}$), which are closely related to the search friction as they govern the expected returns to engaging in search. We thus estimate these seven parameters, and calibrate the

²⁹Specifically, 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 .

remaining two, (σ, p_l) . We discuss this calibration in Appendix Table A9.

Moments Since the ultimate goal is to use the model to extrapolate from the treatments to consider counterfactuals where we vary the extent of the frictions, 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 Türkiye, and the three treatment effects on mobile money order value post mystery shopping.³⁰

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 begin the relationship with $\mu_1 = \mu_0 \lambda/(1 - \mu_0(1 - \lambda)) < \mu_0$. This intends to capture the 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 among very similar firms. Specifically, they face a modified DICC of the form $(1 - \lambda)\delta V_{t+1} \ge \xi cq_t - (1 - \lambda)\delta NV_{t+1}$, or $(1 - \lambda)\delta V_{t+1} \ge \xi/Ncq_t$, where N = 5 is the ratio of the number of firms that made orders from study suppliers to the number of study suppliers.

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 compute standard errors by bootstrapping the estimation procedure 200 times and calculating the standard

³⁰For the mystery shopping moments, we match the treatment effects expressed in percent of control mean, as we do not know the exact share of transactions covered by the mobile money data. To discipline the levels, we also include the control means as explicit moments, but we put weight them downwards by a factor of 0.2 to ensure they do not drive identification.

deviation of the estimates. We provide a substantially more detailed description of the procedure to solve and estimate the model in Appendix E.

Results and Model Fit We present the estimated parameters in Table 6. The value of $\hat{\mu}_0 = 0.51$ suggests an important adverse selection problem as around half of suppliers are bad types. The value of $\hat{\xi} = 0.89$ is close to a standard model of moral hazard (which would have $\xi = 1$), and the monthly discount factor $\hat{\delta}$ is fairly low at 0.76, although the actual bite of moral hazard depends on how severely the DICCs bind. The search cost of s = 217 USD is quite large, although it has a very high standard error. Overall, the estimation implies that at least one of the frictions—adverse selection, moral hazard, or search cost—must be sizable to explain why firms do not search more, despite the sizable treatment effects observed experimentally.

We plot the targeted moments in Panel A of Appendix Figure A9. The model is able to match the treatment effects fairly well. It cannot match the scale of the complementarity in the profit moments, but it is able to match it in survey relationship moments. As in the experimental data, the model generates a heavy right tail: in Panel B, we show that (not directly targeted) quantile treatment effects computed in the model look similar the experiment ones in Figure 7.

Counterfactuals We now use the estimated model to evaluate counterfactuals in which we reduce the frictions by setting key parameters—governing adverse selection, moral hazard, and search—to half their estimated values. We show the results in Figure 8. We first cut the search cost in half—reducing it to zero is not well-defined because as $s \rightarrow 0$ firms will search indefinitely to obtain arbitrarily good matches. This roughly doubles lifetime discounted profits. The effect is large relative to the experimental treatment effect because lowering the search costs by half—permanently—is significantly stronger than receiving three random draws. That said, it is worth noting that these firms are quite small and thus the absolute magnitude of these gains (around \$1400 in discounted lifetime profits) may be in similar orders of magnitude to some fixed costs of other solutions, such as travel.

We then cut the adverse selection parameter, μ_0 , in half—reflecting a scenario of removing half of the bad types from the market. Similar to the search counterfactual, this roughly doubles profits. Note that this too is much stronger than our experimental treatment: our experiment improved beliefs about three specific suppliers; the counterfactual here is improving beliefs about *all* suppliers. When we cut the moral hazard parameter, ξ , in half, reflecting a scenario where a supplier's incentives are doubled, we find much smaller ef-
fects of 14%. This happens because, in the model, while the discount factor is relatively low, moral hazard is not the binding constraint—firms are already reluctant to order from foreign suppliers because of adverse selection, so the DICC binds infrequently in practice. However, this is no longer the case in the last scenario, where we both remove half of the bad types and double supplier's incentives: as suggested by the experimental results, there is a strong complementarity, as incentives become the binding constraint when firms are less worried initially about bad types.

8 Conclusion

We study the extent to which search and trust frictions limit the ability of small firms to source inputs from foreign markets—and whether social media can help reduce these barriers. Finding and developing relationships with suppliers is a first order issue for firms, often constrained by information frictions. Understanding the magnitude and nature of these frictions is critical, especially given the growing reliance on social media to find and evaluate suppliers.

Our results show that social media can meaningfully reduce both search and trust frictions. Our mystery shopping exercise shows that firms connected to foreign suppliers gain greater access to differentiated and higher-quality foreign inputs. This was not obvious *ex ante*: firms might have preferred buying through local intermediaries or traveling abroad themselves. Instead, the sizable effects point to substantial search costs that direct digital connections help overcome.

However, better access is only part of the story. Using survey and mobile money data, we find that trust interventions make it more likely for these new connections to turn into meaningful and lasting supplier relationships. This, in turn, leads to higher profits for some firms.

Despite substantial demand for online commerce, firms rarely use formal B2B platforms. Our Alibaba training had no effect on purchasing behavior, suggesting that the constraint isn't just a lack of knowledge. The stronger appeal of social media may reflect deeper differences in how it reduces frictions in ways that formal platforms do not—through greater flexibility, familiarity, or trust. We view understanding these differences as a promising direction for future research.

Taken together, our results show that both search and trust frictions meaningfully limit the ability of small firms to buy inputs from foreign markets, and that social media can be used to meaningfully reduce them. These technologies are changing the structure of supply

chains: allowing retailers to import directly and helping wholesalers work with suppliers without costly travel. The rapid growth in access to smartphones and mobile connectivity, as well as the efforts of social media companies themselves to introduce e-commerce features, is likely to further these changes. These findings suggest that the rapidly developing digital landscape in lower- and middle-income countries is likely to meaningfully benefit small firms and require researchers, policymakers, and organizations to update how they think about how firms find, learn about, and develop relationships with suppliers.

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.
- A. Akerman, E. Leuven, and M. Mogstad. Information frictions, internet, and the relationship between distance and trade. *American Economic Journal: Applied Economics*, 14(1): 133–163, 2022.
- T. Allen. Information frictions in trade. *Econometrica*, 82(6):2041–2083, 2014.
- 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.
- 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.
- 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.
- L. F. Bergquist, C. McIntosh, and M. Startz. Search costs, intermediation, and trade: Experimental evidence from ugandan agricultural markets. Technical report, National Bureau of Economic Research, 2024.
- J. Boken, L. Gadenne, T. K. Nandi, and M. Santamaria. Community networks and trade. 2024.
- J. Cai and A. Szeidl. Interfirm relationships and business performance. *The Quarterly Journal of Economics*, 133(3):1229–1282, 2018.
- J. Carballo, M. R. Chatruc, C. S. Santa, and C. V. Martincus. Online business platforms and international trade. *Journal of International Economics*, 137:103599, 2022.
- M. X. Chen and M. Wu. The value of reputation in trade: Evidence from alibaba. *Review of Economics and Statistics*, 103(5):857–873, 2021.

- K.-S. Chin, I.-J. Yoon, and D. Smith. Immigrant small business and international economic linkage: A case of the korean wig business in los angeles, 1968–1977. *International Mi*gration Review, 30(2):485–510, 1996.
- A. Cohen. *Custom and Politics in Urban Africa: A Study of Hausa Migrants in Yoruba Towns*. University of California Press, 1 edition, 1969.
- 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.
- L. Einav, A. Finkelstein, and P. Schrimpf. The response of drug expenditure to nonlinear contract design: Evidence from medicare part d. *The Quarterly Journal of Economics*, 130 (2):841–899, 2015.
- A. M. Fernandes, A. Mattoo, H. Nguyen, and M. Schiffbauer. The internet and chinese exports in the pre-alibaba era. *Journal of Development Economics*, 138:57–76, 2019.
- R. Fisman, D. Paravisini, and V. Vig. Cultural proximity and loan outcomes. *American Economic Review*, 107(2):457–492, 2017.
- Y.-f. Fong and J. Li. Relational contracts, limited liability, and employment dynamics. *Journal of Economic Theory*, 169:270–293, 2017.
- C. L. Freund and D. Weinhold. The effect of the internet on international trade. *Journal of international economics*, 62(1):171–189, 2004.
- 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.
- D. Karlan, M. Mobius, T. Rosenblat, and A. Szeidl. Trust and social collateral. *The Quarterly Journal of Economics*, 124(3):1307–1361, 2009.
- A. Lendle, M. Olarreaga, S. Schropp, and P.-L. Vézina. There goes gravity: ebay and the death of distance. *The Economic Journal*, 126(591):406–441, 2016.
- 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.
- 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. *Working Paper*, 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.
- M. L. Weidenbaum and S. Hughes. *The bamboo network: How expatriate Chinese entrepreneurs are creating a new economic superpower in Asia.* Simon and Schuster, 1996.
- 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

	Horizontal	Vertical				Price
	Has Product ≥ 3 Criteria	High Qual Dummy	Qual Score (/50)	Made in Turkey	Index	Price (USD)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled						
Treatment	0.093***	0.131**	-0.391	0.162**	0.412***	0.710
	(0.030)	(0.067)	(0.562)	(0.067)	(0.135)	(0.699)
Panel B: Individual Treatments						
Search Only	0.140***	0.175**	0.259	0.138*	0.426***	1.360
-	(0.038)	(0.083)	(0.675)	(0.082)	(0.166)	(0.896)
	[0.001]	[0.115]	[0.898]	[0.169]	[0.039]	[0.366]
Search + Adverse Selection	0.053	0.159*	-0.027	0.130	0.398**	0.445
	(0.037)	(0.087)	(0.858)	(0.086)	(0.169)	(0.833)
	[0.153]	[0.156]	[0.975]	[0.169]	[0.040]	[0.822]
Search + Moral Hazard	0.105***	0.099	-1.179	0.183**	0.399**	1.064
	(0.038)	(0.082)	(0.789)	(0.081)	(0.162)	(0.866)
	[0.014]	[0.334]	[0.376]	[0.061]	[0.040]	[0.475]
Search + AS + MH	0.072*	0.083	-0.717	0.211***	0.424**	-0.025
	(0.038)	(0.084)	(0.814)	(0.081)	(0.171)	(0.818)
	[0.102]	[0.334]	[0.707]	[0.032]	[0.040]	[0.975]
Control Mean	0.357	0.431	43.064	0.477	0.000	19.990
% Increase (Pooled)	26.1%	30.4%	-0.9%	34.0%	N/A	3.6%
All Coefs Equal <i>p</i> -val	0.098	0.626	0.309	0.670	0.996	0.303
Adjusted R^2	0.09	0.04	0.34	0.09	0.02	0.40
N	1579	359	359	361	359	642

Table 1: Access to Foreign Goods

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). To test the null that the trust treatments have no effect, we report at the bottom of the table the *p*-value for an *F*-test that all coefficients are equal, computed by permuting the *F*-statistic. 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 any 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 firm has a good that matches at least 3 horizontal criteria, and is missing if the firm 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 the label. See the text for full details of how this outcome is constructed. Column 5 is the Anderson (2008) index combining the vertical outcomes to account for multiple hypothesis testing across 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.

	Regular	Suppliers in	Гürkiye	Previous Suppliers		
	Any Sup in Türkiye	Num Sup in Türkiye	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**	0.121*	-0.102	-0.130	0.062***
	(0.021)	(0.035)	(0.063)	(0.164)	(0.170)	(0.021)
Panel B: Individual Treatments						
Search Only	0.024	0.081	0.085	-0.101	-0.111	0.050*
	(0.027)	(0.051)	(0.081)	(0.216)	(0.222)	(0.028)
	[0.576]	[0.250]	[0.466]	[0.891]	[0.922]	[0.114]
Search + Adverse Selection	0.049*	0.063	0.147*	-0.171	-0.245	0.066**
	(0.029)	(0.047)	(0.086)	(0.218)	(0.226)	(0.028)
	[0.197]	[0.298]	[0.202]	[0.841]	[0.665]	[0.052]
Search + Moral Hazard	0.003	0.003	0.009	-0.118	-0.064	0.076**
	(0.026)	(0.041)	(0.079)	(0.201)	(0.212)	(0.029)
	[0.914]	[0.927]	[0.919]	[0.891]	[0.922]	[0.038]
Search + AS + MH	0.075***	0.188***	0.250***	-0.014	-0.096	0.053*
	(0.028)	(0.057)	(0.087)	(0.222)	(0.225)	(0.028)
	[0.031]	[0.002]	[0.012]	[0.957]	[0.922]	[0.114]
Control Mean	0.167	0.222	0.000	3.700	3.213	0.135
% Increase (Pooled)	22.2%	37.4%	N/A	-2.8%	-4.0%	45.9%
All Coefs Equal <i>p</i> -val	0.072	0.015	0.044	0.925	0.880	0.820
Adjusted R^2	0.14	0.12	-0.02	0.32	0.26	0.05
N	1680	1680	1680	1681	1681	1671

Table 2: Supplier Relationships (Followup Survey)

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). To test the null that the trust treatments have no effect, we report at the bottom of the table the *p*-value for an *F*-test that all coefficients are equal, computed by permuting the *F*-statistic. 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 any 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 says that they have a regular supplier in Türkiye. Column 2 is the number of regular suppliers in Türkiye. Column 3 is the Anderson (2008) index combining the previous two columns to account for multiple hypothesis testing across outcomes. 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 firm has ended a relationship with a regular supplier in the past 3 months. A regular supplier is defined as a supplier from whom the firm has made two or more orders with the intention to continue the relationship.

	Post N	lystery Sho	opping	Entire Period		
	Total Order Value (USD)	Num Orders	Avg Order Value (USD)	Total Order Value (USD)	Num Orders	Avg Order Value (USD)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled						
Trust Treatment	7.21*	0.04	2.22***	6.95	0.02	2.08**
	(4.65)	(0.07)	(0.74)	(5.04)	(0.08)	(1.00)
Panel B: Individual Treatments	:					
Search + Adverse Selection	3.71	-0.02	3.12**	4.47	-0.00	2.98*
	(3.06)	(0.07)	(1.60)	(3.68)	(0.10)	(1.74)
	[0.313]	[0.898]	[0.062]	[0.348]	[0.999]	[0.152]
Search + Moral Hazard	6.49**	0.11	2.73**	6.50*	0.06	3.46**
	(3.63)	(0.09)	(1.20)	(4.36)	(0.11)	(1.67)
	[0.066]	[0.513]	[0.030]	[0.189]	[0.911]	[0.076]
Search + AS + MH	11.53	0.05	0.77	9.95	0.00	-0.24
	(12.39)	(0.11)	(0.62)	(12.95)	(0.13)	(0.89)
	[0.440]	[0.898]	[0.233]	[0.596]	[0.999]	[0.796]
Control Mean	3.56	0.14	0.90	9.05	0.33	3.70
% Increase (Pooled)	202.6%	32.4%	247.5%	76.8%	6.4%	56.3%
All Coefs Zero p-val	0.180	0.556	0.033	0.343	0.948	0.067
Adjusted R^2	-0.00	0.01	-0.00	0.01	0.04	0.00
N	1500	1500	1500	1500	1500	1500

Table 3: Order Value (Mobile Money Data)

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). To test the null that the trust treatments have no effect, we report at the bottom of the table the *p*-value for an *F*-test that all coefficients are equal, computed by permuting the *F*-statistic. 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 any 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 the total value of orders from study suppliers after the mystery shopping period finished. Column 2 is the number of orders from study suppliers. Column 3 is the average order size from study suppliers after the mystery shopping period finished. Column 4 is the total value of orders from study suppliers over the entire period. Column 5 is the number of orders from study suppliers over the entire period. Column 5 is the number of orders from study suppliers over the entire period. Column 6 is the average order size from study suppliers over the entire period. Mystery shopping took place during the first 3 months of the study, from 16 November 2023 to 22 February 2024. The entire period covered by the data comprises 18 months, from 16 November 2023 to 31 May 2025. All values are in USD.

		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**	245.2**	0.314**	45.5**	121.5	0.174**	
	(31.6)	(108.4)	(0.107)	(21.2)	(78.6)	(0.076)	
Panel B: Individual Treatments							
Search Only	31.7	311*	.197*	22.1	178	.21*	
	(28.8)	(177)	(.114)	(25.4)	(112)	(.108)	
	[0.547]	[0.154]	[0.181]	[0.619]	[0.239]	[0.121]	
Search + Adverse Selection	43.5	80.8	.193	30.8	73.1	.124	
	(37.7)	(114)	(.129)	(30.7)	(100)	(.102)	
	[0.547]	[0.709]	[0.226]	[0.619]	[0.686]	[0.371]	
Search + Moral Hazard	15.8	-24	.036	8.21	-25.3	6.2e-03	
	(25.1)	(106)	(.085)	(23.2)	(87)	(.078)	
	[0.568]	[0.830]	[0.697]	[0.737]	[0.771]	[0.939]	
Search + AS + MH	254***	636***	.88***	128***	269**	.371***	
	(89.3)	(267)	(.315)	(39.2)	(124)	(.137)	
	[0.005]	[0.029]	[0.003]	[0.002]	[0.093]	[0.021]	
Control Mean	188.3	609.5	0.000	188.3	609.5	0.000	
% Increase (Pooled)	43.8%	40.2%	N/A	24.2%	19.9%	N/A	
All Coefs Equal <i>p</i> -val	0.055	0.041	0.038	0.024	0.048	0.026	
Adjusted R^2	0.13	0.12	-0.01	0.25	0.29	-0.02	
N	1351	1378	1298	1351	1378	1298	

Table 4: Profit and Sales

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). To test the null that the trust treatments have no effect, we report at the bottom of the table the *p*-value for an *F*-test that all coefficients are equal, computed by permuting the *F*-statistic. 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 any 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 to account for multiple hypothesis testing across outcomes. 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 to account for multiple hypothesis testing across outcomes. Profit is measured using the survey question from De Mel, McKenzie, and Woodruff (2009). Sales is measured using a similar survey question.

	Heard of Alibaba (1)	Searched on Alibaba (2)	Compared Prices with Supplier (3)	Bought on Alibaba (4)
e-Commerce Treatment	0.065** (0.025)	0.113** (0.052)	0.087* (0.049)	0.014 (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

Table 5: Effect of Alibaba Training

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. Conventional robust standard errors are reported in parentheses. This table shows the effect of the Alibaba training treatment on Alibaba usage.

Description	Parameter	Estimate
Share of bad types	μ_0	0.51 (0.12)
Moral hazard cost saving	ξ	0.90 (0.17)
Sequential search cost (USD)	s	217 (94)
Match productivity (mean)	ψ_{μ}	-5.17(0.89)
Match productivity (std dev)	ψ_{σ}	1.48 (0.38)
Discount factor	δ	0.76 (0.15)
Bad type quality probability	λ	0.25 (0.06)

Table 6: Estimated Structural Parameters

Notes: This table shows the estimated structural parameters. We estimate the parameters using Simulated Method of Moments, where we simulate the treatment in the model and match the model-implied treatment effects to the experimental treatment effects on profit, whether the firm has a supplier in Turkey, and orders from mobile money. We weight the moments by the inverse of the empirical variance-covariance matrix, and run the optimiser on 100 randomized starting points. We present bootstrapped standard errors in parentheses, which we calculate as the standard deviation of the results of the estimation procedure run on 200 sets of treatment effects obtained by bootstrapping the experimental data. Please see Appendix E for more detail on the procedure to solve and estimate the model.

Figures

Figure 1: Experimental Design Tree



Note: This figure presents the design tree of the experiment, which identifies three frictions—search, adverse selection, and moral hazard—via randomized variation in information and messaging. Treatments are described in Section 4.1. Among recruited firms, 80% of them are randomly assigned to the Search treatment, in which they are matched via WhatsApp groups to three suppliers of Turkish-made goods. Within the Search group, 50% are cross-randomized into the Adverse Selection treatment, where they are placed in peer groups to share supplier experiences and receive seeded information from independent recommenders. Separately, 50% of firms in the Search group are cross-randomized into the Moral Hazard treatment, in which firms are told that suppliers are rated and subject to removal, increasing perceived incentives for supplier effort. Two additional sub-treatments—Alibaba training and placebo peer groups—are cross-randomized within the pure control group to test secondary hypotheses within the control group. The first provides a brief Alibaba training to assess whether limited platform knowledge constrains B2B e-commerce adoption. The second creates peer WhatsApp groups without supplier matches to test whether connecting firms alone drives outcomes, serving as a placebo for the Adverse Selection treatment. Neither sub-treatment yields meaningful effects.





Note: These pictures show two typical supplier WhatsApp groups used in our study context, as described in Section 2.3. These groups function as virtual storefronts: the supplier regularly posts photos or videos showcasing available inventory, often with prices and product details. Clients—usually 50 to 100 per group— observe the posts but can rarely interact publicly; instead, they inquire or negotiate privately with the supplier via direct messages. Such groups could help reduce search costs by giving buyers access to a steady stream of updated product information and facilitating seamless communication. They might also reduce trust frictions to a lesser extent, as visible group membership and repeated engagement raise the reputational costs of cheating.



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

(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

Note: This figure shows the results of two exercises in which surveyors showed various images of garments to respondants, randomising whether they stated that the good was made in Türkiye 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 the surveyor stated 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



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 months since the study begun (16 November 2023). As the group sizes are slightly different due to integer indivisibility issues, we normalize group-level totals by N_j/N_{avg} , where *j* is the number of units in treatment group N_j and N_{avg} is the average number of units across all four treatment groups. Pure control is omitted as they were not connected to any study suppliers.





(a) Profit





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.

Figure 8: Counterfactuals



Note: This figure shows the results of counterfactuals that alleviate various combinations of the frictions. The first column is the baseline scenario. The second column reduces the search cost parameter *s* by half. The third column reduces the share of bad types, μ_0 , which governs the size of the adverse selection problem, by half. The fourth column reduces ξ , the parameter governing how much a supplier can reduce the marginal cost by if they choose low effort, which scales moral hazard, by half. Finally, the last column reduces both μ_0 and ξ by half each. All values are expressed in the expected present discounted value of lifetime profits for the firm.

Appendix A – Additional Tables and Figures

Tables

	Control	Search	Search AS	Search MH	Search AS	Joint <i>p</i> -value
					MH	
Variable	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.49	0.51	0.49	0.53	0.49	0.7
Online Only	0.67	0.66	0.68	0.7	0.65	0.02
Business Age	4.85	4.74	5.2	4.72	5.21	0.5
Share Cust Turkey	0.43	0.45	0.45	0.44	0.45	0.95
Any Reg Supp Turkey	0.22	0.21	0.24	0.21	0.19	0.52
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	0.42
Bought Alibaba Ever	0.16	0.13	0.15	0.18	0.13	0.22
N	362	373	379	381	367	

Table A1: Balance Table

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 *p*-value of 0.600.

	Indicator for Answ	vers 1 or 2 (out of 5)
	Find Products From Turkey	Trust Supplier
	(1)	(2)
Panel A: Pooled		
Treatment	0.110***	0.020
	(0.029)	(0.030)
Panel B: Individual Treatments		
Search Only	0.078**	
2	(0.038)	
	[0.048]	
Search + Adverse Selection	0.140***	0.009
	(0.037)	(0.036)
	[0.001]	[0.794]
Search + Moral Hazard	0.087**	-0.025
	(0.037)	(0.036)
	[0.048]	[0.688]
Search + AS + MH	0.136***	0.075**
	(0.038)	(0.038)
	[0.001]	[0.103]
Control Mean	0.453	0.303
% Increase (Pooled)	24.3%	6.6%
All Coefs Equal <i>p</i> -val	0.237	N/A
All Coefs Zero <i>p</i> -val	N/A	0.050
Adjusted R^2	0.06	0.06
N	1636	1201

Table A2: One Week Survey

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). To test the null that the trust treatments have no effect, we report at the bottom of the table the *p*-value for an *F*-test that all coefficients are equal, computed by permuting the *F*-statistic. 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 any 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 they can find Turkish products. Column 2 is 1 if they reported a trust level 1 or 2, 0 if they reported a trust level of 3 or 4 or 5, and missing otherwise.

	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	0.029	0.101***	0.361***	0.357**	
	(0.025)	(0.035)	(0.124)	(0.140)	
Panel B: Individual Treatments					
Search Only	0.029	0.144***	0.664***	0.705***	
-	(0.031)	(0.043)	(0.161)	(0.178)	
	[0.665]	[0.005]	[0.000]	[0.001]	
Search + Adverse Selection	0.014	0.067	0.196	0.182	
	(0.031)	(0.044)	(0.160)	(0.182)	
	[0.665]	[0.129]	[0.323]	[0.322]	
Search + Moral Hazard	0.045	0.099**	0.379**	0.308*	
	(0.030)	(0.043)	(0.158)	(0.175)	
	[0.345]	[0.069]	[0.046]	[0.197]	
Search + AS + MH	0.025	0.092**	0.204	0.232	
	(0.032)	(0.044)	(0.158)	(0.179)	
	[0.665]	[0.070]	[0.323]	[0.322]	
Control Mean	0.781	0.457	1.650	2.111	
% Increase (Pooled)	3.7%	22.1%	21.9%	16.9%	
All Coefs Equal <i>p</i> -val	0.731	0.357	0.017	0.018	
Adjusted R^2	0.10	0.04	0.05	0.03	
N	1579	1269	1579	1269	

Table A3: Horizontal Outcomes (Detailed)

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). To test the null that the trust treatments have no effect, we report at the bottom of the table the *p*-value for an *F*-test that all coefficients are equal, computed by permuting the *F*-statistic. 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 any 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 agrees to sell or search for the product, and is missing if the firm never replied to the mystery shopper or was otherwise unreachable. Column 2 is 1 if the firm has 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 firm 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, for a product, and is 0 if the firm agreed but never sent any product.

		From Turkey	
-	Made in Turkey (Label)	Made in Turkey (Tailor Judgement)	Made in Turkey (Label + Tailors)
	(1)	(2)	(3)
Panel A: Pooled			
Treatment	0.210**	0.099	0.159**
	(0.081)	(0.081)	(0.072)
Panel B: Individual Treatments			
Search Only	0.169*	0.134	0.138
	(0.096)	(0.095)	(0.086)
	[0.147]	[0.424]	[0.189]
Search + Adverse Selection	0.172*	0.010	0.107
	(0.100)	(0.102)	(0.089)
	[0.147]	[0.909]	[0.231]
Search + Moral Hazard	0.241**	0.132	0.192**
	(0.094)	(0.095)	(0.084)
	[0.033]	[0.424]	[0.081]
Search + AS + MH	0.271***	0.096	0.189**
	(0.097)	(0.097)	(0.084)
	[0.023]	[0.499]	[0.081]
Control Mean	0.489	0.581	0.544
% Increase (Pooled)	42.9%	17.0%	29.2%
All Coefs Equal <i>p</i> -val	0.509	0.479	0.605
Adjusted R^2	0.10	-0.00	0.07
N	255	287	330

Table A4: Vertical Outcomes (Detailed)

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). To test the null that the trust treatments have no effect, we report at the bottom of the table the *p*-value for an *F*-test that all coefficients are equal, computed by permuting the *F*-statistic. 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 any 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.

	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	-0.090***	-0.047	-0.037	-0.054	-0.103***
·	(0.019)	(0.040)	(0.033)	(0.032)	(0.032)	(0.036)	(0.039)
	[0.309]	[0.913]	[0.018]	[0.395]	[0.584]	[0.226]	[0.025]
Search + Adverse Selection	-0.029	-0.028	-0.063*	-0.006	-0.038	-0.033	-0.088**
	(0.019)	(0.045)	(0.034)	(0.032)	(0.033)	(0.035)	(0.038)
	[0.309]	[0.913]	[0.061]	[0.939]	[0.584]	[0.335]	[0.041]
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)
	[0.405]	[0.942]	[0.025]	[0.480]	[0.839]	[0.208]	[0.041]
Search + AS + MH	-0.023	-0.013	-0.103***	-0.011	-0.012	-0.058	-0.109***
	(0.020)	(0.045)	(0.033)	(0.033)	(0.033)	(0.035)	(0.039)
	[0.405]	[0.942]	[0.006]	[0.939]	[0.839]	[0.226]	[0.020]
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%
All Coefs Equal <i>p</i> -val	0.902	0.826	0.610	0.480	0.769	0.835	0.882
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

Table A5: Supplier Relationships (Further Results on Substitution)

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). To test the null that the trust treatments have no effect, we report at the bottom of the table the *p*-value for an *F*-test that all coefficients are equal, computed by permuting the *F*-statistic. 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 any 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 says that they have a regular supplier in China. Column 2 is 1 if the firm says that they have a regular supplier in China. Column 3 is 1 if the firm says that they use Facebook to learn about suppliers. Column 4 is 1 if the firm says that they use TikTok to learn about suppliers. Column 5 is 1 if the firm says that they use Instagram to learn about suppliers. Column 6 is 1 if the firm 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 firm says that they have forwarded a photo or video from a supplier group to a regular supplier to try to obtain a supplier group to a regular supplier to try to obtain 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 firm has made two or more orders with an intention of continuing the relationship.

	Post N	lystery Sho	opping	Entire Period		
	Total Order Value (USD)	Num Orders	Avg Order Value (USD)	Total Order Value (USD)	Num Orders	Avg Order Value (USD)
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Pooled						
Trust Treatment	1.087	0.272	1.258***	0.553	0.079	0.442*
	(0.550)	(0.417)	(0.356)	(0.356)	(0.250)	(0.213)
Panel B: Individual Treatments	1					
Search + Adverse Selection	0.667	-0.194	1.509***	0.369	0.022	0.575*
	(0.522)	(0.532)	(0.480)	(0.316)	(0.301)	(0.289)
	[0.457]	[0.902]	[0.017]	[0.659]	[0.994]	[0.131]
Search + Moral Hazard	1.029	0.584	1.417***	0.533	0.175	0.660**
	(0.494)	(0.453)	(0.423)	(0.324)	(0.303)	(0.272)
	[0.269]	[0.550]	[0.013]	[0.407]	[0.904]	[0.077]
Search + AS + MH	1.432	0.273	0.614	0.732	0.028	-0.066
	(0.901)	(0.649)	(0.428)	(0.709)	(0.389)	(0.245)
	[0.457]	[0.902]	[0.226]	[0.659]	[0.994]	[0.799]
Control Mean	3.56	0.14	0.90	9.05	0.33	3.70
% Increase (Pooled)	196.5%	31.3%	251.8%	73.8%	8.2%	55.6%
All Coefs Zero p-val	0.386	0.533	0.015	0.558	0.952	0.039
Adjusted R^2	0.03	0.03	0.04	0.01	0.04	0.06
N	1500	1500	1500	1500	1500	1500

Table A6: Order Value (Mobile Money Data) – Poisson Regressions

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). To test the null that the trust treatments have no effect, we report at the bottom of the table the *p*-value for an *F*-test that all coefficients are equal, computed by permuting the *F*-statistic. 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 any 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 the total value of orders from study suppliers after the mystery shopping period finished. Column 2 is the number of orders from study suppliers. Column 3 is the average order size from study suppliers after the mystery shopping period finished. Column 4 is the total value of orders from study suppliers over the entire period. Column 5 is the number of orders from study suppliers over the entire period. Column 5 is the number of orders from study suppliers over the entire period. Column 6 is the average order size from study suppliers over the entire period. Mystery shopping took place during the first 3 months of the study, from 16 November 2023 to 22 February 2024. The entire period covered by the data comprises 18 months, from 16 November 2023 to 31 May 2025. All values are in USD.

		Profit	Sales		
	Profit 30 Days	Profit 30 Days	Sales 30 Days	Sales 30 Days	
	(USD)	Winsorized 1% (USD)	(USD)	Winsorized 1% (USD)	
	(1)	(2)	(3)	(4)	
Panel A: Pooled					
Treatment	0.307*	0.188	0.217	0.118	
	(0.132)	(0.112)	(0.142)	(0.126)	
Panel B: Individual Treatments					
Search Only	0.088	0.074	0.315	0.198	
	(0.147)	(0.133)	(0.210)	(0.158)	
	[0.924]	[0.832]	[0.398]	[0.487]	
Search + Adverse Selection	0.097	0.119	-0.041	-0.026	
	(0.189)	(0.162)	(0.167)	(0.159)	
	[0.924]	[0.832]	[0.814]	[0.865]	
Search + Moral Hazard	-0.047	-0.009	-0.151	-0.147	
	(0.138)	(0.143)	(0.174)	(0.143)	
	[0.924]	[0.954]	[0.646]	[0.518]	
Search + AS + MH	0.865***	0.480***	0.475**	0.260	
	(0.214)	(0.142)	(0.212)	(0.162)	
	[0.001]	[0.012]	[0.142]	[0.359]	
Control Mean	188.3	188.3	609.5	609.5	
% Increase (Pooled)	35.9%	20.7%	24.2%	12.5%	
All Coefs Equal <i>p</i> -val	0.005	0.019	0.065	0.058	
Adjusted R^2	0.34	0.38	0.46	0.48	
N	1351	1351	1378	1378	

Table A7: Profit and Sales (Poisson Regression)

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). To test the null that the trust treatments have no effect, we report at the bottom of the table the *p*-value for an *F*-test that all coefficients are equal, computed by permuting the *F*-statistic. 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 any 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 A8: 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	-0.001	-0.009	-0.020	-0.005	-0.022	-0.004
	(0.049)	(0.039)	(0.026)	(0.017)	(0.013)	(0.004)
	[0.976]	[0.929]	[0.854]	[0.986]	[0.277]	[0.696]
Search + Adverse Selection	0.007	-0.011	-0.012	-0.001	-0.025**	0.004
	(0.041)	(0.034)	(0.022)	(0.018)	(0.012)	(0.008)
	[0.975]	[0.929]	[0.854]	[0.995]	[0.144]	[0.842]
Search + Moral Hazard	-0.051	-0.050*	-0.014	-0.008	-0.015	0.010
	(0.038)	(0.030)	(0.020)	(0.017)	(0.014)	(0.008)
	[0.410]	[0.284]	[0.854]	[0.977]	[0.298]	[0.587]
Search + AS + MH	-0.052	-0.043	-0.016	-0.001	-0.022	0.001
	(0.037)	(0.032)	(0.020)	(0.020)	(0.014)	(0.006)
	[0.410]	[0.397]	[0.854]	[0.995]	[0.277]	[0.885]
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%
All Coefs Equal <i>p</i> -val	0.164	0.252	0.986	0.981	0.873	0.168
Adjusted R^2	0.18	0.28	0.09	0.04	0.07	0.01
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 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). To test the null that the trust treatments have no effect, we report at the bottom of the table the *p*-value for an *F*-test that all coefficients are equal, computed by permuting the *F*-statistic. 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 any 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 internationally 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 internationally 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 China in the past 3 months. Column 6 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 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 A9: Calibrated Parameters

Parameter	Value	Origin
σ	3.02	Average markup from baseline survey.
p_l	12.29	Average input price from baseline survey among firms without a foreign supplier.

Note: This table presents the values of the parameters that are calibrated in the model.

Figures



Figure A1: Social Media Usage (Physical Store Only)

(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 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 + Adverse Selection, Search + Moral Hazard, Search + AS + MH) pooled for visual ease. To be consistent with the regression in the table, we first residualise quality using stratum fixed effects and any covariates selected by PDS Lasso 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: Treatment Effects on Transition Matrix

Note: This figure shows the treatment effects for the Search + Adverse Selection + Moral Hazard treatment on indicators for whether a firm transitioned between different percentile profit buckets between baseline and endline.


Figure A8: Placebo Check: Quantile Treatment Effects on Baseline Profit

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.







(b) Fit on Quantile Treatment Effects

Note: Panel A shows the 13 targeted moments in the structural estimation, with confidence internals on the empirical moments (computed by bootstrapping the experimental treatment effects) shown to highlight the relative precisions. We weight moments by the inverse of the variance-covariance matrix of the treatment effects, except for the two control mean moments, which only have the diagonal term and are downweighted by receiving an additional multiplicative penalty of 0.2. Panel B compares quantile treatment effects between the model and the data, which were not explicitly targeted moments, for the group with both trust treatment. The Data moments are the same as in Figure 7.

Figure A9: Model Fit



Figure A10: The role of incentives at different levels of baseline beliefs

Note: This figure shows the value of the Lagrange multiplier, ρ_t , on the Dynamic Incentive Compatability Constraint on the equilibrium path at both $\hat{\mu}_0$ and $\hat{\mu}_0/2$. We average the multipliers across the estimated ψ distribution and, to avoid selection issues, only include cases where the firms orders any positive amount in both cases. We normalize the values by ζ_t^b .

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)

$$\max_{q} \lambda r(q) - (1 - \xi)cq < \max_{q} r(q) - p_l q$$

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., $(\max_q \lambda r(q) - (1 - \xi)q)/(1 - \delta)$. 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 $\tau_t \ge cq_t > (1 - \xi)cq_t$ 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:

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

The modified program is as follows.

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

The FOCs are as follows

$$(1 - \mu_t (1 - \lambda))r'(q_t)(1 + \nu_t^b) = \rho_t \xi(c - c_0) + (\gamma_t + \nu_t^s)c = 0 \qquad (q_t)$$

$$1 + \nu_t^b = \gamma_t + \nu_t^s \tag{(\tau_t)}$$

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

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

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

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

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.

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

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

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 $(1 \delta)\overline{U}$, 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 $q_{t+1} \le q_t$. The FOC for q_t is

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

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

$$\delta(1-\lambda)V_{t+1} < (1-\lambda)V_{t+1} = (1-\lambda)\left(R_{t+1} - cq_{t+1} + \delta V_{t+2}\right) = \delta(1-\lambda)V_{t+2} = \xi cq_{t+1} \le \xi cq_t.$$

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

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

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

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

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., $U_{t+2} \ge \overline{U}$. 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 $U_t < \overline{U}$. 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

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

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

Lemma 3. If any trade occurs, then there exists finite $T^* \ge 1$ such that (i) LL binds for all $t < T^*$ and is slack for all $t \ge T^*$, 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 $s \ge T^*$. 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 $t \ge 1$ 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 $t \ge T^*$, Corollary 2 implies that DEC binds for all $t > T^*$.

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

Proof. This follows from Remark 1 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

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

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

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 $\gamma_{t+1} = 1 + \nu_t^b \ge 1$. If $\gamma_t \in (0, 1)$, this implies $\gamma_{t+1} > \gamma_t$, which is a contradiction.

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 $LHS \ge 1$. 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.

Proposition 3. *The period-0 value of the relationship is decreasing in* μ_0 *and* ξ *.*

Proof. We first show the case for ξ . Using the Envelope Theorem to differentiate the original infinite horizon Lagrangian with respect to ξ gives

$$\frac{dL}{d\xi} = -\sum_{t=0}^{\infty} \delta^t \rho_t \sum_{s=t+1}^{\infty} cq_s.$$

Since Proposition 2 established that the DICC binds for at least some t, we know that $\rho_t > 0$ for at least some t. So this is strictly negative except for the trivial case where $q_t = 0$ for all t, in which case it is zero.

Second, for μ_0 , it is easier to work with the recursive formulation. Differentiating the period-0 program gives

$$\frac{dW_0}{d\mu_0} = \underbrace{-(1-\lambda)r(q_0)(1+\nu_0^b) - \delta(1-\lambda)\nu_0^b(U_1-\bar{U}) - \delta(1-\lambda)W_1}_{\equiv \phi_0} + \delta(1-\mu_0(1-\lambda))\frac{d\mu_1}{d\mu_0}\frac{\partial W_1}{\partial \mu_1}$$

The first three terms are all negative. The only term that we cannot immediately sign is $\partial W_1/\partial \mu_1$. However, since the problem is recursive, we can simply repeatedly lead ϕ_t to obtain

$$\frac{dW_0}{d\mu_0} = \sum_{t=0}^{\infty} \delta^t \prod_{s=0}^t (1 - \mu_s (1 - \lambda)) \frac{d\mu_{s+1}}{d\mu_s} \phi_t.$$

This is strictly negative, except in the trivial case where the firm never purchases anything with the foreign supplier. \Box

Appendix C – Heterogeneity Analysis

	Horizontal Dummy (1)	High Quality Dummy (2)	Any Supp in Türkiye (3)	Profit Sales Index (1%) (4)
Search Only	0.194***	0.158	0.021	0.004
	(0.047)	(0.097)	(0.034)	(0.104)
Search + AS	0.049	0.221**	0.053	0.013
	(0.047)	(0.105)	(0.036)	(0.121)
Search + MH	0.131***	0.119	-0.006	0.004
	(0.047)	(0.100)	(0.033)	(0.091)
Search + AS + MH	0.036	0.058	0.054	0.024
	(0.049)	(0.110)	(0.037)	(0.103)
S Only * Wholesaler	-0.158*	-0.071	-0.011	0.469
	(0.083)	(0.194)	(0.067)	(0.324)
S + AS * Wholesaler	0.048	-0.179	0.001	0.255
	(0.083)	(0.193)	(0.067)	(0.323)
S + MH * Wholesaler	-0.024	-0.139	0.016	-0.144
	(0.084)	(0.190)	(0.066)	(0.231)
S + AS + MH * Wholesaler	0.114	0.005	0.023	1.209***
	(0.083)	(0.187)	(0.068)	(0.437)
All Inter Zero <i>p</i> -val	0.016	0.796	0.989	0.035
Adjusted R^2	0.01	-0.01	0.00	0.05
N	1579	359	1680	1298

Table C1: Heterogeneity by Retailer vs Wholesaler

Note: This table shows the main results with treatment interacted with an indicator for whether the firm sells at least partially wholesale, defined as having a positive share of sales that are wholesale to other firms. *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. The outcome in Column 1 is the same as Column 1 of Table 1. The outcome in Column 2 is the same as Column 2 of Table 1. The outcome in Column 4 is the same as Column 6 of Table 4.

	Horizontal Dummy (1)	High Quality Dummy (2)	Any Supp in Türkiye (3)	Profit Sales Index (1%) (4)
	0.000***	0.112	0.040	0.059
Search Only	0.208***	0.113	0.040	-0.058
	(0.048)	(0.097)	(0.038)	(0.076)
Search + AS	0.115**	0.139	0.065*	-0.017
	(0.047)	(0.104)	(0.039)	(0.098)
Search + MH	0.182***	0.074	0.005	-0.053
	(0.047)	(0.097)	(0.036)	(0.060)
Search + AS + MH	0.130***	0.033	0.081**	-0.026
	(0.049)	(0.104)	(0.040)	(0.066)
S Only * Physical Store	-0.205**	0.090	-0.066	0.653**
	(0.082)	(0.200)	(0.056)	(0.337)
S + AS * Physical Store	-0.168**	0.048	-0.035	0.519
	(0.081)	(0.194)	(0.060)	(0.376)
S + MH * Physical Store	-0.222***	-0.079	-0.036	0.165
	(0.082)	(0.209)	(0.056)	(0.285)
S + AS + MH * Physical Store	-0.152*	0.098	-0.031	1.582***
	(0.082)	(0.185)	(0.061)	(0.497)
All Inter Zero <i>p</i> -val	0.062	0.912	0.833	0.027
Adjusted R^2	0.04	-0.01	0.02	0.12
N	1579	359	1680	1298

Table C2: Heterogeneity by Physical Store vs Online Only

Note: This table shows the main results with treatment interacted with an indicator for whether the firm had a physical store at baseline. *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. The outcome in Column 1 is the same as Column 1 of Table 1. The outcome in Column 2 is the same as Column 2 of Table 1. The outcome in Column 4 is the same as Column 6 of Table 4.

	Horizontal Dummy	High Quality Dummy	Any Supp in Türkiye	Profit Sales Index (1%)
	(1)	(2)	(3)	(4)
Search Only	0.149***	0.114	0.050*	0.074
	(0.048)	(0.104)	(0.030)	(0.121)
Search + AS	0.067	0.120	0.070**	0.008
	(0.048)	(0.106)	(0.031)	(0.102)
Search + MH	0.106**	0.045	0.024	-0.135*
	(0.048)	(0.106)	(0.028)	(0.075)
Search + AS + MH	0.081*	0.023	0.072**	0.416**
	(0.049)	(0.108)	(0.031)	(0.192)
S Only * Experience	-0.021	0.076	-0.104	0.180
	(0.083)	(0.177)	(0.068)	(0.278)
S + AS * Experience	-0.013	0.110	-0.053	0.246
	(0.082)	(0.190)	(0.069)	(0.329)
S + MH * Experience	0.044	0.091	-0.080	0.285
	(0.082)	(0.178)	(0.067)	(0.230)
S + AS + MH * Experience	-0.014	0.114	-0.014	0.319
	(0.083)	(0.179)	(0.072)	(0.430)
All Inter Zero <i>p</i> -val	0.938	0.973	0.511	0.760
Adjusted R^2	0.01	-0.01	0.05	0.02
N	1579	359	1680	1298

Table C3: Heterogeneity by Direct Import Experience

Note: This table shows the main results with treatment interacted with an indicator for whether the firm has direct online importing experience, defined as having bought goods from another country online directly at least once in the past 12 months. *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. The outcome in Column 1 is the same as Column 1 of Table 1. The outcome in Column 2 is the same as Column 2 of Table 1. The outcome in Column 3 is the same as Column 1 of Table 2. The outcome in Column 4 is the same as Column 6 of Table 4.

	Horizontal Dummy	High Quality Dummy	Any Supp in Türkiye	Profit Sales Index (1%)
	(1)	(2)	(3)	(4)
Search Only	0.063	0.180	0.024	0.038
	(0.063)	(0.145)	(0.043)	(0.173)
Search + AS	-0.006	0.121	0.052	-0.033
	(0.061)	(0.155)	(0.044)	(0.176)
Search + MH	0.034	0.200	-0.020	-0.168
	(0.061)	(0.150)	(0.038)	(0.125)
Search + AS + MH	0.034	0.354**	0.041	0.255
	(0.063)	(0.141)	(0.045)	(0.261)
S Only * High Share	0.137*	-0.053	-0.012	0.161
	(0.081)	(0.178)	(0.059)	(0.236)
S + AS * High Share	0.110	0.046	0.015	0.204
	(0.080)	(0.187)	(0.061)	(0.248)
S + MH * High Share	0.128	-0.177	0.023	0.214
	(0.080)	(0.182)	(0.056)	(0.178)
S + AS + MH * High Share	0.067	-0.427**	0.037	0.420
	(0.082)	(0.176)	(0.062)	(0.359)
All Inter Zero <i>p</i> -val	0.409	0.073	0.943	0.696
Adjusted R^2	0.02	0.01	0.01	0.01
N	1542	353	1636	1261

Table C4: Heterogeneity by Baseline Turkish Share

Note: This table shows the main results with treatment interacted with an indicator for whether the share of the firm's customers in the past 30 days that buy Turkish-made products is above the median. *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. The outcome in Column 1 is the same as Column 1 of Table 1. The outcome in Column 2 is the same as Column 2 of Table 1. The outcome in Column 4 is the same as Column 6 of Table 4.

	Horizontal Dummy (1)	High Quality Dummy (2)	Any Supp in Türkiye (3)	Profit Sales Index (1%) (4)
Search Only	0.028	0.111	-0.028	0.439**
	(0.056)	(0.128)	(0.037)	(0.216)
Search + AS	-0.047	0.148	-0.010	0.240
	(0.054)	(0.127)	(0.038)	(0.206)
Search + MH	0.044	-0.033	-0.056	0.005
	(0.056)	(0.120)	(0.036)	(0.152)
Search + AS + MH	-0.036	0.040	0.039	0.904***
	(0.054)	(0.122)	(0.041)	(0.321)
S Only * Female	0.226***	0.052	0.090	-0.574**
	(0.078)	(0.173)	(0.058)	(0.232)
S + AS * Female	0.222***	0.017	0.134**	-0.313
	(0.077)	(0.177)	(0.060)	(0.242)
S + MH * Female	0.148*	0.200	0.097*	-0.047
	(0.078)	(0.170)	(0.056)	(0.182)
S + AS + MH * Female	0.232***	0.046	0.060	-0.822**
	(0.078)	(0.173)	(0.062)	(0.343)
All Inter Zero <i>p</i> -val	0.014	0.769	0.216	0.028
Adjusted R^2	0.02	-0.01	0.02	0.05
N	1579	359	1680	1298

Table C5: Heterogeneity by Gender of Owner

Note: This table shows the main results with treatment interacted with an indicator for whether the firm owner is female. *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. The outcome in Column 1 is the same as Column 1 of Table 1. The outcome in Column 2 is the same as Column 3 is the same as Column 1 of Table 2. The outcome in Column 4 is the same as Column 6 of Table 4.

	Horizontal Dummy	High Quality Dummy	Any Supp in Türkiye	Profit Sales Index (1%)
	(1)	(2)	(3)	(4)
Search Only	0.135***	0.184*	0.053**	0.177
	(0.045)	(0.097)	(0.027)	(0.127)
Search + AS	0.037	0.245**	0.054**	0.196
	(0.044)	(0.099)	(0.028)	(0.148)
Search + MH	0.104**	0.133	0.028	-0.044
	(0.044)	(0.096)	(0.026)	(0.093)
Search + AS + MH	0.061	0.125	0.095***	0.438**
	(0.045)	(0.100)	(0.030)	(0.182)
S Only * In Turkish Group	0.019	-0.173	-0.144	-0.221
	(0.095)	(0.191)	(0.089)	(0.314)
S + AS * In Turkish Group	0.072	-0.295	-0.033	-0.547**
	(0.096)	(0.210)	(0.088)	(0.239)
S + MH * In Turkish Group	0.092	-0.221	-0.116	0.042
	(0.096)	(0.200)	(0.088)	(0.267)
S + AS + MH * In Turkish Group	0.072	-0.237	-0.126	0.421
	(0.096)	(0.192)	(0.090)	(0.564)
All Inter Zero <i>p</i> -val	0.854	0.660	0.407	0.099
Adjusted R^2	0.01	-0.00	0.10	0.01
N	1562	357	1662	1286

Table C6: Heterogeneity by Baseline Turkish Supplier Group

Note: This table shows the main results with treatment interacted with an indicator for whether the firm was in a supplier WhatsApp group with a supplier based in Türkiye at baseline. *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. The outcome in Column 1 is the same as Column 1 of Table 1. The outcome in Column 2 is the same as Column 2 of Table 1. The outcome in Column 3 is the same as Column 1 of Table 2. The outcome in Column 4 is the same as Column 6 of Table 4.

Appendix D – 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:

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

	Horizontal	tal Vertical				
	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	0.101***	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
Ν	1579	359	359	361	361	642

Table D1: Access to Foreign Goods (No Covariates)

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.

	Regular S	uppliers in Turl	ĸey	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	
Ν	1680	1680	1680	1681	1681	1671	

Table D2: Supplier Relationships (Followup Survey) (No Covariates)

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.

	Any Order	Value Post My	stery 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	0.000	4.684*	1.246*	4.382	0.428	
	(0.020)	(2.402)	(0.508)	(3.202)	(0.300)	
Panel B: Individual Treatments						
Search + Adverse Selection	0.003	1.713	0.645	2.173	0.235	
	(0.025)	(1.547)	(0.531)	(2.934)	(0.310)	
	[0.896]	[0.504]	[0.425]	[0.745]	[0.686]	
Search + Moral Hazard	0.013	6.386***	1.476**	5.335	0.501	
	(0.025)	(2.908)	(0.502)	(3.870)	(0.331)	
	[0.837]	[0.023]	[0.062]	[0.357]	[0.438]	
Search + AS + MH	-0.017	5.983	1.426	5.674	0.525	
	(0.024)	(6.303)	(0.876)	(7.263)	(0.549)	
	[0.825]	[0.504]	[0.420]	[0.745]	[0.686]	
Control Mean	0.134	1.891	1.891	8.209	8.209	
% Increase (Pooled)	0.0%	247.7%	247.6%	53.4%	53.4%	
Adjusted R^2	-0.00	0.00	0.03	-0.00	0.01	
N	1500	1500	1500	1500	1500	

Table D3: Order Value (Mobile Money Data) (No Covariates)

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.

Column 1 is an indicator for whether the firm 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.

	Raw			W	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	

Table D4: Profit and Sales (No Covariates)

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 E – 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 completely 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, Pearce, and Stacchetti (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

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

subject to

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

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

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

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.

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

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

$$h_0^b(q_t, \tau_t, \mu_t) \equiv \left[(1 - \mu_t (1 - \lambda)) r(q_t) - \tau_t + \delta \mu_t (1 - \lambda) \bar{U} \right]$$
$$h_1^b(q_t, \tau_t, \mu_t) \equiv -\bar{U}$$
$$h_0^s(q_t, \tau_t) \equiv \tau_t - cq_t$$
$$h_1^s(q_t, \tau_t) \equiv -\xi cq_t$$

 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,

$$W(\zeta_{t}^{b},\zeta_{t}^{s},\mu_{t},\beta_{t}) = \min_{\eta_{t},\rho_{t},\gamma_{t}} \max_{q_{t},\tau_{t}} (\zeta_{t}^{b}+\eta_{t})\beta h_{0}^{b}(q_{t},\tau_{t},\mu_{t}) + \eta_{t}\beta h_{1}^{b}(q_{t},\tau_{t},\mu_{t}) + (\zeta_{t}^{s}+\gamma_{t})h_{0}^{s}(q_{t},\tau_{t}) + \rho_{t}h_{1}^{s}(q_{t},\tau_{t}) + \delta W(\zeta_{t+1}^{b},\zeta_{t+1}^{s},\mu_{t+1},\beta_{t+1}).$$

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, 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 linearly interpolate over the grid.³¹ 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 when constructing the path.

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 $(\mu_0, \xi, c, \delta, \lambda)$, where *c* is the cost of foreign supplier (varying *c* and ψ are isomorphic and one can simply map between them by setting $\psi = 1/c$), and linearly interpolate over the grid. This gives *functions* for all of the relevant theoretical objects, such as $y_t(\mu_0, \xi, c, \delta, \lambda)$, which means that we do not need to further solve the model as we can simply evaluate these functions at a given $(\mu_0, \xi, c, \delta, \lambda)$. We use a grid of 0.01 : 0.01 : 0.91 for μ_0 and ξ , a grid of 1.0 : 0.1 : 13.0for *c* (all that matters in practice is the ratio c/p_l , and we calibrate $p_l = 12.29$), a grid of 0.39 : 0.1 : 0.99 for δ , and a grid of 0.1 : 0.2 : 0.9 for λ . The model satisfactorily converges for 91% of the parameter combinations, and we impute the values for the combinations that do not converge using a simple nearest-neighbour means algorithm.

Second, we use the above functions for estimation. We define 15 theoretical moments. The

³¹We experimented with both linear interpolation and cubic splines, and generally found that linear interpolation performed better.

first four are the treatment effects on winsorized profit, i.e., the coefficients in Column 4 of Table 4. The next there are the post-mystery shopping total value ordered treatment effects in Column 1 of Table 3. For these three moments, because the mobile money data does not capture all transactions (and is likely top-censored), we do not match the coefficients directly but rather the relative treatment effects—i.e., the coefficients divided by the control mean reported in the table. The next four moments are the four treatment effects on the survey outcome of whether a firm has a regular supplier in Turkey, i.e., the coefficients in Column 1 in Table 2. This gives 11 treatment effect moments which provide exogenous variation to identify the parameters governing the frictions. As these moments only identify differences, we also include the control mean of the profit treatment column and the control mean of the supplier in Turkey column, which ensures that the estimated distribution of match-specific productivity does not match the treatment effects while deviating significantly from baseline levels. We calculate the variance-covariance matrix of these 13 empirical moments by drawing 1,000 bootstrap sub-samples of the experimental data. We then zero out the off-diagonal terms that capture the covariance between the two means moments and the treatment effect moments, making the matrix block diagonal. We then invert the matrix and multiply the two diagonals covering the means moments by a factor α to ensure they do not dominate the treatment effects moments—the means are typically more precised estimated, but of less value in terms of identification. We set $\alpha = 0.2$, although the results are fairly similar and the model does a reasonable job of matching the treatment effects without straying too far from the baseline levels for a fairly wide range of α .

The two dimensions of heterogeneity are the firm-specific productivities, z, and the matchspecific productivities, ψ . Since baseline profit (at least for firms interacting with a domestic supplier) is given by $\pi = \frac{(\sigma-1)^{\alpha-1}}{\sigma^{\sigma}} z^{\sigma} p^{1-\sigma}$, once we have calibrated σ and p we have a one-to-one mapping of baseline profits (which we observe in our baseline survey) to z. We thus take this as the empirical distribution of z. As firms often report profit in round numbers, there are several mass points in this distribution, so to improve numerical stability in the optimization routine, whenever we draw z we jitter it by adding $u \sim U[-1, 1]$. For distribution of match-specific productivity, we directly estimate the parameters of the LogNormal, $(\psi_{\mu}, \psi_{\sigma})$. We winsorize the upper tail such that $c = 1/\psi$ never falls below 1. Note that the absolute level is not meaningful—the model is homogenous of degree $1 - \sigma$ in c and p_l , so all that matters is the ratio c/p_l , and since we calibrate $p_l = 12.29$, a relative price of less than one tenth is already very low. We do this both because heavy tails make estimation more sensitive to implementation details of the SMM and because the value function becomes very large as c becomes small (causing numerical issues with convergence). We actually compute the moments by Monte Carlo integration using $\Omega = 10,000$ draws of ψ .³²

As this is a non-stationary dynamic model, the questions of what is baseline and how changes in the environment affect existing relationships are non-trivial. We assume that the baseline equilibrium is the long-run steady state of the model. That means that all firms with $z \geq \overline{z}$, where \overline{z} is the endogenous cutoff above which a firm will engage in search for a foreign supplier, have already searched and are in the "terminal" phase of the relationship. Recall from Section 3 that the optimal contract features a unique T^* such that the firm earns exactly their outside option for all $t > T^*$, which we refer to as the terminal phase. Thus, all firms with $z \leq \bar{z}$ earn $U_l = \frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}} z^{\sigma} p_l^{1-\sigma}/(1-\delta)$, while all firms with $z > \bar{z}$ earn $\overline{U}(z) > U_l$, where $\overline{U}(z)$ is the expected value of participating in the search game, defined implicitly by $\int_{\psi} \max\{U(\psi, z) - \overline{U}(z), 0\} dF(\psi) = s$. For simplicity, we simply assume that these firms earning $\overline{U}(z)$ also have access to a technology that allows them to frictionlessly purchase at a constant price $p_e(z)$ defined such that $\bar{U}(z) = \frac{(\sigma-1)^{\sigma-1}}{\sigma^{\sigma}} z^{\sigma} p_e(z)^{1-\sigma}/(1-\delta)$, i.e., the constant price that gives the same utility as they are currently earning. This simplifies the model as they can then combine orders from this technology with orders from the new supplier without us having to take a stand on how or when they might renegotiate their existing dynamic contract. Thus, the baseline is defined by firms with $z < \overline{z}$ having existing price $p_e = p_l$, while firms with $z > \overline{z}$ have existing price $p_e(z) < p_l$. Using the fact that the model is homogeneous of degree $1 - \sigma$ in p_e and c, we can then evaluate the interpolated objects from the grid of solved model runs.

For the theoretical moments, we calculate the theoretical profit in each treatment group, with the treatments implemented as described in the main text. For profit, we calculate the mean profit in the first three periods, $\frac{1}{3}\sum_{t=0}^{2} y_t$, as the survey asked about profits over the past 30 days and typically took place within 2-3 months after recruitment. For the binary outcome relationships, we set this to 1 if they already have an existing foreign supplier (i.e., firms with $z > \bar{z}$) and 0 if they have $z < \bar{z}$ and they choose $q_{f0}(c_{\min}) = 0$, where c_{\min} is the minimum of the three free cost draws from the search treatment (or, equivalently, the maximum of the three productivity draws, $\psi = 1/c$). For firms with $q_{f0} > 0$ and $z < \bar{z}$, we set the moment equal to the probability that the relationship reaches the third period (as a function of μ_0). Finally, for mobile money, we calculate $\sum_{t=3}^{18} \tau_t$, reflecting the total value ordered after mystery shopping finished (and, as discussed above, we divide the treatment means by the control–i.e., search only–mean). The final moments are shown in Figure A9.

³²We experimented with different numbers of draws, and found that the value of the objective function becomes fairly stable across seeds starting at around Ω of 5,000.