Female Entrepreneurship and Trust in the Market

Nava Ashraf   Alexia Delfino

Edward Glaeser   Kim Sarnoff*

Abstract

Commerce requires trust, but trust is difficult when one group can expropriate another due to differences in power. This can lead the weaker group to self-segregate into industries and activities; female-led businesses, for example, tend to be small and clustered in a small number of industries where collaborators are also female. We present a model which relates this economic segregation to rule of law, and predicts that female trust depends on the protective preferences of adjudicators in weak rule of law environments. We then show that effective dispute resolution in Lusaka, Zambia, especially as administered by "market chiefs," enables trusting behavior by female entrepreneurs, both in cross-section correlations and in two artefactual field experiments. Such trust generates increased economic returns. We find considerable heterogeneity across market chiefs in their preferences for protecting more vulnerable women.

1 Introduction

Collaboration is fundamental to economic life, but a central theme of the long and distinguished literature on repeated games is that trust depends on punishment (Abreu, 1986, 1988; Fundenberg and Maskin, 1990; Friedman, 1971; Dal Bó and Fréchette, 2018). People become vulnerable to hold-up and expropriation in joint undertakings (Hart, 2009), such as business partnerships and scientific coauthorships. Without protective institutions "the strong contract with the strong, the weak with the weak, and gains from trade between parties of different legal strengths are not realized" (Behrer et al., 2021). When actors are strangers or mobile, power asymmetries loom larger—particularly in settings where enforcement gaps give stronger parties greater leverage (Grossman and Hart, 1983; Hart and Moore, 1988). In this paper, we study how power asymmetries rooted in gender shape the risks of expropriation in collaborative activities.[1]

Our empirical work focuses on female entrepreneurship in Zambia, but the framework in this paper applies to many male-dominated domains, including technology firms, STEM academic fields and large American corporations in the 1960s and 1970s. We first present a model that details the conditions under which the more vulnerable avoid engagement with the more powerful, and how institutional intermediaries can either exacerbate or overcome this economic segregation. We then turn to the microdata we collected across entrepreneurs in Lusaka, Zambia, and explore the central implication of our model: that effective dispute adjudication can particularly benefit the weak, allowing them to trade with the strong, to the economic benefit of all. We also design and implement two artefactual field experiments (List and Metcalfe, 2014; List and Rasul, 2011) with a large set of entrepreneurs, which reveal that customary institutions to resolve disputes can lead women to trust more, engage more in risky collaboration and earn higher returns. We study the institution of the market chief, who delivers a personalized form of justice that depends on the idiosyncratic tastes of the arbitrator- either exacerbating or overcoming existing power asymmetries.

Collaborating with more powerful partners creates risks. One example of these risks is the prevalence of sexual harassment in male-dominated workplaces: in Sweden, 13% of women report having experienced harassment in the past year, a figure that rises to 25% in the most male-dominated jobs (Folke and Rickne, 2022). Moreover, male perpetrators tend to face weaker economic consequences when they hold more power than their female victims (Adams-Prassl et al., 2023). In lower-income contexts, these risks may be even more pronounced, as unequal social norms and insecure rights—both in the household and the marketplace—further heighten women's vulnerability (Lim et al., 2018). These dynamics may help explain why women consistently report lower levels of trust than men.[2] In our own data, the most common reason entrepreneurs cite for why women decline potentially profitable collaborations with men is fear of being being exploited or mistreated.

We begin by providing facts about entrepreneurship around the world, which motivate our model and our empirical work in Zambia. Business ownership beyond small scale self-employment remains a male-dominated activity in almost every country in the world (OECD, 2012; Halim, 2020). Female entrepreneurs cluster in a small number of less profitable female-dominated industries (Nordman et al., 2011; Cirera and Qasim, 2014; Fairlie et al., 2017;

---

[1] Asymmetries of power can arise from differences in social power and/or physical strength. Despite the change in the relative importance of physical strength in economic activity, women remain vulnerable to expropriation and violence, particularly in male-dominated fields (Wilson and Herrnstein, 1985). Sviatschi and Trako (2024), Eswaran and Malhotra (2011) and Aizer (2010) provide research documenting gender-based violence and examining policies to reduce it.

[2] Trust is particularly low among women in weak rule-of-law countries. For example, in Africa and South America, typically between ten and fifteen percent of respondents say that most people can be trusted, but that number falls on average by 6 percent among women (World Value Survey, 2004–2014).

Alibhai et al., 2017; Singer et al., 2018; Berge and Pires, 2020; Campos et al., 2019; Essers et al., 2021; Hardy and Kagy, 2020), such as apparel and food production.This segregation may explain why women appear to receive lower benefits from loans (De Mel et al., 2008, 2009) and business training (De Mel et al., 2014; Berge et al., 2015), especially when norms are more restrictive (Field et al., 2010; Jayachandran, 2020).

We then present a model in which the under-representation of more vulnerable groups in economic activity persists because fear of expropriation leads women to avoid male-dominated fields. Such missed collaboration harms the vulnerable, but also creates an efficiency loss for society as a whole. Appropriate contract enforcement, particularly with trusted dispute adjudicators, can mitigate the losses that come from missed collaboration. This implication differentiates our model from other theories of gender segregation, such as discriminatory tastes or gender homophily in social networks.[3]

In our model, women choose whether to become entrepreneurs and whether to partner with men. If their partner cheats, they have access to an adjudicator, who resembles the institution we study in the Zambian setting: the urban market chief. Adjudicators in our model, like those chiefs, may disproportionately favor their powerful friends or they may see their mission as protecting the weak. Alternatively, entrepreneurs can rely on their own "social power" to punish misbehavior, perhaps by harming a cheater's reputation. In strong rule of law environments, adjudicators regularly intervene and adjudicator preference do not matter. In weak rule of law environments, adjudicators' taste end up determining the likelihood they will intervene and if the weak get protection.

Evidence from the World Justice Project suggests that courts generally favor men[4]. However, there is a paternalistic strain that often appears in traditional leaders and that can protect women. When adjudicators favor the powerful, then women will not partner with strong men.In our treatment of that case, we extend Behrer et al.'s (2021) result that when adjudicators are weak, the strong trade only with the strong and the weak trade only with the weak.

The model suggests that adjudicator preferences can be assessed by knowing whether women want their gender to be revealed during an adjudication process. In our model, unsurprisingly, women prefer to reveal their gender when adjudicators favor the less powerful. We consequently expect to see more partnership between women and men in markets where women prefer to reveal their gender to the adjudicating chief.[5] The model's key implication is that better contract-enforcement institutions can overcome female vulnerability in entrepreneurship, particularly in male-dominated fields. Consequently, our model predicts more gender segregation in weak rule of law environments. Better adjudication increases the returns to female entrepreneurship since anarchy tends to reinforce power imbalances that disadvantage women.[6]

---

[3]We neither dispute that these forces are powerful, nor do we claim that fear of expropriation is more important than these forces in explaining either gender segregation across industries or the prevalence of within-gender collaboration. Moreover, the link between fear of expropriation and discriminatory tastes and gender homophily in networks is likely to be complex. For example, male discrimination may provide men with more "social power" than women and that will make male expropriation more common. Fear of expropriation may also play a role in explaining gender homophily within networks.

[4]All over the world, women also face disadvantages in accessing fair legal resolution. For instance, 28% of respondents in the World Justice Project (http://data.worldjusticeproject.org/) say that being a woman represents a disadvantage with the local police. This share goes from a high of 48%, 33% and 32% in Latin America, MENA and Africa to a low of 14% in Europe and North America.

[5]These results relate to the pathbreaking work of Goldin and Rouse (2000), who document that shrouding women's gender is useful in some, but not all, parts of the audition process.

[6]Our framework can account for the rich empirical evidence on inter-gender relationships in group work and scientific production (Boschini and Sjögren, 2007; Lissoni et al., 2013; Sarsons et al., 2021; Coffman, 2014; Aman-Rana et al., 2021; Shan, 2021). Previous work has highlighted that men (are able to) get more credit than women for joint work and that women's performance worsens in mixed-gender settings, implying that

We test our model using micro-data that we collected - the Census of Urban Entrepreneurs in Lusaka, Zambia - which we overlay with artefactual field experiments embedded in the city's urban markets. Zambia is a country with both weak rule of law and discriminatory gender norms.[7] Yet it is also a country in which chiefs often protect the weak. We test whether women partner more when they have a chief who favors the vulnerable.

To empirically study power, gender, and collaboration in business, we collect geocoded data on more than 2000 firms, which represents sixty percent of all the small-scale manufacturers in Lusaka. Our qualitative work suggests that economies of scale can generate large returns to collaboration for these entrepreneurs, in activities such as lending machines to each other or subcontracting workers (as also suggested by Bassi et al. (2022) and Hardy et al. (2024)).[8] Being located in dense urban marketplaces—particularly in a rapidly growing city—expands opportunities for collaboration with strangers, but also heightens the risks associated with such interactions.

In our sample, twenty-seven percent of the entrepreneurs in manufacturing are women, and women earn slightly more than one-half of male earnings. In Lusaka, three-fourths of female entrepreneurs make apparel and eighteen percent make food. Between one-half and three-fourths of the gender earnings gap for Lusaka entrepreneurs can be explained, in an accounting sense, by the clustering of female entrepreneurs into the least profitable industries. While many factors contribute to the segregation of women in these industries, in our qualitative work, Lusaka's female entrepreneurs repeatedly emphasized the difficulties of trusting men.[9]

In Section 4, we present our survey measures of trusting behavior, such as working collaboratively to fill an order or jointly buying inputs or even giving advice. As the model predicts and our qualitative work suggests, women are less likely to take actions that require trust. Perhaps most strikingly, we find that women are much less likely to learn their trade from incumbent workers. Instead, they turn more often to formal educational institutions.[10]

We then test whether female entrepreneurs in Lusaka trust and collaborate more when contract enforcement is stronger. In this context, three institutions can potentially adjudicate commercial disputes: the Small Claims Court (SCC), the police, and local "Market Chiefs" who exercise authority within their local market areas. In practice, very few entrepreneurs in our sample know about the existence of the SCC or view it as accessible. The police are more commonly known, but still rarely used because of negative perceptions. Instead, most disputes are resolved informally through Market Chiefs, and we therefore focus our analysis on this institution. We measure institutional strength by location within a market that is led by a chief. We find that female-led businesses located inside formal markets cooperate more, even controlling for business density, marketplace and industry fixed effects, and individual characteristics. The gender gap in sales is also smaller within markets.

Building on this evidence, we test our model's prediction that women will disproportion-

---

women will seek to work with other women (either through the choice of fields or co-workers). We propose an overarching framework and a mechanism through which gender gaps may be mitigated.

[7]According to the World Justice Project, Zambia ranks among the bottom 40 countries globally for rule of law, placing it around the regional average for Sub-Saharan Africa.

[8]Bassi et al. (2022), for example, show that inter-firm cooperation is key to achieving scale in small manufacturing in Uganda. While their data do not focus on gender, the sectors they study — carpentry, grain milling, metalwork — are likely male-dominated. Our findings suggest that in male-dominated sectors, female entrepreneurs may be less able to enter or scale through partnerships, not because of skill or capital constraints, but because collaboration entails higher enforcement risk.

[9]Lusaka's female entrepreneurs both say that they trust others less, and indeed are less trusting in standard laboratory measures.

[10]Alfred Marshall (1890) emphasized that in dense clusters, "the mysteries of the trade become no mystery but are, as it were, in the air," but it seems as if female entrepreneurs cannot access these human capital spillovers in Lusaka, partially because they cannot trust the men that surround them.

ately benefit from access to adjudicators who will protect them from expropriation. This implies that the average impact of institutions may mask substantial heterogeneity depending on adjudicators' preferences, and that women should benefit the most when institutions are unbiased or more likely to protect the weak. First, we collect individual measures of gender bias, legal competence and managerial duties among Lusaka's market chiefs. Consistent with the model, we show that women inside markets with unbiased chiefs cooperate as much as men do, while a large gender gap in cooperation persists in markets with a chief characterized by high gender bias.[11]

We then move beyond observational evidence and conduct two artefactual field experiments with our population of entrepreneurs. We use an adapted version of the trust game (Berg et al., 1995), framed as an opportunity to invest in another person's business, and involve real adjudicators in the game. The two experiments allow us to ask whether institutional access enables women to trust in business interactions depending on how adjudicators use their authority. In Experiment 1, we examine the effect of access to a gender-blind chief on trust and cooperation, benchmarking it against two alternatives: the absence of any enforcement mechanism and access to a gender-blind formal legal institution. In Experiment 2, we allow chiefs to observe the participants' gender, reintroducing the possibility of personalized enforcement. Together, the experiments explore two institutional paths to inclusion: one through formal neutrality, and one through contextualized protection. We combine behavioral outcomes, survey measures of adjudicator attitudes and participants' incentivized expectations of adjudicator's behavior to understand which institutional features most effectively support inter-gender economic collaboration.

In the first experiment, the game replicates our survey results and echoes the model: the control group shows a significant gender gap in trust, with women sending fewer tokens than men in our game. Sending money in the trust game is positively correlated with our actual measures of cooperation by the participants in their real lives. In the treatment group, providing access to a gender-blind chief significantly increases women's trusting behavior, increasing the surplus for both parties and earnings for both women and men. Given participants' limited awareness and connection with the the formal legal system, being randomized to get adjudication by a gender-blind court does not achieve the same result for women. An indigenous institution appears far more effective than a legal court initially transplanted from the European system.

Yet while formal courts and the Common Law are meant to reduce the importance of the preferences of any one actor, market chiefs are human and their tastes may influence their judgments. Market chiefs often know the individuals they serve and may factor social identities, including gender, into their decisions. This fact leads to two more questions: how much do chief's preferences impact outcomes and do their preferences typically help or harm women?

In our second experiment, we test whether revealing subjects' gender to a chief improves or worsens female trust and outcomes. We find that when an investor's gender is revealed to their own market chief, women send more compared to when the chief is gender-blind. This effect — the revelation trust premium — is larger for women who express a preference for gender revelation and in markets where chiefs display more favorable behavior toward women in the game.

But we find significant variation across markets. We combine both experimental behavioral data with survey measures to identify chiefs who are gender biased, and also perceived to

---

[11]The chief's level of gender bias is not correlated with the share of female manufacturers within a market (correlation coefficient = -0.05, $p = 0.66$), with female turnover or observable characteristics, indicating that we are not simply capturing a selection effect of women-friendly market chiefs being elected in markets with more - or different - female businesses.

be so by participants in their markets. We find that women are less likely to want their gender to be revealed to such chiefs, and the trust premium associated with revealing gender to the adjudicator declines. This fact highlights the downside of informal adjudication: outcomes are dependent on the preferences of one person and these preferences will vary. Sometimes, to be effective, justice must be personal but personal justice is inevitably idiosyncratic.

Together, the two experiments illustrate a broader point: when formal legal systems are weak or inaccessible, local institutions can fill the gap. In their gender-blind form, they can mitigate power asymmetries and facilitate trust. But when designed—or perceived—to protect the vulnerable, identity-aware institutions can go even further, enabling high-risk, high-return collaboration for those most in need of institutional support. Indeed, we find that in markets where women have greater social power, women behave in the experiments much like men do- and chiefs are correspondingly less likely to protect them as they are not in need of protection (Gneezy et al., 2009).

The results of our experiments reveal that local informal institutions can play a vital role in fostering economic collaboration—particularly when they act as protectors of the weak. Women, who face greater risks in mixed-gender business interactions, are not simply seeking neutrality—they are, in some cases, seeking allies. Chiefs, though embedded in local norms, can function as benevolent adjudicators, and women respond strategically to this possibility. In contrast to the conventional wisdom that anonymity protects the vulnerable, our findings show that being visible to an unbiased institution can itself be empowering.

This paper connects to several literatures. By positing that the risks associated with incomplete contracts and imbalances in social power can deter women from engaging in potentially profitable activities, it relates to research on occupational gender segregation and female vulnerability in predominantly male sectors (Cortes and Pan, 2018; Folke and Rickne, 2022; Adams-Prassl et al., 2023; Ductor and Prummer, 2024). Our findings particularly follow the seminal work of Udry (1996) on intra-household ineffies, as we find that women's lower engagement in high-return activities may reflect rational responses to risks of expropriation or exploitation. Our work also relates to research on the role of courts in shaping firm outcomes (Johnson et al., 2002; Ponticelli and Alencar, 2016; Chemin, 2020; Aberra and Chemin, 2021; Sadka et al., 2024; Rao, 2024; Cahuc et al., 2024) and on differences in the use of formal institutions by social standing (Sandefur and Siddiqi, 2013; Britto et al., 2025).[12]

Our paper suggests that fear of expropriation in environments characterized by asymmetric power may help explain the persistent gender gaps in entrepreneurship and industrial segregation by gender in low-income countries (McKenzie and Woodruff, 2017; Campos et al., 2019; Jayachandran, 2020; Ubfal, 2024). Hardy and Kagy (2018, 2020) show that women-owned businesses in Ghana tend to cluster in highly competitive, low-return sectors. Our paper provides a potential explanation: if women disproportionately avoid industries or occupations where contract enforcement is critical, they may be left with a narrower—and more crowded—set of viable opportunities. We do not want to claim that the strategic avoidance of collaboration or growth explains these gaps more than talents or tastes, but we hope that future work can explore the relative importance of all of these factors.

The paper proceeds as follows. Section 2 provides some motivating facts about female entrepreneurship around the world. Section 3 presents our theoretical model. Section 4 describes our Zambian empirical setting and data. Section 5 presents evidence on gender gaps in entrepreneurship among small-scale entrepreneurs and the potential mitigating role of local adjudicators. Section 6 presents experimental variation in institutional quality, which allows us to show the causal impact of access to contract enforcement mechanisms on female business collaboration, and how informal justice can work to protect weaker groups. Section

---

[12]See also Marx et al. (2019) and Balán et al. (2022) on the role of local intermediaries in shaping outcomes for citizens and local governments in urban areas of low-income countries.

6.3 includes a discussion on external validity and the conditions needed for our framework to apply in other settings. Section 7 concludes.

2 Motivating Facts about Female Entrepreneurship Around the World

Global data on female entrepreneurship point to three recurring patterns: women are underrepresented among entrepreneurs, face large earnings gaps, and are highly concentrated in a few industries. These facts motivate our theoretical model and empirical work.

Globally, only one in three small, medium, or large businesses is owned by a woman (Halim, 2020). In high-income countries such as the U.K., France, Germany, and the U.S., more than two men select into entrepreneurship for every nascent female entrepreneur (Singer et al., 2018; Fairlie et al., 2017). In developing countries, gender gaps in entrepreneurship are of similar magnitude but display greater heterogeneity—particularly across business size, sector, and geographic region. While women are overrepresented in small-scale, informal self-employment, they remain significantly underrepresented among owners of larger or more formal businesses (Calderon et al., 2017). For instance, female ownership of formal businesses ranges from just 18% in South Asia to around 50% in Latin America and the Caribbean (Halim, 2020). Limited female entrepreneurship may not only reflect underused talent, but also carries significant aggregate costs (Chiplunkar and Goldberg, 2024).

We illustrate these patterns using three World Bank datasets that capture different business segments: the Regular Enterprise Survey (WBES), which covers formal firms with more than five employees; the Micro-Enterprise Survey, focused on smaller firms with five or fewer employees; and the Informal Enterprise Survey, which targets unregistered businesses. We define a firm as female-owned if women hold a majority or full ownership stake.[13]

Women own just over 10% of formal firms in the median WBES country, rising to about 30% among micro and informal firms (Figure 2). The rates of female entrepreneurship differ substantially across countries, although very few have a clear majority of firms led by women. Some of the lowest rates are found in North Africa and the Middle East, perhaps reflecting an important influence of religion (Ashraf et al., 2025).

Among formal firms, women's sales are 60.5% lower than men's on average, narrowing to 51.3% after controlling for industry, country, and year. As seen in Figure C.2, the distributions differ especially in their right tail, suggesting that women hit a ceiling to their earnings. The sales gender gap is narrower among smaller and informal firms, suggesting that structural barriers—like limited access to capital or networks—become more binding as businesses grow (Campos and Gassier, 2017; Jayachandran, 2020; Ashraf et al., 2025).

Strikingly, female entrepreneurship is highly concentrated in a few lower-return industries—a pattern observed in several countries at different levels of development (Campos et al., 2014; Alibhai et al., 2017; Goldstein et al., 2019; Hardy and Kagy, 2020). The regular WBES data show that the highest shares of female owners are in retail, food and beverage manufacturing, and hospitality. While women own only 14% of manufacturing firms overall, they account for 21% in food and apparel. The cross-country average of the

---

[13]Firms with equal male-female ownership are excluded. Hardy et al. (2022) compare gender shares in entrepreneurship across World Bank enterprise datasets and nationally representative household surveys that include non-farm enterprise modules. They find that female shares are systematically lower in the World Bank data—largely because the World Bank surveys target more formal businesses or those with a physical structure (i.e., those operated outside the household). Since our study focuses on small-scale urban businesses in Lusaka—most of which operate outside the household—these enterprise surveys provide a relevant benchmark for our setting.

Herfindahl–Hirschman index (HHI) of industrial concentration is 0.13 for female-led businesses, significantly greater than the 0.04 average HHI for male-led businesses (p=0). In the micro-enterprise survey, over 56% of female-led businesses are in textiles, hospitality, retail, or food production—compared to just 40% of male-led firms.

The industries women enter tend to offer greater exposure to other women, as peers, customers, or employees. Female-led firms employ twice as many women on average as male-led firms (50% vs. 28%) and are far more likely to have a female top manager (60% vs. 7%). This gender segregation has been explained by preferences over fields of work and flexibility that certain industries offer, often shaped by social norms (Goldin and Katz, 2011; Campos and Gassier, 2017). In this paper, we explore an additional explanation: working in female-led industries shield female entrepreneurs from workplace interactions with men, which can make them vulnerable to expropriation.

3 A Model of Female Entrepreneurship, Gendered Power and Contract Enforcement

Urban density can enable cooperation by reducing distances, but that advantage can be lost if distrust erodes collaboration. We now present a model where female entrepreneurs undertake collaborative tasks with partners who may cheat or expropriate them. Our goal is to explore the link between cooperation and protective institutions, and to understand the determinants of female entrepreneurship more broadly.

In our model, gender interacts with collaboration because it impacts the ability to punish. As in many settings, in our model, mistreated individuals have both private and public tools to punish an offender. If women have less social power, then they may have less ability to punish privately, by using violence or verbal abuse or spreading remarks that can damage a reputation. "Public punishment" can mean a lawsuit or going to the police or, in our setting, seeking help from a market "chief." That chief may discriminate against women, or paternalistically intervene to protect the weak, which could benefit women.

Formal judicial systems are often seen as less accessible to women and to the poor. The World Justice Project (2023) reports that "in nearly 90 percent of countries, people living in poverty tend to face greater barriers to justice than those not living in poverty," and "women face greater barriers to justice than men in nearly 70 percent of the countries surveyed." However, courts are less relevant in our setting. Globally, The World Justice Project (2024) ranked Zambia 119th out of 142 countries in whether "people can access and afford civil justice." Market chiefs, rather than local magistrates, are the natural source of justice. This fact does not make the Zambian case unique or even unusual. Informal sources of justice appear everywhere from the ombudspeople of professional associations to empowered religious leaders in many immigrant communities.

Market chiefs are a local institution that have taken on much of the role of a traditional and widespread rural institution (the "chief"), adapted to an urban setting.[14] While "Zambian chiefs have few legally recognized powers," they "have unofficial courts in which they adjudicate cases" (Baldwin, 2018). Palagashvili (2018) writes that "the chief is also called upon to resolve legal disputes regarding violations of contracts," even though "contracts in indigenous societies were not written, but were formed on the basis of social norms or customs." Logan (2013) uses Afrobarometer data and reports that "the influence of traditional leaders is most evident with respect to solving local disputes, where they are reported to play

---

[14]The most likely origin of the role of "market chiefs" is in traders' associations, and the term "chief" is a colloquial reference to market chairmen rather than a formal link to rural chieftaincy. While— to the best of our knowledge—there is no official connection to traditional chiefs, both serve similar coordination and dispute-resolution roles in their respective settings.

a greater role than either local or central government" and that in 16 out of 19 countries (including Zambia), a majority of respondents "think that the amount of influence traditional leaders have in governing your local community should increase."

We now turn to the model. In stage 0, a female entrepreneur "E" chooses whether to pay a fixed cost of f_i to enter industry i. She knows the share of co-workers who are men (m_i) and the average character of adjudicators within the industry. If "E" does not join the industry, the model ends.

If E enters, then in stage 1 she is offered a business opportunity to make and sell a product of value π. E is also matched with a randomly chosen potential partner "P". Neither E nor P can make the product on their own at a cost less than π. If the parties do not partner, E can either pay π to produce the good or do nothing. In either case, her payoff is zero. If the parties do partner, E can make P a take-it-or-leave-it offer to collaborate.[15] Like E, P's opportunity cost of time is 0, and so the contract must just compensate P for production costs.

In stage 2, P chooses whether to work or shirk. P can work and fulfill the contract by paying a cost of $q < .5\pi$ for effort and materials. P can also breach the contract and pay only $q - b$. This breach might take the form of P doing shoddy work or of P stealing E's inputs (in which case b may be greater than q). We assume that $.5\pi > b - q$. E incurs her own production cost, which also equals q, and works during this period.[16]

In stage 3, E pays P the contracted wage for his or her work. This payment is not contingent upon observing whether P has shirked. This assumption can be justified either because P needs the money *ex ante* to perform the task or because E doesn't immediately observe the shirking. We make this payment non-contingent both to match the experiments discussed later, and because the ability to make contingent payments would complicate the model and the role of the adjudicator.

If P has fulfilled the contract, then E can sell the good and receive π. If P breached the contract, then E must pay a remediation cost of $b + \Delta$ in order to deliver the product. The value of b and Δ are both known at the time of the contract and $\Delta > 0$. If E remediates the harm, then she receives a payment of π. We assume that $2q + \Delta > \pi > b + \Delta$, which ensures both that E will always remediate if work has begun, and that if shirking always occurs, a partnership will not generate a positive social surplus and will not occur.

We have chosen an "entrepreneurship" frame for this model, because that captures the examples that initially motivated our work, but this model is also equivalent to a standard investment frame (which maps onto the trust game used in our experiments). In that frame, E decides whether to invest $2q$ in P in period 1, some of which is meant to be P's salary. In period 2, P's investment increases in value to π. In period 3, the investment is meant to return to E, but P can decide to keep "b" of value for himself, which reduces the value of the investment to E by $b + \Delta$. While the two frames differ because E pays P in period 1 in the investment frame but in period 3 in the entrepreneurship frame, as long as the period 3 payments are unconditional, they are functionally equivalent to period 1 payments.

If P shirks, then E has two potential means of punishment. E can punish P with private "social punishment" and E can appeal to a "public" adjudicator, who may be willing to intervene on behalf of the injured party. There is no double jeopardy, so that if the adjudicator intervenes, there is no possibility of applying social punishment. The size of the social punishment that a person can inflict and the probability that they have access to the adjudicator are functions of the "power gap" between the injurer and the injured party. We

---

[15]Gender-based power asymmetries may also shape the ex-ante bargaining stage. However, since our empirical setting does not provide variation at this stage, we abstract from it in the current model. A previous version shown in Ashraf et al. (2019) includes gender differences in bargaining power.

[16]It is never optimal for E to shirk as she is the residual claimant of the product.

let θ denote the difference between the power of P and E. Social punishment costs zero to impose and applies a penalty of $\max[k - \theta, 0]$, independent of the harm done. The adjudicator will intervene with probability $\delta + \alpha\theta$, where α reflects the "protective" instincts of the adjudicator. If the adjudicator intervenes, then the adjudicator will force P to pay E the damages of $b + \Delta$. Consequently, E will always prefer the adjudicator's intervention to social punishment. We will typically assume that θ is known, at least by E and P, but the model would be unchanged if we treat θ as the expected power difference. While we will typically think of settings in which a male P has more power than a female E ($\theta > 0$), it is possible for θ to be either positive or negative.

If the parameter α is negative, then the adjudicator favors the mighty and that seems to be true in much of the world. The World Justice Project reports that access to justice is systematically more limited for low-income groups. Chanock (1998) reports that in colonial Zambia, chiefs often seemed to be biased against women: "The decisions of the chiefs were usually in favor of the man in a marital dispute," and "women, especially in the mixed Chewa/Ngoni areas, and Lomwe migrant, who distrusted the Yao headmen under whom they had settled, used the courts," that had been set up by the British.

Yet there is also evidence that α is positive for at least some African chiefs. Logan (2013) shows, through modern Afrobarometer data, that support for chiefs declines sharply with education, and that "the lack of a gender difference is especially noticeable," which seems to reject that view that the chiefs particularly favor either elites or men. Matsvayi (2012) studies three chiefs' courts in Zimbabwe and finds that two of the courts illustrate "how we can build our traditional systems into a model that ensures many more women, especially rural women, access justice." Baldwin (2016) reports that in Zambia, "40 percent of the chiefs' visitors are seeking assistance in resolving disputes, whereas almost one quarter are looking for economic or financial assistance." She also finds "chiefs benefit, rather than harm, local public goods provision," and that "the deaths of chiefs significantly reduce [...] the building of new schools and the sinking of boreholes." The fact that chiefs supply both charity and public services suggests that many chiefs see helping the weak as part of their job.

We let $\hat{\alpha}$ denote the expectation of α by both E and P. If the adjudicator's identity is known, then $\hat{\alpha}$ will equal α. Otherwise, $\hat{\alpha}$ will equal the population average of α across adjudicators. In either case, the expected probability of intervention is $\delta + \hat{\alpha}\theta$. If the adjudicator does not know θ, then his probability of intervention is based on its expectation, which is assumed to equal zero when adjudication is anonymous. Proposition 1 assumes that the power differential is known by the adjudicator and everyone else. Proposition 2 allows for the possibility that the adjudication is anonymous, and in that case the probability of intervention will be δ. We will interpret δ as the overall quality of the legal environment, although its value may also be shaped by the preferences of the adjudicator and the circumstances of the case. Murky evidence will generally make intervention less likely, but some adjudicators may be more tolerant of imperfect knowledge than others.

Proposition 1 describes the returns to partnering as a function of the relative power of E and P (θ), of the the overall quality of the legal environment (δ) and of the adjudicator's propensity to protect the weak (α):

Proposition 1. *E will always partner with P if $k - b > \theta$, but if $k - b < \theta$, then there exists a value of δ denoted δ^*, which is between $-\hat{\alpha}\theta$ and $1 - \hat{\alpha}\theta$, such that E will partner with P if and only if $\delta > \delta^*$. The value of δ^* is always declining with Δ, and increasing with b. The value of δ^* is decreasing with $\hat{\alpha}$ if and only if $\theta > 0$, and increasing with $\hat{\alpha}$ if and only if $\theta < 0$. If $k - b < \theta$ and $k > \theta$, then the value of δ^* is always decreasing with k and increasing with θ if and only if $\frac{\Delta}{(\Delta + b - k + \theta)^2} > \hat{\alpha}$. If $k - b < \theta$ and $k < \theta$, then the value of δ^* is always independent of k and increasing with θ if and only if $0 > \hat{\alpha}$.*

The proposition first notes that when θ, P's power advantage relative to E, is less than $k - b$, then the adjudicator becomes irrelevant. In that case, E is relatively strong and the strong don't need adjudicators. But when E is weak, she can only stop P from shirking (which is necessary for cooperation to occur) when the adjudicator is sufficiently likely to intervene, which occurs when $\delta > \delta^*$. The minimum threshold level for intervention, δ^*, will determine whether a partnership is to take place. We expect to see more cooperation and female entrepreneurship in settings with a lower intervention probability threshold.

The minimum intervention threshold needed to generate cooperation is always falling with Δ, the harm done by misbehavior, since that damage drives the size of the punishment meted out by the adjudicator. As long as $k > \theta$, so that E actually uses private punishment, then k determines the maximum private punishment that E can generate. As k increases, the threat from private punishment increases and the minimum intervention probability threshold falls. The threshold rises with the gains from cheating, b, since a higher value of b make shirking harder to deter. The threshold is falling with α if E is weaker than P and rising with α otherwise, since α determines the response of the adjudicator to strength or weakness. If E is weak, then she wants an adjudicator who is more likely to protect the vulnerable. If E is strong, then she wants an adjudicator who respects strength.

Two of the chiefs discussed by Matsvayi (2012) seem to see their role as protecting women, such as wives who are physically abused by their spouses. That fact suggests a positive α, as we suspect that husbands have a power edge over wives in rural Zimbabwe. By contrast, the colonial chiefs described by Chanock (1998) who favored men and were shunned by female litigants, appear to have a negative α.

If $\alpha < 0$ and the adjudicator favors the strong, then an increase in E's social power increases both her ability to punish outside the court and her likelihood of getting help from the arbitrator. Both factors make it easier to stop P from shirking, which makes partnership more attractive. If $\alpha > 0$, and the adjudicator favors the weak, then an increase in E's social power raises her ability to punish outside the court, but makes the adjudicator less likely to intervene on her behalf. When both private punishment and adjudication are used, these two forces work against each other, and if $\frac{\Delta}{(\Delta + b - k + \theta)^2} < \alpha$, then the threshold actually falls as E becomes weaker, because her weakness increases her support from a sympathetic adjudicator more than it decreases her ability to punish outside the court. If private punishment is not used and $\alpha > 0$, then E's social power is a liability, and the weak are more likely to partner than the strong.

The ratio $\frac{b - k + \theta}{\Delta + b - k + \theta}$ is the ratio of the incentive to cheat if there is no adjudication ($b - k + \theta$) to the increase in punishment size associated with adjudication relative to private punishment ($\Delta + b - k + \theta$). It is negatively related to the benefit associated with a higher probability of adjudicator intervention (δ). The expression $\frac{\Delta}{(\Delta + b - k + \theta)^2}$ captures the derivative of that ratio with respect to θ. The inequality $\frac{\Delta}{(\Delta + b - k + \theta)^2} > \hat{\alpha}$ provides the condition under which a higher θ does more to reduce the relative advantage of adjudication relative to private punishment (which is captured by $\frac{\Delta}{(\Delta + b - k + \theta)^2}$) or more to the probability of adjudication, which is determined by $\hat{\alpha}$.

We now turn to the question of hiding power imbalances, which can be done artificially in a lab setting, and could occur naturally in a court setting where the adjudicator just doesn't know the parties involved.[17] Adjudication can either be anonymous or identified. These differences will correspond to the literally blind or unmasked adjudications in our experiments, and they relate to the demand for masking that we measure. In the lab setting, we also blind the players so that they do not know each other's power.

We assume that E believes that P believes that the cost that E can impose on P with

---

[17]During piloting across Lusaka markets, we learned that it is not uncommon for individuals to send delegates to the market chief in order to mask their identity when lodging a complaint.

private punishment is $k - \hat{\theta}$, and that E also believes that the gap in power is $\hat{\theta}$. In completely blind settings, it would be natural to assume that E has a non-degenerate distribution of possible P beliefs, but we instead assume that E is sure of the social gap and sure of P's belief. Assuming a non-degenerate distribution of beliefs complicates the analysis but does not change our results materially. Moreover, in most real world settings, partners will know each other and consequently have good information about each other's social power.

Proposition 2. *If $k - b < \theta$, and E and P are anonymous so that θ is hidden from the adjudicator, and if P believes that θ equals $\hat{\theta}$ and that E shares that belief, then there exists a value of δ denoted δ^{**}, between $\hat{\theta}$ and 1, such that E will partner with P if and only if $\delta > \delta^{**} = \delta^* + \hat{\alpha}\hat{\theta}$. The value of δ^{**} is declining with Δ and k and increasing with b and $\hat{\theta}$. If $\hat{\theta} > 0$, then E believes that she will be weakly better off if the adjudicator learns the power gap between E and P if and only if $\hat{\alpha} > 0$.*

Proposition 2 shows that hiding one's power is advantageous if you are powerful and the courts protect the weak, or if you are weak and the courts prefer the strong. If the revelation of a female E's gender reveals a power gap that favors P, then she should only want to reveal her gender to an adjudicator who she expects to protect the vulnerable. As long as the adjudicator displays more care for the weak, then revelation will be appealing because it will increase the probability that the adjudicator acts, which makes partnership and profits more likely. If the adjudicator is hostile to women, which we interpret as $\hat{\alpha} < 0$, then E will prefer anonymity. This proposition motivates our examination of heterogeneity across adjudicators in our experiments. We predict that weaker women will want revelation with courts that protect the weak, and that stronger women will not.

We finally turn to E's decision to enter industry or occupation i. We assume that entering into the industry means that E receives an order and is matched with a P. Matching between P and E is random so that the probability of receiving a male potential partner in industry i is m_i. E will enter into the industry if and only if her expected returns is greater than f_i. For our next proposition, we assume that social power is homogenous within genders, and that the difference across genders is denoted θ_G. We assume that $k > b$ so that there is no shirking when partners share the same gender, but that $b > k - \theta_G$, so that benefit of shirking will exceed the private punishment imposed by a female E on a male P, then men will cheat women without the presence of an adjudicator. We also assume that the expected value of α in the industry is $\hat{\alpha}_i$, and that the genders of E and P are common knowledge.

Proposition 3. *If $\delta + \hat{\alpha}_i\theta_G < \frac{b+\theta_G-k}{\Delta+b+\theta_G-k}$, then women in industry i will not partner with men and they will enter the industry if and only if $(1-m_i)(\pi-2q) > f_i$. If $\delta + \hat{\alpha}_i\theta_G > \frac{b+\theta_G-k}{\Delta+b+\theta_G-k}$, then a male P will not shirk, a female E in industry i will partner with a male P and women will enter the industry if and only if $\pi - 2q > f_i$.*

If E's estimate of the probability that the adjudicator will intervene, $\delta + \hat{\alpha}_i\theta_G$, is lower than $\frac{b+\theta_G-k}{\Delta+b+\theta_G-k}$, then women cannot trust men. Consequently, women will only join industries with enough women. The quotient, $\frac{b+\theta_G-k}{\Delta+b+\theta_G-k}$, divides the incentive to cheat with no adjudication by if there was only private punishment $(b - k + \theta_G)$ with the same term augmented with the punishment associated with adjudication $(\Delta + b - k + \theta_G)$.

If $\delta + \hat{\alpha}_i\theta_G$ is high, then men will expect intervention by the adjudicator, and they will not shirk. Consequently, the gender composition of the industry becomes irrelevant. Women will be more likely to enter industries with better protection for women, and when they enter, women will be more likely to work together with men than in industries with less reliable adjudication.

The model doesn't formally explain why some industries or occupations are better paid than others, but it can explain why women may not be present into some better paid occupations. Consider an open economy with a hierarchy of industries, some of which are

inherently more profitable than others. If there was a time in which only men were in the labor market, then those men would choose to work only in the most inherently profitable industries. If women then entered into the labor market, they would avoid industries where men are abundant and adjudication is unreliable. We would then observe a cross-sectional pattern in which men work in more profitable industries and women work in less profitable industries.

The model allows for male misbehavior to do reputational harm, but it does not allow particular men to specialize in providing trustworthy partnerships for women. For those men, the reputation for good behavior could become a valuable asset, effectively increasing the power of private punishment for women. Yet if there are few women in the field, then the incentive to create such a reputation is limited, and expropriation may yield higher returns than reputation-building.

We also do not consider the social power that can be borrowed from male family members or other allies. Our focus group participants frequently discuss being accompanied by their husbands to negotiate for better prices, but they also note that the man's presence can come at the cost of lost control over the business or a share of the profits. We could model this by adding an extra stage to the model where E has the option to access extra private punishment power at a price, which might be a new risk of expropriation.[18]

4 The Zambian Context: Data and Institutions

In this section, we introduce the Zambian context and our main source of entrepreneurship data, the Census of Small-Scale Manufacturers. We then describe the main adjudicating institutions for entrepreneurs' disputes. Zambia is a natural setting to study weak institutions, gender power asymmetries and female entrepreneurship. According to the OECD's 2019 SIGI index (Ferrant et al., 2020), Zambia ranks 99th out of 179 countries in gender discrimination, and 94th out of 132 countries in restrictions on women's physical integrity. Sixty-eight percent of the 2007 World Values Survey respondents in Zambia say that it is justifiable for men to beat their wives in some circumstances, which is the highest share in sub-Saharan Africa. Zambian rule of law is neither particularly good nor particularly bad for sub-Saharan Africa, but it is in the bottom half of countries worldwide. The country's score on the World Justice Project's Rule of Law Index is below Ghana and South Africa, but above Zimbabwe and Nigeria, and about the same as Russia and Mexico. Zambia's overall index of gender equality and index of gender equality in courts from the World Justice Project are among the lowest in Africa. As of 2024, the country ranks 119 over 142 countries in whether people can access and afford civil justice. Not surprisingly, Zambian firms cite institutional constraints as being more important than those related to finance to their expansion and growth (Bloom et al., 2014). Just as in many Sub-Saharan African countries, entrepreneurship is an important activity for many households in urban Zambia. Despite gender discrimination and weak rule of law, our data still documents a sizable number of female entrepreneurs.

4.1 The Lusaka Census of Small-Scale Manufacturers

Between May and September 2016, we collected the Lusaka Census of Urban Entrepreneurs (hereafter, "Census"), which is a spatial mapping of all the firms in Lusaka. For each establishment operating from a fixed location, across all industrial sectors, the Census includes

---

[18]For instance, a focus group participant mentioned: "At the beginning of the business a man took on credit and she had to chase him at the Intercontinental Hotel with someone else. I was scared that he would take advantage of me."

geocoded location, industrial classification at the North-American Industry Classification System (NAICS) 4-digit level, number of employees and structural description (e.g., stand-alone building).[19]

The Census data describe the distribution, size and characteristics of economic activities in a fast-urbanizing environment and supplements the Central Statistical Office's 2012 Economic Census of the whole country. Our Lusaka Census includes a total of 48,163 establishments. As there were 16,063 businesses listed in the Lusaka district in the 2012 Economic Census, our data suggests either a 200 percent growth rate over 4 years or differences in methodology or comprehensiveness. Our data includes far more small firms than the 2012 Economic Census. Ninety percent of the firms in our Census have fewer than 5 employees, six percent have between 5 and 10 employees, and fewer than one percent have 50 or more employees.[20] The spatial distribution of businesses in the Census is shown in Figure C.3.

Table C.1 presents the distribution of businesses across industries at the NAICS 2-digit level and their main characteristics. The largest sectors (by the number of businesses) are retailing, accommodation and food services, and other services (e.g., hairdressers). Retailing firms typically have fewer employees compared to any other industry. The Census also shows that Zambian manufacturing enterprises are smaller– in terms of employment– than those in both neighboring and developed countries (Hsieh and Klenow, 2009; Bloom et al., 2014).

We complemented the Census with a short survey of business owners with less than 20 employees belonging to manufacturing, mining, and construction, which we refer to as the "Census of Small-Scale Manufacturers". This survey had a total of 2,216 respondents, which accounts for 58.3% of the total number of businesses in these sectors.[21] The survey includes questions on business practices, sales and history, levels of trust, collaborative behavior with other businesses, and demographics. We know the gender of the owner for 97% of businesses eligible to participate in the "Census of Small-Scale Manufacturers".[22] We focus on manufacturing because it is traditionally male, offers the possibility of exploiting economies of scale through partnerships, and has been a priority for the Zambian industrial and development strategy for decades. Nevertheless, for external validity, we also include other industries in

---

[19]A business was considered to operate from a fixed location if 1) the business operated from a permanent structure with concrete foundations, 2) the business maintained either stock or machinery on-site overnight, and 3) the permanent structure in question was not exclusively used for storage. We split Lusaka into subregions called Census Supervisory Areas (CSAs), and our census covered all the businesses in 90 percent of all CSAs in the Lusaka district. For security reasons, we excluded 8 CSAs in the region surrounding Chibolya compound (Harry Mwaanga Nkumbula ward). We also excluded areas of low population density in the following wards: Kabulonga, Lubwa, Lilayi, Munkolo, Mwebeshi, Kamulanga, Munali, Roma and Mpulungu.

[20]In the 2012 Economic Census, the percentages were respectively 71, 11 and 3 percent. In total 84 percent of businesses disclosed their employment figures in our Census, so the percentages in the body of the paper refer to about the 40,517 respondents to this question. Some larger businesses were unwilling to share their employee numbers with our census takers.

[21]While collecting the Census, we tried to interview all the business owners with less than 20 employees belonging to manufacturing, mining, and construction (N=3803). Table C.2 shows differences in the characteristics of respondents and non-respondents. Respondents to our Census of Small-Scale Manufacturers are smaller than non-respondents, with about half as many employees. Non-respondents are more often from Construction/Mining, part of a chain, or in brick structures, suggesting they may operate more successful businesses and face higher opportunity costs of time. This selection was also reflected in refusal reasons (40% "too busy," 27% "uncomfortable discussing business"). Controlling for industry, the likelihood to participate in the survey is the same for men and women.

[22]If the owner was not available, the interview was conducted with the main manager. Gender is available also for non-respondents as it was coded by our surveyors when asking about the owner's availability to participate in the survey. However, we do not have data on gender for businesses outside our target industries and business size. Table C.2 shows that selection into the survey is very similar for men and women. Importantly, for women we do not find any differences in the type of business location between respondents or non-respondents, including whether they work from a residential building.

our artefactual field experiments discussed in Section 6.

While we did not intentionally exclude home-based enterprises, our mapping approach likely under-represents businesses operating entirely within private residences without visible signage.[23] These tend to be smaller, more informal, and disproportionately female-owned (Hardy et al., 2022). However, our focus is on firms embedded in market-like environments, where interaction with relative strangers - particularly men -, fear of expropriation and adjudicator's role are most salient. We therefore believe our sample captures the relevant population for the institutional dynamics we study.

4.2 Adjudicating Institutions in Lusaka

There are three main sources of legal support for small-scale firms in Lusaka: the Small Claims Court (SCC), the police and more informal local adjudicators. Few entrepreneurs in our sample heard or made any use of the Small Claims Court. Instead, many of them rely on the police, which is deemed to be quite slow and resource-constrained, or local adjudicators known as "Market Chairmen/Chairwomen" or "Market Chiefs", who exercise authority over transactions that occur within their own local market areas. We describe these institutions in the following paragraphs, and provide a more comprehensive account of the chiefs institution. Figure 3 shows the location of formal markets, the Small Claims Court and police stations.

Market chiefs. There are approximately 80 formal markets in Lusaka, covering 30 percent of firms across all industries and 59 percent (N=1324) of manufacturers. Cooperation appears easier within markets, and somewhat surprisingly, rents also tend to be lower relative to spaces in high-traffic commercial areas, potentially due to restricted access, shorter opening hours, and limited business space.

Markets fall under two broad categories: council and cooperative. Historically, Lusaka's markets were governed by the Markets Act of 1937, which gave local authorities control but delegated day-to-day management to appointed Market Masters or, in practice, party-linked "chairmen" who had significant influence over stall allocation and fee collection. Following political reforms and the Markets and Bus Stations Act of 2007, councils formally regained control, yet trader-elected leadership persisted in cooperative and most council markets. These leaders, typically called market chairmen or chiefs, are elected by market members for multi-year terms (usually 2–5 years) to represent traders' interests, enforce market rules (e.g., fines for prohibited behavior, such as drinking or insulting others), manage finances, and maintain order and sanitation. The chairman's exact functions are guided by legislation and market unit guidelines, but, importantly, all market chairmen are involved in resolving disputes among marketeers and/or customers.[24]

Given their central role in everyday market governance—effectively handling issues ranging from unpaid loans to consumer complaints—these chairmen are widely recognized as local authorities in their own right, despite lacking statutory power in council-run markets after

---

[23]According to the Zambian Labor Force Survey, 18% of business owners (25% among female owners) operate from their household, compared to 15% of businesses located in a residential building in our mapping exercise.

[24]Arrangements may vary. In Council markets, which are regulated by the Lusaka City Council, there are sometimes two figures in charge: a market "master" and a market chairperson (the chief). The former is a civil servant appointed by the Lusaka city council who usually has a short mandate (around one or two years). The market chairperson (the chief) is instead elected by marketeers and owners of market slots to manage the market and represent them where necessary. Moreover, the market chief is the main person in charge of dispute resolutions. See Table D.1 for some descriptive statistics on market chiefs. Gender-biased chiefs are no less likely to be elected than non-biased ones, indicating that the method of selection does not appear to explain variation in chief bias.

2007. Traders and customers alike often prefer to bring disputes before them rather than to the police or courts. As one trader explained, "Police say: 'Bring a record from your market committee first'; that's why we sort things here before courts get involved." This de facto institutional authority has earned market chairmen the colloquial title of "market chiefs." The term reflects how entrenched and respected these figures are: they are first-instance adjudicators, peacekeepers, and gatekeepers of access to market stalls, functioning as an informal but well-established governance structure within Lusaka's commercial ecosystem.

For example, one typical dispute involved a marketeer selling a stall to two different buyers. The chief intervened, ensuring that funds were redirected to compensate one buyer. In another case, a customer claimed a marketeer had failed to repair a television as promised; the chief set a two-day deadline or imposed a fine. As one chair said, "We want to handle disputes internally and peacefully first, only if this does not work do we go to the police." Such examples underscore the institutionalized trust and authority these leaders command.

Small Claims Court. To address the weaknesses of the formal court system, a Small Claims Court was founded in 2008 by an act of Parliament. No legal representation is required, as the court is intended to hear minor cases relating to employment, borrowing and lending, insurance, and fraud. The court only permits cases involving amounts up to 20,000KW (around 1,962 USD), which represents 150% of average sales in a good month in our sample of businesses. In our data, only 21% of respondents had heard of the SCC, and of those, 56% did not believe "people like them" had access.

The police. Alongside formal courts, the police is another common institution where small-scale businesses bring their disputes. Since the late 1990s, the Zambia Police Service has being particularly focused on establishing a connection between its forces and the communities they serve. For instance, in 1999, a Community Services Directorate (CSD) was established with the explicit goal to improve policing within communities and commercial centers. Community Liaison Officers were created to help the vulnerable and disadvantaged people in the community, including women.

A comparison of contract enforcement institutions. Tables C.8 and C.9 compare the three dispute-resolution institutions available to Lusaka entrepreneurs—market chiefs, police, and the Small Claims Court—in terms of knowledge, perceived fairness and speed, and actual usage. Respondents are the entrepreneurs who participated in our artefactual field experiments, described in Sections 6.1 and 6.2.

When presented with hypothetical business disputes, the vast majority of respondents indicated that they would seek help from their market chief: 73 percent of men and 68 percent of women would turn to this local authority (Table C.8, Panel A). By contrast, only 17 percent reported that they would approach the police, and fewer than 10 percent would use the Small Claims Court. The market chief also has the strongest reputation among the three institutions: 55 percent of men and 50 percent of women consider their chief to be always or usually fair, while only 18 percent of men and 24 percent of women perceive the chief as always or usually slow (Table C.8, Panel B). The police are seen as less fair than both chiefs and courts, and both the police and the SCC are perceived as slower than chiefs. Men and women are equally pessimistic about the fairness of police on average, but women are more likely than men to think that either the market chief or the police are slow in dispute resolution (see Table C.8).[25] A striking pattern in Table C.8 is respondents' limited knowledge of the formal legal system. Only one in five entrepreneurs has ever heard of the Small Claims Court, and just one in four can name any other institution for dispute

---

[25]Anecdotally, delaying dispute resolution is one way in which Market Chiefs try to favor one party.

resolution besides the chief and the police. Among those who can, only half mention "courts."

These patterns are mirrored in actual behavior. As Table C.9 shows, only 21% percent of male and 13% of female entrepreneurs have ever used a court to resolve a dispute (p<0.05 for the gender difference). About half of respondents know someone who has used a court. In practice, however, fewer than one in five business owners seek help from any institution when faced with a dispute. Men are more likely than women to report work-related disputes and are more likely to seek assistance from a third party (see Table C.9). Among those who seek help, both men and women primarily turn to their market chief rather than to formal state institutions. For example, in our second experiment, 76% of women and 50% of men sought assistance from their market chief (N=57).

Taken together, this evidence suggests that market chiefs function as the primary and most trusted adjudicators for Lusaka's small manufacturers. They are well-known, widely used, and perceived as relatively fair and efficient, making them a good empirical proxy for the imperfect third-party enforcers described in our theoretical model. We therefore focus on market chiefs in the main analysis, while presenting complementary results for the other institutions in the Appendix.

5 Gender Gaps in Entrepreneurship and Collaboration

In this section, we use our Census of Small-Scale Manufacturers to describe widespread gender gaps in entrepreneurial activity. While some of these gaps have been documented in other contexts, we show additional facts on gender gaps in trust, business collaboration and peer-to-peer learning. We then show that these gender gaps in economic collaboration are smaller when female entrepreneurs have access to the chief's (unbiased) adjudication.

5.1 Segregation and Earnings

Women-led businesses represent twenty-six percent of the businesses in manufacturing, construction and mining (N=3,723), which is unsurprising given that manufacturing is often a male-dominated activity (Campos et al., 2014; Goldstein et al., 2019). Women's firms have 0.38 fewer full-time employees and 0.7 fewer part-time employees than male-led firms.

Women earn less than men. On average, women's sales value in good weeks is 2,356 Kwacha (KW), as compared to 4,085KW for men, or 180 and 311 dollars in the current exchange rates for women and men respectively (1 KW = 0.076 USD). In bad weeks, women earn on average 599.9KW as compared to 1,313KW for men, or 45 and 100 USD respectively. Figure C.4 shows the kernel densities of log sales in good and bad weeks by gender, where the sales variables have been winsorised at the top 1% to remove the influence of outliers.[26] The distributions are significantly different between men and women (p=0.00, Kolmogorov-Smirnov test). Moreover, men have a higher variance in sales in good weeks than women's (p=0.00, variance ratio test), which is driven primarily by a thicker right tail.

Figure 2 shows that Lusaka women make different sectoral choices than men. Ninety-three percent of women operate in apparel and food manufacturing, while women represent a minority in wood, metal manufacturing and printing. Women appear to select into non-complex industries, despite having on average the same qualifications as men.[27]

---

[26]During piloting, we found that recalling exact sales digits or for periods longer than a week was challenging for most of the respondents. We thus asked the sales in the previous working day, the sales in a good week and the sales in a bad week. If the person could not provide an exact number, we also asked for an upper and lower bound. We then imputed the average of the bounds to the sales variables.

[27]We define non-complex industries as those that have less than the mean number of skilled occupations associated with the corresponding NAICS3 code (following Minondo and Requena-Silvente, 2013). Our index of complexity is correlated with owner's education in the data.

In our survey of manufacturers, industry choice, not observable human capital, explains much of the gender gap in earnings. Male and female entrepreneurs have similar levels of education, as shown in Table C.3. Women are more likely to have participated in management or entrepreneurship training than men and the two groups do not differ in terms of practices such as record keeping.

Column (2) of Table C.4 shows the raw gender gap in sales variables (average, in a good week and a bad week), both in levels and in logs, and employment level and growth. Column (3) shows that controlling for education does not reduce the gender gap in sales. These results are unchanged when including alternative proxies for skills, such as literacy, numeracy, and social skills. Column (4) of Table C.4 adds controls for industry and Column (5) adds household constraints, including marital status and work time. Selection into different industries explains between two-thirds and three-fourths of the gender gap in average sales, in logs and levels respectively. The coefficient on the female dummy for log sales in a good week is reduced from -0.59 to -0.18 when controlling for education and industry, and the R-squared goes from 0.045 to 0.26. The coefficient on log sales in a bad week is similarly reduced by one-half when controlling for education and industry, and the R-squared goes from 0.045 to 0.25. The average level of sales is the lowest in apparel manufacturing– where most women operate– and the highest in food manufacturing. The other sectors lie between these two. Employment growth differences between the two genders are completely explained by the distribution of the two genders across different industries. When controlling for sector, education, working time and marital status, the coefficient on the female dummy becomes insignificant for sales in both levels and logs, as well as for employment variables.[28] One explanation for female industrial segregation may be that, when asymmetries of power are large and institutions are absent or weak, women cannot trust men.

5.2 Trust and Collaboration

Partnerships in Lusaka resemble the world of incomplete contracts described by our theory. Written contracts are used only in one out of five partnerships. Collaborations are mostly triggered by time constraints (e.g., in fulfilling big orders in a short time frame), which limits entrepreneurs' ability to search for or repeat partners. In seventy-five percent of cases, partnerships form between entrepreneurs that are located close to each other. We first examine whether female entrepreneurs are less trusting and form fewer partnerships than their male counterparts. We also discuss our limited information on the transfer of knowledge across Lusakan entrepreneurs.

We use standard survey measures of trust taken from the World Values Survey (WVS) and General Social Survey (GSS).[29] We couple these general trust questions with questions about business behavior that involves trust, including the formation of business partnerships. We conducted extensive piloting to identify the most common cooperative activities that

---

[28]This evidence is compatible with previous studies also finding that women have lower profits and sales growth even controlling for extensive observable characteristics (Bardasi et al., 2011; Klapper and Parker, 2011; Campos et al., 2014; Nix et al., 2014; Nordman and Vaillant, 2014; McKenzie and Woodruff, 2017; Hardy and Kagy, 2018, 2020). Recent evidence by the World Bank similarly finds that one-quarter of the gender gap in profits in the Democratic Republic of Congo can be explained by industry (Campos et al., 2019). For a recent review on the several constraints faced by female entrepreneurs, see Ubfal (2024).

[29]See tables D.1 and D.2 for validity checks of the survey measures of trust and trustworthiness using our experimental data. The GSS Trust measure is an indicator for the answer to "Do you think that most people can be trusted or you cannot be too careful?." Trust in strangers/neighbors is taken from the WVS questions: "How much do you trust people you meet for the first time / your neighbors?," measured on a 4-points Likert scale. We ask for trust in business as agreement to the statement: "I am comfortable leaving my shop unattended during the day if I need to do something for 30 minutes,"measured on a 5-points Likert scale. The Trust in strangers, neighbors and in business variables were converted into dummy variables.

small-scale manufacturers engage in, across all industries. We then adapted the language used by our piloting participants to create questions on collaboration with other businesses in terms of shared orders, joint buying of materials, lending, and offering advice.[30] In our analyses, we use both indicator variables that take on a value of one if an individual ever engaged in the last 12 months in a particular activity and an index of cooperative behavior from their responses created as the average of the four indicator variables.

Panel A of Table 1 shows the mean differences in trust between men and women. Women are less trusting across all three direct survey measures. Women are also more likely to disagree with the statement that they would be comfortable leaving their shop unattended. Additionally, Panel B of Table 1 shows that women are less likely to engage in any of the four cooperative behaviors that we measure.

Table C.5 shows that low levels of trust are associated with a lower frequency of cooperation among entrepreneurs. These results support the view that the trust questions are capturing something real about the trustworthiness of the environment. Finally, Table C.6 shows that cooperative behavior is also correlated with our three measures of sales. While these correlations do not imply any causal relationship, they are compatible with the view that entrepreneurial activity benefits from the ability to form partnerships with others.

In the previous section, we documented the selection of female entrepreneurs into less profitable industries. Our model suggests that this self-selection occurs because women choose industries with other women, because they are able to trust and partner with those women, either as customers or collaborators. Figure 5 shows that women collaborate more and similarly to men in apparel, which is the only gender-balanced industry in our data. In contrast, women in other industries cooperate less than female tailors and substantially less than men.[31] Figure C.5 replicates this evidence by looking at more general social interactions with other entrepreneurs. Women talk less about the business with other entrepreneurs than men, but that tendency disappears in industries that are not male-dominated, so that women who work in disproportionately female industries have the same frequency of cooperative and social interactions as men in those industries.[32]

Nevertheless, even in more gender balanced industries, we document strong gender segregation in collaborative activities. In our second experimental sample - which includes tailors as well as hairdressers and restaurant owners (as described in section 6) - respondents reported the gender and number of their collaborators from the previous year. Both men and women overwhelmingly worked with same-gender partners, with only about 1 in 5 collaborating across genders. In addition, women also report having collaborated with

---

[30]After each of the following statements, participants were asked whether they had engaged in the described activity with a similar business in the past 12 months:

1. Sometimes two or more businesses participate in a common order from a client, or one business subcontracts to other businesses part of an order. (share order)

2. Sometimes businesses make joint orders of materials from suppliers. (joint buy)

3. Sometimes businesses ask for advice (or give advice) to other firms doing their same activity, for instance on topics like: the production process, the market conditions, new technologies, business practices, suppliers. (advice)

4. Sometimes businesses borrow (or lend) machines, materials or other assets from firms doing their same activity. Sometimes they hire (subcontract) employees who come from other firms doing their same activity for a short period of time. (lending)

[31]In contrast to women, men tend to cooperate slightly less in apparel manufacturing than they do in other industries. This suggests that female tailors' cooperation level is not driven by a higher industry-specific need for collaboration.

[32]In terms of the correlates of cooperation, we find that businesses which cooperate tend to be older, and their owners are more likely to be married and to have received had capital from their spouse, they are less likely to be recent immigrants and are more educated.

fewer partners. These patterns align with the power asymmetries in our model, suggesting that women have smaller networks and may avoid more powerful collaborators. When asked about why a female entrepreneur may turn down an opportunity to collaborate with a male peer, 50% of this sample mentions men mistreating women and/or gendered trust issues as the main reasons for this rejection.

Perhaps the most important form of cooperation occurs when one urbanite shares knowledge with another. In these informal industries, many skills are often passed along from one worker to another. Yet chains of knowledge also require trust. In many cases, a skilled worker who teaches a newcomer expects that student to serve as an apprentice. For centuries, the relationship between mentor and apprentice has been open to abuse. Apprentices, including Benjamin Franklin, run away to avoid promised service. Mentors typically gain the power to punish their students either physically or by harming their reputations, and that power also creates the potential for misuse, especially between a man and a woman.

Panel C of Table 1 shows that male entrepreneurs are far more likely to have been taught their trade by another entrepreneur or a family member. Female entrepreneurs are usually formally trained. Female entrepreneurs are also less likely to have taught others their trade and slightly less likely to talk with others frequently about their business, but this last difference is not statistically different from zero.

Table 2 shows these results using a linear probability model with controls for business density and owner characteristics. Regression (1) shows that women are 18.4 percent less likely to learn their trade from another owner. Regression (2) shows that they are 13.2 percent less likely to learn from another owner in the same industry. Regression (3) shows that they are 15 percent less likely to learn from family and friends, which seems more plausibly related to gender discrimination within the household than trust. The last three regressions in the table show the relationship between the source of the owner's knowledge, sales and employment. Regression (4) shows that in a good week, women have sales that are 0.6 log points lower than men, but this difference is largely reduced if women have learned from another owner or family member. Regression (5) repeats this regression for sales in a bad week and finds similar results. Regressions (6) and (7) repeat the regression using employment and employment growth as the dependent variables. The pattern for employment is similar but the interaction between owner gender and source of skill is not significant. If human capital externalities rely on trust, then women may be unable to access those externalities and benefit fully from working in a dense urban environment.

5.3   Female Collaboration and Chiefs' Support

In this section, we ask whether institutions can mitigate the gender gap in collaboration and earnings among Lusaka manufacturers. The model predicts that female entrepreneurship is more viable when institutions provide them with reliable contract enforcement (i.e., a higher δ or protective adjudicators with $\alpha > 0$). Yet conditional upon entry, stronger institutions alone are sufficient to increase cross-gender collaboration.[33]

Empirical model. Our main measure of institutional strength is having access to the justice offered by the chief of a formal market. We first test whether being able to resort to the chiefs' adjudication increases female collaboration. To this goal, we compare the outcomes of female and male entrepreneurs who are located within a market with their counterparts who

---

[33]Proposition 1 shows that partnerships between an entrepreneur (E) and a partner (P) occur when the expected institutional intervention rate $\delta + \hat{\alpha}\theta$, exceeds the threshold at which shirking becomes profitable for P. As long as the returns to partnership and the returns and costs of cheating are held constant, the main prediction is that improvements in institutional enforcement (higher δ or a more protective adjudicator, i.e., higher $\hat{\alpha}$ when $\theta > 0$) increase the likelihood of cross-gender collaboration.

are located just outside the same market, controlling for marketplace fixed-effects, business density and a variety of entrepreneurs' characteristics. Through the lens of our model, assuming that women are always on average less powerful than men (i.e. $\theta > 0$), a reduction of the gender gap in collaboration in markets indicates that $\delta + \hat{\alpha}\theta$ is on average higher than partners' shirking incentives. But if market chiefs are thoroughly biased, then physical proximity to them will not engender cooperation by women. The simplest way of thinking about the mapping between these measures and our model occurs when women have no ability to enforce contracts against men without these adjudicators, but the likelihood of the chiefs' intervention in support of women increases to some positive number in markets.

Table 3 shows the following regression where cooperative behaviors are regressed on access to the market chief's adjudication and the interaction between this institution and gender. For business i in sector s within 1 km from market m we run the following specification:

$$y_{ism} = \beta_1 Fem_{ism} + \beta_2 InMarket_i + \beta_3 Fem_{ism} * InMarket_i + X'_{ism}\zeta + \lambda_m + \epsilon_{ism}$$

The set of controls X_{ism} includes the logarithm of the total number of businesses within 100 meters of business i, the logarithm of the total number of businesses in the same industry (NAICS3) within 100 meters around business i, a dummy for whether the business is within 100 meters of a market, the population density in the administrative area where the business is located and whether the business has a below-median distance from the police or the SCC ("density controls"). We add marketplace fixed effects λ_m and standard errors are clustered at the marketplace level.[34]

Variable Fem_{ism} is an indicator for a female-owned business and $InMarket_i$ is an indicator variable equal to one if the business i is located inside a formal market, and thus can refer to the market chief for disputes.

Our empirical strategy compares the effect of having access to the chief's support for businesses of different genders that are exposed to similar demand conditions and density effects, controlling for unobservable characteristics of the marketplaces where businesses are located. We control for both population and business density at a granular level as increased business interactions can be directly affected by supply (other businesses) or demand (customers) factors. We add marketplace fixed effects to control for idiosyncratic differences in the quality of institutions across marketplaces. Our coefficient of interest β_3 captures the systematic effect of the chief's support on female owners' outcomes.

Drivers of business location. Our empirical strategy relies on variation in business location inside/outside markets. The identifying assumptions are that location is not correlated with chiefs' quality and it is not differential by gender. To assess the sensibility of these two assumptions, we asked participants in our second experiment why they chose to locate their business within the current market. The great majority of respondents cited business opportunities, available services, or proximity to home—regardless of gender (see Figure C.6). Only a small share selected "quality of market management," suggesting that institutional quality is not a primary driver of location. We also asked why firms chose to be inside their current market rather than just nearby. The main reasons were better services (e.g., waste management, toilets), security, and lower rents, because market stalls are typically cheaper than nearby outside ones. Again, fewer than 10% mentioned management quality. These findings align with our qualitative interviews with entrepreneurs, who noted that locations outside markets offer more customers and longer hours, but at the cost of

---

[34]A marketplace is defined as the set of businesses located inside a formal market or within one kilometer from the market. Each business located outside a formal market is assigned to the closest market within a kilometer. Businesses located further than one kilometer from a formal market (N=103) are assigned to one extra category ("Far from market").

higher rents and fewer basic services. Inside markets, interviewees emphasized the value of amenities. In terms of gender-based selection into markets, Table C.7 shows that the observable differences between women who locate inside versus outside markets closely mirror those for men, suggesting that selection patterns are similar across genders. Consistently, Table C.12 indicates that the average share of women located inside markets is nearly identical to that of men.

Results on collaboration. Panel A of Table 3 shows that being in a market disproportionately increases women's probability of enacting any cooperation activity with other businesses, as well as average cooperation.[35] The dependent variables y_{ism} in columns 2-5 of Panel A of Table 3 are the indicator variables for cooperative behavior: whether the respondent said that they had lent/borrowed capital, given/received advice, participated in a common order from a client, or placed a joint order of materials with another business like their own. The dependent variable in column 1 is their mean.

Being inside a formal market increases women's collaboration by 16 percentage points, which is double the effect of being in a market for men and represents 30% of the mean collaboration levels. This is consistent with the hypothesis that strong market leaders might provide support to women's interactions with other businesses, overcoming their disadvantage in power. Figure 4 shows that the increase in women's cooperation is not driven by markets where women's manufacturers are the majority, but also happens in markets where men's manufacturers are the majoritarian group.

Results on economic performance. Panel B of Table 3 turns to sales and employment, which are our primary measures of economic success. We use the same empirical specification described at the beginning of this section, but we further control for sector s fixed effects given their explanatory power for gender differences in performance. The outcome variables in Columns (2) and (3) are average sales, in levels and logs, while Columns (4) and (5) focus on full-time employment level and growth. Column (1) shows results on an index of economic performance, which averages standardized measures of sales and employment (in levels).[36] On average women sell less than men, but the interaction term is positive across all the columns, suggesting that female businesses particularly benefit from locating within a market and having access to a market chief. The coefficient on the interaction $Fem_{ism} * InMarket_i$ is statistically significant for mean sales in levels but not in logs. Yet, it is large in magnitude and similar to the coefficient on the female dummy. Similarly, while women tend to have smaller businesses that do not grow, the gender gap in both employment levels and change is attenuated inside markets. The coefficient on the market dummy is negative across specifications, and significant for sales outcomes, indicating that male-owned businesses perform worse inside markets than their counterparts outside the same market.

These results suggest that the economic pie in markets is smaller for everyone, but relatively smaller for men with respect to women, with a resulting attenuation of the gender gaps in economic performance uncovered in previous sections. Given the central role of the chief in dispute resolution, these results may reflect women's increased ability to trust and redistribute the pie in their favor when there is a stronger enforcing institution. The worse economic performance of businesses inside markets also suggests that owners may be negatively selected into markets. However, the mitigated gender gap may still be the result

---

[35]The relationship is substantially unchanged when including business owner controls (including: trust, business and owner's age, days worked, marital status, educational achievement and ethnicity) and industry fixed effects. See Table C.11.

[36]As in previous specifications using sales, the variables have been winsorised asymmetrically at the top 1% to control for outliers and reduce noise. Figure C.8 shows the robustness of the coefficients on the interaction term $Fem_{ism} * InMarket_i$ for different cleaning procedures and sales variables.

of the chief's support as long as selection is not differential by gender, as Table C.7 suggests.[37]

Proximity to formal institutions. Given how few small-scale businesses in Lusaka know or use formal institutions, we explore in the Appendix whether simply being located closer to them matters. Figure C.7 shows cooperation rates by gender across distance quintiles from the SCC or a police station. For men, cooperation is fairly stable except in the most remote areas. For women, however, cooperation declines steadily with distance.

Table C.10 confirms this pattern using our main regression model, where we replace the market location dummy with indicators for whether a business is located closer than the median distance to either the Small Claims Court or a police station. Being closer to formal institutions is associated with more cooperation by women, but this relationship is statistically significant only for distance from the SCC (Panel A). However, we do not find evidence that proximity to these institutions differentially affects business performance for women (Table C.10, Panel B). One likely explanation is that the Market Chief remains the primary and most trusted source of dispute resolution—unlike the Small Claims Court, which is largely unknown, or the police, which is seen as slow and ineffective. Since the SCC is located in a central area near other government offices, the positive association between proximity and cooperation could also reflect other factors, such as a higher security from institutional presence or higher opportunities for larger orders, e.g., for public procurement.

Summary. While the findings presented in this section are robust to different controls, unobservable characteristics of entrepreneurs located inside markets may still affect our estimates. First, businesses decide where to locate. If there is endogenous selection of more cooperative female businesses closer to local institutions, our estimates would capture this selection rather than the effect of institutional support per sè. Second, if the marginal female marketeer is better than the marginal male marketeers inside markets, higher cooperation and sale may be attributed to women's better quality instead of enforcement.

Our next sections address these concerns. First, we will show that the market effect on cooperation is heterogeneous depending on the chief's quality and gender-biasedness of arbitration. Then, we will experimentally isolate the impact of having access to chiefs with varying preferences.

5.4 Chief Bias Mediates Female Collaboration

Our model predicts that women – who suffer from a power gap θ – should disproportionately benefit from competent and redistributive adjudicators, who are characterized by a high parameter δ and are sensitive to the gender power gap through $\alpha > 0$. Yet, weak (low δ) or gender biased (α low or negative) adjudicators may even discourage female collaboration above and beyond the effect of power asymmetries. Thus the average impact of institutions uncovered in the previous section may mask substantial heterogeneity depending on the degree of adjudicator's likelihood of acting in disputes and bias.[38]

Measuring chiefs' gender bias. In 2021, we went back to Lusaka's markets and tried to interview all the market chiefs who were in charge of markets during 2016. We found a

---

[37]The effect of being in a formal market on mean sales remains the same when including business owner controls (as defined above) in addition of industry fixed effects. See Table C.11.

[38]When asked about a hypothetical dispute between a man and a woman, market chiefs varied in their adjudications and views. Some voiced pro-women sentiments (*"Most women are vulnerable and men take advantage of that"*), others bias (*"It is easier to deal with men as most women are quarrelsome"*), while some highlighted that gender would not be a factor in their adjudication (*"We rule in favor of one who provides enough evidence"*)

total of 76 respondents, of whom 68% are the original 2016 chiefs. Appendix D contains details on the survey collection.

Chief's sensitivity to gender power gaps (α) is linked to gender bias, as chiefs are less likely to want to protect women if they are biased against them. We create an index of the chief's gender bias, which captures whether the chief is unlikely to help a woman in a business dispute against a man and/or has traditional gender attitudes.

The gender bias index aggregates survey answers about the chief's use of subjective decision rules, whether the chief thinks that a woman is unlikely to win a dispute against a man, whether s/he is more likely to catch a thief for a male rather than a female marketeer, whether s/he holds more traditional gender attitudes related to women's participation in jobs, business or politics (using questions from the World Value Survey, WVS) and whether s/he reports that women - but not ethnic or religious minorities - tend to be disadvantaged with the police or market leaders (using questions from the World Justice Project).[39]

We further construct an index which captures the overall quality of the chief's adjudication process (of "legal competence") and another index for the chief's degree of involvement in market affairs (of "managerial duties"). The former can be thought as a determinant of δ (e.g., how easy is to find evidence to be used in the adjudication so that the chief takes action), thus we expect better legal competence to positively affect business collaboration. We measure the chief's managerial burden as a possible confounder in the relationship between the chief's characteristics and female outcomes.

Do women sort across markets based on chiefs' gender bias? Several empirical exercises provide reassuring evidence that we are not simply capturing a selection effect of more - or different types of - women locating into markets with women-friendly chiefs. First, the chief's level of gender bias is not correlated with the share of female manufacturers within a market (correlation coefficient = -0.05, $p = 0.66$). Second, the observable characteristics of men and women under more or less biased chiefs are similar. Third, women in markets with biased chiefs are not more likely to relocate or exit (Figure C.16). Two factors likely contribute to this lack of patterns between female business (re-)location and chief bias. First, chiefs tend to have shorter tenures (2 to 5 years) compared to the average business age (9 years), reducing incentives to move in response to a specific leader. Second, mobility is low overall, with 61% of businesses remaining in the same location since founding.

Chiefs' gender bias and the gender gap in collaboration. Figure 6 shows the main results on the relationship between the chief's gender bias and female trust in business using the raw data. The gender gap in average cooperation between women and men goes from -10 to -7 percentage points moving from outside markets to markets with a highly biased chief (i.e. with an above-median index). The gap becomes indistinguishable from zero only in markets with an unbiased chief. While access to a chief improves collaboration, chiefs without gender bias are most strongly associated with female collaboration, consistent with our theoretical predictions.

We next restrict our attention to businesses located inside markets and compare the collaboration behavior of women and men, allowing for differences in the degree of chief's gender bias across markets. Figure 7 shows the coefficients from the following specification:

$$coop_{ism} = \beta_1 Fem_{ism} + \beta_2 HighIndex_{im} + \beta_3 Fem_{ism} * HighIndex_{im} + X'_{ism}\delta + \epsilon_{ism}$$

---

[39]The exact components of the index are listed in Table D.2. Questions on the way in which the chief arbitrates disputes were asked referring to two hypothetical scenarios involving a female and a male marketeers, reported in Appendix D. The WVS questions we chose (e.g., "Men make better business leaders than women") have a strong and significant correlation with the rates of female entrepreneurship across countries.

where $coop_{ism}$ is average cooperation for business i in sector s located inside market m, Fem_{ism} is equal to 1 for being a female owner and $HighIndex_{im}$ is a dummy variable for having a chief who scores above median in one of the indexes constructed from the chiefs' survey (gender bias, legal competence and managerial duties). The set of controls X_{ism} is defined in section 5.3.

Figure 7a shows two facts. First, as previously discussed, women inside markets with unbiased chiefs cooperate as much as men (top panel). Second, consistently with our model, women's cooperation is substantially lower in markets with a chief characterized by high gender bias (β_3 is always negative, and significant with individual and industry controls).

Figure 7b shows that female cooperation is also related to the degree of legal competence of the chief. Women in markets with a legally competent chief cooperate significantly more than women in markets with an adjudicator of low quality ($p \leq 0.10$ when controlling for industry fixed effects). This is in line with better institutions – even in a gender-neutral way (i.e. higher δ) – having the potential to benefit women in business.[40]

Figure D.1 shows similar - albeit noisier - patterns for the sample of women only, by comparing their average cooperation levels when located inside versus outside markets with varying levels of chief's gender bias, legal competence and managerial duties. Figure D.1a shows that there is no significant difference in the cooperative behavior of women located outside markets which differ in the degree of chief's bias, which provides reassuring evidence that the index is not a proxy for broader gender attitudes in the area where the chief operates.

Since chiefs are not randomly assigned to markets, we cannot exclude that they become less biased in areas where women and men cooperate more together. To overcome remaining endogeneity concerns, we turn to our artefactual field experiments.

6 Experimental Variation in Institutional Access and Quality

Institutional support can reduce the risks of economic collaboration—especially for women—as shown by the positive relationship between having access to a market chief and female cooperation. But this effect varies, as adjudicators do not function uniformly. Chiefs are embedded in local social structures and may differ in how they exercise authority—some may reinforce prevailing gender hierarchies, others may act as egalitarians or as protectors of vulnerable groups like women.

In this section, we experimentally isolate the causal impact of institutional access on women's trust in business interactions. Specifically, we ask whether such access disproportionately benefits women and whether it depends on how adjudicators use their authority. In Experiment 1, we examine the effect of access to a gender-blind chief on trust and cooperation, benchmarking it against two alternatives: the absence of any enforcement mechanism and access to a formal legal institution. In Experiment 2, we allow chiefs to observe the participants' gender, reintroducing the possibility of personalized enforcement. In light of the potentially powerful effects of blinding institutions to gender, the second experiment allows us to understand whether, when institutions do see gender, this visibility helps overcome or simply reinforces existing power asymmetries.

Together, the experiments explore two institutional paths to inclusion: one through formal neutrality, and one through contextualized protection. We use both behavioral outcomes and original measures of adjudicator attitudes and participant expectations to understand which institutional features most effectively support inter-gender economic collaboration.

---

[40]Figure 7c also shows that a busy chief seems to favor male rather than female cooperation. This effect is not coming from a positive relationship between managerial duties and gender bias, as the correlation coefficient is -0.14 between the two indexes ($p = 0.21$).

Common design features. We conduct two artefactual field experiments (List and Metcalfe, 2014; List and Rasul, 2011), with a similar design. First, both experiments adapt the investment game introduced by Berg et al. (1995) and – in line with our model – frame it as two entrepreneurs having the opportunity to partner for business. Second, they involve entrepreneurs operating inside or around formal markets in Lusaka. To limit participant burden, minimize attrition, and preserve anonymity, the games were conducted inside participants' shops, with surveyors relaying choices electronically to each other in real time. Third, we embedded real local institutions into the experimental setting—namely, Market Chiefs and Small Claims Court judges. While we build on a rich literature that uses trust games to study institutional and interpersonal dynamics (e.g., Bohnet et al., 2008; Fehr and Rockenbach, 2003; Ederer and Schneider, 2022; Bartoš and Levely, 2021), to our knowledge, the involvement of real adjudicators from the local institutional environment is a relatively unique feature of our design.

6.1 Experiment 1: Testing Gender-Blind Adjudication

Our first experiment tests an extreme case of our model: an unbiased institution with no information on entrepreneurs' power and/or gender ($\theta = 0$, i.e. hidden to the adjudicator).

6.1.1 Design

Game rules and structure. The experiment consists of a modified version of the trust game (Berg et al., 1995), framed as an opportunity to invest in another person's business opportunity. Player A (the Investor) is given 10 tokens that can be kept or invested in the business of Player B (the Trustee). The Trustee receives three times the number of invested tokens and must decide how many to return to the Investor. We used the strategy method to elicit the Trustees' choices: before receiving the Investor's tokens, Player B completed a matrix indicating how many tokens they would like to return for any given amount received. Once the Investor's choice was made, the corresponding decision from the matrix was implemented.

Participants played the game for two rounds in total, but they switched roles and partner between the first and the second round. Players who started playing as an Investor (Trustee) would play as a Trustee (Investor) in the second round, and be matched with a different partner. This structure was common knowledge to players.

Experimental conditions. We cross-randomized two experimental conditions. See Table C.13 for sample sizes by experimental condition, and Tables C.14 and C.15 for balance checks.

Randomized access to gender-blind adjudicators. The first condition tests the effect of gender-blind institutions on trust. We randomly assigned pairs of players to three groups: a control group that receives no access to institutions, a first treatment group that has access to the Small Claims Court, and a second treatment group that has access to the market chief. In any of these two treatment groups, if the Investor thinks that the amount received by the Trustee is not fair, then they can ask the experimenters to go on their behalf to the court or chief to adjudicate. The trustee is told that the Investor can consult the chief or SCC clerk, and that their ruling will determine the distribution of tokens if the Investor complains.

This condition brings real institutions into a controlled setting, with the complaint process designed to closely reflect how disputes are handled in actual markets. However, the treatment introduces an important difference with respect to reality: adjudicators receive only information about the players' choices—not their gender or their identity. By ensuring anonymity, we remove the possibility that decisions are influenced by gender, allowing

us to test whether institutional access, in the absence of gender-based power differences, disproportionately benefits women.[41]

The implementation of the complaint required us to ensure that participants trusted that the experimenters would call the institution to complain, and that the Chief and the Small Claims Court Senior Clerk would understand the game and act as if it were a real-life dispute. We provide details on our solutions to these challenges in Appendix B.

Randomized match with market outsiders. Our second experimental condition allows us to explore whether in-group vs out-group dynamics increase cooperation within markets independently of contract enforcement (Kranton et al., 2020). Players were randomly assigned into two types of pairs. In the first treatment group, both players were selected from within the same market. In the second treatment group, one player was selected from inside the market and the other from outside. In all cases, participants were informed whether their partner was from the same market or an outsider, but received no additional information about their partner—only whether they operated inside or outside the market. Because the chief cannot adjudicate disputes involving businesses outside the market, we only assign pairs of insiders to the chief condition. We present results on this condition in the Appendix.

Payment. After the game was complete, players could exchange tokens for real money. In addition, all players received a participation fee of 35KW (3.5 dollars) as a token of appreciation for their time. The sessions were on average 90 minutes long.[42]

Sampling and randomization. Using data from the Lusaka Census of Urban Entrepreneurs, we randomly selected participants for the experiment, stratifying by whether the business was located within a market. The sample selection procedure was designed to oversample manufacturers and women in a selection of Lusaka's 80 dense urban marketplaces. We chose sixteen markets that contained the largest numbers of (female) manufacturers. For each market, we used the Census data to assemble one list of entrepreneurs located inside the market and a different list of entrepreneurs outside the market, but within 500 meters of its borders. We oversampled the manufacturing sector by placing all of the manufacturers at the top of the list in random order and all other businesses at the bottom of the list in random order. If a marketeer could not be located or refused to participate, we replaced the marketeer with the next female on the list. As soon as all women had been exhausted, the recruiters simply moved to the next marketeer on the list. The survey team visited businesses on the lists in this way until they had valid responses for 24 businesses inside each market and 6 businesses outside each market. We recruited a total of 480 participants across the 16 marketplaces.[43] We randomly paired players together and assigned them to a certain experimental condition based on a pre-determined order.

Benchmarking the effect of the chief. We implemented a Pre-Games Survey with questions about access and opinions about different institutions, dispute resolution and exposure to crime.[44] Our Pre-Games Survey helped us uncover that awareness of the Small Claims

---

[41]By submitting complaints on participants' behalf, we also simplify access and eliminate any image concerns associated with filing a complaint.

[42]In 7 (9) percent of chief-treated (court-treated) games the Investor asked for arbitration. Average earnings for all games before arbitration were 11.3 KW for the Investor and 6.7 KW for the Trustee. Together with the participation fee, earnings represent between 30% and 40% of the daily median sales value.

[43]Manufacturers are 78% of the final sample and we could include between 1 and 17 women in each market. We excluded 3 responses in the analysis because of limited understanding of the respondent. Because of the limited female representation, we do not have power to look at market-by-market treatment effects by gender. We overcome this limitation in experiment 2.

[44]Answers to these questions are presented in Table C.8 and discussed in previous sections. The pre-game

27

Court's existence was very limited. Eighty percent of our sample had not heard of the Small Claims Court before. Among those who had heard of it, half thought that "business owners like them had no access to the Small Claims Court". Five women out of 25 (and seven men out of 76) mention that the Court has high costs and "terribly" long procedures. This limited and faulty knowledge, which we discovered only after setting the experiment in motion, led us to conclude that the Small Claims Court was not understood as an institution and we therefore present our results on the Small Claims Court only in Appendix C. We therefore interpret the SCC treatment only as a useful benchmark: a formal institution offering legal protection on paper, but lacking credibility or awareness among entrepreneurs.

Theoretical predictions. The main prediction we test in this experiment is that the weak should benefit more from blind adjudication ($\theta = 0$) than the strong, since their ability to employ social punishment is lower.

Multiple hypotheses testing (MHT). We consider trust (i.e., tokens sent) our primary outcome; trustworthiness and investor/trustee earnings are secondary outcomes. We apply MHT corrections separately within the primary and secondary outcome families, and follow the approach of List et al. (2019).[45]

6.1.2 Results

Women's trust increases with gender-blind chief. Table 4 shows our primary results. Panel A, Columns (1) and (2) use the amount sent in the trust game as our experimental measure of trust. In this regression, we find that women without access to the Market Chief trust much less than men (see also Figure 8a). This trust gap is consistent with all of the previous results in this paper showing that female entrepreneurs collaborate less than men.

The gender difference in trust is eliminated when we introduce the possibility of complaining to a gender-blind market chief. This effect comes from women sending significantly more tokens when being randomly assigned to having access to the chief than in control. This result suggests that the experimentally generated differences in access to contract enforcement encourage cooperation disproportionately for women. This is in line with our theoretical prediction that weaker groups benefit the most from power-blind adjudicators.

Figure 9 shows that the increased women's investment in the chief condition occurs mainly in markets where the chief has above-median gender bias.[46] Thus women used to unfavorable chiefs benefit the most from the improvement in local institutions offered in the game. Women located in markets where the chief has a good reputation also slightly increase the tokens sent as Investors, probably as access is also easier compared to reality.

Access to gender-blind chiefs discourages bad male behavior. Columns (3) and (4) examine trustworthiness using the average return ratio, as in Glaeser et al. (2000). We divide the number of tokens the Trustee would return by the number of tokens available (return ratio) for each possible amount of tokens received, and then average over all return ratios. In the control group women return less than men, even if this result is no longer statistically significant when controlling for individual characteristics. In the chief condition, neither men nor women's trustworthiness change on average.

---

survey might have primed participants to think about their past experiences with institutions before the play. This should not be an issue for the interpretation of the results as long as this effect is the same across experimental conditions.

[45]We pre-registered this classification for experiment 2, and apply the same logic to experiment 1.

[46]A similar pattern appears when splitting by marketeers' opinion of the chief (fairness and speed): the increase comes from markets with below-median reputation.

While there is no significant average impact on trustworthiness, Figure 8b shows that access to the chief's ruling discourages extremely low return ratios and reduces the variance in the return to trust, in particular among men. Effectively, the Market Chief coordinates trustees' behavior around the norm considered fair by the marketeers (a 50/50 split) and reduces the risk involved in trusting others. Lower risk together with the chief's protection, in turn, disproportionately helps women trust others. Women also seem aware of the value of the chief's support. As an exploratory question, we asked players at the end of the experiment how much they would be willing to pay for institutional access should the activity be repeated. Women's distribution of answers is shifted to the right of men's (although not statistically different from each other).

Higher female trust increases the economic pie in the game. Panel B of Table 4 shows the final earnings. Female and male Investors in the control group tend to receive the same number of tokens at the end of the game. The chief raises the Investors' total earnings, with the effect being statistically the same for entrepreneurs of both genders. The last columns of Table 4 show that the chief has little impact on the Trustees' earnings.[47] As a result of higher female trust, investors' earnings increase, and the economic pie becomes bigger. Focusing into rounds of the game with male or female investors, Figure 10 indeed shows that women's increased trust in the chief condition leads to an increase in the economic surplus produced in the experiment of around 19%, with a greater share distributed to both the Investor and the Trustee.

A gender-blind court does not affect trust. Table C.18 shows that neither female nor male investors randomized into having access to the gender-blind adjudication of the court significantly change their behavior in sending tokens. This is consistent with the limited awareness of the SCC revealed in our surveys: without knowledge of the court's existence or procedures, access to its adjudication does not reduce perceived uncertainty around the investment. Interestingly, however, trustworthiness does increase in the Court condition—Trustees return more on average—even though Investors do not appear to anticipate this change.[48] This suggests that entrepreneurs do not internalize the potential deterrence effect of the court on others' behavior. Overall, this indicates that legal reforms may be difficult to implement effectively in informal, low-income and low-education settings: simply mentioning the possibility of formal punishment is unlikely to influence behavior unless institutions are well-known, trusted, and understood by the community.

Women are less trustworthy towards outsiders. Tables C.18 and C.19 show results for whether being randomly matched with a market outsider—rather than an insider—affects trust and trustworthiness behavior in the absence of institutional support (Column 3). We find no statistically significant difference in trust, suggesting that participants expect similar returns from sending tokens to fellow marketeers or to business owners outside the market. However, interestingly, we do find that women - but not men - return significantly fewer tokens when matched with an outsider. Since no institutional enforcement is available, women may rely more heavily on informal expectations of reciprocity—expectations that are weaker with outsiders.

---

[47]See Appendix C for results splitting the sample by all the different institutional treatments.

[48]Investors neither change their behavior nor revise their expectations of trustworthiness in the court condition, according to non-incentivized beliefs questions.

6.1.3 Summary

Our results show that introducing a gender-blind Market Chief closes the trust gap between men and women—especially benefiting women in markets with biased chiefs. A gender-blind court does not increase female trust, indicating that women's greater responsiveness is not simply due to preferences for third-party enforcement in general. Instead, women appear particularly sensitive to the Market Chief's (impartial) role embedded in their local context, highlighting that the identity and perceived legitimacy of the adjudicator matter.

Nevertheless, the observed trust increase among women may reflect several overlapping factors: confidence in the blindness of adjudication within the game; reliance on trusted local actors due to weaker means of social punishment; and potentially greater sensitivity to risk reduction, as the chief lowers the uncertainty of interactions relying on trust.

These findings raise an important question: is it the (game-induced) impartiality of the chief that matters most to women, or the broader institutional role the chief plays in local power structures? To answer this, we turn to a second experiment that compares the effects of gender-blind versus gender-aware chiefs on women's trust.

6.2 Experiment 2: Gender-Aware vs Gender-Blind Chiefs

The first experiment demonstrated that access to a gender-blind adjudicator—particularly one embedded in familiar, informal structures—can increase trust and close the gender gap in risky collaboration. However, real-world institutions rarely operate in the blind. Market chiefs often know the individuals they serve and may factor social identities, including gender, into their decisions. This raises a critical question: when institutions do see gender, does this visibility reduce or reinforce power asymmetries?

In this second experiment, we explore how women's trust responds when gender is visible to the adjudicator, and whether that visibility helps or harms them. Specifically, we test whether institutions that recognize group identity—in this case, gender—can improve trust for vulnerable populations, and under what conditions this recognition is perceived as protective versus biased. This experimental variation allows us to test the full theoretical model, particularly the predictions of Proposition 2 regarding women's preferences for adjudicators who may or may not be aware of their power gaps (θ), may or may not be protectors of the weak (α), and women's expectations of this sensitivity to power ($\hat{\alpha}$).

6.2.1 Design

Game rules and structure. The rules of the game mirror that of Experiment 1: the Investor (or Player A) decides how many of 10 tokens to send to the Trustee (or Player B), with the amount tripled upon transfer. The Trustee pre-commits to how much to return for each possible amount received. As in the chief condition of Experiment 1, all Investors in Experiment 2 have the option to file a complaint if they feel the return is unfair. Again, if they complain, their real market chief will adjudicate the dispute and may reallocate the final payoff. Unlike in Experiment 1, where participants play in both roles across rounds, here each participant plays only once and in a single role (Investor or Trustee) to avoid potential order effects.

Experimental treatments. In this experiment, all the investors have access to the chief's adjudication. The key experimental variation lies in what information the adjudicator receives about the Investor. Participants were randomly assigned to one of two treatments (see Table C.16 for balance checks):

- *Chief: Gender-Blind.* Complaint forms submitted to the chief reveal only the choices made by the investor and the trustee, without any identifying information.

- *Chief: Gender-Aware.* Complaint forms submitted to the chief reveal the choices made by the investor and the trustee, as well as the investor's gender.

Unlike in Experiment 1, where players only knew about the treatment they were assigned to, here all Investors are informed that two distinct types of complaint processes exist—one where gender is revealed and one where it is not. This choice ensures that gender is made equally salient across both treatment arms, avoiding a confound where gender salience alone could drive behavior in the Gender-Aware condition.

Trustees were informed that complaints could be submitted to the chief but were unaware that multiple versions of the complaint form existed. Investors, in turn, knew that Trustees were blind to the treatment. This design ensures that expectations about trustworthiness and potential returns from trust remain constant across treatment arms. What varies is only the ex-post likelihood and nature of institutional intervention, depending on whether the chief observes the investor's gender.[49]

This design reintroduces the power asymmetries that were shut down in Experiment 1. The key question becomes: do women trust more—or less—when their gender is visible to their market chiefs?

Providing information about Investor's gender offers the most parsimonious way to introduce variation in perceived power across treatment arms. When the adjudicator is blind to both players' gender, the power gap (θ) is implicitly treated as zero. In contrast, when Investor's gender is revealed, the adjudicator can form an expectation of θ based on possible pairings, without the need to observe the Trustee's identity. For a female investor, this expectation averages over Female–Female ($\theta = 0$) and Female–Male ($\theta > 0$) matches, leading to a positive expected power gap assuming that women are on average less powerful than men. For a male investor, the mix of Male–Male and Male–Female pairings implies a negative expected θ. Thus, even partial information is sufficient to generate meaningful variation in institutional response as long as the chief attaches a positive - even small - probability to inter-gender pairings.

Sampling and randomization We selected 12 markets from among Lusaka's dense urban marketplaces, five of which were also in Experiment 1. The market sampling procedure had two aims. First, to achieve variation in market size, market type (council or cooperative), and attitudes of the chief, based on our chief survey.[50] Second, to ensure that the sampled markets had a sufficient number of businesses in the industries we planned to recruit.

To increase the external validity of our findings, the second trial broadens the set of industries and stratifies both participant sampling and treatment assignment by gender and industry group (manufacturing vs. non-manufacturing). Within non-manufacturing, we focus on restaurants and hairdressers—two sectors that represent a large share of market vendors and, based on piloting, frequently engage in collaborative activities that exploit returns to scale.

Because our primary outcome is trust,[51] we stratify sampling by gender and industry group only for Player As. Player Bs are drawn from any industry ineligible to be a Player A,

---

[49]The script read to each player can be found in Appendix B.2.

[50]Figure C.9 shows that the markets where we conducted the experiment have an average chief gender bias index similar to those of the non-selected markets.

[51]Our pre-registration (AEA RCT number AEARCTR-0014859) lists trust as the primary outcome, and trustworthiness and earnings as secondary outcomes. Our main pre-registered heterogeneity is by gender and by chief's gender bias.

predominantly retail shops. In particular, within each market, we recruited an equal number of men and women, with overall targets determined by market size from a business listing. For Player As, the split between manufacturers and non-manufacturers was set based on availability, aiming for equal representation of manufacturers across genders. For Player Bs, we recruited an equal number of men and women without distinguishing industries.[52]

Treatment assignment was randomized within markets and Investor gender–industry cells. Trustees were not treated (i.e., they were unaware of the experimental treatments) but were paired to Investors within strata defined by the Investor's gender–industry and the Trustee's gender. Real-time assignment protocols, based on pre-randomized slots and randomized business lists, ensured that treatment was orthogonal to participant characteristics.

Further details about sampling and field protocols are included in Appendix B.2.

Theoretical predictions. Our empirical tests rely on two key assumptions: (i) that women are on average weaker than men, and (ii) that participants hold informed beliefs about the type of their market chief (i.e., whether they are protective of the weak or not).[53] Given these assumptions, the model yields two testable predictions that allow us to identify adjudicators who act as protectors of the weak (i.e., those with $\alpha > 0$):

- *Prediction 1: The weak benefit from revealing their gender if and only if the adjudicator is protective ($\alpha > 0$).*
  This implies three empirical patterns when the chief is protective:

  (i) Women are more likely to prefer revealing their gender;

  (ii) Women exhibit a higher level of trust in the gender-aware condition compared to the gender-blind one. We call the difference in trust between these two conditions the "revelation trust premium".

  (iii) Both the preference for gender-aware adjudication and the revelation trust premium should be negatively correlated between the weaker and more powerful groups within a market (i.e., between women and men), if choices are consistent with strategic updating about adjudicator type.

- *Prediction 2: Gender-aware chiefs are more likely to intervene in favor of women if i) they are protectors of the weak, and ii) the gender-based power gap in the market is large.*
  To test this, we examine how the gender gap in preferences for revelation and the revelation trust premium - which is our main proxy for a chief's protective preferences - relates to:

  (i) Chiefs' arbitration behavior - and participants' expectations of such behavior - in the trust game;

  (ii) Survey-based measures of chief decision-making (e.g., gender bias, salience of gender in decision making);

  (iii) Female power within the market (e.g., available means of social punishment and historical empowerment).

---

[52]Participants are between 28 and 72 per market. Overall, 48% of Player As are manufacturers.

[53]We find empirical support for both assumptions. In Experiment 2, only 17% of female participants report having a police contact they could turn to in case of a dispute, compared to 32% of male participants. Figure C.14 shows gender gaps in this support in every market. 87% of respondents report knowing their market chief at least somewhat, with only 13% saying they did not know the chief at all when asked, "How well do you know the chief?" This suggests that participants are broadly familiar with their chief and can form informed expectations about his behavior.

6.2.2 Results

On average, women prefer to reveal their gender. Figure C.10 shows that, on average, women prefer a complaint form that reveals their gender. According to our first theoretical prediction, this suggests that women generally perceive chiefs as protectors of the weak (i.e. $\hat{\alpha} > 0$). This preference is especially pronounced in markets with low gender bias among chiefs, where nearly 70% of women favor contextualized (gender-aware) adjudication, compared to fewer than 20% who prefer neutral (gender-blind) adjudication. By contrast, men tend to be indifferent between revealing or withholding their gender. However, the share of men who prefer to reveal their gender is highest in markets with chiefs who are biased against women (at around 50%)—consistent with the idea that more powerful groups benefit from being identifiable when adjudicators are not protective of the weaker party.

Women's trust increases with gender-aware chiefs. Panel A of Table 5, Columns (1) and (2), show treatment effects on trust. On average, women send more tokens when assigned to the gender-aware chief condition. In contrast, men send slightly less tokens when the chief knows their gender compare to a blind chief, but the effect is small and not statistically significant. This suggests that, in our games, women prefer being visible to being anonymous, and confirms that on average women perceive the chief to be supportive towards weaker parties.

In this second experiment, a gender gap in trust emerges even when participants are assigned to a gender-blind chief. Women send about 15% fewer tokens than men on average—a smaller gap than the 25% observed in the previous experiment —and the difference becomes statistically insignificant once we control for individual characteristics. Yet, this stands in contrast to the first experiment, where women trust as much as men under a gender-blind chief. One likely reason for this discrepancy is that, in the second experiment, participants are aware of the existence of both the gender-aware and gender-blind conditions, making gender more salient. If women expect that only a gender-aware chief would actively protect their interests, they may interpret the blind condition as less favorable, leading to lower trust. In the first experiment, the neutral chief is presented without comparison, and participants may simply assume fairness.

Effects on trust vary across markets and are negatively correlated between genders. The average effect of assigning a gender-aware chief masks variation across markets. Figure 11 ranks markets by the difference in tokens sent by women under the gender-aware versus the gender-blind chief condition, or the "revelation trust premium." In three markets, women send slightly more tokens under the gender-blind compared to the gender-aware condition, though these differences are not statistically significant. In the remaining eight out of twelve markets, women send more tokens when the chief is gender-aware, with large and significant effects in three of them. The figure also reports the corresponding effects for men. Three main patterns emerge. First, men's revelation trust premium is generally negative, while the opposite holds for women: in seven markets, men trust more in the gender-blind condition. Second, the magnitude of the treatment effect for men is smaller than for women, suggesting that men are less sensitive to variation in adjudication. This is consistent with the idea that men, due to greater social power, may be less dependent on third-party enforcement or less attentive to institutional features. Third, markets where women's revelation trust premium is larger are those where men's trust premium is lower (see also Figure C.15). This negative correlation between women's and men's revelation trust premia generates a positive difference-in-differences across markets (i.e. the difference between women's and men's revelation trust premia).

Gendered trust premia reflect adjudication behavior and market context. Gender differences in the revelation trust premium across markets are correlated with several other experimental outcomes, as shown in Table C.22. First, in markets with a larger difference in the revelation trust premium between women and men, women are also more likely to state a preference for adjudication forms that reveal their gender (Panel B). Chief's arbitration behavior and participants' incentivized expectations further support the interpretation of the revelation trust premium - and its difference between genders - as a measure of chief's protective preferences. In markets with a larger gender gap in the revelation trust premium, chiefs are more likely to take action when the complainant is a woman rather than a man (Panel C, Table C.22). This is mirrored in marketeers' beliefs: when asked to guess how the chief would behave in hypothetical complaints within the game, respondents expect chiefs to redistribute more tokens in favor of women in markets with higher trust differentials across conditions by gender (see Panel D, Table C.22). Taken together, these results suggest that the heterogeneity across markets reflects meaningful variation in how chiefs are perceived and how they actually behave in gendered disputes. Moreover, it suggests that women's greater trust in gender-aware adjudication stems from context-specific expectations of protection, and their awareness of the chief's type.

The variation in treatment effects across markets captures more than just participants' behavior in the experiment—it also reflects key features of the markets themselves (Table C.23 and Figure C.12). In markets with a higher revelation trust premium for women, real-world female collaboration is also significantly higher ($p < 0.05$). These are also the markets where women report more frequent disputes and are more likely to cite security concerns as a reason for choosing their business location. This pattern suggests that women respond more strongly to protective adjudication in environments where they feel more exposed or vulnerable, and where institutional support is especially valued.

Market heterogeneity is explained by female power and chiefs gender bias.

The second prediction of our theoretical model further tells us that, ceteris paribus, women's preference for gender-aware adjudicators should be weaker when the chief is biased against women (a lower or negative α) and when female power is higher (a lower θ).

Figure 12a shows that the gender difference in the revelation trust premium is smaller in markets where the chief exhibits greater bias against women, based on our gender bias index (defined in Section 5.4). In contrast, women send relatively more tokens in the gender-aware treatment when the chief is expected to arbitrate in favor of women (according to incentivized measures) or scores higher on an implicit gender-salience task (a "Triad picture task," see Appendix D.2), which captures whether the chief uses gender as a basis for categorizing people. This suggests that women are more willing to trust when revealing their gender to chiefs who are attentive to gender yet not biased against their group.

To proxy female power, we rely on two measures. First, we use self-reported data on social support in the event of a dispute, collected through our survey.[54] This measure maps closely to our model, where power is the ability to impose social punishment. Our second measure captures variation in gender norms across ethnic groups. Drawing on the methodology of Ashraf et al. (2020), we link the ethnicity of business owners in our sample to Murdock's Ethnographic Atlas, which provides indicators of matrilineality, matrilocality, and women's participation in agriculture. We combine these variables into a single index of "historical female empowerment," which reflects long-standing cultural norms around female authority and autonomy. This allows us to test whether, as Gneezy et al. (2009) find, female power as proxied by tribal matrilineality would enable women to respond to adjudication as men do.

---

[54]In the survey implementation, this question was asked only to manufacturing businesses. For missing observations in other sectors, we impute the market-gender-level mean.

Figure 12b shows the relationship between our measures of female power and gender gap in the revelation trust premium at the market level. Consistent with the model's predictions, markets with higher female power—whether measured through social support or historical empowerment—exhibit a smaller trust gain for women under the gender-aware chief condition, relative to the blind condition. In this way, women become less sensitive to, and needing of, the protection of the chief. In other words, when they have other sources of power, women behave more like men (Gneezy et al., 2009). They adjust their willingness to trust based on their expectations of adjudicator's behavior, rather than fixed preferences or risk attitudes. This evidence points to the importance of institutional design in shaping gendered economic outcomes.

Trustworthiness and earnings results. Trustees in this experiment were blind to the existence of any treatment, and to the best of our knowledge, this protocol was consistently followed in the field. Among female trustees, we find no significant differences in the average return ratio across treatment arms. However, Figure C.11 shows that male trustees assigned to the gender-aware chief condition commit to returning significantly more tokens than their counterparts in the blind condition, as well as more than women in any condition. While this effect is statistically precise, it is modest in size—around 10% of the mean return ratio—and concentrated in just two small markets, comprising a total of 15 male trustees (see Figure C.11c).[55] Figure C.11c orders markets by the difference in tokens sent by women in the gender-aware versus blind chief treatments. Notably, the observed differences in male trustworthiness across treatment groups do not arise in the same markets where we observe larger treatment effects on women's trust. This suggests that the patterns in trustee behavior are not driven by the same structural market characteristics that appear to explain investor responses, and which are the main focus of the analysis.[56]

We find no significant effects of either treatment on investor's or trustees' earnings (Panel B, Table 5). Nevertheless, we replicate the result of our previous experiment on total earnings: when female investors trust more, the game economic pie is larger (see Figure C.13).

6.3 Discussion

The results of our experiments reveal that local informal institutions can play a vital role in fostering economic collaboration—particularly when they act as protectors of the weak. While previous theoretical frameworks emphasize the rule of law, and indeed ours began with considering different types of contract enforcement including legal courts, our experimental results suggest that informal intermediaries can play a large role in increasing protection of minorities' investment. This makes even more sense in a world with pervasive contract incompleteness, as informal intermediaries may be more flexible in their support compared to the legal system. Women, who face greater risks in mixed-gender business interactions, are not simply seeking neutrality—they are, in some cases, seeking allies. Chiefs, though embedded in local norms, can function as benevolent adjudicators, and women respond strategically to this possibility. In contrast to the conventional wisdom that anonymity protects the vulnerable, our findings show that being visible to a fair institution can itself be empowering.

Together, the two experiments illustrate a broader point: when formal legal systems are weak or inaccessible, local institutions can fill the gap. In their gender-blind form, they can mitigate power asymmetries and facilitate trust. But when designed—or perceived—to

---

[55]While we aimed to recruit an equal number of male and female trustees within each market, logistical constraints occasionally prevented us from achieving perfect gender balance across treatment arms.

[56]The overall distribution of the average return ratio in this second experiment is indistinguishable from the distribution from the previous experiment for both genders (KS test $p > 0.30$).

protect the vulnerable, identity-aware institutions can go even further, enabling high-risk, high-return collaboration for those most in need of institutional support.

Beyond contract enforcement institutions, what else could women resort to for protection from expropriation? In our qualitative work, we repeatedly asked female entrepreneurs why they don't bring their husbands, brothers or any other trusted male family member to help in disputes and negotiations. Their replies consistently pointed to one issue: if women get help, they have to give away decision rights over the business to their men. Thus the insecurity of property rights, beyond incomplete contracts, limits alternative solutions. A similar situation has been shown for female farmers, whose plots are less productive – not because they are on worse soil or because women are less skilled, but because there is less investment in inputs on those lands (Udry, 1996). Maximizing production by rearranging inputs would benefit the whole household, but insecure property rights for women mean the wife would lose some of her decision, and potentially property, rights to the land if her husband started investing in her plots. As in our case of collaborative market exchange, production is sub-optimal because of insecure contracting rights, with insecurity being highly gendered.

Our findings raise broader questions about which types of economic cooperation are most constrained by weak enforcement and trust—particularly for women. One such area is input sourcing and trade. Startz (2016) documents that Nigerian entrepreneurs often travel long distances to source goods in the face of search and contract frictions. Future work might investigate whether female entrepreneurs avoid sourcing arrangements that require travel or high-stakes intermediation due to contracting risk, and whether this limits their access to more profitable supply chains. More broadly, insights from the trade literature on which sectors are most vulnerable to enforcement frictions may help explain patterns of female under-representation across industries. Relatedly, our experiment does not have variation in the type of output generated by pairs of players, which would depend on their skills, effort and coordination ability. However, it's natural to think that the economic losses of missed collaboration opportunities will be larger in activities that hinge on partnerships between people with diverse skills, such as entrepreneurial innovation and scientific knowledge.

We also speak to recent work documenting gender segregation in buyer-seller matches (Hardy and Kagy, 2018, 2020) and employees hiring (Chiplunkar and Goldberg, 2024), with women predominantly selling to or hiring other women. These patterns may reflect mutual concerns about trust and enforceability in cross-gender transactions—on both the supply and demand side. Viewed through this lens, our findings may contribute to a deeper understanding of how institutional frictions shape not only firm outcomes, but also the gender structure of entire markets, including spillovers into wage labor markets.

External validity. We assess the external validity of our findings following the framework proposed by List (2020). Our sample consists of small-scale entrepreneurs operating in Lusaka. Using nationally representative Demographic and Health Survey (DHS) data, we benchmark our participants against both the general adult population and the self-employed in Zambia (Tables C.24 and C.25). Our sample is broadly similar in terms of marital status and religion, but somewhat more urban, younger, and more educated—differences that reflect the urban and sectoral focus of our study. Comparable patterns emerge when comparing to populations across other Sub-Saharan African (SSA) countries.

In terms of attrition, in Section 4.1 we compare the small-scale manufacturers interviewed in our Census to the 40% of eligible firms that were not reached (Table C.2). Compliance in the games was universal. We also compare participants in the experimental games to the Census of Manufacturers and find that, despite differences in industry composition and location, gender gaps in trust, cooperation and peer learning are to some extent present across all samples (Table 1) and owners' characteristics are comparable (e.g., in age, education).

36

Considering naturalness, the experimental tasks were designed to reflect everyday business behaviors familiar to participants, including lending, sharing orders, and resolving disputes. Participants consistently reported that the trust game and adjudication scenarios reminded them of situations they encounter in their work. The mechanisms for resolving disputes were modeled on real institutions, which enhances the naturalism of the design. While the games maintain the control of a lab experiment, they bring together a representative population of entrepreneurs and the institutions they regularly engage with.

Finally, to understand how our findings might extend to settings beyond Zambia, we draw on both external datasets and prior literature. First, studies across low-income contexts consistently document gender gaps in entrepreneurship that echo those we observe in Lusaka. Second, comparisons using the World Values Survey (WVS) and Afrobarometer show that gender differences in trust and access to justice in Lusaka mirror broader regional patterns (Table C.26). Finally, because we adapted several of our survey questions from the WVS and the World Justice Project - Global Population Poll (GPP), we are able to show that our market chiefs' gender attitudes are broadly comparable to national and regional averages. However, our chiefs tend of have slightly more progressive gender attitudes related to women in business and also seem to be more aware of female disadvantage with institutions than citizens in Zambia, SSA or the US (Table C.27).

While our study is set in Lusaka, the mechanisms we uncover—particularly the role of adjudicator discretion in shaping trust under power asymmetries—are likely to be relevant beyond this context. Our framework applies to settings that share three key conditions. First, production processes with contractual incompleteness and/or complexity, which create environments where cooperation relies heavily on trust, and where formal legal enforcement is limited or prohibitively costly. These frictions are present also in high-income settings, including professional services, creative industries, and early-stage ventures. Second, institutional arrangements involving intermediated enforcement—where disputes are resolved not through formal courts but via informal gatekeepers or delegated adjudicators. These intermediaries are common, e.g. in decentralized work environments, digital platforms, and community-level dispute resolution. Third, societal structures marked by power differentials—whether by gender, ethnicity, or class—can influence both how collaboration unfolds and how fairness is perceived by adjudicators.

Although our parameter estimates are not intended to be representative, the model and heterogeneity analysis offer a framework for identifying the types of variation that matter when studying trust and collaboration in unequal partnerships. In particular, our measures of social power, institutional discretion, and local norms can help researchers think about where and when similar dynamics are likely to emerge. While we explore variation across these dimensions within our sample, we see them as relevant for understanding broader patterns across settings, including in environments where women's participation is constrained not by formal rules but by unequal bargaining power and limited recourse.

These conditions are not unique to Lusaka or low-income settings. Indeed, they characterize many other domains of economic collaboration—such as scientific production and co-authorship, where tasks are interdependent, power may be uneven, and enforcement is informal. In such contexts, one may ask: what is the equivalent of the adjudicator in our model? In academia, roles such as department chairs or seminar moderators can potentially have a role to play in protecting the inclusion of women and minorities and their investment in science. In the context of seminars, Dupas et al. (2025) and Schmidt et al. (2017) mention session chairs as a potentially important "institution" for guiding seminar interactions and maintaining a professional environment against the hostile treatment of female presenters. Lissoni et al. (2020) document that women are more likely to be excluded from patents' inventorship in mixed-gender and mixed-seniority teams, but different institutional settings

can mitigate (or exacerbate) the gap. Goldin and Rouse (2000) show that blind auditions increase women's chances of advancing from preliminaries and being hired—but only when there is no semifinal round. In semifinals, the effect of the screen is negative, which they interpret as committees using partial gender information to exercise a form of affirmative action: advancing women only when confident in their quality. This mirrors our setting, where impartial adjudication can be less effective in protecting women when it prevents adjudicators from acting on their preferences for helping those with lower social power.

7 Conclusion

Economic opportunities arise through the density of potential collaborators and ideas. Yet the ability to interact positively is limited if one partner consistently fears expropriation by the other. Consequently, contract enforcement complements urbanization, which creates dense possibilities for interaction, not just because effective public order reduces activities that generate negative externalities, like water pollution, but also because adjudication enables a wider range of positive interactions. The advantages of contract enforcement are particularly high when one set of actors has more power in a state of anarchy than the other.

When one group systematically holds more social power, partnerships involving members of the weaker group may break down or never form—especially in the absence of institutions that protect against exploitation. In Lusaka, we believe that the industrial segregation of female entrepreneurs partially reflects the safety of being able to trade within genders.

A central theme of this paper, and an implication of our work, is that even gender-neutral applications of contract enforcement can have benefits that accrue disproportionately to women, particularly when they help level the playing field for those with lower baseline power. When rule of law is absent, unequal power dynamics can undermine even seemingly innocuous market transactions.

While our focus is on female entrepreneurship, the framework also offers insight into broader labor market dynamics. Our model implies that enforcement quality should matter most in self-employment, where individuals must manage their own contracts and partnerships. In contrast, discriminatory norms can affect outcomes across both wage work and entrepreneurship. While gender norms are not explicitly incorporated in our theoretical framework, our empirical analysis highlights how institutional support interacts with underlying social expectations around gender roles. This suggests that improvements in contract enforcement alone may not shift women's labor force participation over the course of development (Olivetti and Petrongolo, 2016), but may change the relative expected utility of self-employment with respect to wage work, affecting the occupational structure by gender (Ashraf et al., 2022). Future research may shed further light on this topic.

Finally, our findings point to the importance of informal institutions and local adjudicators in shaping business dynamics—often in ways that formal legal systems overlook. A promising avenue for future research is to explore how these localized mechanisms interact with country-level social norms and enforcement institutions. In particular, understanding whether social norms and contract enforcement - both formal and informal - operate as complements or substitutes in shaping gender gaps in entrepreneurship would help assess the external validity of our findings across settings.

Bibliography

Aberra, A. and Chemin, M. (2021). Does legal representation increase investment? evidence from a field experiment in kenya. *Journal of Development Economics*, 150:102612.

Abreu, D. (1986). Extremal equilibria of oligopolistic supergames. *Journal of economic theory*, 39(1):191–225.

Abreu, D. (1988). On the theory of infinitely repeated games with discounting. *Econometrica: Journal of the Econometric Society*, pages 383–396.

Adams-Prassl, A., Huttunen, K., Nix, E., and Zhang, N. (2023). Violence against women at work. *The Quarterly Journal of Economics*, 139(2):937–991.

Aizer, A. (2010). The gender wage gap and domestic violence. *American Economic Review*, 100(4):1847–59.

Alibhai, S., Buehren, N., Papineni, S., and Pierotti, R. (2017). Crossovers: female entrepreneurs who enter male sectors - evidence from ethiopia. Technical report, Washington, D.C.

Aman-Rana, S., Minaudier, C., Álvarez Pereira, B., and Chaudry, S. (2021). Gender and choice over co-workers: Experimental evidence. *Working paper*.

Ashraf, N., Bandiera, O., Minni, V., and Quintas-Martinez, V. (2022). Gender roles and the misallocation of labour across countries. *Working Paper*.

Ashraf, N., Bau, N., Nunn, N., and Voena, A. (2020). Bride price and female education. *Journal of Political Economy*, 128(2):591–641.

Ashraf, N., Delfino, A., and Glaeser, E. L. (2019). Rule of law and female entrepreneurship. Technical report, National Bureau of Economic Research.

Ashraf, N., Delfino, A., Glaeser, E. L., and Solmone, I. (2025). The inverted u-shaped relationship between female entrepreneurship and economic development. *AEA Papers and Proceedings*, 115:500–508.

Balán, P., Bergeron, A., Tourek, G., and Weigel, J. L. (2022). Local elites as state capacity: How city chiefs use local information to increase tax compliance in the democratic republic of the congo. *American Economic Review*, 112(3):762–797.

Baldwin, K. (2016). *The paradox of traditional chiefs in democratic Africa*. Cambridge University Press.

Baldwin, K. (2018). When politicians cede control of resources: Land, chiefs, and coalition-building in africa. *Comparative Political Studies*, 51(3):253–291.

Bardasi, E., Sabarwal, S., and Terrell, K. (2011). How do female entrepreneurs perform? evidence from three developing regions. *Small Business Economics*, 37(417).

Bartoš, V. and Lively, I. (2021). Sanctioning and trustworthiness across ethnic groups: Experimental evidence from afghanistan. *Journal of Public Economics*, 194:104347.

Bassi, V., Muoio, R., Porzio, T., Sen, R., and Tugume, E. (2022). Achieving scale collectively. *Econometrica*, 90(6):2937–2978.

Behrer, A. P., Glaeser, E. L., Ponzetto, G., and Shleifer, A. (2021). Securing property rights. *Journal of Political Economy*, 129(4):1157–1192.

Berg, J., Dickhaut, F., and McCabe, K. (1995). Trust, reciprocity, and social history. *Games and Economic Behavior*, 10(1):122–142.

Berge, L., Bjorvatn, K., and Tungodden, B. (2015). Human and financial capital for microenterprise development: Evidence from a field and lab experiment. *Management Science*, 61(4):707–722.

Berge, L. and Pires, A. (2020). Gender, formality, and entrepreneurial success. *Small Business Economics volume*, 55:881–900.

Bloom, N., Fisher, G., Rasul, I., Rodríguez-Clare, A., Suri, T., Udry, C., Verhoogen, E., Woodruff, C., and Zane, G. (2014). *Firm Capabilities and Economic Growth*. IGC Evidence Paper. International Growth Center (IGC).

Bohnet, I., Greig, F., Herrmann, B., and Zeckhauser, R. (2008). Betrayal aversion: Evidence from brazil, china, oman, switzerland, turkey, and the united states. *American Economic Review*, 98(1):294–310.

Boschini, A. and Sjögren, A. (2007). Is team formation gender neutral? evidence from coauthorship patterns. *Journal of Labor Economics*, 25(2):325–365.

Britto, D., Germinetti, L., Gerard, F., Naritomi, J., and Sampaio, B. (2025). Access to justice and social protection. In *AEA Papers and Proceedings*, volume 115, pages 329–334. American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

Cahuc, P., Carcillo, S., Patault, B., and Moreau, F. (2024). Judge bias in labor courts and firm performance. *Journal of the European Economic Association*, 22(3):1319–1366.

Calderon, G., Iacovone, L., and Juarez, L. (2017). Opportunity versus necessity: Understanding the heterogeneity of female micro-entrepreneurs. *World Bank Economic Review*, 30:86–96.

Campos, F., Coleman, R., Conconi, A., Donald, A., Gassier, M., Goldstein, M., Chavez, Z., Mikulski, J., Milazzo, A., Paryavi, M., Pierotti, R., O'Sullivan, M., and Vaillant, J. (2019). *Profiting from Parity: Unlocking the Potential of Women's Businesses in Africa. Main Report (English)*. World Bank Group, Washington, D.C.

Campos, F. and Gassier, M. (2017). *Gender and Enterprise Development in Sub-Saharan Africa: A Review of Constraints and Effective Interventions*. Number 8239 in Policy Research Working Paper. World Bank, Washington DC.

Campos, F., Goldstein, M., McGorman, L., Munoz Boudet, A., and Pimhidzai, O. (2014). *Breaking the metal ceiling: female entrepreneurs who succeed in male-dominated sectors in Uganda*. Africa region gender practice policy brief 9. World Bank Group, Washington, D.C.

Chanock, M. (1998). *Law, custom and social order: The colonial experience in Malawi and Zambia*. Heinemann.

Chemin, M. (2020). Judicial efficiency and firm productivity: Evidence from a world database of judicial reforms. *Review of Economics and Statistics*, 102(1):49–64.

Chiplunkar, G. and Goldberg, P. K. (2024). Aggregate implications of barriers to female entrepreneurship. *Econometrica*, 92(6):1801–1835.

Cirera, X. and Qasim, Q. (2014). *Supporting Growth-Oriented Women Entrepreneurs: A Review of the Evidence and Key Challenges*. Number 5 in Innovation, technology and entrepreneurship policy note. World Bank, Washington DC.

Coffman, K. B. (2014). Evidence on self-stereotyping and the contribution of ideas. *The Quarterly Journal of Economics*, 129(4):1625–1660.

Cortes, P. and Pan, J. (2018). Occupation and gender. *The Oxford handbook of women and the economy*, pages 425–452.

Dal Bó, P. and Fréchette, G. R. (2018). On the determinants of cooperation in infinitely repeated games: A survey. *Journal of Economic Literature*, 56(1):60–114.

De Mel, S., McKenzie, D., and Woodruff, C. (2008). Returns to capital in microenterprises: Evidence from a field experiment. *The Quarterly Journal of Economics*, 123(4):1329–1372.

De Mel, S., McKenzie, D., and Woodruff, C. (2009). Are women more credit constrained? experimental evidence on gender and microenterprise returns. *American Economic Journal: Applied Economics*, 1(3):1–32.

De Mel, S., McKenzie, D., and Woodruff, C. (2014). Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from sri lanka. *Journal of Development Economics*, 106:199–210.

Ductor, L. and Prummer, A. (2024). Gender homophily, collaboration, and output. *Journal of Economic Behavior & Organization*, 221:477–492.

Dupas, P., Handlan, A., Modestino, A. S., Niederle, M., Sere, M., Sheng, H., Wolfers, J., and the Seminar Dynamics Collective (2025). Gender differences in economics seminars. *American Economic Review, forthcoming.*

Ederer, F. and Schneider, F. (2022). Trust and promises over time. *American Economic Journal: Microeconomics*, 14(3):304–20.

Essers, D., Megersa, K., and Sanfilippo, M. (2021). The productivity gaps of female-owned firms: Evidence from ethiopian census data. *Economic Development and Cultural Change*, 69(2):645–683.

Eswaran, M. and Malhotra, N. (2011). Domestic violence and women's autonomy in developing countries: theory and evidence. *Canadian Journal of Economics/Revue canadienne d'économique*, 44(4):1222–1263.

Fairlie, R. W., Morelix, A., and Tareque, I. (2017). *2017 Kauffman Index of Startup Activity: National Trends*. Ewing Marion Kauffman Foundation.

Fehr, E. and Rockenbach, B. (2003). Detrimental effects of sanctions on human altruism. *Nature*, 422(6928):137–140.

Ferrant, G., Fuiret, L., and Zambrano, E. (2020). The social institutions and gender index (sigi) 2019: A revised framework for better advocacy.

Field, E., Jayachandran, S., Pande, R., and Rigol, N. (2010). Do traditional institutions constrain female entrepreneurship? a field experiment on business training in india. *American Economic Review*, 100(2):125–29.

Folke, O. and Rickne, J. (2022). Sexual harassment and gender inequality in the labor market. *The Quarterly Journal of Economics*, 137(4):2163–2212.

Friedman, J. W. (1971). A non-cooperative equilibrium for supergames. *The Review of Economic Studies*, 38(1):1–12.

Fundenberg, D. and Maskin, E. (1990). Evolution and cooperation in noisy repeated games. *The American Economic Review*, 80(2):274–279.

Gentner, D. and Brem, S. (1999). Is snow really like a shovel? distinguishing similarity from thematic relatedness. *Proceedings of the Twenty-First Annual Meeting of the Cognitive Science Society.*

Glaeser, E., Laibson, D., Scheinkman, J., and Soutter, C. (2000). Measuring trust. *Quarterly Journal of Economics,*, 115(3):811–846.

Gneezy, U., Leonard, K. L., and List, J. A. (2009). Gender differences in competition: Evidence from a matrilineal and a patriarchal society. *Econometrica*, 77(5):1637–1664.

Goldin, C. and Katz, L. F. (2011). The cost of workplace flexibility for high-powered professionals. *The Annals of the American Academy of Political and Social Science*, 638:45–67.

Goldin, C. and Rouse, C. (2000). Orchestrating impartiality: The impact of "blind" auditions on female musicians. *American economic review*, 90(4):715–741.

Goldstein, M., Gonzalez Martinez, P., and Papineni, S. (2019). Tackling the global profitarchy: Gender and the choice of business sector. *World Bank Policy Research Working Paper*, (8865).

Grossman, S. J. and Hart, O. D. (1983). Implicit contracts under asymmetric information. *The quarterly journal of economics*, pages 123–156.

Halim, D. (2020). Women entrepreneurs needed-stat! World Bank Blogs. `https://blogs.worldbank.org/en/opendata/women-entrepreneurs-needed-stat`.

Hardy, M. and Kagy, G. (2018). Mind the (profit) gap: Why are female enterprise owners earning less than men? *AEA Papers and Proceedings*, 108:252–255.

Hardy, M. and Kagy, G. (2020). It's getting crowded in here: Experimental evidence of demand constraints in the gender profit gap. *The Economic Journal*, 130(631):2272–2290.

Hardy, M., Kagy, G., and Jimi, N. (2022). The enterprise gender data gap: Evidence from sub-saharan africa. Technical report.

Hardy, M., Kim, S., McCasland, J., Menzel, A., and Witte, M. (2024). Allocating labor across small firms: Experimental evidence on information constraints. *Journal of Development Economics*, 171:103345.

Hart, O. (2009). Hold-up, asset ownership, and reference points. *The Quarterly Journal of Economics*, 124(1):267–300.

Hart, O. and Moore, J. (1988). Incomplete contracts and renegotiation. *Econometrica: Journal of the Econometric Society*, pages 755–785.

Hsieh, C. and Klenow, P. (2009). Misallocation and manufacturing tfp in china and india. *The Quarterly journal of economics*, 124(4):1403–1448.

Jayachandran, S. (2020). *Social Norms as a Barrier to Women's Employment in Developing Countries.* Number 27449.

Johnson, S., McMillan, J., and Woodruff, C. (2002). Courts and relational contracts. *Journal of Law, Economics, and organization*, 18(1):221–277.

Klapper, L. F. and Parker, S. C. (2011). Gender and the business environment for new firm creation. *World Bank Research Observer*, 26(2):237–257.

Kling, J., Liebman, J., and Katz, L. (2007). Experimental analysis of neighborhood effects. *Econometrica*, 75(1):83–119.

Kranton, R., Pease, M., Sanders, S., and Huettel, S. (2020). Deconstructing bias in social preferences reveals groupy and not-groupy behavior. 117(35):21185–21193.

Lim, S. C., Ghani, F., and Remme, M. (2018). Sexual harassment: A global problem. *Policy Brief*.

Lin, E. and Murphy, G. (2001). Thematic relations in adults' concepts. *J Exp Psychol Gen.*, 130(1):3–28.

Lissoni, F., Montobbio, F., and Zirulia, L. (2013). Inventorship and authorship as attribution rights: An enquiry into the economics of scientific credit. *Journal of Economic Behavior & Organization*, 95:49–69.

Lissoni, F., Montobbio, F., and Zirulia, L. (2020). Misallocation of scientific credit: the role of hierarchy and preferences. An extension of Lissoni et al. (2013). *Industrial and Corporate Change*, 29(6):1471–1482.

List, J. A. (2020). Non est disputandum de generalizability? a glimpse into the external validity trial. Technical report, National Bureau of Economic Research.

List, J. A. and Metcalfe, R. (2014). Field experiments in the developed world: An introduction. *Oxford Review of Economic Policy*, 30(4):585–596.

List, J. A. and Rasul, I. (2011). Field experiments in labor economics. In *Handbook of Labor Economics*, volume 4, pages 103–228. Elsevier.

List, J. A., Shaikh, A. M., and Xu, Y. (2019). Multiple hypothesis testing in experimental economics. *Experimental Economics*, 22(4):773–793.

Logan, C. (2013). The roots of resilience: Exploring popular support for african traditional authorities. *African Affairs*, 112(448):353–376.

Markman, E. M. and Hutchinson, J. E. (1984). Children's sensitivity to constraints on word meaning: Taxonomic versus thematic relations. *Cognitive Psychology*, 16(1):1–27.

Marshall, A. (1890). *Principles of Economics.* Macmillan, London.

Marx, B., Stoker, T. M., and Suri, T. (2019). There is no free house: Ethnic patronage in a kenyan slum. *American Economic Journal: Applied Economics*, 11(4):36–70.

Matsvayi, A. (2012). *The role of chiefs' courts in women's access to justice: A case study of the courts of Chief Makoni, Chief Seke and Chief Chimoyo in Zimbabwe.* PhD thesis, University of Zimbabwe.

McKenzie, D. and Woodruff, C. (2017). Business practices in small firms in developing countries. *Management Science*, 63(9):2967–2981.

Minondo, A. and Requena-Silvente, F. (2013). Does complexity explain the structure of trade? *Canadian Journal of Economics*, 46(3):928–955.

Mirman, D., Landrigan, J.-F., and Britt, A. E. (2017). Taxonomic and thematic semantic systems. *Psychological bulletin*, 143(5):499–520.

Nix, E., Gamberoni, E., and Heath, R. (2014). Bridging the gap: Identifying what is holding self-employed women back in ghana, rwanda, tanzania, the republic of congo, and uganda. *World Bank Policy Research Working Paper*, (6946).

Nordman, C. J. and Vaillant, J. (2014). *Inputs, Gender Roles or Sharing Norms? Assessing the Gender Performance Gap Among Informal Entrepreneurs in Madagascar.* Number 8046 in IZA Discussion Papers.

Nordman, J., Robilliard, A., and Roubaud, F. (2011). Gender and ethnic earnings gaps in seven west african cities. *Labour Economics*, 18(S1):S132–S145.

OECD (2012). Gender equality in entrepreneurship. In *Closing the gender gap: Act now*, chapter part IV, pages 273–343. OECD Publishing.

Olivetti, C. and Petrongolo, B. (2016). The evolution of gender gaps in industrialized countries. *Annual Review of Economics*, 8:405–434.

Palagashvili, L. (2018). African chiefs: comparative governance under colonial rule. *Public Choice*, 174(3):277–300.

Ponticelli, J. and Alencar, L. S. (2016). Court Enforcement, Bank Loans, and Firm Investment: Evidence from a Bankruptcy Reform in Brazil *. *The Quarterly Journal of Economics*, 131(3):1365–1413.

Rao, M. (2024). Front-line courts as state capacity: Evidence from india. *Available at SSRN 4854161*.

Sadka, J., Seira, E., and Woodruff, C. (2024). Information and bargaining through agents: Experimental evidence from mexico's labour courts. *Review of Economic Studies*, 91(6):3677–3711.

Sandefur, J. and Siddiqi, B. (2013). Delivering justice to the poor: Theory and experimental evidence from liberia. In *World Bank Workshop on African Political Economy*, volume 20, pages 1–61.

Sarsons, H., Gërxhani, K., Reuben, E., and Schram, A. (2021). Gender differences in recognition for group work. *Journal of Political Economy*, 129(1):101–147.

Schmidt, S. J., Douglas, S., Gosnell, N. M., Muirhead, P. S., Booth, R. S., Davenport, J. R., and Mace, G. N. (2017). The role of gender in asking questions at cool stars 18 and 19. *arXiv preprint arXiv:1704.05260*.

Shan, X. (2021). Does minority status drive women out of male-dominated fields? *Unpublished manuscript*.

Singer, S., Herrington, M., and Menipaz, E. (2018). *Global Entrepreneurship Monitor 2017/18*. Global Entrepreneurship Research Association, London Business School.

Sviatschi, M. M. and Trako, I. (2024). Gender violence, enforcement, and human capital: Evidence from women's justice centers in peru. *Journal of Development Economics*, 168:103262.

The World Justice Project (2023). Access to justice remains elusive, new data shows. `https://worldjusticeproject.org/news/access-justice-remains-elusive-new-data-shows`. Accessed: 2025-07-29.

The World Justice Project (2024). Rule of law index 2024: Zambia – civil justice. `https://worldjusticeproject.org/rule-of-law-index/country/2024/Zambia/Civil%20Justice/`. Accessed: 2025-07-29.

Ubfal, D. (2024). Supporting women-led businesses: Narrative review of recent causal evidence. *The World Bank Research Observer*, page lkae012.

Udry, C. (1996). Gender, agricultural production, and the theory of the household. *Journal of political Economy*, 104(5):1010–1046.

Wilson, R. and Herrnstein, R. (1985). *Crime and Human Nature: The Definitive Study of the Causes of Crime*. Simon and Schuster, New York.

Xu, Y., Burns, M., Wen, F., Thor, E. D., Zuo, B., Coley, J. D., and Rhodes, M. (2022). How culture shapes social categorization and inductive reasoning: A developmental comparison between the united states and china. *Journal of Cognition and Development*, 23(5):644–659.

Tables

Table 1: Trust and Cooperation Among Small-Scale Businesses in Lusaka, by Gender

	Census Data			Experiment 1			Experiment 2		
	Male N=1579	Fem N=637	Diff	Male N=347	Fem N=128	Diff	Male N=290	Fem N=304	Diff
Panel A: Trust									
Trust Strangers	0.29	0.20	-0.09***	0.18	0.07	-0.11***			
	(0.45)	(0.40)	(0.02)	(0.02)	(0.02)	(0.04)			
Trust Neighbours	0.54	0.43	-0.11***	0.69	0.62	-0.07			
	(0.50)	(0.50)	(0.02)	(0.02)	(0.04)	(0.05)			
Trust GSS	0.10	0.04	-0.06***	0.17	0.10	-0.07*	0.07	0.05	-0.02
	(0.31)	(0.19)	(0.01)	(0.02)	(0.03)	(0.04)	(0.02)	(0.01)	(0.02)
Feel safe leaving shop unattended				0.44	0.32	-0.12**	0.26	0.15	-0.10***
				(0.03)	(0.04)	(0.05)	(0.03)	(0.02)	(0.03)
Panel B: Cooperation									
Joint Buy	0.35	0.27	-0.08***	0.28	0.26	-0.02	0.24	0.22	-0.02
	(0.48)	(0.44)	(0.02)	(0.02)	(0.04)	(0.05)	(0.03)	(0.02)	(0.03)
Lent	0.55	0.44	-0.11***	0.69	0.62	-0.06	0.35	0.37	0.02
	(0.50)	(0.50)	(0.02)	(0.02)	(0.04)	(0.05)	(0.03)	(0.03)	(0.04)
Advice	0.76	0.71	-0.05**	0.91	0.88	-0.03	0.23	0.22	-0.01
	(0.43)	(0.45)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)	(0.03)
Share Order	0.58	0.54	-0.04*	0.62	0.54	-0.08	0.39	0.32	-0.06
	(0.49)	(0.50)	(0.02)	(0.03)	(0.04)	(0.05)	(0.03)	(0.03)	(0.04)
Ever cooperated in any activity	0.86	0.81	-0.06***	0.95	0.95	0.00	0.71	0.63	-0.08**
	(0.34)	(0.40)	(0.02)	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.04)
Coop Average	0.56	0.49	-0.07***	0.63	0.58	-0.05*	0.30	0.28	-0.02
	(0.33)	(0.34)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)	(0.02)
Ever cooperated with a woman							0.22	0.53	0.31***
							(0.02)	(0.03)	(0.04)
Ever cooperated with a man							0.68	0.26	-0.42***
							(0.03)	(0.03)	(0.04)
Share of partners of different gender							0.11	0.18	0.06***
							(0.01)	(0.02)	(0.02)
Number of Partners							1.67	1.28	-0.39***
							(0.09)	(0.07)	(0.12)
Panel C: Knowledge sharing									
Learnt from family or friends	0.42	0.25	-0.17***	0.42	0.20	-0.22***			
	(0.49)	(0.43)	(0.02)	(0.03)	(0.04)	(0.05)			
Learnt from entrepreneur	0.35	0.17	-0.18***	0.40	0.23	-0.17***			
	(0.48)	(0.38)	(0.02)	(0.03)	(0.04)	(0.05)			
Learnt from entrepreneur in same sector	0.26	0.13	-0.13***	0.31	0.16	-0.15***			
	(0.44)	(0.33)	(0.02)	(0.02)	(0.03)	(0.05)			
Learnt through formal training	0.15	0.52	0.36***	0.12	0.49	0.38***			
	(0.36)	(0.50)	(0.02)	(0.02)	(0.04)	(0.04)			
Ever taught others	0.72	0.64	-0.08***	0.83	0.68	-0.15***			
	(0.45)	(0.48)	(0.02)	(0.02)	(0.04)	(0.04)			
Talk frequently with others about business	0.52	0.49	-0.04	0.62	0.55	-0.07			
	(0.50)	(0.50)	(0.02)	(0.03)	(0.04)	(0.05)			

Stars denote statistical significance of two-sided t-test by gender. *** p<0.01, ** p<0.05, * p<0.1.

Panel A reports gender differences in trust, Panel B in cooperation, and Panel C in knowledge sharing. 'Trust in Strangers' and 'Trust in Neighbors' are scaled 1–4 and coded as dummies (1 = scores 3 or 4). 'Feel safe leaving shop unattended' equals 1 if respondents agree/strongly agree they can leave their shop unattended for 30 minutes. 'Trust GSS' equals 1 if people say that most people can be trusted. 'Joint Buy', 'Lent', 'Advice' and 'Share Order' are dummies that indicate whether a person engaged in the relevant activity in the past 12 months. 'Coop Average' is the average of these four. 'Ever cooperated with a male/female' equals 1 if any past-year partner was of that gender. 'Share of partners of different gender' is the share of opposite-gender partners among the six most important collaborators. 'Number of partners' counts business partners in the past year. 'Learnt from' variables reflect responses to 'Who taught you this job?'. 'Formal training' typically refers to classroom instruction by the government, NGOs, or associations. 'Ever taught others' equals 1 if the respondent trained someone who later started a business. 'Talk frequently with others about business' is based on reported frequency of business discussions, coded as 1 if more than once a week or daily. The sample in the first three columns comes from the Census of Small-Scale Manufacturers in Lusaka (N=2,216), for columns (4) to (6) from participants in experiment 1 (N=477) and for columns (7) to (9) from participants in experiment 2 (N=594). The samples from the two experiments include both manufacturing and non-manufacturing firms, and mainly businesses located inside markets.

Table 2: Learning History and Sales

	Learnt from			Logged Sales		Employment	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Owner	Owner in same sector	Family or friends	Good Week	Bad Week	Level	Growth
Female	-0.184***	-0.132***	-0.153***	-0.623***	-0.733***	-0.533***	-0.0680
	(0.0215)	(0.0191)	(0.0236)	(0.0961)	(0.116)	(0.168)	(0.0640)
Informal learning				0.0473	0.0612	-0.0783	0.0563
				(0.0804)	(0.0956)	(0.152)	(0.0648)
Female*Informal learning				0.361***	0.352**	0.299	-0.0290
				(0.125)	(0.157)	(0.202)	(0.0841)
Observations	2081	2081	2072	1962	1752	2079	2004
Adjusted R^2	0.051	0.035	0.074	0.162	0.133	0.176	0.020
Mean Dep Var	.3	.22	.37	7.08	5.86	.91	.15
Controls: Business Density	✓	✓	✓	✓	✓	✓	✓
Controls: Individual	✓	✓	✓	✓	✓	✓	✓
Controls: Industry FE							

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

The variables 'Owner', 'Owner in same sector', 'Family or Friends' are dummies that indicate whether an entrepreneur learnt the job from each category. 'Informal learning' is a dummy for whether the owner learnt the business from any of the previous three categories. Business Density Controls include a dummy variable for whether the business is located within 100 meters of a market, the total number of businesses within 100 meters, the number of business from the same sector within 100 meters and population density in the ward where the business is located. Individual Characteristics Controls includes owner's age, business age, how many days the business owner spends working in the business, educational dummies and whether business owner is married or not. In columns (4) and (5), the dependent variables are the log of the answers given when asked the sales in a good week and the sales in a bad week. If the person could not provide an exact number, we also asked for an upper and lower bound. We then imputed the average of the bounds to the sales variables. In Column (6), the dependent variable is the number of full-time employees. In Column (7), the dependent variable is the growth rate in the number of full-time employees since the start of the business. The sample comes from our Census of Small-Scale Manufacturers in Lusaka (N=2,216). Differences in observations across columns are due to missing controls or outcomes.

Table 3: Local Adjudication, Cooperation and Economic Performance

	(1)	(2)	(3)	(4)	(5)
			Collaboration Activities		
	Avg coop	Share order	Lent	Advice	Joint buy
Panel A: Cooperation					
Female	-0.102***	-0.078*	-0.140***	-0.082**	-0.106**
	(0.029)	(0.040)	(0.040)	(0.037)	(0.043)
	[0.000]				
In market	0.081**	0.116**	0.108**	-0.016	0.114**
	(0.031)	(0.052)	(0.053)	(0.047)	(0.048)
	[0.016]				
Female × In market	0.081**	0.072	0.109*	0.064	0.080
	(0.040)	(0.051)	(0.056)	(0.047)	(0.057)
	[0.024]				
P-value (In + Fem.*In = 0)	0.00	0.01	0.00	0.47	0.00
P-value (Fem + Fem.*In = 0)	0.41	0.83	0.40	0.54	0.43
Observations	2215	2215	2215	2215	2215
Adjusted R^2	0.072	0.049	0.053	0.044	0.038
Mean Dep Var	0.54	0.57	0.52	0.75	0.32
Controls: Business Density	✓	✓	✓	✓	✓
Controls: Individual Characteristics					
Controls: Industry FE					
Controls: Marketplace FE	✓	✓	✓	✓	✓

	(1)	(2)	(3)	(4)	(5)
			Mean Sales		Employment
	Index	Levels	Logs	Levels	Growth
Panel B: Economic Performance					
Female	-0.134**	-776.935**	-0.270**	-0.221*	-0.143**
	(0.057)	(309.117)	(0.107)	(0.127)	(0.067)
	[0.002]				
In market	-0.220**	-988.149*	-0.414**	-0.367	-0.009
	(0.095)	(527.903)	(0.177)	(0.252)	(0.109)
	[0.000]				
Female × In market	0.151***	767.364**	0.140	0.324***	0.197***
	(0.054)	(311.017)	(0.123)	(0.121)	(0.067)
	[0.111]				
P-value (In + Fem.*In = 0)	0.52	0.68	0.09	0.88	0.16
P-value (Fem + Fem.*In = 0)	0.52	0.94	0.07	0.09	0.16
Observations	2213	2088	2080	2211	2103
Adjusted R^2	0.326	0.204	0.301	0.284	0.047
Mean Dep Var	0.02	1872.35	6.66	1.94	0.15
Controls: Business Density	✓	✓	✓	✓	✓
Controls: Individual Characteristics					
Controls: Industry FE	✓	✓	✓	✓	✓
Controls: Marketplace FE	✓	✓	✓	✓	✓

Clustered standard errors in parentheses. Multiple hypothesis testing corrected p-values in brackets.
*** p<0.01, ** p<0.05, * p<0.1.

The variables in Panel A 'Joint Buy', 'Lent', 'Advice' and 'Share Order' are dummies that indicate whether a person engaged in the relevant activity in the past 12 months. The variable 'Avg coop' is an index of cooperative behavior, calculated as the average of the four dummies. The variable 'In market' is an indicator for the business being located inside a formal market. In Panel B the Index of economic performance is obtained by standardizing mean sales in levels and employment levels (separately) and then by taking the average of the two standardized values. The sample comes from our Census of Small-Scale Manufacturers in Lusaka (N=2,216). Clustered standard errors at the marketplace level in parentheses. MHT corrected p values following (List et al., 2019) are shown in brackets and correct across the two models shown in Column 1, Panels A and B.

Table 4: Experiment 1 - Games' Behavior and Earnings: No Chief vs Gender Blind Chief

Panel A: Behavior	(1) Tokens Investor Sent	(2)	(3) Average Return Ratio	(4)
Female	-1.243***	-0.925*	0.000	0.013
	(0.430)	(0.520)	(0.026)	(0.031)
	[0.026]	[0.016]	[0.996]	[0.422]
	{0.006}	{0.083}	{0.996}	{0.657}
Chief Gender Blind	-0.085	-0.224	0.025	0.013
	(0.382)	(0.410)	(0.020)	(0.020)
	[0.873]	[0.606]	[0.150]	[0.546]
	{0.828}	{0.594}	{0.215}	{0.535}
Female × Chief Gender Blind	1.610**	1.706**	-0.052	-0.028
	(0.682)	(0.753)	(0.042)	(0.044)
	[0.052]	[0.052]	[0.201]	[0.547]
	{0.024}	{0.025}	{0.227}	{0.534}
P-value(Ch Blind + Fem.*Ch Blind = 0)	0.01	0.02	0.44	0.69
P-value(Fem + Fem.*Ch Blind = 0)	0.49	0.19	0.10	0.64
Observations	189	189	190	190
Mean Dep Var	3.96	3.96	0.43	0.43
Controls: Individual		✓		✓

Panel B: Earnings	(1) Investor's earnings	(2)	(3) Trustee's earnings	(4)
Female	-0.210	-0.238	0.355	0.430
	(0.400)	(0.351)	(1.202)	(1.339)
	[0.412]	[0.451]	[0.847]	[0.659]
	{0.608}	{0.474}	{0.770}	{0.742}
Chief Gender Blind	0.505	0.406	-0.373	-0.616
	(0.317)	(0.343)	(0.668)	(0.710)
	[0.101]	[0.220]	[0.665]	[0.398]
	{0.106}	{0.233}	{0.572}	{0.370}
Female × Chief Gender Blind	0.743	0.929	1.055	0.414
	(0.694)	(0.711)	(1.552)	(1.590)
	[0.310]	[0.221]	[0.547]	[0.805]
	{0.326}	{0.230}	{0.495}	{0.788}
P-value(Ch Blind + Fem.*Ch Blind = 0)	0.04	0.04	0.62	0.89
P-value(Fem + Fem.*Ch Blind = 0)	0.34	0.25	0.15	0.44
Observations	189	189	190	190
Mean Dep Var	11.17	11.17	6.75	6.75
Controls: Individual		✓		✓

Robust standard errors in parentheses. Randomization inference and multiple hypothesis testing corrected p-values in brackets and curly brackets respectively.
* p<0.10, ** p<0.05, ***p<0.01

Panel A: Columns (1)–(2) report regressions on the number of tokens sent by the Investor and Columns (3)–(4) on the Trustee's average return ratio. Panel B: Columns (1)–(2) report regressions on the Investor's earnings and Columns (3)–(4) on the Trustee's earnings. The variable "Chief Gender Blind" is an indicator equal to one for being assigned to the treatment with access to the (gender-blind) chief, and zero for having no institutional access. Individual controls include: sector (manufacturing), age and age squared, education dummies (primary, secondary), marital status, days and hours worked per week, number of collaborations in the past year, business age, trust (agreement with "most people can be trusted" and "comfortable leaving shop unattended for 30 minutes"), smartphone ownership, number of full-time employees, awareness of the Small Claims Court, ethnicity (Bemba, Chewa, Nsenga), and whether the business experienced a dispute in the past 12 months. All regressions include a dummy for the round of play. Data are from experiment 1, and the number of observations reflects all participants assigned to the Control or Chief treatment arms (N = 191). Each participant appears once, based on their decision in either the investor or trustee role. Participants assigned to the Small Claims Court treatment (N = 192) or to the inside/outside treatments are excluded here, thus the control group only includes businesses inside markets playing against other insiders. Results for the other treatment groups are reported in Appendix tables. Multiple Hypotheses Testing (MHT) corrections (following List et al. (2019)) adjust separately for trust (primary outcome) versus the others. Each of these families is estimated twice, once without and once with controls. Thus the MHT-corrected p-values are adjusted within the following four groups of coefficients: (i) Panel A, Col 1; (ii) Panel A, Col 2; (iii) Panel A, Col 3 and Panel B, Cols 1 and 3; (iv) Panel A, Col 4 and Panel B, Cols 2 and 4.

Table 5: Experiment 2 - Games' Behavior and Earnings: Gender Blind vs Gender Aware Chief

Panel A: Behavior	(1) Tokens Investor Sent	(2) Tokens Investor Sent	(3) Average Return Ratio	(4) Average Return Ratio
Female	-0.904**	-0.717	-0.007	-0.000
	(0.401)	(0.440)	(0.017)	(0.021)
	[0.030]	[0.040]	[0.990]	[1.000]
	{0.055}	{0.193}	{0.708}	{0.987}
Chief: Gender Aware	-0.335	-0.272	0.044**	0.045**
	(0.409)	(0.411)	(0.018)	(0.019)
	[0.480]	[0.570]	[0.040]	[0.040]
	{0.420}	{0.502}	{0.099}	{0.127}
FemaleXChief: Gender Aware	1.286**	1.127**	-0.063***	-0.068***
	(0.551)	(0.568)	(0.023)	(0.025)
	[0.020]	[0.050]	[0.010]	[0.010]
	{0.058}	{0.115}	{0.054}	{0.051}
P-value (Fem + Fem*Ch Aware = 0)	0.314	0.347	0.000	0.000
P-value (Ch Aware + Fem.*Ch Aware = 0)	0.011	0.032	0.211	0.139
Observations	297	297	297	297
Mean Dep Var	4.431	4.431	0.442	0.442
Controls: Individual		✓		✓

Panel B: Earnings	(1) Investor's Earnings	(2) Investor's Earnings	(3) Trustee's Earnings	(4) Trustee's Earnings
Female	-0.452	-0.648*	1.156*	1.359*
	(0.316)	(0.382)	(0.685)	(0.707)
	[0.350]	[0.260]	[0.220]	[0.120]
	{0.551}	{0.415}	{0.431}	{0.292}
Chief: Gender Aware	0.346	0.450	0.379	0.596
	(0.370)	(0.359)	(0.583)	(0.606)
	[0.390]	[0.200]	[0.530]	[0.380]
	{0.837}	{0.639}	{0.761}	{0.690}
FemaleXChief: Gender Aware	0.375	0.373	-0.704	-1.078
	(0.482)	(0.482)	(0.891)	(0.908)
	[0.470]	[0.450]	[0.420]	[0.230]
	{0.829}	{0.685}	{0.867}	{0.619}
P-value (Fem + Fem*Ch Aware = 0)	0.834	0.521	0.427	0.649
P-value (Ch Aware + Fem.*Ch Aware = 0)	0.021	0.015	0.629	0.461
Observations	297	297	297	297
Mean Dep Var	11.939	11.939	6.923	6.923
Controls: Individual		✓		✓

Robust standard errors in parentheses. Randomization inference and multiple hypothesis testing corrected p-values in brackets and curly brackets respectively.
* p<0.10, ** p<0.05, ***p<0.01

Panel A: Columns (1)–(2) report regressions on the number of tokens sent by the Investor and Columns (3)–(4) on the Trustee's average return ratio. Panel B: Columns (1)–(2) report regressions on the Investor's earnings and Columns (3)–(4) on the Trustee's earnings. The variable "Chief: Gender Aware" is an indicator equal to one for being assigned to the treatment with access to the gender-aware chief, and zero for being assigned to the gender-blind chief. Individual controls include: age, age squared, dummies for primary and secondary education, marital status, business age, days per week and hours per day worked, a dummy for having had any collaboration in the past 12 months, the number of collaborations in that period, trust in others ("most people can be trusted"), comfort leaving the shop unattended for 30 minutes, dummies for Bemba, Chewa, or Nsenga ethnicity, and a dummy for having had a business dispute in the past 12 months. Data are from experiment 2 (N=594). Each participant played only once as Investor or Trustee. Multiple Hypotheses Testing (MHT) corrections (following (List et al., 2019)) adjust separately for trust (primary outcome) versus the others. Thus the MHT-corrected p-values are adjusted within the following four groups of coefficients: (i) Panel A, Col 1; (ii) Panel A, Col 2; (iii) Panel A, Col 3 and Panel B, Cols 1 and 3; (iv) Panel A, Col 4 and Panel B, Cols 2 and 4.

Figures

Figure 1: Theoretical Predictions
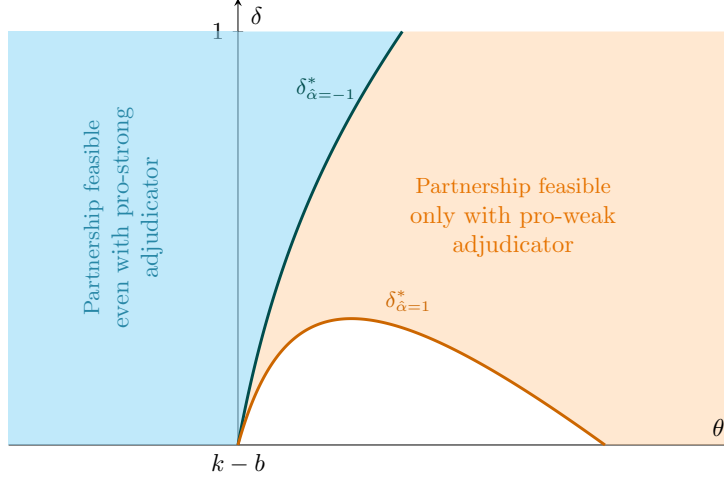
This figure illustrates the predictions of Proposition 1. Partnership is sustainable when legal quality δ exceeds the threshold $\delta^*(\theta) = \frac{b-k+\theta}{\Delta+b-k+\theta} - \hat{\alpha}\theta$. In the plot we fix $b = k = 1$ and $\Delta = 0.2$, so $\delta^*(\theta) = \frac{\theta}{0.2+\theta} - \hat{\alpha}\theta$. **Curves.** The blue solid curve shows $\delta^*(\theta)$ when the adjudicator is biased in favor of the powerful ($\hat{\alpha} = -1$); the orange solid curve shows the threshold when the adjudicator favors the weak ($\hat{\alpha} = +1$). **Shaded regions.** For power gaps $\theta \leq k - b = 0$ (blue strip on the left of $k - b = 0$) the entrepreneur is strong enough to deter cheating on her own, so partnership always forms. For $\theta > k - b = 0$ partnership is feasible for (θ, δ) pairs lying *above* the relevant curve. The orange wedge marks situations in which partnership can be sustained only by a protector of the weak, whereas the blue wedge above the blue curve indicates values where even a pro-strong adjudicator suffices.

Figure 2: Distribution across Industries by Gender
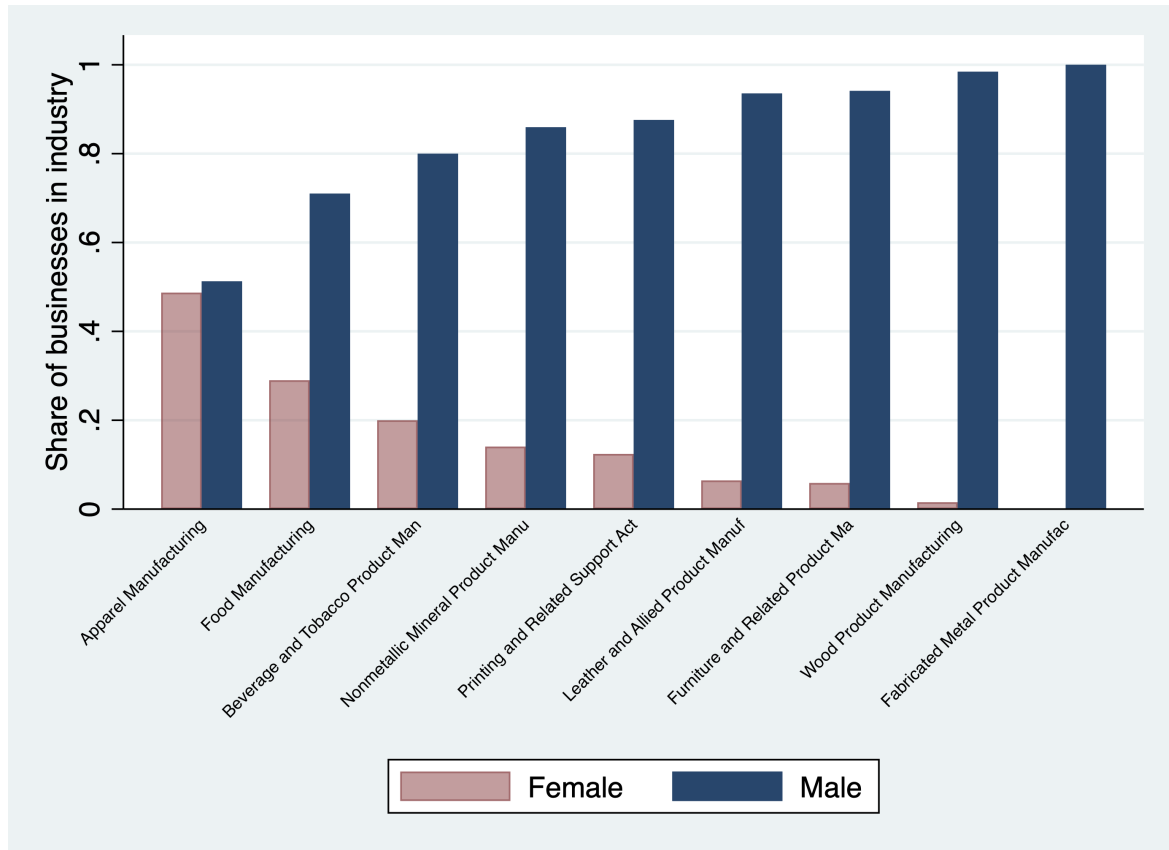
This figure shows the industries in which manufacturers from the Lusaka Census of Urban Entrepreneurs operate, by gender. Gender data are available for 97% of businesses in manufacturing, construction and mining with less than 20 employees (N=3723). The graph shows only industries with more than 10 businesses in our data.

Figure 3: Local Institutions in Lusaka

◆ Police Stations

▮ Markets
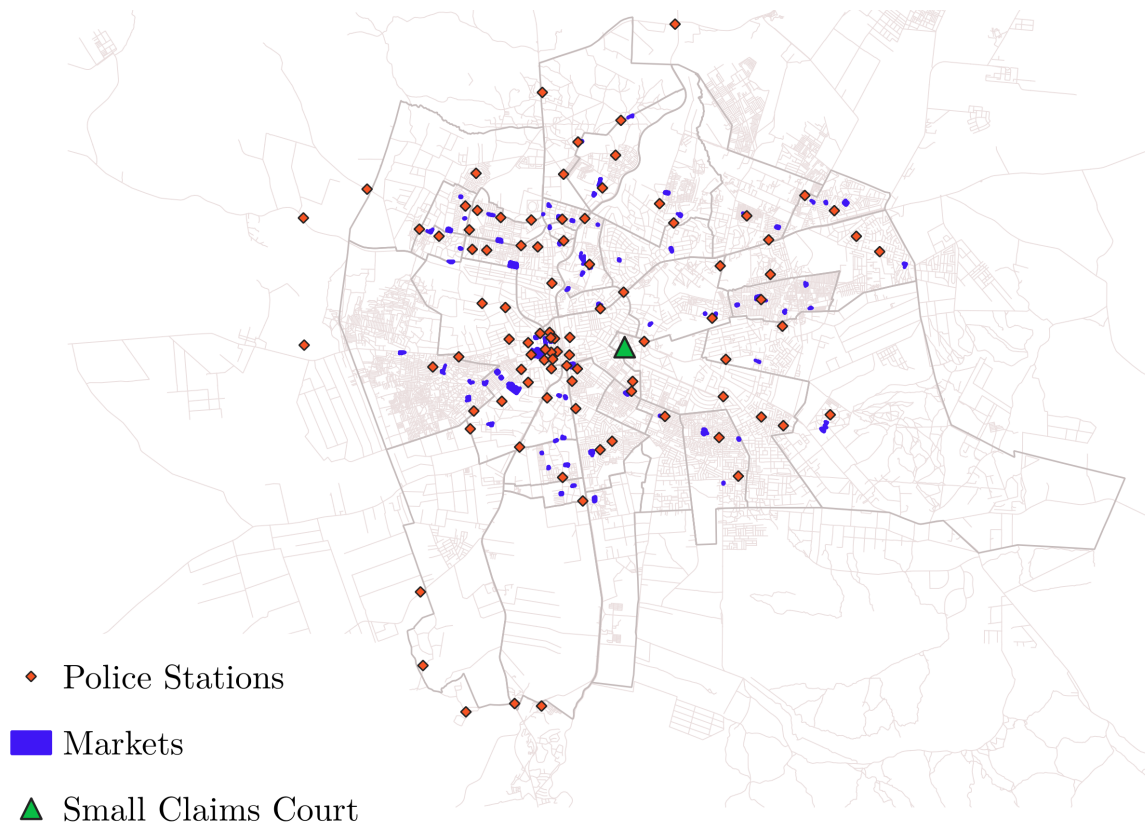
▲ Small Claims Court

The figure shows the spacial distribution of local institutions in Lusaka. The blue dots show where markets are located, the green triangle shows the Small Claims Court and the red diamonds show the location of police stations.

Figure 4: Cooperation by Gender and Market Location
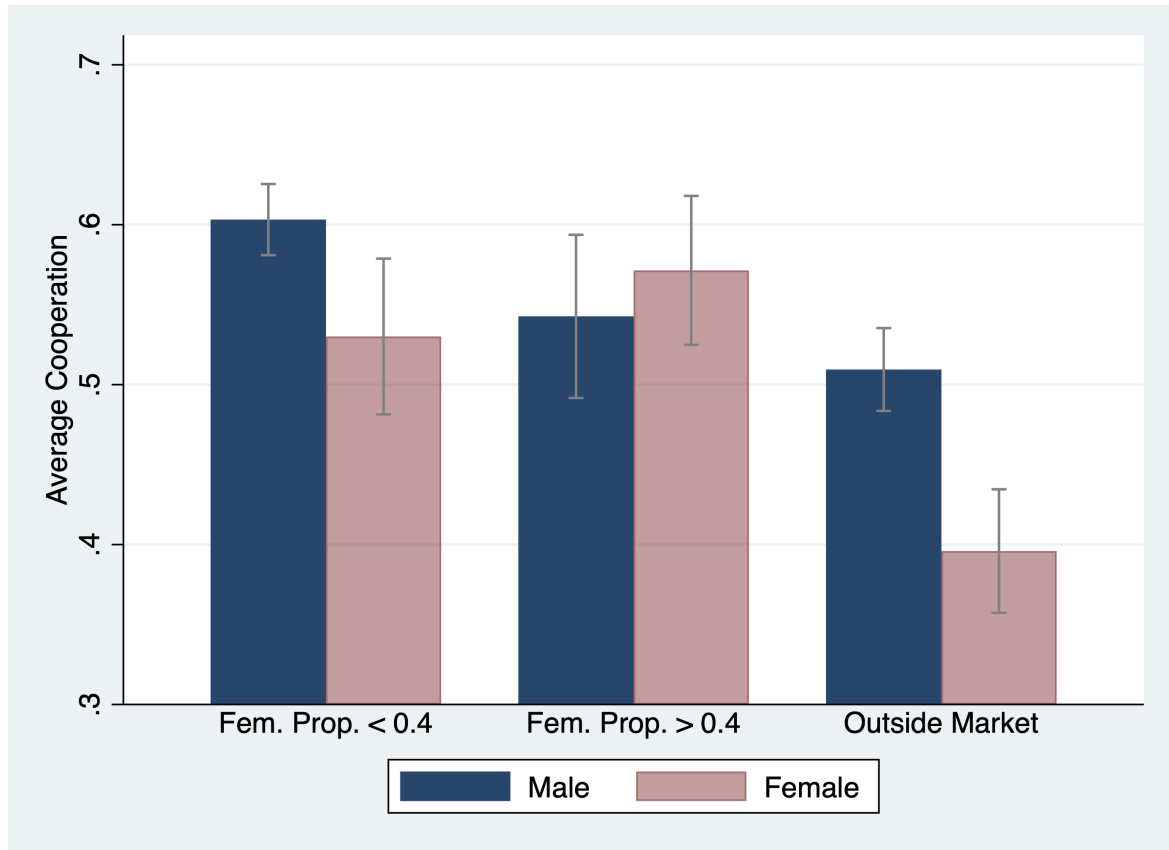
This figure shows the average cooperation of female and male-led businesses in markets with a large or a small share of female businesses, or outside of markets. The threshold of 0.4 corresponds to the 75th percentile of the distribution of gender shares across markets. The sample comes from our Census of Small-Scale Manufacturers in Lusaka (N=2,216).

Figure 5: Cooperation by Gender and Industry
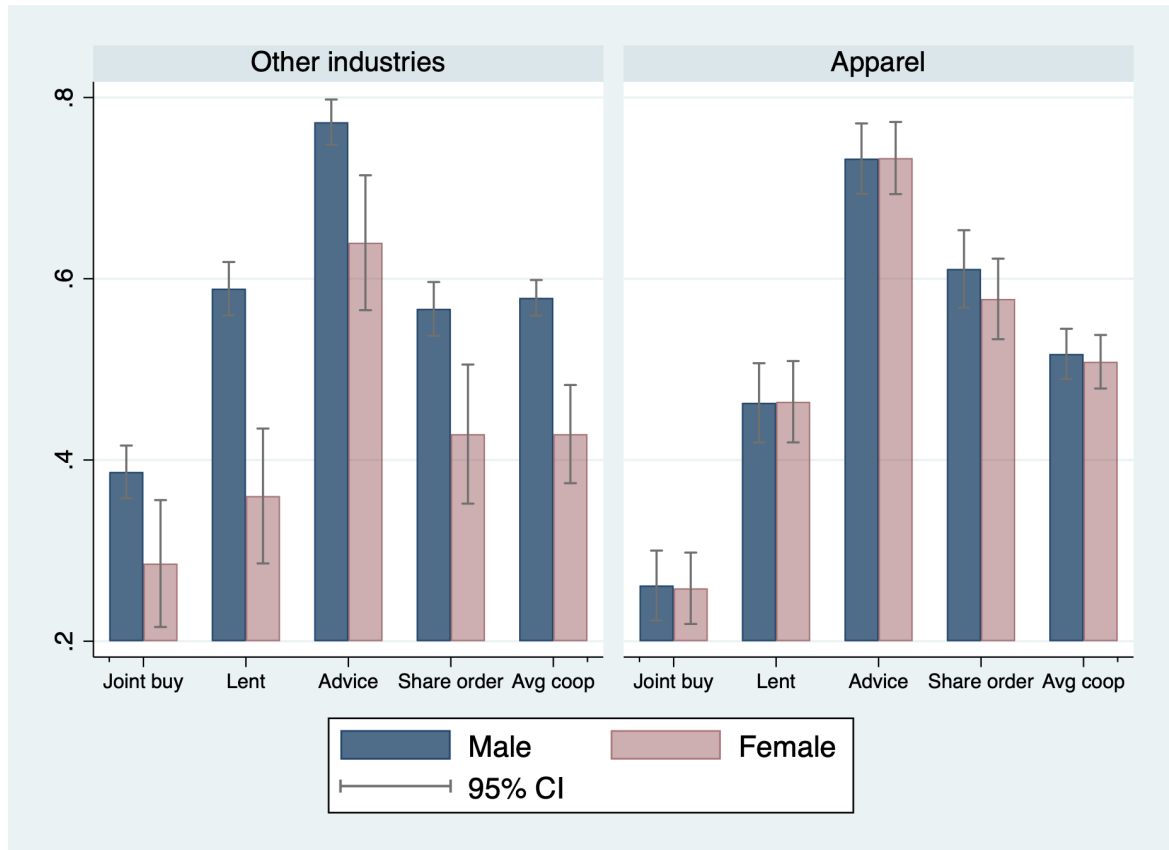
This figure shows the proportion of men and women engaging in each of the collaboration activities we asked about, and their level of average cooperation (last bar). The variables "Joint Buy", "Lent", "Advice" and "Share Order" are dummies that indicate whether a person ever engaged in the relevant activity. The variable "Coop Average" is an index of cooperative behavior, calculated as a simple average of the four dummies "Joint Buy", "Lent", "Advice" and "Share Order". The sample comes from our Census of Small-Scale Manufacturers in Lusaka (N=2,216).

Figure 6: Gender Gap in Cooperation by Chief's Gender Bias
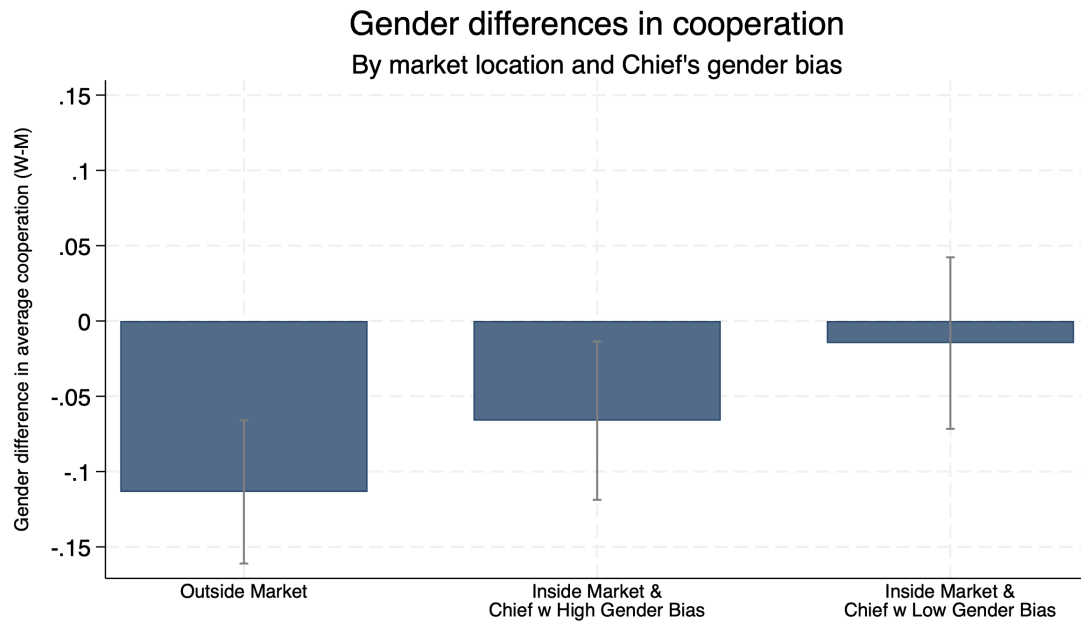
This figure shows the difference in average cooperation between women and men outside markets, inside markets with a chief with above median gender bias and inside markets with a chief with below median gender bias. Average cooperation is an index of cooperative behavior, calculated as the average of the four dummies "Joint Buy", "Lent", "Advice" and "Share Order" defined in the main tables. Chief's gender bias is an index constructed from the following survey variables: whether the chief follows subjective rules to arbitrate disputes, whether the chief thinks that a woman does not usually win in inter-gender disputes, the difference in the chief's effectiveness in catching perpetrators of thefts against men vs women, World Value Survey (WVS) and World Justice Project (WJP) gender attitudes. More details on the construction of the gender bias index are given in Appendix D. Error bars show 95% confidence intervals.

Figure 7: Cooperation by Chief's Characteristics: Women vs Men Inside Markets

(a). Cooperation and Chief's Bias

Chief effect on average cooperation
Women vs men: inside markets

SE clustered at the market-area level, N=1297
P-value (Female+Index*Female=0): 0.10 / 0.90 / 0.40
P-value (Index+Index*Female=0): 0.60 / 0.20 / 0.10
P-value (Female=Index*Female): 0.70 / 0.10 / 0.10

(b). Cooperation and Chief's Legal Competence

Chief effect on average cooperation
Women vs men: inside markets

SE clustered at the market-area level, N=1297
P-value (Female+Index*Female=0): 0.40 / 0.10 / 0.00
P-value (Index+Index*Female=0): 0.10 / 0.00 / 0.00
P-value (Female=Index*Female): 0.50 / 0.20 / 0.20

(c). Cooperation and Chief's Managerial Duties

Chief effect on average cooperation
Women vs men: inside markets

SE clustered at the market-area level, N=1297
P-value (Female+Index*Female=0): 0.20 / 0.90 / 0.60
P-value (Index+Index*Female=0): 0.90 / 0.30 / 0.30
P-value (Female=Index*Female): 0.30 / 0.00 / 0.00

Each figure in Panels (a), (b) and (c) show the coefficients of three regressions, which differ only in the controls used. All the specifications regress average cooperation on a dummy variable for being a female owner ("Female"), a dummy variable for having a chief with above average score on a certain index and the interaction between the two. In Panel (a), the dummy is for above-median score in the index of gender bias. In Panel (b), the variable is for above-median score in the index of legal competence and in Panel (c) for above-median managerial duties. The top coefficient (green) is for a regression with area density controls, the intermediate (yellow) coefficient is for a regression with density controls and industry fixed effects, and the bottom coefficient (purple) is for a regression with density controls, individual controls and industry fixed effects (NAICS3). Average cooperation is an index of cooperative behavior, calculated as the average of the four dummies "Joint Buy", "Lent", "Advice" and "Share Order" defined in the main tables. Error bars show 95% and 90% confidence intervals. The construction of the indices on gender bias, legal competence and managerial duties is explained in Appendix D.

Figure 8: Experiment 1: Trust and Trustworthiness, Control vs Gender Blind Chief

(a). Trust
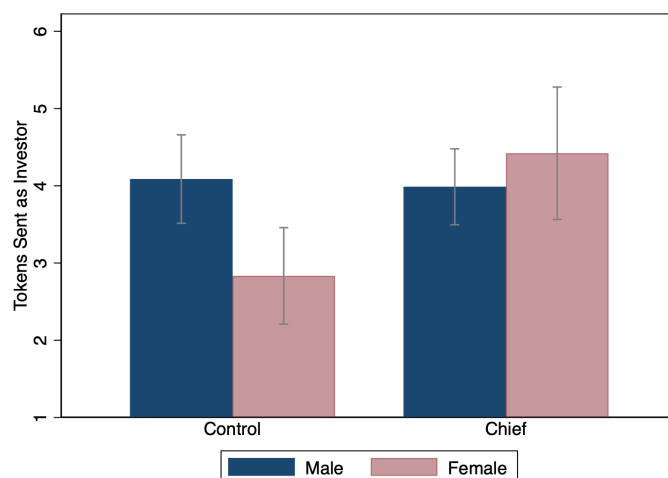
(b). Trustworthiness
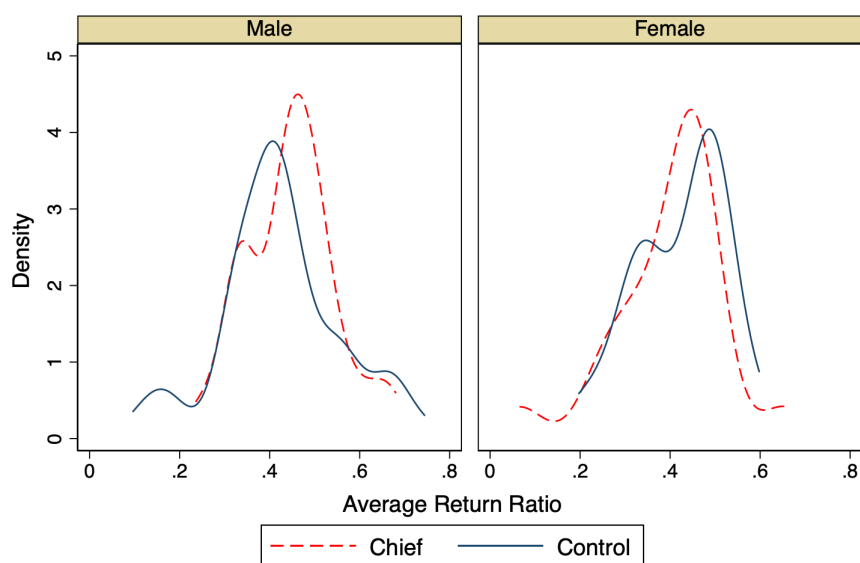
Pooled Genders SD-test Control vs Chief: p=.06
Pooled Genders test of mean differences Control vs Chief: p=.46

Figure (a) shows the number of tokens sent by gender and treatment group. Figure (b) shows the distribution of average return ratios by gender and treatment. For both Figures, we exclude participants assigned to the SCC and outside/inside treatments, thus the total number of observations is N=191 and the control group only includes businesses inside markets playing with other insiders.

Figure 9: Experiment 1: Number of Tokens Sent by Market Chief's Gender Bias: Control vs Gender Blind Chief

(a). Tokens Sent by Women
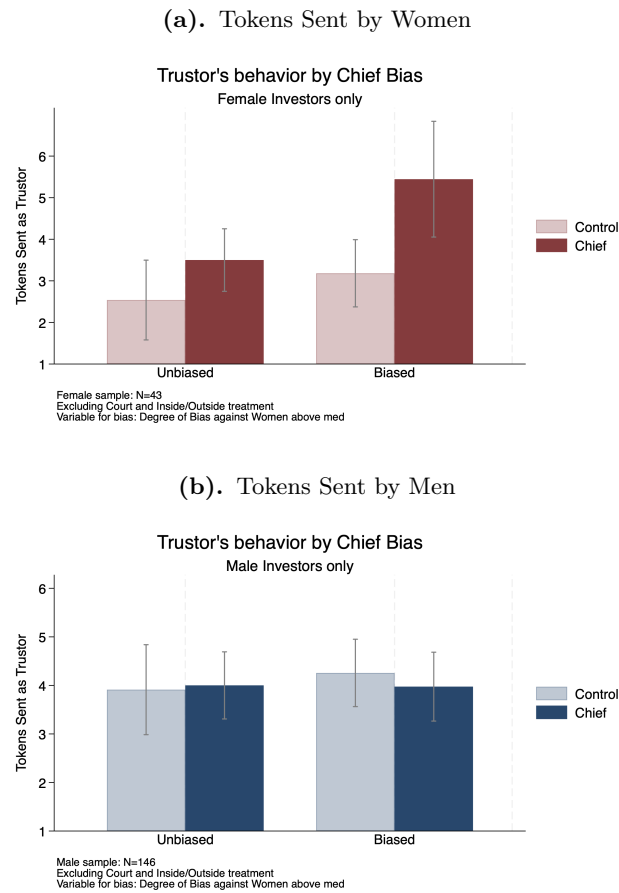
(b). Tokens Sent by Men

This figure shows the number of tokens sent by women (panel a) and men (panel b), splitting the sample of markets depending on the median market chief's gender bias. The index construction is described in Appendix D. Data are from experiment 1 (N=477) and the market chiefs survey. For this Figure, we exclude participants assigned to the SCC and outside/inside treatments, thus the total number of observations is N=191 and the control group only includes businesses inside markets playing with other insiders.

Figure 10: Experiment 1: Distribution of Total Earnings between Investor and Trustee by Investor's Gender: Control vs Gender Blind Chief
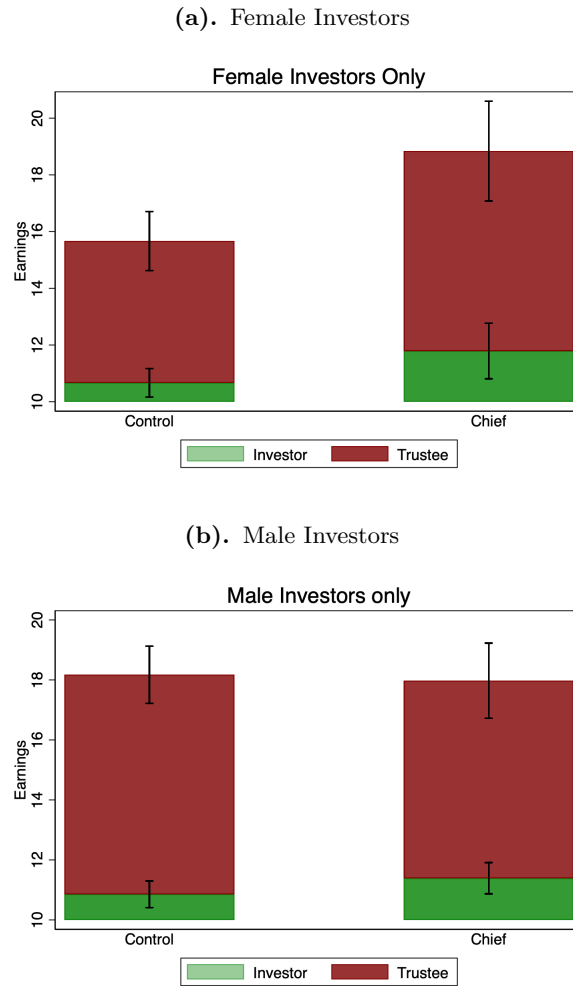
(a). Female Investors

(b). Male Investors

This figure shows the distribution of total final earnings in rounds with a female Investor (Panel a) and male investor (Panel b). Data are from experiment 1 (N=477). For this Figure, we exclude participants assigned to the SCC and outside/inside treatments, thus the total number of observations is N=191 and the control group only includes businesses inside markets playing with other insiders.

Figure 11: Experiment 2: Revelation Trust Premium across Markets, by Gender

(a). Raw Differences in Tokens Sent

Difference in tokens sent by the investor
in Chief: Gender Aware vs Gender Blind

Women in Chief: Gender Aware - Gender Blind
Men in Chief: Gender Aware - Gender Blind
DID: (F Ch:Aware - F Ch:Blind) - (M Ch:Aware - M Ch:Blind)

90% confidence intervals
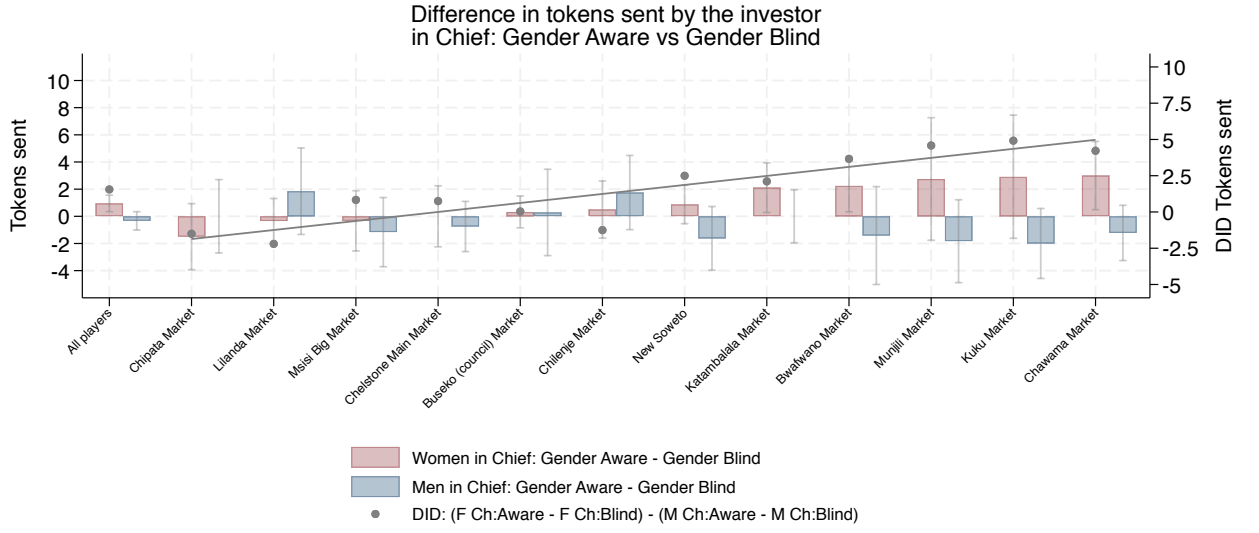
(b). Regression Coefficients

Market specific regressions

Coeff: Female X Chief Gender Aware
Coeff: Chief Gender Aware

95% Confidence intervals.
Regression of coefficient for the interaction term on the coefficient for the Chief Gender Aware dummy:
Coefficient= -1.738; P-value = 0.000
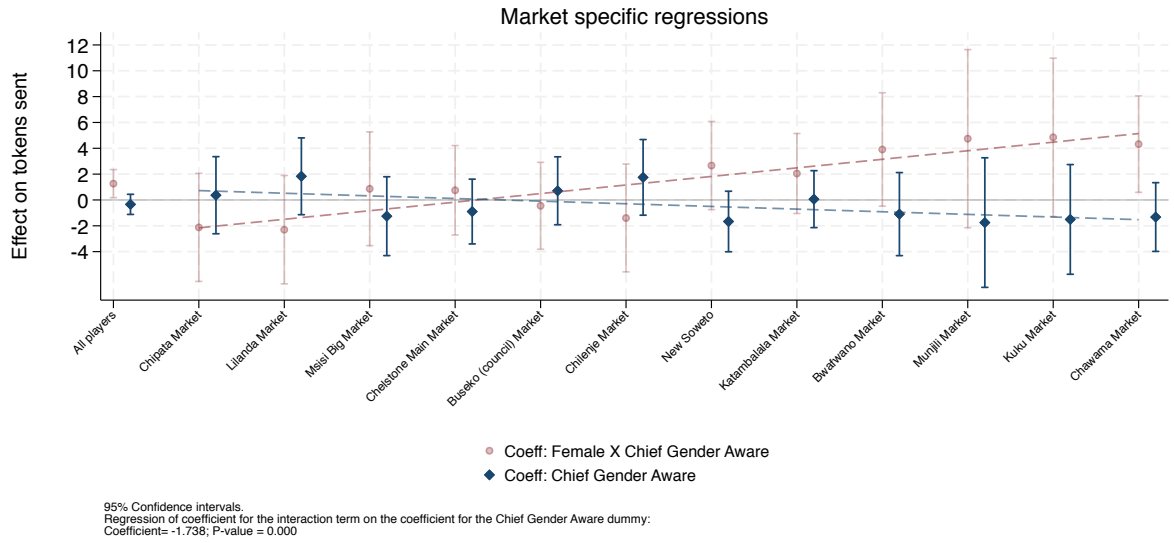
Panel (a) plots the difference in tokens sent by Investors between the "Chief: Gender Aware" and "Chief: Gender Blind" treatments (i.e. the "Revelation Trust Premium"), separately by gender. The first set of bars shows the average differences; the remaining twelve show market-specific differences, with markets ordered by the size of the Gender Aware–Gender Blind difference for female investors. The solid grey line and dots indicate the difference-in-differences. Panel (b) reports coefficients from a regression of tokens sent on a Chief: a female dummy, a "Chief: Gender Aware" dummy (blue coefficients), and the interaction of this treatment indicator with the female dummy (pink coefficients). The first two coefficients show average effects; the remaining twelve show market-specific effects, with markets ordered as in Panel (a). Data are from Experiment 2 ($N = 153$ women, $N = 144$ men, investors only).

61

Figure 12: Experiment 2: Gender Differences in the Revelation Trust Premium, Chief Gender Bias and Female Power

(a). Chief's gender bias and gender salience

90% confidence intervals.
P-values shown in the columns on the right.
Coefficients in the first columns on the right unadjusted p-values in the second column, MHT corrected p-values in the third column
MHT corrected p-values in black are calculated across indices, while those in gray are corrected within indices.
All variables in the lefthand side column are standardized.

(b). Female Power

All regressors are standardized. 90% confidence intervals.
Coefficients in the first columns on the right unadjusted p-values in the second column, MHT corrected p-values in the third column
All variables in the lefthand side column are standardized

The graphs show the coefficients of several univariate market-level regressions of the gender difference in the revelation trust premium (i.e. tokens sent by women in (Chief:Aware - Chief:Blind) - men in (Chief:Aware - Chief:Blind)) with variables related to the chief or market-level female power. Panel (a) shows correlations of the double difference in trust on chief gender bias, expected chief's behavior in the game and gender salience and Panel (b) on proxies for female power. The construction of the indices is described in Appendix D. "Chief Gender Bias" is the index defined in previous figures. "Chief Gender Salience" is constructed from a *triad picture task* and equals 1 if the chief uses gender (in more than half of scenarios) as a criterion to categorize people. "Participants' expected chief behavior" is an index of market-level differences in beliefs about how chiefs treat complaints from women versus men, from incentivized guesses about the chief's likelihood of intervening ("cutoff") and the amount they would reallocate ("change in return ratio"). "Means of Social Punishment" is the female market-level average of three indicators: whether a woman reports family support, police support, or support from another business owner in case of a dispute. "Historical Female Empowerment" is the female market-level average of three ethnicity-based measures (patrilineal vs. matrilineal descent, patrilocal vs. matrilocal residence, and gendered agricultural participation), using Murdock's Ethnographic Atlas. All regressors are standardized, and all variables are at the market level (N=12). The three columns on the right of each figure report the coefficients, p-values and MHT-adjusted p-values from each regression. For the main indices, we perform MHT corrections within each panel across indices (so we correct for three hypotheses within panel (a) and two hypotheses within panel (b)). We also report MHT-corrected p-values within each index (e.g., for the chief gender bias index, we correct across its five components).

Appendices

A   Proofs

We begin by recapitulating the structure and timing of the model:

Stage 0: The entrepreneur E chooses whether to enter a market, knowing the share of men m_i and the available adjudicators.

Stage 1: E can make and sell a product of value π, but it will cost π to make it alone. If E partners with P, production costs will be collectively lower, but E must make an offer to compensate P for production costs, and the payment of this compensation is unconditional.

Stage 2: P decides whether to work or shirk. If P works, P pays $q < \frac{1}{2}\pi$. If P shirks, P pays $q - b$. E pays q and works.

Stage 3: E pays P the contracted payment of q.

Stage 4: If P worked, E sells the good and receives π. If P shirked, E must pay $b + \Delta$ to undo the harm and receive the payment π. The condition $b + \Delta < \pi$ ensures that E will always remediate if work has begun. The condition $\pi < 2q + \Delta$ (total costs with shirking) ensures that sharking means that the partnership has produced a social loss.

If P shirks, E can use "social punishment" to punish P. The size of the punishment is related to θ, the power gap between E and P where $\theta > 0$ means P has more power than E. The punishment is $\max[k - \theta, 0]$.

The adjudicator intervenes with probability $\delta + \alpha\theta$. If the adjudicator intervenes, then P must pay E the damages $b + \Delta$.

Proof of Proposition 1.

We proceed recursively starting in Stage 2, with P's decision about whether or not to shirk. The benefit of shirking is b. The cost of shirking is $b + \Delta$ if the arbitrator intervenes, and that intervention occurs with probability $\delta + \hat{\alpha}\theta$. If the arbitrator does not intervene, then shirking carries a cost of $\max[k - \theta, 0]$ and that occurs with probability $1 - \delta - \hat{\alpha}\theta$.

If $k - \theta > b$, then shirking will never occur, as the benefit is less than both the private and public punishments that can occur. Consequently, P will always partner with E.

If $0 < k - \theta < b$, then shirking provides P with a benefit of b and expected costs of $(1 - \delta - \hat{\alpha}\theta)(k - \theta) + (\delta + \hat{\alpha}\theta)(b + \Delta)$. In this region, shirking occurs if and only if $(1 - \delta - \hat{\alpha}\theta)(k - \theta) + (\delta + \hat{\alpha}\theta)(b + \Delta) < b$ or $(\delta + \hat{\alpha}\theta)\Delta < (1 - \delta - \hat{\alpha}\theta)(b - k + \theta)$ or $\frac{b - k + \theta}{\Delta + b - k + \theta} - \hat{\alpha}\theta > \delta$ and we define $\delta^* = \frac{b - k + \theta}{\Delta + b - k + \theta} - \hat{\alpha}\theta$. Hence, if $0 < k - b < \theta$, E will partner if and only if $\delta > \delta^*$. As $b - k + \theta > 0$, δ^* is greater than $-\alpha\theta$. As $\Delta > 0$, δ^* is less than $1 - \hat{\alpha}\theta$. Differentiation gives us that δ^* is always declining with Δ and k and increasing with b, decreasing with $\hat{\alpha}$ if and only if $\theta > 0$ and increasing with θ if and only if $\frac{\Delta}{(\Delta + b - k + \theta)^2} > \hat{\alpha}$.

If $0 > k - \theta$, then shirking provides P with a benefit of b and expected costs of $(\delta + \hat{\alpha}\theta)(b + \Delta)$. Consequently, in this region, shirking only occurs if $(\delta + \hat{\alpha}\theta)\Delta < (1 - \delta - \hat{\alpha}\theta)b$ or $\frac{b}{\Delta + b} - \hat{\alpha}\theta > \delta$. In this region, we define $\delta^* = \frac{b}{\Delta + b} - \hat{\alpha}\theta$, and so δ^* remains greater than $-\alpha\theta$ and less than $1 - \hat{\alpha}\theta$. Hence, if $0 > k - \theta$, E will also partner if and only if $\delta > \delta^*$. In this region, δ^* is always declining with Δ and increasing with b, decreasing with $\hat{\alpha}$ if and only if $\theta > 0$ and increasing with θ if and only if $0 > \hat{\alpha}$.

Proof of Proposition 2.

If θ is unknown to the adjudicator, so that the probability of intervention is δ, then E believes that shirking provides P with a benefit of b but a cost of $b + \Delta$ with probability

δ and an expected cost of $k - \hat{\theta}$ with probability $1 - \delta$, so that total expected costs are $(1 - \delta)(k - \hat{\theta}) + \delta(b + \Delta)$. Hence shirking occurs if and only if $(1 - \delta)(k - \hat{\theta}) + \delta(b + \Delta) < b$ or $\delta\Delta < (1 - \delta)(b - k + \hat{\theta})$ or $\frac{b - k + \hat{\theta}}{\Delta + b - k + \hat{\theta}} = \delta^{**} > \delta$, where $\delta^{**} = \delta^* + \hat{\alpha}\hat{\theta}$. Differentiation gives us that δ^{**} is always declining with Δ and k and increasing with b and $\hat{\theta}$.

If $\hat{\alpha} > 0$ and $\hat{\theta} > 0$, and $b + \Delta$ is greater than $k - \hat{\theta}$, then E believes that she will be weakly better off when the gap in social power between P and E is revealed. If there is revelation, E believes that P believes that P's expected punishment from shirking is $(1 - \delta - \hat{\alpha}\theta)(k - \hat{\theta}) + (\delta + \hat{\alpha}\hat{\theta})(b + \Delta)$, which is greater than $(1 - \delta)(k - \hat{\theta}) + \delta(b + \Delta)$ if $\hat{\alpha} > 0$ and $\hat{\theta} > 0$. If $k - \hat{\theta}$ is greater than $b + \Delta$, then E believes that there will be no shirking with or without revelation, and consequently, revelation is equivalent to non-revelation. Consequently, there are parameter values for which E will partner with P when she expects the gap in social power to be revealed to the judge, but not when there is no revelation.

Proof of Proposition 3.

With probability $1 - m_i$, E is partnered with a woman and there is no shirking and, consequently, her returns are $\pi - 2q$.

With probability m_i, E is partnered with a male P, who will shirk if and only if b is greater than the expected penalty from shirking: $(1 - \delta - \hat{\alpha}\theta_G)(k - \theta_G) + (\delta + \hat{\alpha}\theta_G)(b + \Delta)$ or $(\delta + \hat{\alpha}_i\theta_G)\Delta < (1 - \delta - \hat{\alpha}_i\theta_G)(b - k + \theta_G)$ or $\hat{\alpha}_i\theta_G < 1 - \delta - \frac{\Delta}{\Delta + b + \theta_G - k}$.

If shirking doesn't occur with either gender, then the returns to E will equal $(\pi - 2q)$ and entry will only occur if $(\pi - 2q) > f_i$. If shirking occurs with a male P, the expected returns to E will drop to $(1 - m_i)(\pi - 2q)$ and entry will only occur if $(1 - m_i)(\pi - 2q) > f_i$.

B   Games Procedures

B.1   Experiment 1

Participant sampling and recruitment. Six surveyors and two recruiters were hired to conduct the games and were managed by a research assistant. Typically two days were spent in each market. We created two lists of randomly-ordered businesses for each market; one list for businesses located inside the market and one for businesses located outside the market. Manufacturers were placed at the top of each list. The two recruiters met early and were given the randomized lists of entrepreneurs; they then set up appointments with potential participants, following the order of the list. If a marketeer could not be located or refused to participate, the marketeer was replaced by the next female on the list. As soon as all women had been exhausted, the recruiters simply moved to the next marketeer on the list. In markets where the response rate was low, or we failed to find many businesses, or many businesses were closed, we also surveyed and played the games with non-manufacturers. All entrepreneurs on the outside list were screened extensively to ensure that they truly did not belong to the market.

Treatment assignment. In general, the first day at a market, the games were played with inside-inside pairs, wherein both players worked within the market. The corresponding three treatments arms for these pairs were control, Court and Chief. Generally, the second day at a market, the games were played with inside-outside pairs, wherein one player was a marketeer and the other an entrepreneur located outside of the market. The corresponding two treatment arms for these pairs were control and Court. Surveys were conducted first, and then the games. However, if there were delays in reaching a participant or if a player refused to play the games after having completed the survey, the order was switched. The order in which each treatment arm was conducted was randomized each market to limit selection bias due to logistical factors (e.g., time of day, eagerness of participants).

Players were never told who they were playing against, however, they knew whether the player was located inside or outside the market. Furthermore, after playing the first round as Investor or Trustee, they played the second round as Trustee (Investor respectively) against a new player, so that no two players played twice together. It was made clear to them that they would be randomly assigned to a new partner after the first round and they were reminded whether this player was inside or outside the market and if they (or the opposing player) could complain to the Chief or Court or neither.

As some of the markets were small, it is possible that participants had heard of the games we were conducting before we reached them. Thus, we cannot fully exclude the possibility that some players may have played with more information to begin with. However, this is more likely for the second day in a specific market, as the games conducted in a specific day followed back-to-back, thus there was not much time for entrepreneurs to talk in-between. We present results from day 2 only in the Appendix of the paper, as the chief treatment was never implemented in that day. Furthermore, the field team sought to minimize the attention drawn to themselves by sending the two recruiters into the depths of the markets to find participants, whilst the surveyors would generally just directly be taken to the participants.

Field procedures. All protocols were approved by the local IRB and the researchers trained and supervised the field team. We ran the games on groups of six entrepreneurs at time, with one surveyor assigned to one entrepreneur. For each market, a switch matrix was constructed to inform surveyors whether their assigned entrepreneur was to be Player A (Investor) or Player B (Trustee) first, and who the entrepreneur was to play in each round. For logistical reasons, surveyors whose respondents were located outside of the market always

started the games with the respondent being the Trustee, therefore all outsiders played the games in the same order, first as Trustees then as Investors. Once finishing a survey with a respondent, a surveyor sent a text to their corresponding surveyor to inform that they were about to start the instructions for the games in the following format: "surveyornum-initials-CensusID-AB/BA-start (e.g. 4-D-6230-AB-start)".

The instructions for the games were given to the respondent in written as well as oral form in the respondent's preferred language, and surveyors asked the respondent check questions to ensure that they fully understood the pay-out rules, who they were playing against, and the possibilities for complaint when relevant. Once both surveyors in a pair had received the "ready" text message from their partner surveyor for that round and had ensured that the participant understood the rules, they started with the game. Everyone had to correctly answer understanding checks about the rules of the games with the surveyors before proceeding.

The surveyor of the Investor sent the surveyor of the corresponding Trustee the number of tokens his player had chosen to send in the following format "SurveyorNum-Surname-Game-Player-TokensSent (e.g. 4-AD-R1-A3)". The surveyor of the Trustee then responded with the number of tokens the Trustee had decided to send back in the following format "SurveyorNum-Surname-Game-Player-TokensReturned (e.g. 2-PB-R1-B2)". The surveyors were instructed to always use neutral language to inform the respondents of the amount that had been sent (or sent back) to them.

Complaint process. In the Court and Chief treatments, the Investor, upon being told how many tokens the Trustee had sent back, was asked if they wanted to complain to the Court (chief) or not. The surveyor then messaged the corresponding surveyor whether or not the Investor intended to complain (format: A-Comp OR A-NoComp).

In the case of the chief treatment, the recruiter would ask the market's chief in real-time how s/he wanted to settle the complaint and sent their response in the same format to the surveyor of the Investor. The chief her/himself was given information (both oral and written) prior to the games commencing, which explained the game and made clear her/his role as an arbitrator during the games. Similar to the respondents, they were asked questions to check whether they had understood the games and their role in the games. If/ when a complaint reached the chief, the recruiter who was assigned to her/him would announce the complaint and ask for her/his decision in the following way: "There has been a complaint from a player. In this game, player A (the Investor) sent XX number of tokens to B (the Trustee), which means B received XX number of tokens. B sent back XX tokens. A has complained to you. Do you wish to redistribute the tokens? If so, how?"

In the case of the Court treatment, the recruitment officer already had a completed matrix of decisions from the Court (obtained from the Small Claims Court before the start of the games field work). The recruitment officer examined the matrix and sent the Court's decision to the surveyor of the Investor in the following format: "ANumTokensReceived-BNumTokensReceived. E.g. A6-B3". A similar message was sent to communicate the chief's decision.

Instructions. Excerpts from the instructions of Players A and B about the complaint process follow:

> *From Player A's instructions*
> If you think that the number of tokens sent back by Player B is not fair, you can ask us to call the chief (senior clerk at the Small Claims Court) on your behalf, to decide how many tokens each of you should get. The chief (senior clerk) will then decide how many tokens each of you should get. The chief's (senior clerk's)

ruling is based only on your choices and the choices of player B, and we will tell you their decision accordingly. The chief (senior clerk) does not know anything about you and the other player, only your choices. Player B will know that you can complain to the chief (senior clerk). In the case that you complain the final division of tokens will be determined by the ruling of the chief (Small Claims Court).

From Player B's instructions

Player A can ask us to consult the chief (senior legal clerk at the Small Claims Court) on his/her behalf, to decide how many tokens each of you should get. The chief's (senior clerk's) ruling is based only on your choices and the choices of player A, and we will tell you their decision accordingly. The chief (senior clerk) does not know anything about you and the other player, only your choices. In the case that player A complains, the final division of tokens will be determined by the ruling of the chief (Small Claims Court).

B.2   Experiment 2

Market sampling. Sampling proceeded in four steps. First, we constructed a preliminary set of 63 eligible markets by dropping

- markets where we ever piloted surveys and/or experiments (12 markets)

- markets where we did not have a chief bias measure from the previous chief survey (5 markets, none of which were pilot markets)

We dropped pilot markets to avoid contamination, and markets without a chief bias measure because our sampling strategy uses this as a balance variable.

Second, we paired markets in the preliminary set based on observables. The matching was stratified by the market's type (cooperative, council, or street) and size, measured by the number of businesses. There were six size groups: the largest two markets, the smallest two markets, and quartiles of the remaining markets. We separated out the largest and smallest markets because they are substantially different in size from the next closest ones. The two largest markets have 1,695 and 1,051 businesses, and the next largest has 841. The two smallest markets have 7 and 8 businesses, and the next smallest has 22. Within each market type-size group, we formed pairs to globally minimize within pair distance on the following characteristics: chief management quality index, chief bias index, number of manufacturing firms, number of female-owned manufacturing firms, number of tailors, number of firms in accommodation and food services (NAICS 2), and number of retailers (NAICS 2).[57] The first two variables are from the previous chief survey. The rest are from the census of businesses. If there was an odd number of markets in the stratification group, we left out the market that, when excluded, led to the minimum within pair distance amongst the remaining markets. This process yielded 29 pairs and 5 unpaired markets.

Third, we selected one eligible market from each pair according to the following procedure:

- If both markets in the pair were not in the first experiment, we randomly selected one.

- If one market in the pair was in the first experiment and the other market was not, we automatically selected the new one.

- If both markets in the pair were in the first experiment, we randomly selected one.

---

[57]We implemented the minimum distance matching using Blossom's algorithm.

We paired all of the markets to ensure we had a fully representative sample of unused markets for future follow-up work. However, for the current experiment, we excluded the smallest pair, the largest pair, and all markets in the first quartile of size. These exclusions were based on practical considerations: our participant recruitment strategy (discussed below) would be infeasible in the largest markets, and the smallest markets were unlikely to have sufficient businesses in our target participant industries. After these exclusions, we were left with 22 markets: 12 cooperative markets selected from pairs, 6 council markets selected from pairs, 3 unpaired cooperative markets, and 1 unpaired council market. Excluding the first quartile of size dropped all 4 street markets. These markets, however, are not appropriate for our experiment, given that they have a less cohesive governance structure.

Fourth, we listed the number of businesses in our target industries – manufacturing, hairdressing, and restaurants – in the 18 paired markets, as well as the unpaired council market. Our final experimental sample of 12 markets was selected based on the density of target industries. We have a total of five markets where we conducted both experiments, and seven new markets where we conducted only experiment 2.[58]

Participant sampling. For Player A, sampling was stratified by gender and sector (manufacturing vs. non-manufacturing). In each market, we recruited an equal number of men and women, with an overall target number based on the market's size from the listing. We set the split between manufacturers and non-manufacturers based on availability, aiming for the same share of manufacturers for men and women. For Player B, we sampled an equal number of men and women, but did not distinguish between industries.

A consent team visited the market to get consent from the target number of Player As in each gender-industry, as well as the target number of Player Bs (half women and half men). Particularly for manufacturers, it was sometimes the case that the target was equal to the number of available businesses. However, when there was a surplus of businesses, a consenter began at one end of the market and alternated sections for recruitment. For example, if the market was organized in aisles, as many are, they recruited from every other aisle. Each consenter was assigned a particular gender-industry to recruit so they could keep track of this. One to two days later, the data collection team visited the market to conduct the actual study activities.

Treatment assignment and A-B pairing. The main body of the paper describes the chief treatment (gender-blind vs. gender-aware). In addition, at the end of the trust game we elicited incentivized belief measures on how the participant's chief would arbitrate a dispute brought by a man or a woman. The gender of the hypothetical complainant was itself randomly assigned, as described here.

For Player A's, the chief-belief treatment pair was stratified by gender-industry. For Player B's, the belief treatment was stratified by gender, and pairing was stratified by Player A-gender-industry and Player B gender.

To achieve stratification, team supervisors ran a real time assignment protocol based on two instruments. For Player A, these were:

1. The target number for each gender-industry, with a chief-belief treatment pair pre-assigned to each slot (i.e. first participant, second participant, and so on). For example, if slot #1 for female manufacturers was assigned *Gender-Blind* and belief questions about men, then the first female manufacturer who participated would be assigned to *Gender-Blind* and belief questions about men.

2. Gender-industry specific lists of consenting businesses, in random order. Businesses

---

[58]Surveyors trained with real participants in 2 of the 7 remaining markets.

in a given gender-industry were to be visited in the randomized sequence, so that treatment (which was preset along visit order) would be orthogonal to the participant.

For Player B, these were:

1. The target number for each Player A-gender-industry and Player B gender pair (i.e. female manufacturer A's with female B's, female manufacturer A's with male B's, and so on), with a belief treatment pre-assigned to each slot.

2. Gender specific lists of consenting businesses, in random order. Businesses in a given gender-industry were to be visited in the randomized sequence.

In a given round of data collection, the supervisors told each surveyor which specific business to visit. If that business was unavailable, the surveyor asked the supervisors for a replacement business until they had a participant finalized. Once the participant was finalized, the supervisors texted the surveyor the assigned treatment.

In each round of data collection, the team tried to recruit Player A's from each gender-industry in proportion to the target, so that time of interview would be orthogonal to the participant. If a participant was unavailable at the time of initial visit, the data collection team made one additional attempt to include them, either by making an appointment or visiting again later.

Field procedures. All protocols were approved by the local IRB and the researchers trained and supervised the field team. Within each market, we first conducted a listing of potential participants within our target industries. Next, enumerators returned to the markets and gathered consents from a targer number of businesses listed in phase 1. The third phase was the entrepreneurs' survey and the chief's survey. Enumerator teams visited the markets after the consent had been collected and approached marketeers that had accepted to take part of the survey. Surveys were carried out following the strict sampling and randomization protocols described above. While the main team carried out these surveys, a designated enumerator surveyed the market chief.

The pre-games survey for marketeers included socio-demographic information, business activity and collaboration, harassment and institutions related questions. After the survey, the trust game followed. Additionally, at the end of the experiment, we elicited incentivized beliefs about the chief's choices when s/he has to arbitrate complaints coming from men or women within the context of the game. The assignment of complaints brought by men or women was randomized between subjects, stratifying by the main strata (market-gender-industry) and treatment assignment (as described above).

The chiefs' survey was a shortened version of the one conducted by the research team in 2021 (described in Appendix D). In addition, after explaining the trust game to the chiefs, we asked them how they would behave in hypothetical game complaints coming for both female and male players.

Instructions. Excerpts from the instructions of Players A and B about the complaint process follow:

> *From Player A's instructions*
> If you think that the number of tokens sent back by Player B is not fair, you can ask us to go to the market chair, Mr/Ms. [Name], on your behalf. The chair will then make a ruling about how many tokens each of you will get. The chair can only reallocate the earnings from the business opportunity, if any, and cannot reallocate the tokens you keep. The chair can decide that you should get more,

less or the same amount of earnings. The decision made by the chair is final. If you complain, your identity will be completely anonymous. Your complaint will be submitted via a form. This protects your privacy.

There are two types of forms that you may be using. [Surveyor: take out blue form and red form for the participant's gender]

For some of the participants in this activity, the chair will not know anything about you or Player B, only your choices and Player B's choices. Here is the complaint form that the chair will see. [Surveyor: show BLUE form]. This form says the number of tokens you send, the earnings and the number of tokens Player B returns.

For some other participants in this activity, the form will contain more information about the participant complaining. Here is the complaint form that the chair will see in this case. [Surveyor: show RED form].This form says the number of tokens you send, the earnings and the number of tokens Player B returns. In addition, the form shows that the complaint comes from a business [man/woman].

In any case, Player B knows you can complain to the chair, but does not see the form used for the complaint.

Whether you are going to use the blue form or the red form is going to be determined randomly. You will press a button on the screen to randomly select which form you will use.

From Player B's instructions

If Player A thinks the number of tokens you send back to him or her is unfair, he or she can ask us to complain to the chair, Mr./Ms. [insert name here], on his or her behalf. The chair will then make a ruling about how many tokens each of you will get. The chair can only reallocate the earnings from the business opportunity, if any, not the tokens Player A keeps. The chair can decide that Player A should get more, less or the same amount of earnings than you sent him/her. The decision made by the chair is final.

If Player A complains, your identity will be completely anonymous. Player A will complain by filling in a form, so that the chair doesn't find out his or her identity.

C   Additional Figures and Tables

Figure C.1:   Female-owned Firms across Countries

(a). Regular Enterprise Survey

Source: Regular Survey. Number of countries: 180.
Horizontal line at the median
For each country, female ownership is computed as the average % across years in the data.
On average, each country has data for 2 different years.

(b). Informal Business Survey

Source: Informal Survey. Number of countries: 33.
Horizontal line at the median
For each country, female ownership is computed as the average % across years in the data.
On average, each country has data for 1 different years.

(c). Micro Entrepreneur Survey

Source: Micro Survey. Number of countries: 37.
Horizontal line at the median
For each country, female ownership is computed as the average % across years in the data.
On average, each country has data for 1 different years.

This Figure shows the share of female-led businesses across countries. Female-led firms are defined as firms with a majority of female owners. The top panel shows share of female-led businesses for the regular WBES survey, the middle one for the informal business survey and the bottom one for the micro business survey.

Figure C.2: Business Earnings by Owner's Gender Across Countries

(a). Regular Enterprise Survey

Source: Regular
Residuals controlling for country, year of data collection and industry FE
K-S test: p-val=0
Number of country-year observations: 261

(b). Informal Business Survey

Source: Informal
Residuals controlling for country, year of data collection
K-S test: p-val=.0139
Number of country-year observations: 31

(c). Micro Entrepreneur Survey

Source: Micro
Residuals controlling for country, year of data collection
K-S test: p-val=.0064
Number of country-year observations: 22
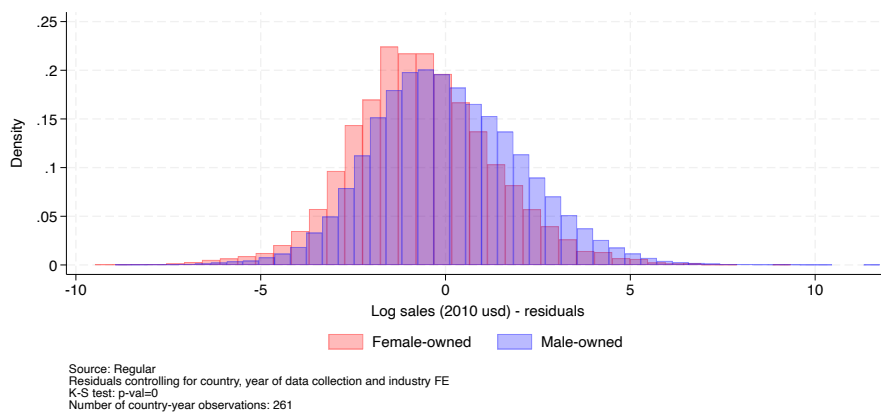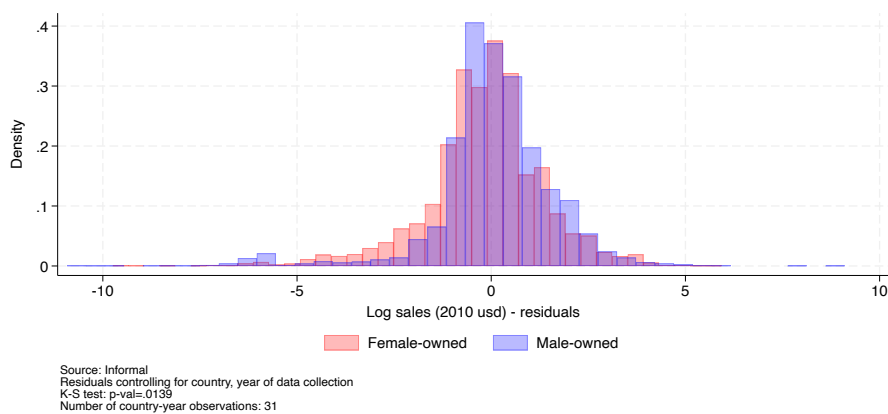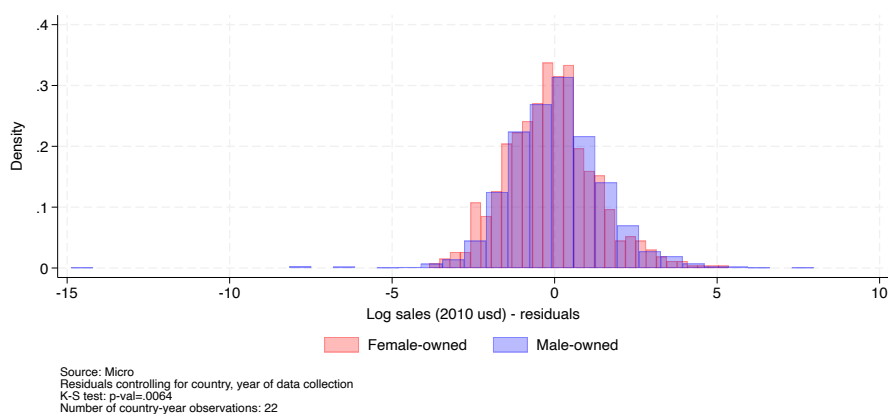
This Figure shows the gender-specific distributions of logsales in 2010 USD after controlling for country, year of collection and, for the regular WBES, industry fixed effect. The top panel shows the distribution for the regular WBES survey, the middle one for the informal business survey and the bottom one for the micro business survey. Female-led businesses are defined as firms with a strict majority of female owners.

Table C.1: Lusaka Census of Urban Entrepreneurs: Descriptive Statistics

	Mean	Observations
Number of Employees	2.27	40,517
Industry (Naics2)		
Retailing	0.51	48,163
Manufacturing	0.08	48,163
Accomodation/food services	0.14	48,163
Other Services	0.13	48,163
All Other Industries	0.13	48,163
Number of Employees by Industry		
Number Emp. Retail	1.12	20,472
Number Emp. Manufacturing	3.86	3,625
Number Emp. accom/food services	2.34	5,854
Number Emp. other services	1.20	5,541
Number Emp. other ind	7.05	4,873
Business is part of a chain	0.04	48,670
Location		
Business has standalone structure (bricks)	0.11	48,670
Business in building with multiple businesses	0.71	48,670
Business has standalone structure (cardboard)	0.03	48,670
Business is in residential house	0.15	48,670

This table presents industry composition and business characteristics from the 2016 Lusaka Census of Urban Entrepreneurs. Observations may be dropped because of missing values.

Figure C.3: Lusaka Census of Urban Entrepreneurs: Coverage
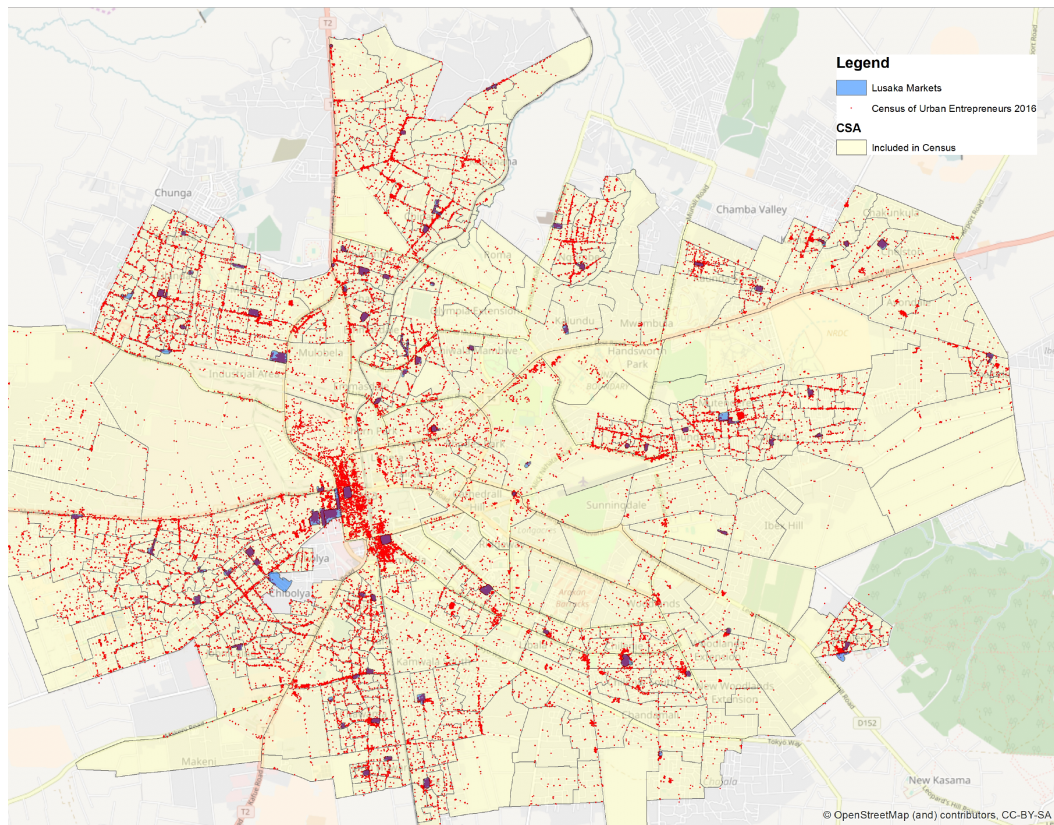
This figure shows the spacial distribution of businesses in the Lusaka Census of Urban Entrepreneurs. It covers 48,163 establishments in Lusaka. The blue squares show where markets are located.

Table C.2: Census of Small-Scale Manufacturers: Comparison of Respondents vs Non-Respondents

	All genders					Female			Male		
	(1) All Businesses Mean (SD)	(2) Eligible Businesses Mean (SD)	(3) No interview Mean (SD)	(4) Interview completed Mean (SD)	(5) Diff(4)-(3)	(6) No interview Mean (SD)	(7) Interview completed Mean (SD)	(8) Diff(7)-(6)	(9) No interview Mean (SD)	(10) Interview completed Mean (SD)	(11) Diff(10)-(9)
Number of Employees	2.27	1.43	2.15	1.00	-1.15***	1.44	0.67	-0.77***	2.47	1.14	-1.33***
	(11.25)	(2.89)	(3.55)	(2.31)	(0.10)	(2.70)	(1.76)	(0.15)	(3.84)	(2.49)	(0.13)
Retailing	0.51	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.50)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Manufacturing	0.08	0.96	0.92	0.99	0.07***	0.97	1.00	0.03***	0.91	0.99	0.08***
	(0.27)	(0.20)	(0.27)	(0.10)	(0.01)	(0.17)	(0.04)	(0.01)	(0.29)	(0.11)	(0.01)
Accomodation/food services	0.14	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.35)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Other services	0.13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	(0.34)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
All other industries	0.13	0.04	0.08	0.01	-0.07***	0.03	0.00	-0.03***	0.09	0.01	-0.08***
	(0.34)	(0.20)	(0.27)	(0.10)	(0.01)	(0.17)	(0.04)	(0.01)	(0.29)	(0.11)	(0.01)
Business is part of a chain	0.04	0.04	0.07	0.03	-0.04***	0.08	0.03	-0.05***	0.06	0.03	-0.03***
	(0.20)	(0.21)	(0.25)	(0.17)	(0.01)	(0.27)	(0.17)	(0.01)	(0.24)	(0.16)	(0.01)
Business has standalone structure (bricks)	0.11	0.09	0.11	0.08	-0.03***	0.06	0.07	0.01	0.12	0.08	-0.04***
	(0.32)	(0.29)	(0.31)	(0.27)	(0.01)	(0.25)	(0.26)	(0.02)	(0.33)	(0.28)	(0.01)
Business in building with multiple businesses	0.71	0.82	0.79	0.84	0.05***	0.86	0.84	-0.02	0.77	0.85	0.08***
	(0.45)	(0.38)	(0.41)	(0.36)	(0.01)	(0.34)	(0.37)	(0.02)	(0.42)	(0.36)	(0.02)
Business has standalone structure (cardboard)	0.03	0.02	0.02	0.02	-0.00	0.01	0.01	0.00	0.02	0.02	-0.00
	(0.17)	(0.15)	(0.15)	(0.14)	(0.00)	(0.11)	(0.12)	(0.01)	(0.15)	(0.15)	(0.01)
Business is in residential house	0.15	0.06	0.08	0.05	-0.02***	0.06	0.07	0.01	0.08	0.04	-0.04***
	(0.35)	(0.24)	(0.27)	(0.22)	(0.01)	(0.24)	(0.26)	(0.02)	(0.27)	(0.21)	(0.01)
Has visible sign	0.61	0.64	0.70	0.60	-0.10***	0.70	0.60	-0.10***	0.70	0.60	-0.10***
	(0.49)	(0.48)	(0.46)	(0.49)	(0.02)	(0.46)	(0.49)	(0.03)	(0.46)	(0.49)	(0.02)
Observations	48670	3803	1587	2216	3803	354	637	991	1109	1579	2688

Standard deviations and standard errors (for columns 5, 8, 11) are reported in parentheses. *$p < 0.10$, **$p < 0.05$, ***$p < 0.01$.

This Table shows summary statistics of all Businesses, depending on their interview status in the Census of Small-Scale Manufacturers, by gender. Columns (1)–(5) present results for the full sample, while columns (6)–(8) and (9)–(11) restrict the comparison to female- and male-owned businesses, respectively. Gender is available only for businesses eligible to participate in the Census of Small-Scale Manufacturers. Column (1) includes all the businesses geocoded in the Census of Urban Entrepreneurs, Column (2) only businesses eligible to participate in the Census of Small-Scale Manufacturers (i.e. in manufacturing, mining or construction with less than 20 employees), Column (3) eligible businesses that were not interviewed, Column (4) eligible businesses that were interviewed and Column (5) the difference between Columns (4) and (3). Columns (6) and (9) show statistics for female-led or male-led, respectively, eligible businesses which were not interviewed. Columns (7) and (10) show statistics for female-led or male-led, respectively, eligible businesses which were successfully interviewed.

Figure C.4: Census of Small-Scale Manufacturers: Sales in Good and Bad Weeks by Gender

(a). Logged Sales in Good Week

Source: Census; Sales variable asymmetrically winsorized at 1%; Male N = 1464; Female N = 605
K-S test p-val = 0; Variance test p-val = 0
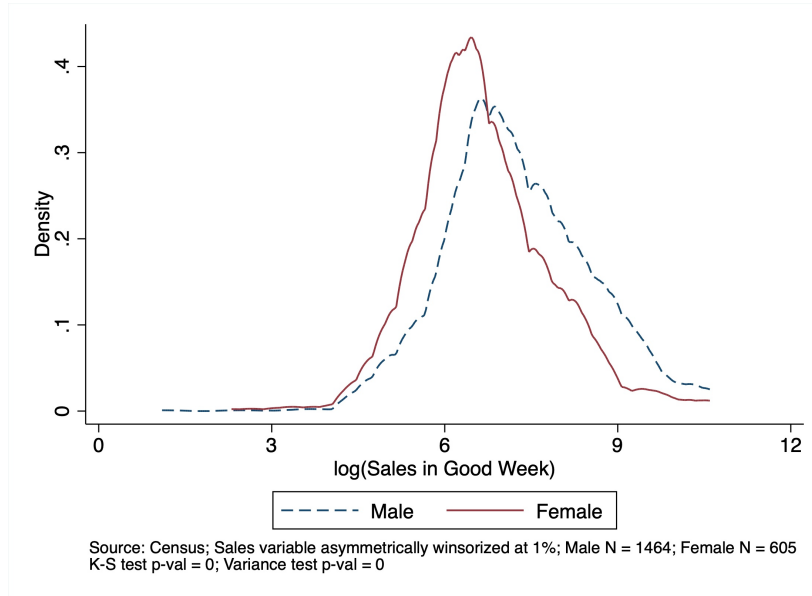
(b). Logged Sales in Bad Week

Source: Census; Sales variable asymmetrically winsorized at 1%; Male N = 1298; Female N = 550
K-S test p-val = 0; Variance test p-val = 1
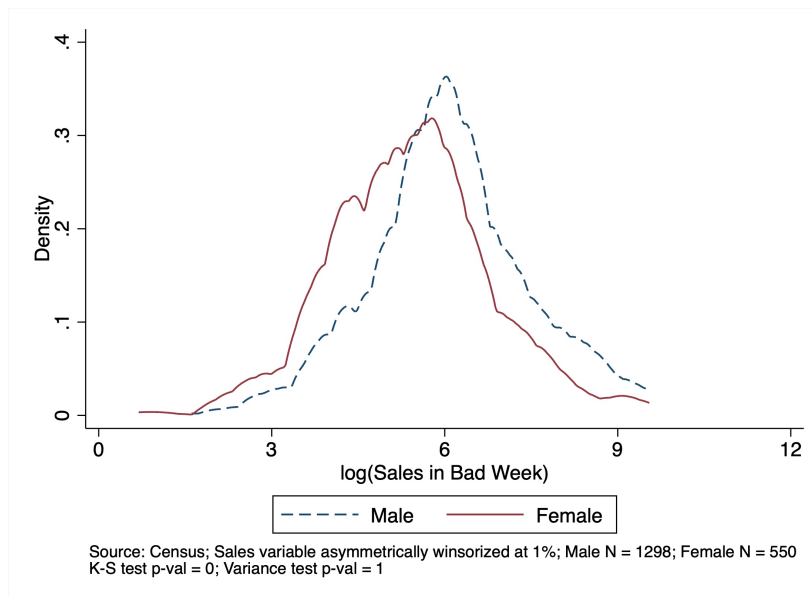
These figures shows the kernel density of logged sales in good and bad weeks by gender. Both sales variables are winsorised at the 99% level to control for outliers before taking the log. The dashed blue line shows the distribution for men and the solid red line for women. The p-value of the Kolmogorov-Smirnov equality of distributions test is 0.00. The p-value of the variance test ratio is 0.001 for sales in good weeks and > 0.50 for sales in bad weeks. The sample comes from our Census of Small-Scale Manufacturers in Lusaka (N=2,216).

Table C.3: Census of Small-Scale Manufacturers: Education, Firm Size and Business Practices by Gender

	Census Data		
	Male N=1579	Fem N=637	Diff
Panel A: Firm Size and Growth			
No. full-time emp.	2.05	1.68	-0.38***
	(0.06)	(0.07)	(0.10)
No. part-time emp.	0.99	0.29	-0.69***
	(0.08)	(0.04)	(0.13)
No. apprentices/unpaid	0.68	0.35	-0.33
	(0.13)	(0.04)	(0.21)
No. family members	0.49	0.30	-0.19***
	(0.03)	(0.03)	(0.04)
No. full-time emp at start	2.06	1.77	-0.30***
	(0.06)	(0.07)	(0.11)
Growth rate in full-time emp.	0.19	0.06	-0.13**
	(0.03)	(0.03)	(0.05)
Positive full-time emp. growth	0.14	0.10	-0.04**
	(0.01)	(0.01)	(0.02)
Panel B: Education			
Training (mgmt/Entrep)	0.21	0.28	0.06***
	(0.01)	(0.02)	(0.02)
University	0.04	0.03	-0.01
	(0.00)	(0.01)	(0.01)
Diploma	0.16	0.19	0.03*
	(0.01)	(0.02)	(0.02)
Completed Secondary	0.23	0.20	-0.03
	(0.01)	(0.02)	(0.02)
No Formal Education	0.02	0.01	-0.01*
	(0.00)	(0.00)	(0.01)
Panel C: Business Practices			
Forecasting Future Budget	0.34	0.32	-0.02
	(0.01)	(0.02)	(0.02)
Keeping Business Records	0.48	0.47	-0.01
	(0.01)	(0.02)	(0.02)
Using Records to Monitor Liquidity	0.87	0.90	0.03
	(0.01)	(0.02)	(0.02)

Stars for two-sided t-test. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Panel A compares firm size and employment growth by gender. The number of full-time employees includes the owner. Panel B compares educational background by gender, and Panel C compares business practices. Data are from the Census of Small Scale Manufacturers (N=2216).

Table C.4: Gender Gap in Performance (Sales and Employment)

(1) Outcome Variable	Regression Controls:				(6) Male Mean of Y	(7) N
	(2) None	(3) +Education	(4) + Industry	(5) + Work/marital		
Mean Sales	−944.72***	−940.17***	−278.02	−166.33	2125.21	2073
Standard Error	(151.848)	(152.822)	(172.949)	(196.013)		
R^2	.014	.071	.141	.142		
Sales Good Week	−1500.2***	−1491.8	−381.91	−243.53	3273.36	2048
	(234.164)	(234.591)	(265.321)	(302.912)		
	.014	.081	.141	.142		
Sales Bad Week	−429.03***	−428.80	−210.49	−149.62	1010.86	2048
	(83.2490)	(85.2699)	(92.8399)	(105.558)		
	.009	.039	.133	.134		
ln(Mean Sales)	-.586***	-.591***	-.199***	-.089	6.82	2066
	(.0579)	(.057)	(.057)	(.0619)		
	.0439	.1040	.25	.2590		
ln(Sales Good Week)	-.586***	-.593***	-.179***	-.083	7.26	2048
	(.057)	(.057)	(.0560)	(.0610)		
	.0450	.1120	.2560	.2630		
ln(Sales Bad Week)	-.658***	-.663	-.321	-.219	6.07	1830
	(.0700)	(.0710)	(.0710)	(.0759)		
	.0450	.0879	.2479	.2579		
Employment Growth	-.128***	-.130***	-.037	-.028	.18	2075
	(.0390)	(.0410)	(.0450)	(.0460)		
	.003	.0170	.0240	.0240		
Employment Level	-.375***	-.377***	.0399	.0460	2.05	2183
	(.0909)	(.0879)	(.0970)	(.1000)		
	.0060	.1159	.1589	.1609		

Robust Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

This table shows explanatory variables of employment and sales for entrepreneurs in the 2016 Census of Small-Scale Manufacturers. The dependent variables are listed in Column 1, and each cell reports the coefficient from a regression of the outcome variable on a dummy for being a female owner. The two rows below the coefficient report the robust standard error and the adjusted r-squared from the same regression. For sales in a good or bad week, if the person could not provide an exact number, we also asked for an upper and lower bound. We then imputed the average of the bounds to the sales variables. The mean sales variable is constructed as the average of sales in a good or bad week. All the sales variables are winsorised at the 99% level to control for outliers. Regressions in Column (2) report the coefficient of a regression of the row variable on the female dummy only. Regressions in Column (3) control for education dummies, which are indicator variables indicating different levels of educational achievement: the omitted category is "illiterate or literate, no formal education", the other categories are "Primary Incomplete", "Primary", "Secondary Incomplete", "Secondary", "Diploma", "University". Regressions in Column (4), in addition to education indicators, also control for two indicator variables for being in food manufacturing or apparel manufacturing (omitted category is "other industries"). Regressions in Column (5), in addition to education and industry indicators, also control for variables regarding the time worked (daily hours and days per week) and whether the owner is married or not. Column (6) reports the mean of the dependent variable in the sample of male manufacturers. Column (7) reports the number of observations from the model run in Column (5): this is a lower bound given that the hours and marital status controls are missing for some respondents.

Table C.5: Correlations between Trust and Cooperation

	Trust ST	Trust NB	Trust GSS	Joint Buy	Lent	Advice	Share Order	Coop Average
Trust in Strangers	1							
Trust in Neighbors	0.313***	1						
Trust GSS	0.129***	0.119***	1					
Joint Buy	0.0516*	0.0577**	0.00748	1				
Lent	0.0773***	0.123***	0.0151	0.259***	1			
Advice	0.0596**	0.0978***	-0.0123	0.258***	0.356***	1		
Share Order	0.0618**	0.0819***	-0.0394	0.349***	0.328***	0.347***	1	
Coop Average	0.0901***	0.129***	-0.0104	0.665***	0.708***	0.683***	0.734***	1
N	2216							

*** p<0.01, ** p<0.05, * p<0.1

This table shows correlations between the trust variables and cooperation variables. The variables "Trust in Strangers" and "Trust in Neighbors" are measured on a scale from 1-4, and have been converted into dummy variables by combining low scores (1 and 2) and high scores (3 and 4). Low scores were given a value of 0, and high scores a value of 1. The variable "Trust GSS" is a dummy measuring whether most people can be trusted or not: a value of 1 indicates that "most people can be trusted", and a value of 0 indicates that "you cannot be too careful in dealing with people". The variables "Joint Buy", "Lent", "Advice" and "Share Order" are dummies that indicate whether a person engaged in the relevant activity in the past 12 months. The variable "Coop Average" is an index of cooperative behavior, calculated as the average of the four dummies "Joint Buy", "Lent", "Advice" and "Share Order". The sample comes from our Census of Small-Scale Manufacturers in Lusaka (N=2,216).

Figure C.5: Talking about the Business with Other Entrepreneurs
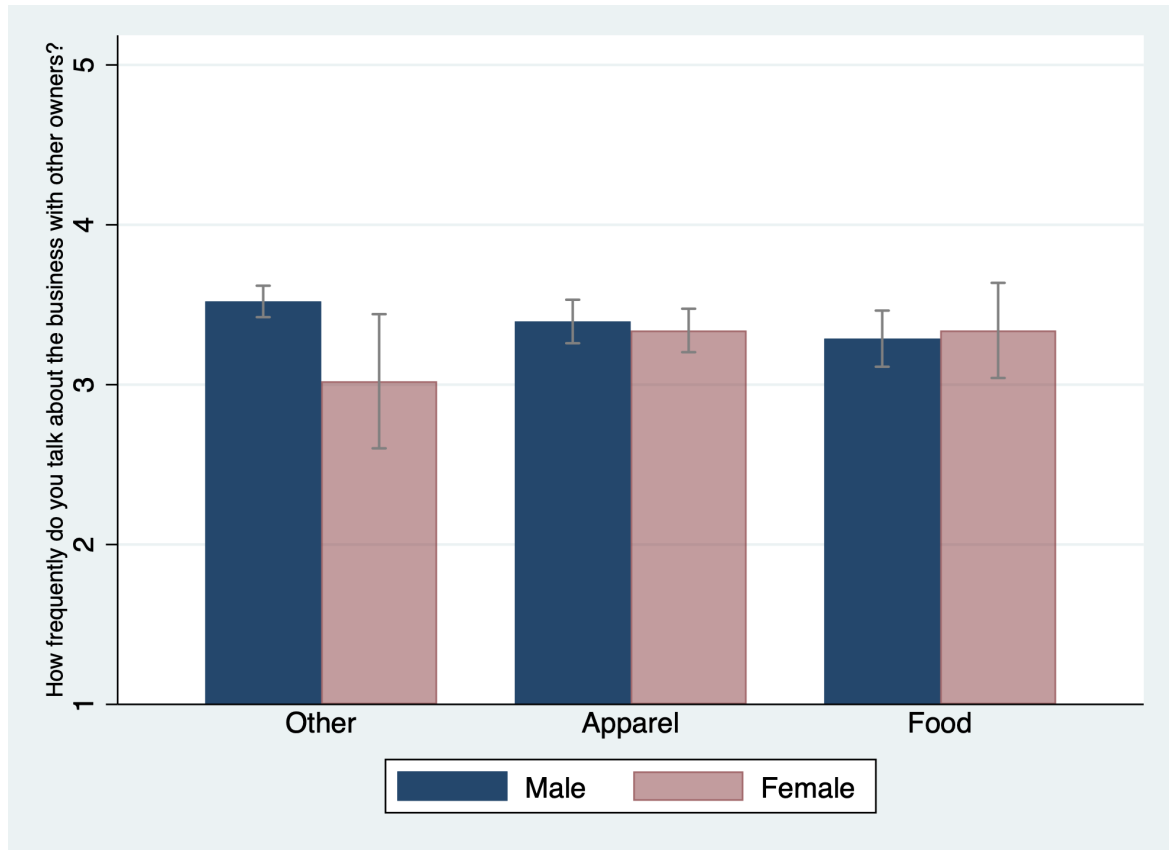
This figure shows how frequently men and women talk with other entrepreneurs about their business. We asked "Consider other business owners in your sector in this neighborhood. How many times do you talk about topics related to the business?" Answers were given on a 5-Likert scale: Never, Once a Month, Between 1 and 4 times a month, More than once a week, Every day. The figure presents raw means. The sample comes from our Census of Small-Scale Manufacturers in Lusaka (N=2,216).

Table C.6: Correlations between Performance, Trust and Cooperation

	Logged sales			Employment	
	Good Week	Bad Week	Yesterday	Level	Growth
Trust NB	0.01	0.0	-0.01	0	0.05**
Trust ST	0.04**	0.04	0.03	0.01	0.01
Trust GSS	0.04*	0.04*	0.01	0.05**	0.03
Coop Average	0.09***	0.02	0.05**	-0.04*	0.02
Lent	0.03	-0.03	0.02	-0.04*	0.0
Advice	0.05***	0.01	0.04*	-0.02	0.02
Share Order	0.06***	0.0	0.03	-0.05**	0.01
Joint Buy	0.1***	0.07***	0.06**	0.01	0.04*

*** p<0.01, ** p<0.05, * p<0.1

This table shows correlations between sales or employment and trust or cooperative behavior. The dependent variables are the log of the answers given when asked the sales in the previous working day, the sales in a good week and the sales in a bad week. If the person could not provide an exact number, we also asked for an upper and lower bound. We then imputed the average of the bounds to the sales variables. All the sales variables are winsorised at the 99% level to control for outliers. The variables "Trust in Strangers" and "Trust in Neighbors" are measured on a scale from 1-4, and have been converted into dummy variables by combining low scores (1 and 2) and high scores (3 and 4). Low scores were given a value of 0, and high scores a value of 1. The variable "Trust GSS" is a dummy measuring whether most people can be trusted or not: a value of 1 indicates that "most people can be trusted", and a value of 0 indicates that "you cannot be too careful in dealing with people". The variables "Joint Buy", "Lent", "Advice" and "Share Order" are dummies that indicate whether a person engaged in the relevant activity in the past 12 months. The variable "Coop Average" is an index of cooperative behavior, calculated as the average of the four dummies "Joint Buy", "Lent", "Advice" and "Share Order". The sample comes from our Census of Small-Scale Manufacturers in Lusaka (N=2,216).

Figure C.6: Main Reasons for Market Location

(a). Reasons for locating inside current market
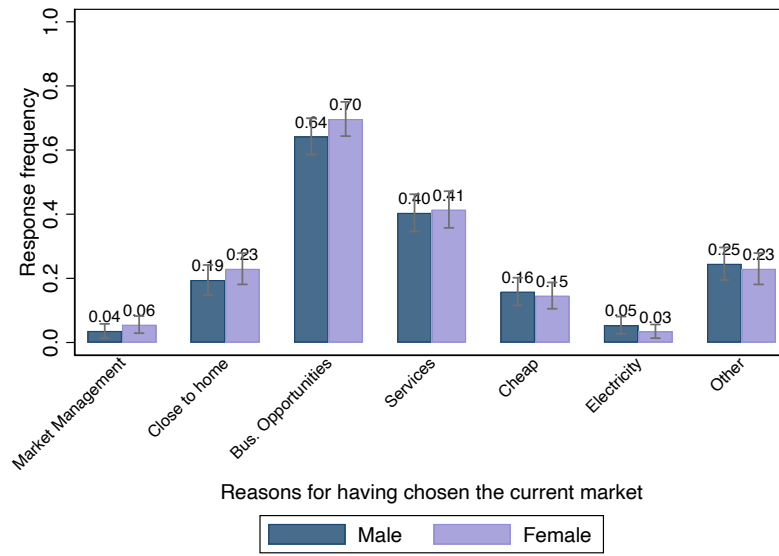
(b). Reasons for locating inside current market instead of just outside
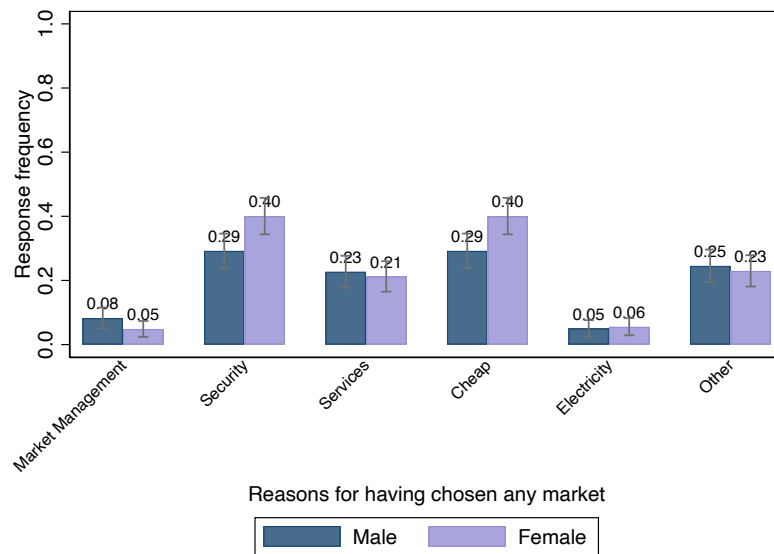
This figure shows the main reasons reported by entrepreneurs for choosing to be located inside their current market, by gender. The sample comes from Experiment 2 (N=594), which includes businesses located inside markets only. Panel (a) shows the reported reasons that participants gave for locating in the specific market they are stationed. Panel (b) shows the reason for locating inside the market, instead of in the same area but just outside the market boundaries.

Table C.7: Correlates of Market Location, by Gender

Variable	(1) Outside Mkt	(2) Inside Mkt	(3) Delta In - Out	(4) Outside Mkt	(5) Inside Mkt	(6) Delta In - Out	(7) DID
		Women			Men		
Age	37.265	40.352	3.087***	37.775	40.117	2.342***	0.745
	(10.331)	(8.551)	(0.764)	(11.144)	(11.049)	(0.575)	(0.980)
Business Age	7.212	11.488	4.276***	8.893	13.503	4.610***	-0.334
	(7.217)	(8.184)	(0.651)	(9.049)	(9.413)	(0.484)	(0.794)
Has clear sign	0.700	0.534	-0.166***	0.709	0.529	-0.180***	0.014
	(0.459)	(0.499)	(0.039)	(0.455)	(0.499)	(0.025)	(0.045)
Not Registered	0.430	0.728	0.299***	0.426	0.753	0.327***	-0.029
	(0.496)	(0.445)	(0.038)	(0.495)	(0.432)	(0.024)	(0.045)
Belongs to Association	0.046	0.040	-0.006	0.030	0.023	-0.007	0.001
	(0.211)	(0.196)	(0.017)	(0.170)	(0.148)	(0.008)	(0.019)
Keeps Records	0.626	0.369	-0.257***	0.634	0.384	-0.250***	-0.007
	(0.485)	(0.483)	(0.039)	(0.482)	(0.487)	(0.025)	(0.046)
Keeps Budget	0.363	0.299	-0.063	0.372	0.319	-0.053**	-0.010
	(0.482)	(0.459)	(0.039)	(0.484)	(0.466)	(0.025)	(0.046)
Married	0.549	0.524	-0.025	0.737	0.806	0.069***	-0.094**
	(0.499)	(0.500)	(0.040)	(0.441)	(0.396)	(0.021)	(0.046)
Spouse gave capital	0.560	0.469	-0.091*	0.173	0.150	-0.022	-0.069
	(0.498)	(0.500)	(0.048)	(0.379)	(0.358)	(0.021)	(0.052)
Recent Immigrant	0.035	0.024	-0.011	0.076	0.033	-0.043***	0.031*
	(0.184)	(0.152)	(0.013)	(0.265)	(0.178)	(0.011)	(0.018)
Born in Lusaka	0.506	0.403	-0.103**	0.364	0.379	0.015	-0.119**
	(0.501)	(0.491)	(0.040)	(0.481)	(0.485)	(0.025)	(0.047)
Days Worked	5.770	5.759	-0.011	6.000	5.980	-0.020	0.009
	(0.926)	(0.806)	(0.069)	(1.000)	(0.802)	(0.046)	(0.085)
Hours Worked	9.941	9.557	-0.384***	10.613	10.334	-0.280***	-0.104
	(1.911)	(1.424)	(0.132)	(1.814)	(1.535)	(0.085)	(0.165)
Education Level	4.767	4.064	-0.703***	4.691	3.872	-0.819***	0.116
	(1.262)	(1.233)	(0.101)	(1.436)	(1.337)	(0.071)	(0.125)
Apparel	0.553	0.884	0.332***	0.181	0.440	0.259***	0.073*
	(0.498)	(0.320)	(0.032)	(0.385)	(0.497)	(0.023)	(0.042)
Food	0.327	0.089	-0.237***	0.296	0.110	-0.186***	-0.051
	(0.470)	(0.286)	(0.030)	(0.457)	(0.313)	(0.019)	(0.039)
Wood	0.012	0.011	-0.001	0.128	0.285	0.157***	-0.159***
	(0.108)	(0.102)	(0.008)	(0.334)	(0.452)	(0.021)	(0.022)
Full time emp at start	1.218	0.475	-0.743***	1.665	0.663	-1.002***	0.260
	(2.243)	(1.352)	(0.147)	(3.255)	(1.730)	(0.129)	(0.217)
Trust (GSS)	1.043	1.037	-0.006	1.103	1.105	0.002	-0.009
	(0.204)	(0.189)	(0.016)	(0.304)	(0.307)	(0.016)	(0.022)
Trusts in neighbors	0.425	0.439	0.015	0.536	0.549	0.014	0.001
	(0.495)	(0.497)	(0.040)	(0.499)	(0.498)	(0.026)	(0.048)
Trusts in strangers	0.207	0.196	-0.011	0.288	0.284	-0.004	-0.007
	(0.406)	(0.397)	(0.032)	(0.453)	(0.451)	(0.023)	(0.040)
Observations	257	380	637	635	944	1,579	2,216

*** p<0.01, ** p<0.05, * p<0.1

This tables shows the main correlates of market location, by gender. Columns (1) and (4) report means and standard deviations (in parenthesis) for each variable for female and male businesses located outside markets, respectively. Columns (2) and (5) report means and standard deviations (in parenthesis) for each variable for female and male businesses located inside markets, respectively. Columns (3) and (6) report the within-gender difference and corresponding standard error between businesses inside or outside markets. Column (7) reports the difference in difference and corresponding standard error. Trust GSS is coded here as 1 for agreeing with "you can't be too careful" and 2 for agreeing with "most people can be trusted". Data comes from the Census of Urban Entrepreneurs (N=2216).

Figure C.7: Cooperative Activities by Distance to Court or Police, by Gender
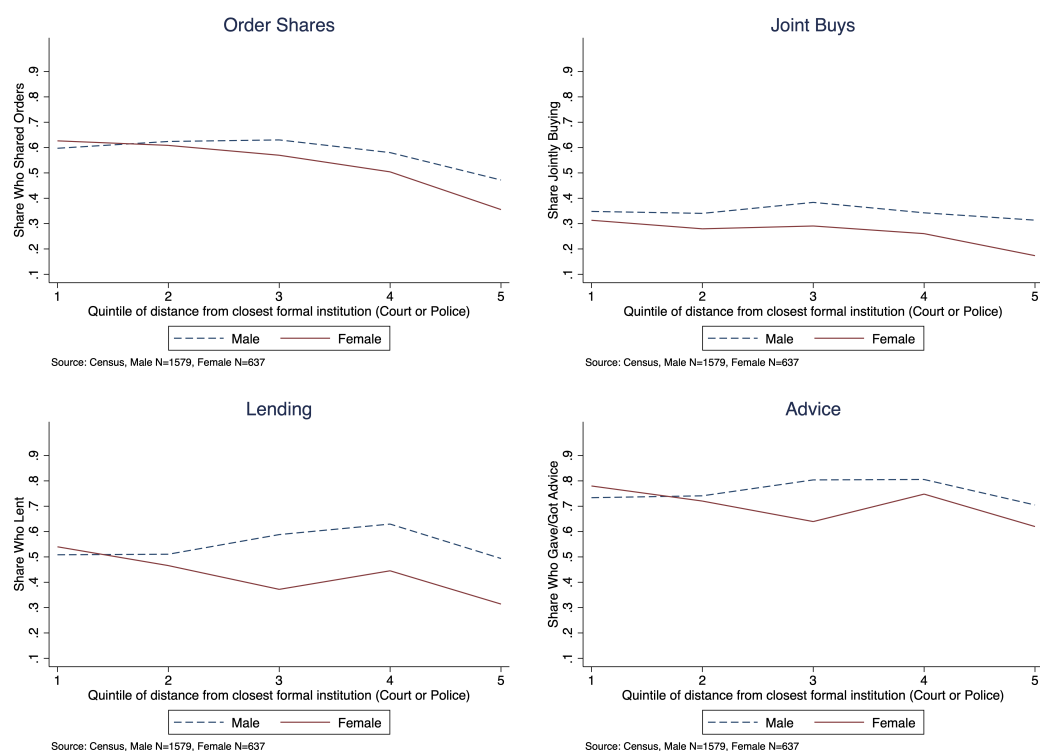
This figure shows the share of entrepreneurs who said they cooperated with a similar business in the given activity by gender and distance from the closest formal institution between the Small Claims Court or the Police. The dashed blue line shows raw average cooperation for men and the solid red line for women. Numbers on the x-axis report the quintile of distance from the closest formal institution. The sample comes from our Census of Small-Scale Manufacturers in Lusaka (N=2,216).

Table C.8: Preferences and Knowledge of Different Institutions by Gender

	Experiment 1			Experiment 2		
	Male N = 347	Female N = 128	Diff F-M	Male N = 290	Female N = 304	Diff F-M
Panel A: Choice of Institutions						
Market Chief	0.73	0.68	-0.05			
	(0.44)	(0.47)	(0.05)			
Small Claims Court	0.04	0.11	0.07***			
	(0.20)	(0.31)	(0.02)			
Police	0.17	0.17	-0.00			
	(0.38)	(0.38)	(0.04)			
If know of court						
Market chief	0.64	0.32	-0.32***			
	(0.48)	(0.48)	(0.11)			
Court	0.12	0.44	0.32***			
	(0.33)	(0.51)	(0.09)			
Police	0.21	0.20	-0.01			
	(0.41)	(0.41)	(0.09)			
Panel B: Opinion on Institutions						
Market Chief Fair	0.55	0.50	-0.05			
	(0.50)	(0.50)	(0.05)			
Small Claims Court Fair	0.51	0.64	0.13			
	(0.50)	(0.49)	(0.12)			
Police Fair	0.21	0.24	0.03			
	(0.41)	(0.43)	(0.04)			
Market Chief Slow	0.18	0.24	0.06			
	(0.39)	(0.43)	(0.04)			
Small Claims Court Slow	0.43	0.42	-0.01			
	(0.50)	(0.50)	(0.11)			
Police Slow	0.29	0.41	0.12**			
	(0.45)	(0.49)	(0.05)			
Panel C: Knowledge of Institutions						
Heard of SCC	0.22	0.20	-0.02			
	(0.41)	(0.40)	(0.04)			
No knowledge of other institutions				0.71	0.76	0.05
				(0.45)	(0.43)	(0.04)
Mention court as institution				0.55	0.55	0.00
				(0.50)	(0.50)	(0.08)
Not know enough about courts				0.29	0.37	0.09**
				(0.45)	(0.48)	(0.04)
Trust somewhat or a lot in courts				0.56	0.57	0.01
				(0.50)	(0.50)	(0.05)
Believes equality before the law				0.50	0.49	-0.01
				(0.50)	(0.50)	(0.04)

*** p<0.01, ** p<0.05, * p<0.1

Panel A shows the share of men (Column 1) and women (Column 2) who would choose a certain institution to solve a business dispute. The survey question was the following: 'Think about an average marketeer in this market. Assume that another person in the market buys items on credit from them and does no pay back what they owe. Which person or group would the person owed money most likely go to?'. The last three rows of panel A show the answers to the same question but limiting the sample to people who have heard of the Small Claims Court (N=101). Panel B shows respondents' opinions of the Market Chief, Police and Small Claims Court fairness and efficiency. For fairness, we gave people the same scenario described above and asked 'If the person owed money goes to the [market chief/Police/SCC], how often do you think the process would be fair?'. We coded the answer as 1 (0) is the person says that the process will be always or usually fair (sometimes, rarely or never fair). For efficiency, we asked how often they think the process would be slow, and coded in the same way. Data are from the Experiment 1 pre-games survey (N=477). Panel C shows men (Column 4) and women's (Column 5) knowledge of formal institutions. Data come from Experiment 2 pre-games survey (N=594), except for the variable Heard of SCC. "Mention court as institution" is a dummy equal to one if the person mentions courts as institutions that they know of for disputes, and it is conditional on the respondent mentioning at least one additional institution besides the market chief and police (N=84 for men and N=73 for women). "Not know enough about courts" is a dummy equal to one if the respondent says that they don't know enough about courts in Zambia to say whether they trust them or not. "Believes equality before the law" is an indicator for saying that people are often or always treated equally under the law in Zambia.

Table C.9: Use of Legal Institutions, Disputes and Harassment by Gender

	Experiment 1			Experiment 2		
	Male N = 347	Female N = 128	Diff F-M	Male N = 290	Female N = 304	Diff F-M
Panel A: Use of Legal Institutions						
Ever used a court				0.21	0.13	-0.08**
				(0.41)	(0.34)	(0.03)
- For business dispute				0.07	0.01	-0.06***
				(0.25)	(0.10)	(0.02)
- For dispute with family member				0.10	0.10	0.00
				(0.30)	(0.30)	(0.02)
- For non-business dispute				0.15	0.12	-0.02
				(0.36)	(0.33)	(0.03)
- For dispute with non-family member				0.04	0.02	-0.02
				(0.20)	(0.15)	(0.01)
Know someone else who used court				0.51	0.46	-0.05
				(0.50)	(0.50)	(0.04)
- Dispute with family member				0.32	0.29	-0.03
				(0.47)	(0.45)	(0.04)
- Dispute with non-family member				0.10	0.13	0.03
				(0.30)	(0.34)	(0.03)
- Business dispute				0.11	0.07	-0.04
				(0.31)	(0.25)	(0.02)
- Non-business dispute				0.40	0.39	-0.01
				(0.49)	(0.49)	(0.04)
Panel B: Work Disputes						
Any Dispute	0.67	0.59	-0.07	0.53	0.40	-0.13***
	(0.47)	(0.49)	(0.05)	(0.50)	(0.49)	(0.04)
Disp. Workplace	0.19	0.12	-0.08*	0.23	0.15	-0.07**
	(0.40)	(0.32)	(0.04)	(0.42)	(0.36)	(0.03)
Disp. Bus. Agree	0.32	0.26	-0.06	0.21	0.10	-0.11***
	(0.47)	(0.44)	(0.05)	(0.44)	(0.30)	(0.03)
Disp. Over Debt	0.52	0.40	-0.12**	0.33	0.27	-0.06
	(0.50)	(0.49)	(0.05)	(0.49)	(0.45)	(0.04)
Disp. Over Goods	0.34	0.25	-0.09*	0.24	0.12	-0.13***
	(0.47)	(0.43)	(0.05)	(0.43)	(0.32)	(0.03)
Sought Help from Third Party	0.25	0.13	-0.12**	0.11	0.08	-0.03
	(0.43)	(0.34)	(0.05)	(0.31)	(0.28)	(0.02)
Panel C: Theft, Assault, Harassment						
Victim of Theft	0.40	0.38	-0.03			
	(0.49)	(0.49)	(0.05)			
Theft Caught	0.22	0.06	-0.16**			
	(0.42)	(0.24)	(0.06)			
Victim of Assault	0.05	0.03	-0.02			
	(0.22)	(0.17)	(0.02)			
Assault Caught	0.22	0.00	-0.22			
	(0.43)	(0.00)	(0.25)			
Police Harass	0.05	0.00	-0.05**			
	(0.22)	(0.00)	(0.02)			
Opposition Abused	0.37	0.40	0.04			
	(0.48)	(0.49)	(0.05)			
Any Harassment				0.77	0.69	-0.08**
				(0.42)	(0.46)	(0.04)
Harassment about Business				0.22	0.18	-0.04
				(0.42)	(0.38)	(0.03)
Verbal Harassment				0.52	0.43	-0.09**
				(0.50)	(0.50)	(0.04)
Physical Harassment				0.27	0.13	-0.14***
				(0.45)	(0.34)	(0.03)
Any Theft				0.64	0.54	-0.10**
				(0.48)	(0.50)	(0.04)
Perpetrator is Police/ from Institution				0.11	0.07	-0.04*
				(0.32)	(0.26)	(0.02)
Panel D: Dispute in Collaborative Activities						
Collaboration with Dispute				0.18	0.10	-0.08**
				(0.39)	(0.31)	(0.04)
Harassment during Collaboration				0.12	0.07	-0.05*
				(0.32)	(0.25)	(0.03)

*** p<0.01, ** p<0.05, * p<0.1

This table shows mean differences between women and men. All the variables are dummy variables. Panel A includes variables related to use of legal institutions, Panel B reports experiences of disputes, Panel C relates to experiences of theft, assault and harassment and Panel D reports disputes in collaboration. Disp. Bus. Agree is dispute over business agreement. Disp. Over Goods is dispute over undelivered or incomplete goods or services. Theft/assault caught is whether the perpetrator of the theft/assault was caught. Opposition Abused is whether the political opposition in the market frequently receives verbal or physical abuse. In Panel B, Sought Help From Third Party is conditional on having any dispute (in Experiment 1, N = 232 for men and N = 76 for women; In Experiment 2, N = 154 for men and N = 123 for women). In Panel C, Theft Caught and Assault Caught are conditional on any theft or any assault, respectively. All the variables related to harassment from experiment 2 ask about experiencing and/or witnessing harassment of a given type in the past 12 months. In Panel D, Collaboration with Dispute is conditional on having at least one collaboration in the past 12 months (N=207 for men and N=192 for women). The Harassment during Collaboration variable is conditional on experiencing or witnessing business harassment(s), verbal harassment(s), physical harassment(s), theft(s) and/or unwanted sexual proposal(s) (N = 217 for men and N = 206 for women).

Table C.10: Local Institutions, Cooperation and Economic Perfomance

	(1)	(2)	(3)	(4)	(5)
			Collaboration Activities		
	Avg coop	Share order	Lent	Advice	Joint buy
Panel A: Cooperation					
Female	-0.116***	-0.099**	-0.158***	-0.089**	-0.117***
	(0.028)	(0.041)	(0.052)	(0.038)	(0.041)
	[0.000]				
Close to SCC	0.050	0.131**	0.018	0.054	-0.004
	(0.035)	(0.054)	(0.054)	(0.069)	(0.043)
	[0.604]				
Female × Close to SCC	0.080**	0.107**	0.067	0.035	0.110**
	(0.035)	(0.048)	(0.058)	(0.047)	(0.046)
	[0.022]				
Close to Police	0.022	0.099***	-0.077**	0.040	0.027
	(0.026)	(0.036)	(0.039)	(0.036)	(0.036)
	[0.020]				
Female × Close to Police	0.024	0.002	0.069	0.036	-0.009
	(0.036)	(0.053)	(0.064)	(0.042)	(0.051)
	[0.510]				
P-value (Close to SCC+ Fem.*Close to SCC = 0)	0.00	0.00	0.19	0.22	0.03
P-value (Close to Police + Fem.*Close to Police = 0)	0.24	0.07	0.88	0.08	0.78
P-value (Fem + Fem.*Close to SCC= 0)	.33	.89	.15	.19	.9
P-value (Fem + Fem.*Close to Police= 0)	.01	.02	.13	.26	0
Observations	2212	2212	2212	2212	2212
Adjusted R^2	0.071	0.049	0.052	0.043	0.037
Mean Dep Var	0.54	0.57	0.52	0.75	0.32
Controls: Business Density	✓	✓	✓	✓	✓
Controls: Individual Characteristics					
Controls: Industry FE					
Controls: Market FE	✓	✓	✓	✓	✓

	(1)	(2)	(3)	(4)	(5)
		Mean Sales		Employment	
	Index	Levels	Logs	Levels	Growth
Panel B: Economic Performance					
Female	-0.030	81.911	-0.155	-0.089	0.011
	(0.068)	(341.652)	(0.095)	(0.143)	(0.058)
	[0.000]				
Close to SCC	0.402***	2558.608***	0.722***	0.344	0.022
	(0.147)	(638.760)	(0.166)	(0.418)	(0.181)
	[0.000]				
Female × Close to SCC	-0.059	-606.621*	-0.056	0.001	-0.019
	(0.070)	(347.405)	(0.105)	(0.179)	(0.068)
	[0.826]				
Close to Police	-0.063	525.657	0.177	-0.389	-0.031
	(0.122)	(642.046)	(0.162)	(0.287)	(0.142)
	[0.142]				
Female × Close to Police	0.032	-71.725	0.005	0.100	-0.036
	(0.086)	(436.133)	(0.102)	(0.156)	(0.067)
	[0.162]				
P-value (Close to SCC+ Fem.*Close to SCC = 0)	0.02	0.00	0.00	0.39	0.99
P-value (Close to Police + Fem.*Close to Police = 0)	0.83	0.57	0.30	0.33	0.62
P-value (Fem + Fem.*Close to SCC= 0)	.28	.21	.05	.61	.92
P-value (Fem + Fem.*Close to Police= 0)	0.97	0.97	0.13	0.94	0.65
Observations	2210	2086	2078	2208	2100
Adjusted R^2	0.325	0.206	0.301	0.283	0.044
Mean Dep Var	0.02	1873.64	6.66	1.94	0.15
Controls: Business Density	✓	✓	✓	✓	✓
Controls: Individual Characteristics					
Controls: Industry FE	✓	✓	✓	✓	✓
Controls: Market FE	✓	✓	✓	✓	✓

Clustered standard errors in parentheses. MHT corrected p in brackets. *** p<0.01, ** p<0.05, * p<0.1

The variables in Panel A 'Joint Buy', 'Lent', 'Advice' and 'Share Order' are dummies that indicate whether a person ever engaged in the relevant activity. The variable 'Avg Coop' is an index of cooperative behavior, calculated as the average of the four dummies. In Panel B the Index of economic performance is obtained by standardizing mean sales in levels and employment levels (separately) and then by taking the average of these two standardized values. The variable 'Close to SCC' is an indicator of value one if the business is located closer than median distance to the SCC and 0 otherwise. The variable 'Close to Police' is an indicator of value one if the business is located closer than the median distance to the Police and 0 otherwise. Standard errors are clustered at the marketplace level. The sample comes from our Census of Small-Scale Manufacturers in Lusaka (N=2,216).

Table C.11: Local Adjudication, Cooperation and Economic Performance: Robustness

	Average cooperation			Mean sales		
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-0.062**	-0.057*	-0.016	-1253.170***	-876.125***	-447.200
	(0.029)	(0.033)	(0.033)	(279.630)	(285.749)	(317.462)
In market	0.071**	0.105***	0.093**	-1503.708**	-1127.219*	-870.560
	(0.030)	(0.037)	(0.036)	(599.340)	(593.046)	(541.574)
Female × In market	0.070*	0.077*	0.066*	880.504***	696.147**	564.326*
	(0.039)	(0.041)	(0.040)	(297.894)	(292.414)	(302.869)
P-value (In + Fem.*In = 0)	0.00	0.00	0.00	0.27	0.43	0.56
P-value (Fem + Fem.*In = 0)	0.72	0.45	0.05	0.03	0.40	0.44
Observations	2213	2042	2038	2093	1950	1944
Adjusted R^2	0.100	0.105	0.128	0.106	0.158	0.237
Mean Dep Var	.54	.55	.55	1876.33	1815.29	1811.24
Controls: Business Density	✓	✓	✓	✓	✓	✓
Controls: Individual Charact		✓	✓		✓	✓
Controls: Industry FE	✓		✓			✓
Controls: Market FE	✓	✓	✓	✓	✓	✓

Clustered standard errors in parenthesis. *** $p<0.01$, ** $p<0.05$, * $p<0.1$

The variable 'Average Cooperation' is an index of cooperative behavior, calculated as the average of the four dummies 'Joint Buy', 'Lent', 'Advice' and 'Share Order' defined in previous tables and in the main body of the paper. The variable 'Mean Sales' is the average of sales in a good week and in a bad week. If the person could not provide an exact number, we also asked for an upper and lower bound. We then imputed the average of the bounds to the sales variables. All the sales variables are winsorised at the 99% percentile to control for outliers. The variable 'In market' is an indicator of value one if the business is located inside a formal market. Density Controls include a dummy variable for whether the business is located within 100 meters of a market, the total number of businesses within 100 meters, the number of business from the same sector within 100 meters, the population density in the administrative area where the business is located (ward) and whether the business has a below-median distance to the police or the Small Claims Court. All the regressions include marketplace level fixed effects. Standard errors are clustered at the marketplace level. The sample comes from our Census of Small-Scale Manufacturers in Lusaka (N=2,216), but the sample in each regression varies because of missing values in the owner controls variables.

Table C.12: Distances to Local Institutions by Gender

	(1)	(2)	(3)
		Coefficient on Female dummy (SE)	
	Male Mean (SD)	OLS (no controls)	OLS (with market FE)
In Market	0.598	-0.001	0.032
	(0.490)	(0.023)	(0.020)
Market within 100m	0.675	-0.022	0.014
	(0.468)	(0.022)	(0.020)
Km to Market	0.128	0.002	-0.012
	(0.229)	(0.011)	(0.011)
< med km to Court	0.468	0.124***	-0.004
	(0.499)	(0.023)	(0.008)
Km to Court	5.646	-0.555***	-0.003
	(2.422)	(0.120)	(0.036)
< med km to Police	0.585	0.032	0.014
	(0.493)	(0.023)	(0.015)
Km to Police	0.485	-0.032*	-0.023*
	(0.400)	(0.019)	(0.013)
< med km to Court or Police	0.735	0.058***	-0.005
	(0.441)	(0.020)	(0.013)

Robust standard errors in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$

This table shows gender differences in the distance to different local institutions. The variable "In Market" is an indicator for whether a business is located inside a formal market. The variable "Market within 100m" is an indicator for whether there is a market in a circle of 100 m radius around the business. The variables "Km to Market," "Km to Court" and "Km to Police" show the distance from a market (conditional on being within one kilometer within a market), from the Small Claims Court or from the Police in kilometers. The variables "< med km to" are indicator variables equal to one if a business has a below-median distance from a given local institution. Column (1) shows the mean and standard deviation of each variable for the sample of male businesses. Columns (2) and (3) report the coefficient of a regression of the row variable on a female dummy, without controls (in Column (2)) and with marketplace fixed effects (in Column (3)).

Figure C.8: Robustness of Coefficients on $Fem_{ism} * InstProx_i$ in Sales Regressions

(a). Mean sales
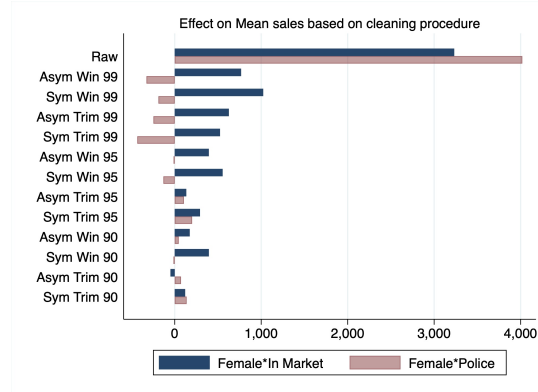
(b). Sales in Bad Week
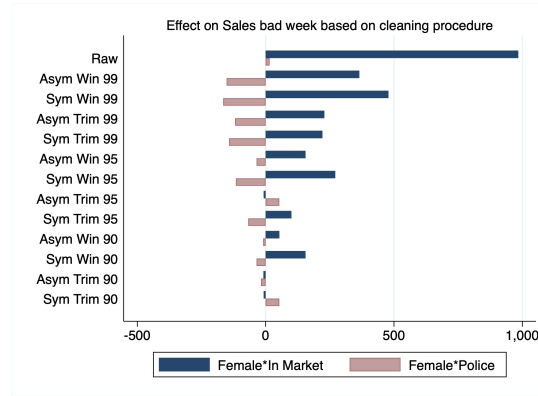
(c). Sales in Good Week
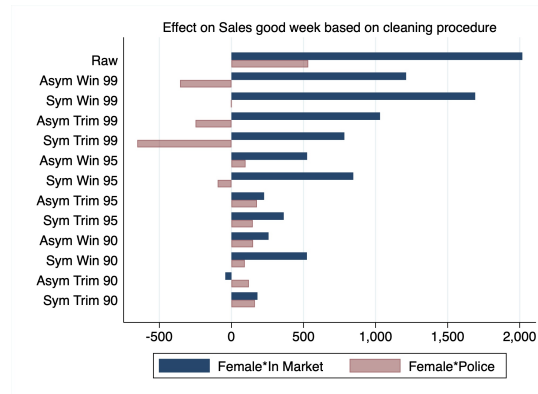
This figure shows the robustness of the coefficients on the interaction term $Fem_{ism} * instProx_i$ in regressions where the dependent variables are sales, for different cleaning procedures of the sales variables. The top figure is for average sales, the middle figure for sales in a bad week and the bottom figure for sales in a good week (all in levels). Blue bars show the coefficients on the interaction of the female dummy with being in a market and pink light bars for the interaction of the female dummy with being closer than median to either the Small Claims Court or the police. The main tables in the paper use the version of the variables winsorised asymmetrically at the top 1%. The raw data version of the sales in good week removes one outlier which is more than 40 times greater than the standard deviation of the sales distribution in a good week.

Table C.13: Experimental Treatment

Experiment 1			Experiment 2	
	Inside/inside	Inside/outside		Inside/inside
Control	94	95	Gender-blind Chief	147
Chief	96	0	Gender-aware Chief	150
SCC	96	96		
Total markets: 16			Total markets: 12	

The numbers represent how many Player As (investors) are in each treatment group. In Experiment 1, each player participated in two rounds: once as an Investor (Player A) and once as a Trustee (Player B), each time paired with a different partner. Therefore, the numbers shown for each group represent the number of participants, each of whom plays once as Investor and once as Trustee. Out of a total of 480 players, three were excluded due to data quality issues. In Experiment 2, participants played only one role—either Investor or Trustee. This means that for each treatment group, there is a separate count of Player Bs. For example, there were 147 Player As and 147 Player Bs (for a total of 294) participating in the "Chief Gender-Blind" treatment. Columns "Inside/inside" denote that players were both from inside the market, and "Inside/outside" denotes that one player was from inside the market and one from outside. In experiment 1, we experimentally manipulated information about these two types of matches. In experiment 2, participants always knew that the other player was from inside the market.

Table C.14: Experiment 1: Balance Checks Pooling Genders

Variable	(1) Mean Control	(2) Mean Chief	(3) Mean Court	(4) Chief vs Control	(5) Court vs Control
Female	0.258	0.198	0.229	-0.060	-0.029
	(0.440)	(0.401)	(0.423)	(0.061)	(0.063)
				[0.998]	[0.999]
Age	41.913	43.021	42.052	1.108	0.139
	(9.108)	(9.973)	(10.876)	(1.395)	(1.466)
				[0.998]	[0.921]
Secondary school	0.344	0.323	0.302	-0.021	-0.042
	(0.478)	(0.470)	(0.462)	(0.069)	(0.068)
				[1.000]	[0.999]
Has a smart phone	0.516	0.479	0.562	-0.037	0.046
	(0.502)	(0.502)	(0.499)	(0.073)	(0.073)
				[1.000]	[1.000]
Business age	15.120	14.926	14.542	-0.194	-0.578
	(10.129)	(8.140)	(9.042)	(1.346)	(1.399)
				[0.987]	[0.998]
Days worked	5.957	6.073	5.990	0.116	0.033
	(0.509)	(0.567)	(0.703)	(0.078)	(0.090)
				[0.927]	[0.994]
Full-time employees	0.355	0.385	0.271	0.031	-0.084
	(0.829)	(1.173)	(0.718)	(0.148)	(0.113)
				[1.000]	[0.999]
Heard of Small Claims Court	0.228	0.240	0.125	0.011	-0.103*
	(0.422)	(0.429)	(0.332)	(0.062)	(0.055)
				[0.998]	[0.709]
Ever lent	0.731	0.698	0.740	-0.033	0.008
	(0.446)	(0.462)	(0.441)	(0.066)	(0.065)
				[1.000]	[0.992]
Ever gave advice	0.903	0.938	0.948	0.034	0.045
	(0.297)	(0.243)	(0.223)	(0.039)	(0.038)
				[0.999]	[0.989]
Ever bought jointly	0.237	0.302	0.292	0.066	0.055
	(0.427)	(0.462)	(0.457)	(0.065)	(0.064)
				[0.997]	[0.998]
Ever shared order	0.613	0.646	0.688	0.033	0.075
	(0.490)	(0.481)	(0.466)	(0.071)	(0.070)
				[1.000]	[0.990]
Trust in Neighbors (1-4 scale)	2.761	2.906	2.906	0.145	0.145
	(0.882)	(0.872)	(0.895)	(0.128)	(0.130)
				[0.971]	[0.973]
Trust in Strangers (1-4 scale)	1.699	1.604	1.646	-0.095	-0.053
	(0.791)	(0.747)	(0.794)	(0.112)	(0.115)
				[0.987]	[0.998]
Manufacturing	0.811	0.948	0.906	0.137***	0.096*
	(0.394)	(0.223)	(0.293)	(0.046)	(0.050)
				[0.057]	[0.681]
Primary Education	0.152	0.208	0.116	0.056	-0.036
	(0.361)	(0.408)	(0.322)	(0.056)	(0.050)
				[0.999]	[1.000]
Ethnicity Bemba	0.172	0.167	0.219	-0.005	0.047
	(0.379)	(0.375)	(0.416)	(0.055)	(0.058)
				[0.922]	[0.999]
Ethnicity Chew	0.129	0.177	0.198	0.048	0.069
	(0.337)	(0.384)	(0.401)	(0.053)	(0.054)
				[1.000]	[0.996]
Ethnicity Nsenga	0.172	0.188	0.188	0.015	0.015
	(0.379)	(0.392)	(0.392)	(0.056)	(0.056)
				[1.000]	[0.985]
Hours Worked	10.410	10.576	10.488	0.166	0.078
	(1.358)	(1.603)	(1.319)	(0.216)	(0.195)
				[0.997]	[0.997]
Observations	95	96	96	191	191

The table shows balance checks for experiment 1. The table excludes participants in the treatment with a player outside the market (which was not implemented for the Chief condition). The sample comes from our first artefactual field experiment conducted in Lusaka in 2017 (N=477). Multiple Hypothesis Testing corrected p-values, computed following List et al. (2019), are reported in brackets.

Table C.15: Experiment 1: Balance Checks by Gender

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
			Women					Men		
Variable	Mean Control	Mean Chief	Mean Court	Chief vs Control	Court vs Control	Mean Control	Mean Chief	Mean Court	Chief vs Control	Court vs Control
Female	1.000	1.000	1.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
				[1.000]	[1.000]				[1.000]	[1.000]
Age	41.261	45.789	40.091	4.529	-1.170	42.130	42.338	42.635	0.207	0.505
	(9.245)	(9.824)	(9.345)	(2.948)	(2.772)	(9.120)	(9.953)	(11.283)	(1.586)	(1.723)
				[0.990]	[0.989]				[0.882]	[1.000]
Secondary school	0.417	0.368	0.455	-0.048	0.038	0.319	0.312	0.257	-0.007	-0.062
	(0.504)	(0.496)	(0.510)	(0.154)	(0.150)	(0.469)	(0.466)	(0.440)	(0.078)	(0.076)
				[0.923]	[0.999]				[1.000]	[1.000]
Has a smart phone	0.500	0.421	0.773	-0.079	0.273*	0.522	0.494	0.500	-0.028	-0.022
	(0.511)	(0.507)	(0.429)	(0.156)	(0.140)	(0.503)	(0.503)	(0.503)	(0.083)	(0.084)
				[1.000]	[0.959]				[1.000]	[0.585]
Business age	11.875	14.263	11.409	2.388	-0.466	16.265	15.093	15.473	-1.171	-0.792
	(8.684)	(8.549)	(7.320)	(2.649)	(2.380)	(10.409)	(8.084)	(9.336)	(1.551)	(1.657)
				[1.000]	[0.999]				[0.999]	[1.000]
Days worked	5.875	5.895	5.636	0.020	-0.239	5.986	6.117	6.095	0.131	0.109
	(0.612)	(0.459)	(1.136)	(0.169)	(0.266)	(0.469)	(0.584)	(0.473)	(0.088)	(0.079)
				[0.932]	[0.949]				[0.913]	[0.995]
Full-time employees	0.417	0.263	0.455	-0.154	0.038	0.333	0.416	0.216	0.082	-0.117
	(0.584)	(0.733)	(0.800)	(0.201)	(0.205)	(0.902)	(1.260)	(0.688)	(0.183)	(0.134)
				[1.000]	[0.999]				[1.000]	[0.999]
Heard of Small Claims Court	0.208	0.105	0.136	-0.103	-0.072	0.235	0.273	0.122	0.037	-0.114*
	(0.415)	(0.315)	(0.351)	(0.115)	(0.114)	(0.427)	(0.448)	(0.329)	(0.073)	(0.064)
				[1.000]	[0.802]				[0.999]	[1.000]
Ever lent	0.625	0.579	0.773	-0.046	0.148	0.768	0.727	0.730	-0.041	-0.038
	(0.495)	(0.507)	(0.429)	(0.154)	(0.137)	(0.425)	(0.448)	(0.447)	(0.073)	(0.073)
				[1.000]	[0.999]				[1.000]	[0.979]
Ever gave advice	0.833	0.895	0.909	0.061	0.076	0.928	0.948	0.959	0.021	0.032
	(0.381)	(0.315)	(0.294)	(0.109)	(0.101)	(0.261)	(0.223)	(0.199)	(0.040)	(0.039)
				[1.000]	[0.999]				[1.000]	[0.999]
Ever bought jointly	0.125	0.263	0.455	0.138	0.330**	0.275	0.312	0.243	0.036	-0.032
	(0.338)	(0.452)	(0.510)	(0.120)	(0.126)	(0.450)	(0.466)	(0.432)	(0.076)	(0.074)
				[1.000]	[0.997]				[0.992]	[0.209]
Ever shared order	0.500	0.526	0.727	0.026	0.227	0.652	0.675	0.676	0.023	0.024
	(0.511)	(0.513)	(0.456)	(0.157)	(0.143)	(0.480)	(0.471)	(0.471)	(0.079)	(0.080)
				[1.000]	[0.998]				[0.999]	[0.784]
Trust in Neighbors (1-4 scale)	2.917	2.895	2.864	-0.022	-0.053	2.706	2.909	2.919	0.203	0.213
	(0.830)	(0.737)	(0.834)	(0.243)	(0.245)	(0.899)	(0.906)	(0.918)	(0.150)	(0.153)
				[0.864]	[0.849]				[0.917]	[1.000]
Trust in Strangers (1-4 scale)	1.458	1.632	1.455	0.173	-0.004	1.783	1.597	1.703	-0.185	-0.080
	(0.588)	(0.684)	(0.800)	(0.194)	(0.206)	(0.838)	(0.765)	(0.789)	(0.133)	(0.136)
				[0.868]	[0.995]				[0.984]	[0.981]
Manufacturing	0.667	0.947	0.909	0.281**	0.242**	0.884	0.948	0.905	0.064	0.021
	(0.482)	(0.229)	(0.294)	(0.120)	(0.119)	(0.323)	(0.223)	(0.295)	(0.046)	(0.052)
				[0.947]	[1.000]				[0.273]	[0.538]
Primary Education	0.083	0.105	0.091	0.022	0.008	0.176	0.234	0.123	0.057	-0.053
	(0.282)	(0.315)	(0.294)	(0.091)	(0.085)	(0.384)	(0.426)	(0.331)	(0.068)	(0.060)
				[1.000]	[0.998]				[1.000]	[0.998]
Ethnicity Bemba	0.125	0.158	0.182	0.033	0.057	0.188	0.169	0.230	-0.020	0.041
	(0.338)	(0.375)	(0.395)	(0.109)	(0.108)	(0.394)	(0.377)	(0.424)	(0.064)	(0.069)
				[1.000]	[1.000]				[1.000]	[0.999]
Ethnicity Chew	0.125	0.158	0.136	0.033	0.011	0.130	0.182	0.216	0.051	0.086
	(0.338)	(0.375)	(0.351)	(0.109)	(0.102)	(0.339)	(0.388)	(0.414)	(0.061)	(0.064)
				[0.999]	[0.997]				[1.000]	[0.999]
Ethnicity Nsenga	0.083	0.053	0.136	-0.031	0.053	0.203	0.221	0.203	0.018	-0.000
	(0.282)	(0.229)	(0.351)	(0.080)	(0.094)	(0.405)	(0.417)	(0.405)	(0.068)	(0.068)
				[0.991]	[0.997]				[1.000]	[1.000]
Hours Worked	9.929	10.300	10.627	0.371	0.698*	10.577	10.644	10.446	0.067	-0.131
	(1.162)	(1.621)	(1.467)	(0.424)	(0.389)	(1.388)	(1.602)	(1.280)	(0.249)	(0.223)
				[0.991]	[0.997]				[1.000]	[1.000]
Observations	24	19	22	43	46	69	77	74	146	143

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The table shows balance checks by gender for the experiment 1. The table excludes participants in the treatment with a player outside the market (which was not implemented for the Chief condition). The sample comes from our first artefactual field experiment conducted in Lusaka in 2017 (N=477). Multiple Hypothesis Testing corrected p-values, computed following List et al. (2019), are in brackets.

Table C.16: Experiment 2: Balance Checks by Gender

Variable	Women			Men		
	(1) Chief: Gender Blind	(2) Chief: Gender Aware	(3) Diff	(4) Chief: Gender Blind	(5) Chief: Gender Aware	(6) Diff
Age	41.132	41.558	0.426	38.648	43.329	4.681 **
	(9.690)	(10.264)	(1.614)	(11.658)	(14.669)	(2.212)
			[1.000]			[0.430]
Primary education or less	0.197	0.247	0.050	0.225	0.219	-0.006
	(0.401)	(0.434)	(0.068)	(0.421)	(0.417)	(0.070)
			[0.998]			[0.936]
Secondary education only	0.684	0.688	0.004	0.662	0.685	0.023
	(0.468)	(0.466)	(0.076)	(0.476)	(0.468)	(0.079)
			[0.960]			[1.000]
Married	0.342	0.494	0.152 *	0.676	0.740	0.064
	(0.478)	(0.503)	(0.079)	(0.471)	(0.442)	(0.076)
			[0.591]			[0.995]
From the Bemba ethnicity	0.303	0.273	-0.030	0.141	0.205	0.064
	(0.462)	(0.448)	(0.074)	(0.350)	(0.407)	(0.063)
			[1.000]			[0.985]
From the Chewa ethnicity	0.039	0.091	0.052	0.155	0.096	-0.059
	(0.196)	(0.289)	(0.040)	(0.364)	(0.296)	(0.055)
			[0.947]			[0.987]
From the Nsenga ethnicity	0.079	0.052	-0.027	0.099	0.151	0.052
	(0.271)	(0.223)	(0.040)	(0.300)	(0.360)	(0.055)
			[1.000]			[0.990]
In the manufacturing sector	0.461	0.481	0.020	0.507	0.493	-0.014
	(0.502)	(0.503)	(0.081)	(0.504)	(0.503)	(0.084)
			[0.999]			[1.000]
Business age	8.147	8.740	0.593	9.972	11.817	1.845
	(7.226)	(7.426)	(1.189)	(7.816)	(11.169)	(1.618)
			[0.999]			[0.984]
Days per Week worked	6.184	6.195	0.011	6.423	6.384	-0.039
	(0.534)	(0.586)	(0.091)	(0.552)	(0.659)	(0.101)
			[1.000]			[1.000]
Hours a day worked	10.534	10.513	-0.021	11.335	11.282	-0.053
	(1.490)	(1.527)	(0.247)	(1.332)	(1.992)	(0.284)
			[0.996]			[1.000]
Has had at least one collaboration (past 12 months)	0.750	0.766	0.016	0.761	0.753	-0.008
	(0.436)	(0.426)	(0.070)	(0.430)	(0.434)	(0.072)
			[0.998]			[0.990]
Number of collaborations (past 12 months)	1.605	1.481	-0.124	1.521	1.493	-0.028
	(1.307)	(1.165)	(0.200)	(1.372)	(1.215)	(0.216)
			[1.000]			[0.998]
General Trust (WVS)	0.053	0.091	0.038	0.085	0.041	-0.044
	(0.225)	(0.289)	(0.042)	(0.280)	(0.200)	(0.040)
			[0.996]			[0.989]
Comfortable leaving the shop unattended	0.197	0.234	0.037	0.394	0.384	-0.010
	(0.401)	(0.426)	(0.067)	(0.492)	(0.490)	(0.082)
			[1.000]			[1.000]
Any business dispute (past 12 months)	0.487	0.403	-0.084	0.535	0.589	0.054
	(0.503)	(0.494)	(0.081)	(0.502)	(0.495)	(0.083)
			[0.989]			[0.998]
Gender of player B: female	0.408	0.597	0.189 **	0.493	0.534	0.041
	(0.495)	(0.494)	(0.080)	(0.504)	(0.502)	(0.084)
			[0.254]			[1.000]
Observations	76	77	153	71	73	144

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, and * p<0.1

The table shows balance checks for Investors in the second experiment. The sample comes from our second artefactual field experiment conducted in Lusaka in 2024/2025 (N=297). Multiple Hypothesis Testing corrected p-values, computed following List et al. (2019), are in brackets.

Table C.17: Experiment 1 - Tokens Sent: Control vs Chief Treatment by Round of Play

	(1) All rounds	(2) Round 1	(3) Round 2
Female	-1.243***	-1.883***	-0.658
	(0.430)	(0.696)	(0.509)
	{0.017}	{0.019}	{0.218}
Chief	-0.085	-0.282	0.125
	(0.382)	(0.621)	(0.467)
	[0.825]	[0.662]	[0.971]
	{0.828}	{0.646}	{0.779}
Female × Chief	1.610**	1.879**	1.998
	(0.682)	(0.887)	(1.346)
	[0.091]	[0.215]	[0.824]
	{0.045}	{0.044}	{0.266}
Observations	189	94	95
Mean Dep Var	4.087	4.429	3.735

Notes: The table reports OLS regressions of tokens sent by player A on a female dummy, a chief-treatment dummy, and their interaction, by round of play. "Round 1" only includes players who were investors in the first round of play, and "Round 2" only includes investors who played in this role in the second round (thus after having played as trustees). Each column corresponds to a different round specification. Robust standard errors in parentheses. Randomization inference p-values in square brackets. Multiple hypothesis testing p-values in curly brackets following (List et al., 2019). Data are from experiment 1, and only include participants assigned to the chief treatment or the control treatment. The SCC treatment and the inside/outside treatments are excluded. * p<0.1, ** p<0.05, *** p<0.01

Table C.18: Experiment 1: Tokens Investor Sends by Institutional Treatment

| | Tokens Investor sent | | | | | |
	(1)	(2)	(3)	(4)	(5)	(6)
Female	-1.243***	-1.030***	-1.236***	-0.925*	-1.111***	-0.949*
	(0.430)	(0.331)	(0.433)	(0.520)	(0.362)	(0.486)
Chief	-0.0851			-0.224		
	(0.382)			(0.410)		
Court		0.252			0.303	
		(0.281)			(0.286)	
In/out			-0.0527			0.0723
			(0.416)			(0.469)
Chief × Female	1.610**			1.706**		
	(0.682)			(0.753)		
Court × Female		0.281			0.0936	
		(0.452)			(0.477)	
In/out × Female			0.462			0.289
			(0.662)			(0.743)
Constant	4.198***	4.101***	4.267***	-2.077	1.450	-1.357
	(0.361)	(0.238)	(0.364)	(3.642)	(1.853)	(3.117)
Observations	189	379	187	189	379	187
Mean Dep Var	4.087	4.076	4.087	4.087	4.076	4.087
Individual Controls				✓	✓	✓

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the number of tokens sent by the Investor. All regressions include a dummy on the Investor's gender, and each column includes a dummy for the different treatment arm (in/out, Court, and Chief), together with the interaction of that treatment arm and the Investor's gender. Columns (1) and (4) only include participants in the Inside/Inside condition, either in Control or Chief arms. Columns (2) and (5) include participants in both the Inside/Inside and Outside/Inside conditions, either in Control or Court arms, and controls for the in/out treatment. Columns (3) and (6) exclude participants in Chief or Court arms. "Individual controls" are the same as those used in Table 4 in the main body of the paper. The sample comes from our first artefactual field experiment conducted in Lusaka in 2017 (N=477).

Table C.19: Experiment 1: Average Return Ratio by Institutional Treatment

			Average Return Ratio			
	(1)	(2)	(3)	(4)	(5)	(6)
Female	0.000115	-0.0427**	0.00144	0.0134	-0.0470**	0.0186
	(0.0263)	(0.0200)	(0.0265)	(0.0310)	(0.0214)	(0.0313)
Chief	0.0247			0.0128		
	(0.0200)			(0.0199)		
Court		0.0340**			0.0262	
		(0.0163)			(0.0162)	
In/out			0.0218			0.0133
			(0.0222)			(0.0240)
Chief × Female	-0.0524			-0.0285		
	(0.0417)			(0.0443)		
Court × Female		0.0284			0.0251	
		(0.0260)			(0.0273)	
In/out × Female			-0.0787**			-0.0773*
			(0.0376)			(0.0410)
Constant	0.416***	0.418***	0.403***	0.529***	0.532***	0.526***
	(0.0208)	(0.0143)	(0.0209)	(0.179)	(0.119)	(0.180)
Observations	190	380	188	190	380	188
Mean Dep Var	0.423	0.435	0.423	0.423	0.435	0.423
Individual Controls				✓	✓	✓

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the Trustee's average return ratio. All regressions include a dummy on the Trustee's gender, and each column includes a dummy for the different treatment arm (in/out, Court, and Chief), together with the interaction of that treatment arm and the Trustee's gender. Columns (1) and (4) only include participants in the Inside/Inside condition, either in Control or Chief arms. Columns (2) and (5) include participants in both the Inside/Inside and Outside/Inside conditions, either in Control or Court arms, and control for the in/out treatment. Columns (3) and (6) exclude participants in Chief or Court arms. "Individual controls" are the same as those used in Table 4 in the main body of the paper. The sample comes from our first artefactual field experiment conducted in Lusaka in 2017 (N=477).

Table C.20: Experiment 1: Investor's Earnings

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
				Investor's Earnings				
Female	-0.0680	-0.202			-0.207	-0.337		
	(0.355)	(0.401)			(0.314)	(0.361)		
Chief	0.566*	0.517	0.620**	0.607**	0.608**	0.537	0.720***	0.683**
	(0.292)	(0.317)	(0.251)	(0.271)	(0.300)	(0.330)	(0.261)	(0.283)
Court	1.107***	1.012***	0.914***	0.888***	1.127***	0.980**	1.012***	0.940***
	(0.282)	(0.366)	(0.221)	(0.292)	(0.282)	(0.381)	(0.229)	(0.302)
In/out	0.208	0.103	0.0577	0.0317	0.246	0.0867	0.190	0.119
	(0.284)	(0.369)	(0.221)	(0.301)	(0.278)	(0.362)	(0.230)	(0.296)
Female × Chief	0.550	0.683			0.614	0.744		
	(0.657)	(0.677)			(0.629)	(0.644)		
Female × Court	-0.588	-0.316			-0.424	-0.148		
	(0.429)	(0.532)			(0.416)	(0.508)		
Female × In/out	-0.256	-0.000892			-0.145	0.114		
	(0.418)	(0.629)			(0.410)	(0.620)		
Female × Court × In/out		-0.510				-0.535		
		(0.845)				(0.804)		
Court × In/out		0.204		0.0516		0.317		0.147
		(0.572)		(0.442)		(0.566)		(0.452)
Constant	10.67***	10.72***	10.73***	10.74***	9.110***	9.175***	8.853***	8.841***
	(0.253)	(0.295)	(0.214)	(0.243)	(1.315)	(1.318)	(1.285)	(1.291)
Observations	475	475	477	477	475	475	477	477
Mean Dep Var	10.855	10.855	10.863	10.863	10.855	10.855	10.863	10.863
Individual Controls					Yes	Yes	Yes	Yes

Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

The dependent variable is the Investor's final earnings. Regressors include the treatment arms (in/out, Court, and Chief). Columns 1, 2, 5 and 6 additionally include a dummy variable indicating the Investor's gender, and the interaction terms of this variable with the treatment arms. Individual controls are the same as those used in Table 4 in the main body of the paper. The sample comes from our first artefactual field experiment conducted in Lusaka in 2017 (N=477).

Table C.21: Experiment 1: Trustee's Earnings

|	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Trustee's Earnings							
Female	-0.119	0.379			-0.457	-0.0195		
	(0.976)	(1.211)			(1.044)	(1.267)		
Chief	-0.202	-0.342	0.114	-0.147	-0.125	-0.360	0.131	-0.212
	(0.596)	(0.660)	(0.536)	(0.598)	(0.619)	(0.679)	(0.550)	(0.620)
Court	-0.393	-0.663	-0.305	-0.824	-0.260	-0.703	-0.170	-0.844
	(0.512)	(0.662)	(0.432)	(0.582)	(0.513)	(0.658)	(0.433)	(0.589)
In/out	0.389	0.0944	0.360	-0.164	0.395	-0.0510	0.279	-0.376
	(0.512)	(0.788)	(0.432)	(0.677)	(0.544)	(0.784)	(0.488)	(0.691)
Female × Chief	1.400	0.907			1.195	0.775		
	(1.390)	(1.566)			(1.422)	(1.608)		
Female × Court	0.239	-0.824			0.306	-0.556		
	(0.971)	(1.404)			(0.969)	(1.391)		
Female × In/out	-0.143	-0.966			-0.319	-0.946		
	(0.987)	(1.582)			(0.996)	(1.603)		
Female × Court × In/out		1.713				1.308		
		(1.949)				(2.058)		
Court × In/out		0.578		1.039		0.967		1.370
		(1.028)		(0.864)		(1.048)		(0.872)
Constant	6.892***	7.038***	6.805***	7.066***	5.852**	5.489*	5.579**	5.437*
	(0.538)	(0.613)	(0.488)	(0.562)	(2.852)	(2.924)	(2.766)	(2.807)
Observations	476	476	477	477	476	476	477	477
Mean Dep Var	6.757	6.757	6.821	6.821	6.757	6.757	6.821	6.821
Individual Controls					Yes	Yes	Yes	Yes

The dependent variable is the Trustee's final earnings. Regressors include the treatment arms (in/out, Court, and Chief). Columns 1, 2, 5 and 6 additionally include a dummy variable indicating the Trustee's gender, and the interaction terms of this variable with the treatment arms. Individual controls are the same as those used in Table 4 in the main body of the paper. The sample comes from our first artefactual field experiment conducted in Lusaka in 2017 (N=477).

Figure C.9: Chief Gender Bias Across Experimental Markets
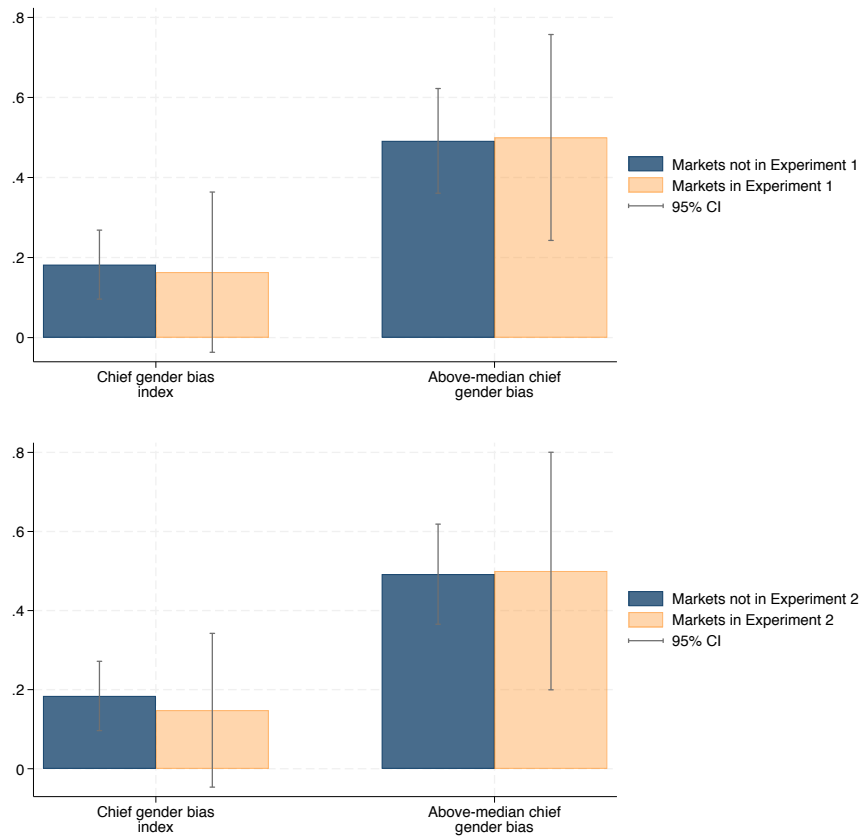
The figure shows the average chief gender bias index (first set of bars) and the share of markets with a gender bias index above median (second set of bars) among markets which were part of the two artefactual field experiments. The top figure is for experiment 1 (where we involved 16 markets) and the bottom figure for experiment 2 (where we involved 12 markets).

Figure C.10: Experiment 2: Preferences for Gender-Revelation, by Gender and Chief's Bias

Graphs by equal to 1 if respondent female
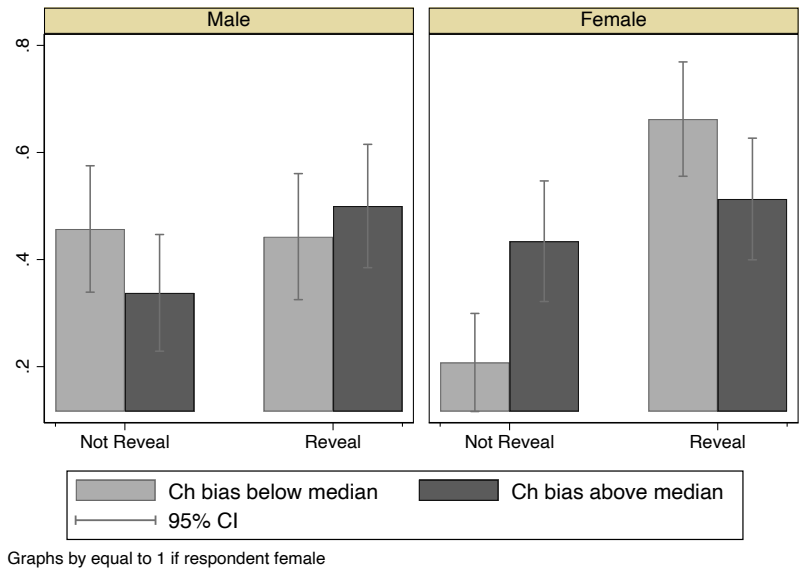
The figure shows investors' preferences for having their gender hidden ("Not Reveal") or revealed ("Reveal") in the complaint form used in the trust game of experiment 2, splitting the sample by respondents' gender and markets with a chief with above or below median chief bias. We ask people to tell us which form they preferred and to justify their answer. Two treatment-blind research assistants independently went through the answers and coded the preferences data into three categories: preference for revealing, for not revealing or no preference. Preferences were elicited after participants found out about their treatment and took their decision on how many tokens to send to the trustee. Around 10% of men and women express no preference, and the share is similar across markets that differ by chief's gender bias (not shown here). Data are from Experiment 2 ($N = 153$ women, $N = 144$ men, investors only).

Figure C.11: Experiment 2: Trust and Trustworthiness by Gender and Treatment
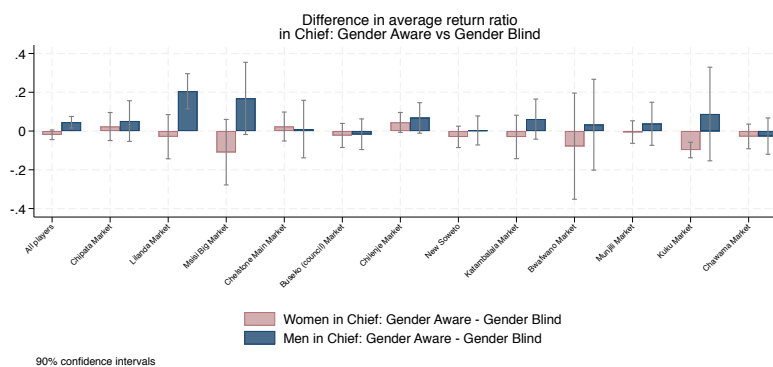
(a). Trust
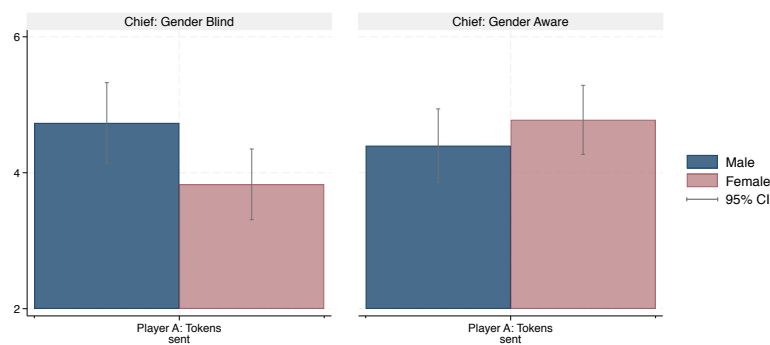
(b). Trustworthiness
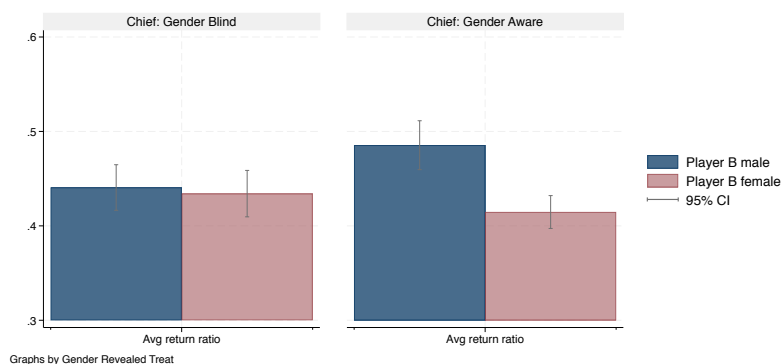
(c). Trustworthiness across markets, by gender

These figures show the average treatment effects of the Chief Gender Blind vs Chief Gender Aware conditions on trust and trustworthiness. Panel (a) shows the average tokens sent by player A in each of the experimental arms, by gender. Panel (b) shows the average return ratio of of tokens of player B, by gender. Panel (c) shows the difference in average return ratio in the gender aware vs gender blind experimental conditions, by gender and market. Data are from experiment 2 (N=594).
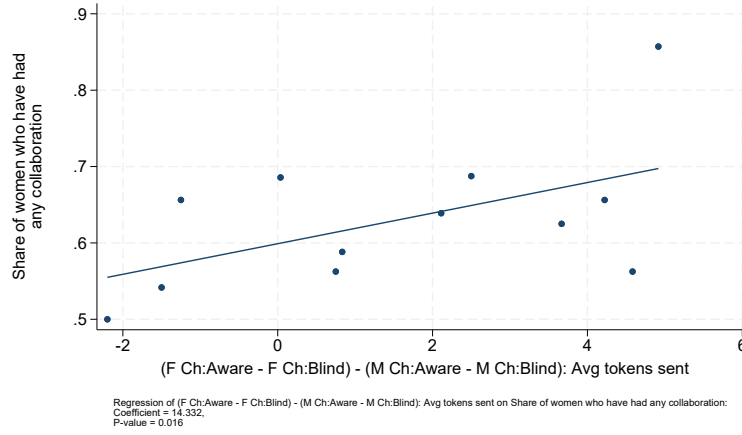
Table C.22: Experiment 2: Gender Differences in the Revelation Trust Premium and Correlation with Other Experimental Outcomes

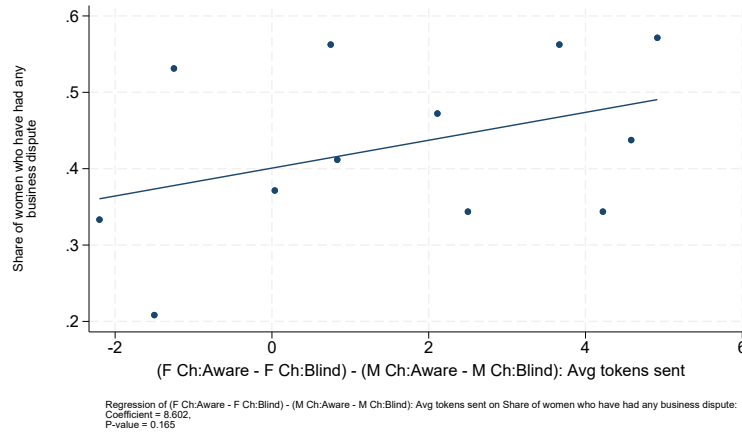|	Correlation	T-stat	N
Panel A: Trust Behavior			
Women in Chief Gender Aware: Avg tokens sent	0.677	3.433	12
Men in Chief Gender Aware: Avg tokens sent	-0.565	-3.511	12
Women (Chief Gender Aware - Chief Gender Blind): Avg tokens sent	0.899	8.083	12
Men (Chief Gender Aware - Chief Gender Blind): Avg tokens sent	-0.863	-7.799	12
Panel B: Preferences			
Preference for Chief Gender Aware: Share of female marketeers	0.486	1.910	12
Preference for Chief Gender Aware: Share of male marketeers	-0.276	-0.931	12
Women (Share Pref Chief Gender Aware - Share Pref Chief Gender Blind)	0.540	2.127	12
Men (Share Pref Chief Gender Aware - Share Pref Chief Gender Blind)	-0.349	-1.304	12
Panel C: Actual Chief Behavior			
Chief Gender Aware: Share of women's complaints with reallocation	0.660	3.438	11
Chief Gender Aware: Share of men's complaints with reallocation	-0.135	-0.345	8
Panel D: Expectations of Chief Behavior			
Incentivized guess about women's complaints: Chief cutoff for action	-0.066	-0.176	12
Incentivized guess about men's complaints: Chief cutoff for action	0.264	1.008	12
Incentivized guess about women's complaints: Avg % change return ratio	0.538	2.319	12
Incentivized guess about men's complaints: Avg % change return ratio	-0.158	-0.685	12
Expectation that chief will favor women: Share of female marketeers	0.371	1.520	12
Expectation that chief will favor women: Share of male marketeers	0.246	1.006	12
Expectation that chief will favor women: Share of all marketeers	0.344	1.315	12
Panel E: Hypothetical Chief's Behavior			
Chief hypo. about women's complaints: Avg cutoff for action	0.234	0.754	12
Chief hypo. about men's complaints: Avg cutoff for action	0.145	0.548	12
Chief hypo. about women's complaints: Avg % change return ratio	-0.311	-1.048	12
Chief hypo. about men's complaints: Avg % change return ratio	0.296	1.006	12

This table shows the correlation between the market-level gender difference in the revelation trust premium and other experimental outcomes at the market or market-gender level. For both genders, the revelation trust premium is obtained as the difference in tokens sent in Chief Gender Aware minus the tokens sent in Chief Gender Blind. We then take the difference between women and men. The third column shows the t-statistic of a univariate regression of this double difference on each of the variables in the left-hand side column. Panel A shows correlations with investors' behavior. Panel B shows correlations with investors' preferences for gender revelation. Preferences are coded from an open-ended question into three categories: "preference for blindness" (gender hidden in the complaint form), "preference for revelation", and "no preference". Panel C shows correlations with the chief's actual redistributive behavior within the games, conditional on the Investor having complained to the chief about the number of tokens the Trustee sent back to them. The number of observations is not equal to 12 for all variables, because Investors in the Chief Gender Aware condition did not complain in all 12 markets. Panel D shows correlations with participants' expected chief behavior in the game. Variables starting with "Incentivized guess" are constructed from incentivized questions asked at the end of the trust game, in which players were presented with multiple scenarios of tokens sent by the Investor and tokens sent back by the Trustee and were asked to guess what was the largest amount of tokens sent back by the Trustee which would lead the chief to reallocate tokens to the Investor, provided the Investor complained. The gender of the Investor in the scenarios was randomized among marketeers. The variable "Expectation that the chief will favor women" is from a non-incentivized question asked to marketeers after they received their treatment assignment about whether they thought the chief's decision would change depending on whether the chief was aware of their gender or not, and why. Panel E describes the Chief's hypothetical behavior. The chief was presented with multiple scenarios of complaints, once for a female complainant and once for a male complainant. Chiefs were asked what was the largest amount of tokens sent back by the Trustee that would lead them to reallocate tokens to the Investor, and how many tokens they would redistribute at such complaint. These are respectively the variables "Chief hypo.: Avg cutoff" and "Chief hypo.: Avg % change return ratio". A higher cutoff means that the chief is likely to take action for a wider range of complaints. Details on the construction of the variables are in Appendix D.

Figure C.12: Experiment 2: Market Heterogeneity in Treatment Effects and Key Market Features
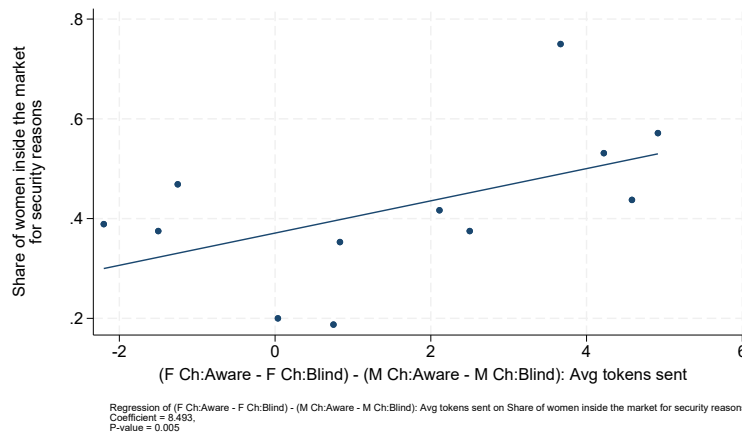
(a). Gender Difference in Revelation Trust Premium and Female Collaboration

Regression of (F Ch:Aware - F Ch:Blind) - (M Ch:Aware - M Ch:Blind): Avg tokens sent on Share of women who have had any collaboration:
Coefficient = 14.332,
P-value = 0.016

(b). Gender Difference in Revelation Trust Premium and Female Disputes

Regression of (F Ch:Aware - F Ch:Blind) - (M Ch:Aware - M Ch:Blind): Avg tokens sent on Share of women who have had any business dispute:
Coefficient = 8.602,
P-value = 0.165

(c). Gender Difference in Revelation Trust Premium and Female Demand for Security

Regression of (F Ch:Aware - F Ch:Blind) - (M Ch:Aware - M Ch:Blind): Avg tokens sent on Share of women inside the market for security reasons:
Coefficient = 8.493,
P-value = 0.005

The graphs show the relationship between the difference in difference in tokens sent ((Female Chief:Aware - Female Chief:Blind) - (Male Chief:Aware - Male Chief:Blind)) on some features of the markets: share of women who have collaborated in the past 12 months (Panel (a)), share of women who have had a dispute in the past 12 months (Panel (b)) and share of women who report being inside a market for security and safety reasons (Panel (c)). All variables are at the market level (N=12). Data are from experiment 2 (N=297, only investors).

104

Table C.23: Experiment 2: Gender Differences in the Revelation Trust Premium and Market Features

|	Correlation	T-stat	N
Panel A: Business Features			
Average number of repeat customers of women	0.463	1.838	12
Mean sales of women (std)	0.133	0.417	12
Average tenure of women in the market	-0.220	-0.701	12
Average number of workers of women	-0.202	-0.840	12
Panel B: Trust and Collaboration			
Share of women inside the market for security reasons	0.524	3.525	12
Share of women who have had any collaboration	0.536	2.878	12
Share of women who have collab. with partner of different gender or age	0.481	2.034	12
Share of women who have collaborated only with women	0.335	1.631	12
Share of women who have had any business dispute	0.396	1.499	12
Share of women having exp/witn any harassment	0.319	1.212	12
General Trust (GSS) of all marketeers	-0.398	-1.296	12
General Trust (GSS) of women	-0.329	-1.591	12
Share of women believing a woman may not collab. with a man due to trust	-0.699	-3.090	12
Panel C: Market Features			
Listing: Share of women owners out of all firms	0.360	1.030	12
Listing: Share of women owners out of manufacturing firms	0.333	0.795	12
Census: Share of women owners out of manufacturing firms	-0.097	-0.317	12
Census: Average number of employees	-0.232	-0.895	12
Census: Number of businesses in the market	-0.594	-3.517	12

This table shows the correlation between the gender difference in the revelation trust premium and market features. For both genders, the revelation trust premium is obtained as the difference in tokens sent in Chief Gender Aware minus the tokens sent in Chief Gender Blind. We then take the difference between women and men. The third column shows the t-statistic of a univariate regression of this double difference on the variables in the left-hand side column. Panel A shows correlations with market-level averages of business features asked in the pre-games survey to participants. The "number of repeat customers" is defined as "customers that come back to buy from your shop goods or services more than once a year". The "mean sales" variable is the mean of the answers by marketeers of the total value of sales for all products and services of the business in a good and in a bad week winsorised at the 99th percentile. "Tenure in the market" is the number of years the business has been operating in the market. Panel B shows correlations with variables on Trust and Collaboration, all measured in the past 12 months. The variable "Inside the market for security reasons" is the share of women who report security to be a reason for choosing to locate their business inside this market. "General Trust (GSS)" variables are dummies that equal 1 if the respondent answers that "most people can be trusted" and 0 if they say that "you can't be too careful in dealing with people". "Share of women believing a woman may not collab. with a man due to trust" uses a question where marketeers were presented with a scenario where woman refuses a collaboration offered by a man. The variables takes 1 if the marketeers gave a reason related to the woman not being able to trust the man (or men in general) and 0 otherwise. Panel C presents market features constructed from the Census of Small-Scale Manufacturers collected in 2016 and the listing of all the businesses belonging to manufacturing, and restaurants and hairdressers within each of the 12 markets in the experiment 2.

Figure C.13: Experiment 2: Distribution of Total Earnings between Investor and Trustee by Investor's Gender

(a). Female Investors
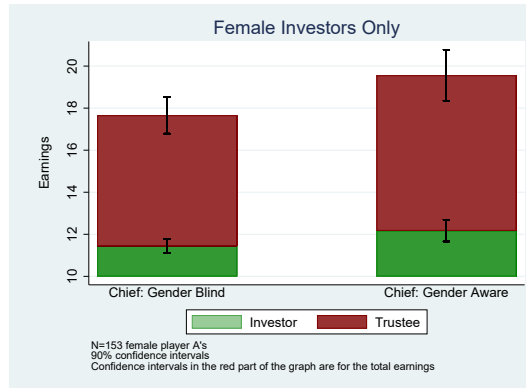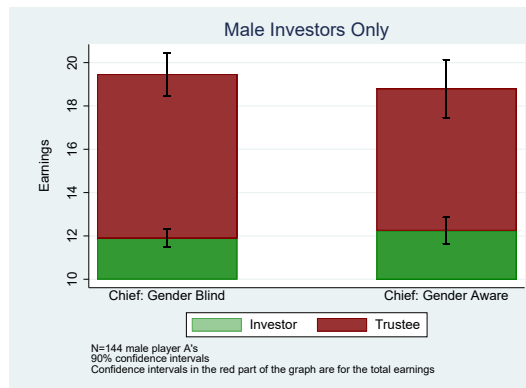
(b). Male Investors

This figure shows the distribution of total final earnings in rounds with a female Investor (Panel a) and male investor (Panel b). Data are from experiment 2 (N=594).

Table C.24: Census and Experiments samples vs DHS – share (%) by category: All respondents

	Data from paper			DHS data	
	Census (2016)	Experiment 1 (2017)	Experiment 2 (2025)	Zambia (2007)	SSA (1999-2009)
Age					
18-24	7.89%	4.03%	11.11%	30.63%	29.93%
25-29	12.71%	10.17%	12.12%	20.06%	19.15%
30-34	14.90%	11.65%	14.81%	16.84%	15.49%
35-39	16.89%	17.16%	12.96%	12.45%	13.13%
40-44	17.54%	19.49%	13.80%	8.54%	10.31%
45-49	14.34%	15.04%	12.63%	7.29%	8.91%
50-54	8.26%	10.59%	11.28%	2.48%	1.88%
55-59	3.90%	5.93%	5.39%	1.72%	1.17%
60+	3.57%	5.93%	5.89%		0.03%
Education					
No education	1.82%	1.27%	1.35%	8.48%	35.81%
Primary	23.34%	28.03%	19.36%	49.16%	30.97%
Secondary	54.47%	57.11%	67.68%	35.15%	28.32%
Higher	20.37%	13.59%	11.62%	7.21%	4.90%
Ethnicity					
Bemba	21.07%	16.49%	23.91%	19.04%	
Chewa	14.08%	17.34%	10.27%	6.73%	
Nsenga	10.83%	16.49%	8.92%	5.14%	
Tonga	7.08%	5.50%	7.74%	11.33%	
Other	46.94%	44.19%	49.16%	57.75%	
Self-employment					
Not self-employed				52.88%	56.43%
Self-employed	100.00%	100.00%	100.00%	47.12%	43.57%
Marital status					
Divorced/Separated	3.71%	3.59%	7.91%	6.93%	5.43%
In a relationship	3.85%	6.13%	7.07%	0.62%	9.53%
Married	71.26%	75.48%	56.23%	66.15%	60.97%
Single	16.27%	8.25%	21.04%	23.00%	21.52%
Widowed	4.49%	6.13%	7.74%	3.30%	2.56%
Other	0.41%	0.42%			
Previous residence type					
Rural	43.58%	50.27%		46.00%	54.78%
Urban	56.42%	49.73%		54.00%	45.22%
Religion					
Catholic	18.65%	19.33%		19.77%	
Muslim	2.62%	1.33%		0.39%	
Protestant	71.96%	70.33%		77.65%	
Other	6.77%	9.00%		2.19%	
Sex					
Female	28.67%	26.85%	51.18%	52.19%	68.51%
Male	71.33%	73.15%	48.82%	47.81%	31.49%

Notes: The table reports the distribution of respondents across demographic categories in our Census of Small Scale Manufacturers (2016), Experiment 1 (2017), Experiment 2 (2025), and the Demographic and Health Surveys (DHS) for Zambia and Sub-Saharan Africa (SSA). Countries in SSA: Benin (2006), Democratic Republic of Congo (2007), Republic of Congo (2005), Ethiopia (2000), Ghana (2008), Guinea (1999), Kenya (2008–2009), Liberia (2007), Madagascar (2008–2009), Mali (2012–2013), Nigeria (2008), Namibia (2006–2007), Rwanda (2000), Sierra Leone (2008), São Tomé and Príncipe (2008–2009), Uganda (2006), Zambia (2007), and Zimbabwe (2005–2006). Percentages indicate the share of individuals in each category within each dataset. For the DHS, estimates refer to adult respondents aged 18–59. Missing cells indicate that the corresponding variable was not collected in that dataset.

Table C.25: Census and Experiments samples vs DHS – share (%) by category: Self-employed respondents

	Data from paper			DHS data	
	Census (2016)	Experiment 1 (2017)	Experiment 2 (2025)	Zambia (2007)	SSA (1999-2009)
Age					
18-24	7.89%	4.03%	11.11%	18.07%	19.65%
25-29	12.71%	10.17%	12.12%	20.07%	19.94%
30-34	14.90%	11.65%	14.81%	19.98%	18.10%
35-39	16.89%	17.16%	12.96%	15.80%	15.91%
40-44	17.54%	19.49%	13.80%	11.02%	12.58%
45-49	14.34%	15.04%	12.63%	9.48%	11.02%
50-54	8.26%	10.59%	11.28%	3.14%	1.64%
55-59	3.90%	5.93%	5.39%	2.45%	1.09%
60+	3.57%	5.93%	5.89%		0.06%
Education					
No education	1.82%	1.27%	1.35%	10.37%	42.31%
Primary	23.34%	28.03%	19.36%	59.72%	33.43%
Secondary	54.47%	57.11%	67.68%	27.64%	22.10%
Higher	20.37%	13.59%	11.62%	2.27%	2.16%
Ethnicity					
Bemba	21.07%	16.49%	23.91%	15.84%	
Chewa	14.08%	17.34%	10.27%	6.33%	
Nsenga	10.83%	16.49%	8.92%	4.80%	
Tonga	7.08%	5.50%	7.74%	9.72%	
Other	46.94%	44.19%	49.16%	63.31%	
Self-employment					
Self-employed	100.00%	100.00%	100.00%	100.00%	100.00%
Marital status					
Divorced/Separated	3.71%	3.59%	7.91%	7.79%	5.87%
In a relationship	3.85%	6.13%	7.07%	0.47%	10.52%
Married	71.26%	75.48%	56.23%	78.18%	71.60%
Single	16.27%	8.25%	21.04%	9.48%	8.93%
Widowed	4.49%	6.13%	7.74%	4.08%	3.07%
Other	0.41%	0.42%			
Previous residence type					
Rural	43.58%	50.27%		57.52%	57.34%
Urban	56.42%	49.73%		42.48%	42.66%
Religion					
Catholic	18.65%	19.33%		17.96%	
Muslim	2.62%	1.33%		0.40%	
Protestant	71.96%	70.33%		78.94%	
Other	6.77%	9.00%		2.69%	
Sex					
Female	28.67%	26.85%	51.18%	48.79%	76.93%
Male	71.33%	73.15%	48.82%	51.21%	23.07%

Notes: The table reports the distribution of respondents across demographic categories in our Census of Small Scale Manufacturers (2016), Experiment 1 (2017), Experiment 2 (2025), and the Demographic and Health Surveys (DHS) for Zambia and Sub-Saharan Africa (SSA), restricted to self-employed individuals. Countries in SSA: Benin (2006), Democratic Republic of Congo (2007), Republic of Congo (2005), Ethiopia (2000), Ghana (2008), Guinea (1999), Kenya (2008–2009), Liberia (2007), Madagascar (2008–2009), Mali (2012–2013), Nigeria (2008), Namibia (2006–2007), Rwanda (2000), Sierra Leone (2008), São Tomé and Príncipe (2008–2009), Uganda (2006), Zambia (2007), and Zimbabwe (2005–2006). Percentages indicate the share of individuals in each category within each dataset. For the DHS, estimates refer to adult respondents aged 18–59. Missing cells indicate that the corresponding variable was not collected in that dataset.

Table C.26: Trust and Courts: Small-Scale Manufacturers vs General Population

|	Small-Scale Entrepreneurs			General Population		
	Census (2021)	Experiment 1 (2017)	Experiment 2 (2025)	Zambia (2007; 2022)	SSA (2005–2014; 2021–2023)	US (2011)
Trust (WVS)						
Generally speaking, would you say that most people can be trusted, or that you can't be too careful in dealing with people?						
Most people can be trusted	6.04%	6.31%	6.06%	11.55%	16.46%	38.17%
Can't be too careful	93.96%	93.69%	93.94%	88.45%	83.54%	61.83%
How much do you trust people you meet for the first time?						
Trust Completely or Somewhat	26.02%	14.83%		16.29%	29.32%	38.73%
Do not Trust Much or Not at All	73.98%	85.17%		83.71%	70.68%	61.27%
Equal treatment in courts (Afrobarometer)						
In your opinion, how often, in this country: Are people treated unequally under the law?						
Never/Rarely			49.83%	45.78%	61.89%	
Often/Always			50.17%	54.22%	38.11%	
How much do you trust each of the following, or haven't you heard enough about them to say: Courts of law?						
Not at all/Just a little			29.12%	31.89%	49.34%	
Somewhat/A lot			37.88%	65.53%	48.29%	
Don't know/Haven't heard enough to say			33.00%	2.59%	2.37%	

The table compares trust and perceptions of equal treatment in courts in the Census of Small-Scale Manufacturers (2016), Experiment 1 (2017), and Experiment 2 (2025), relative to the general population using the Afrobarometer, World Values Survey (WVS), and U.S. data. "Zambia" and "SSA" columns report population-weighted averages from Afrobarometer and WVS for Zambia and Sub-Saharan African countries, respectively. WVS Sub-Saharan Africa countries include: Ghana, Nigeria, Rwanda, South Africa, Zimbabwe, Burkina Faso, Ethiopia, Mali, and Zambia. Afrobarometer Sub-Saharan Africa countries include: Angola, Benin, Burkina Faso, Botswana, Cameroon, Cabo Verde, Central African Republic, Côte d'Ivoire, Ethiopia, Gabon, Ghana, Guinea, Kenya, Lesotho, Liberia, Madagascar, Mauritius, Mali, Malawi, Mozambique, Mauritania, Namibia, Niger, Nigeria, South Africa, Senegal, Seychelles, Sierra Leone, São Tomé and Príncipe, Sudan, Tanzania, Togo, Uganda, Zambia, and Zimbabwe. Years separated by semicolons in headers refer respectively to the WVS and Afrobarometer waves years. Blank cells indicate that the question was not included in that dataset. Percentages represent within-group shares of respondents selecting each answer category.

Table C.27: Gender Attitudes: Chiefs vs General Population

|	Chiefs		General Population		
	Survey (2021)	Experiment 2 (2025)	Zambia (2007; 2014)	Sub-Saharan Africa (2005–2014; 2012–2014)	US (2011; 2014)
WVS					
Men make better political leaders than women do					
Agree	42.11%	50.00%	49.69%	58.93%	18.77%
Disagree	57.89%	50.00%	50.31%	41.07%	81.23%
Men make better business leaders than women do					
Agree	26.32%	8.33%	47.07%	54.68%	11.64%
Disagree	73.68%	91.67%	52.93%	45.32%	88.36%
When job are scarce, men should have more right to a job than women					
Agree	23.68%	33.33%	39.47%	47.96%	7.58%
Disagree	76.32%	66.67%	60.53%	52.04%	92.42%
GPP					
Imagine that the police detain two persons equally suspected of committing a crime. In your opinion, would the following characteristics place one of them at a disadvantage?					
A woman	56.00%	75.00%	44.82%	32.90%	27.79%
A person from an ethnic group or tribe other than that of the police officer involved	50.00%	33.33%	44.42%	41.67%	71.96%
A foreigner (immigrant)	46.67%	75.00%	55.38%	37.99%	77.51%

Notes: The table compares gender attitudes among market chiefs and the general population using data from the Chiefs' Survey (2021), Experiment 2 (2025), the World Values Survey (WVS), and the (GPP). "Zambia" and "Sub-Saharan Africa" columns report population-weighted averages from the respective sources. WVS Sub-Saharan Africa countries include: Ghana, Nigeria, Rwanda, South Africa, Zimbabwe, Burkina Faso, Ethiopia, Mali, and Zambia. GPP Sub-Saharan Africa countries include: Botswana, Burkina Faso, Cameroon, Côte d'Ivoire, Ethiopia, Ghana, Kenya, Liberia, Madagascar, Malawi, Nigeria, Senegal, Sierra Leone, South Africa, Tanzania, Uganda, and Zimbabwe. Years separated by semicolons in headers refer respectively to the WVS and GPP waves years. Blank cells indicate that the question was not included in that dataset. Percentages represent within-group shares of respondents selecting each answer category.

**Figure C.14:** Experiment 2: Gender Gaps in Power by Market
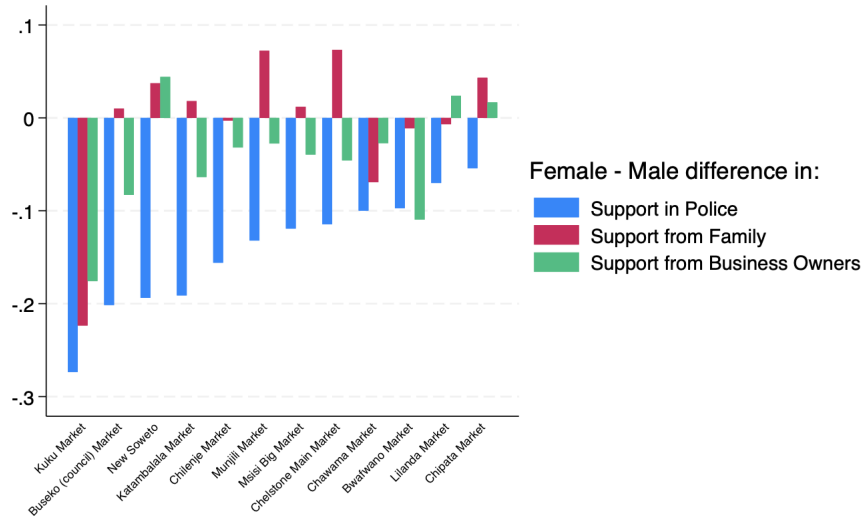
The figure shows the gender gap (F-M) in the components of the power index based on available support from the police (in blue), family (in pink) and other business owners (in green) in case of a dispute, for all the markets in experiment 2 (N=12). For instance, blue bars show the difference in the share of women minus the share of men in each market who report having someone in the police who could support them in case of a dispute.

Figure C.15: Experiment 2: the Revelation Trust Premium is Negatively Correlated between Men and Women across Markets

Coef (SE, pval): -0.48 (0.18, 0.026)
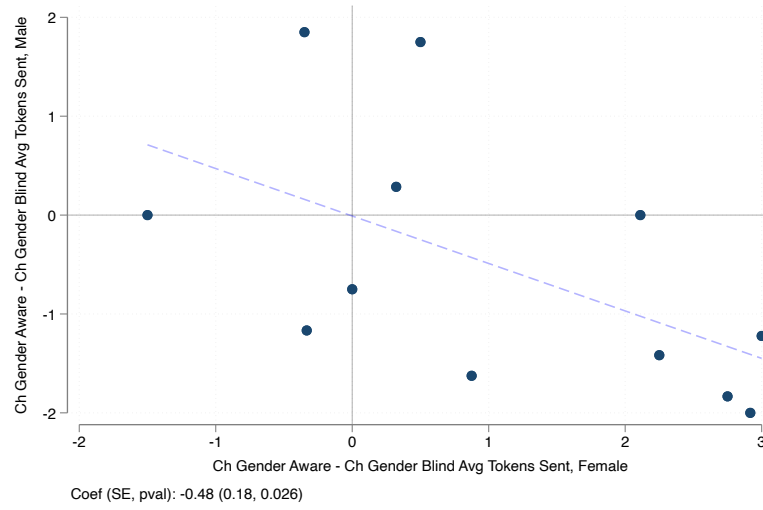
The figure shows the difference in trust in the gender-aware minus gender-blind condition ("revelation trust premium"), for women on the x-axis and men on the y-axis. Each dot is a market from experiment 2 (N=12). The dashed line shows the linear best fit.
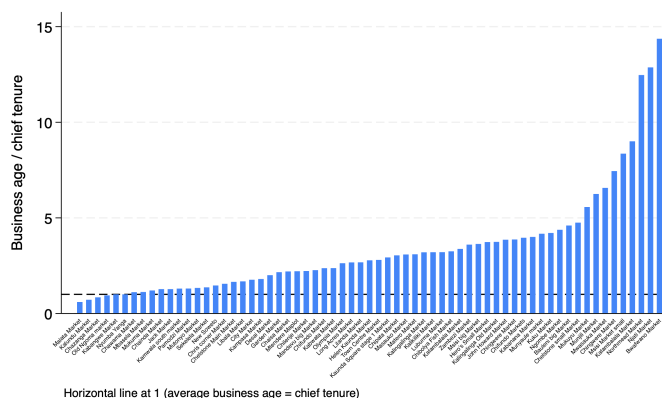
Figure C.16: Business Turnover by Market Location, Gender and Chief's Gender Bias
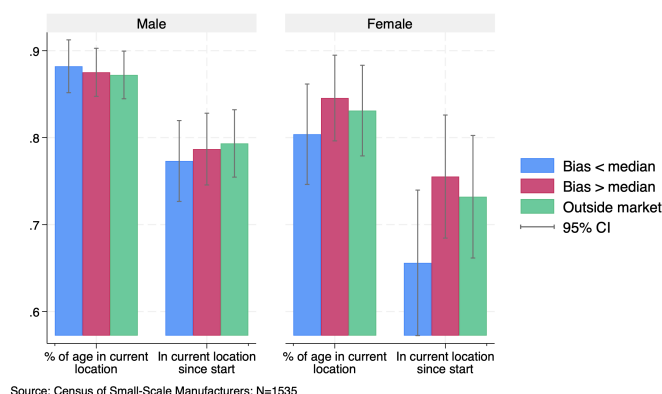
(a). Time spent in current location
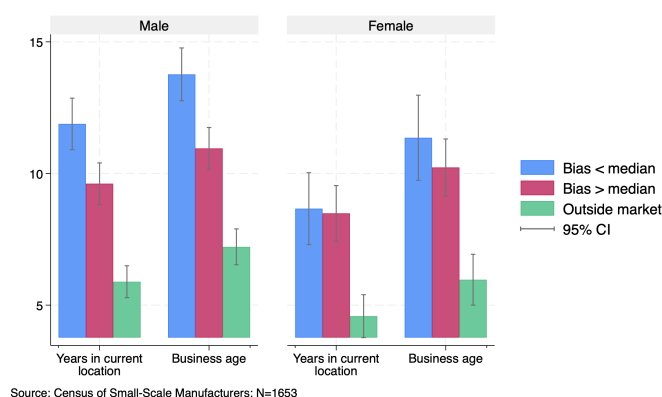
(b). Years in current location and business age

(c). Ratio of average business age and chief tenure

This figure shows business turnover and age across markets in our Census of Small-Scale Manufacturers (N=2216). Panel (a) shows the average proportion of the business age that was spent in the current location at the time of the survey, as well as the proportion of businesses that have been in the current location since their start, by chief bias and gender. The latter variable is available for all the businesses in the dataset (N=2216), while the former is missing for around 25% of the sample. Panel (b) shows the average number of years spent in the current location and business age, by chief bias and gender. The latter variable is available for all the businesses in the dataset (N=2216), while the former is missing for around 25% of the sample. Panel (c) shows the ratio of average business age and chief tenure for each market surveyed. Chief tenure is asked in the chiefs' survey (from 2021) and refers to what is the most common length of a chief's mandate in the market.

D   Chiefs Survey and Indices Construction

D.1   2021 Chiefs survey

The Chiefs survey was collected between the end of July 2021 and the beginning of September 2021. The survey was implemented over the phone in most of the cases to guarantee the safety of the surveyors during the collection. One of the challenges was being able to find chiefs who were operating in the market in 2016 (when our main Census data were collected). Our approach was to visit most markets in person and ask marketeers if they knew who was the market chief in 2016. If we couldn't find the 2016 chief after several attempts, we interviewed the current chief for that particular market. We have a total of 86 respondents in our data, for a total of 76 unique markets. Market chiefs from 2016 are 51. In some markets, we interviewed both the current and old market chief (10 markets), but we consider only the 2016 market chief for the analysis. As we have 81 markets in our data, we were thus able to cover most of the markets with our interviews. In terms of business representativeness, the chiefs interviewed belong to markets were 531 female firms out of 538 in-market female firms are located. If we consider only the 51 chiefs from 2016, they cover 473 female marketeers out of 538.

The survey contained the following main sections:

- Demographics, education, business status, tenure and roles in market

- Market characteristics: chief elections, chiefs' length of mandate, other commitees, written record keeping of market affairs

- Scenario questions on inter-gender dispute resolution

- Gender attitudes and trust

To measure how the chief behaves in dispute resolutions between men and women, we proposed two different scenarios:

Scenario 1: A woman marketeer gets an unusual order and asks a male marketeer to help her. She says that he agreed to produce 50 pieces for 5000KW. But the next day he refuses to help for that quantity. The female marketeer asks you for help, but the man says that they had agreed that he would produce 30 pieces for 5000KW. Now we are going to ask you some questions about the decisions you would take in this scenario.

Scenario 2: A female marketeer leaves the shop unattended for 10 minutes. She comes back and realizes that some of her products have been stolen. She comes to you for help.

We ask follow-up questions belonging to four broad categories: i) institutional support for women, ii) formal and objective rules for adjudication, iii) practice and procedures followed by the chief, iv) likely outcome of the dispute.

Answers to the different questions of the survey feed into the construction of three different indices of gender bias, legal competence and managerial duties. Table D.2 shows the components of the different indices. For all of them, a common methodology was used following Kling et al. (2007). All variables are first consistently signed (e.g. higher value associated with higher bias); (ii) each component of the index is then standardized by subtracting the overall mean and dividing by the overall standard deviation; (iii) the sum of the standardized

114

components is taken and (iv) the sum is standardized again. Finally, a dummy variable is created by taking values above or below median value.

Table D.1 shows descriptive statistics from the first two sections of the survey.

Table D.1: Descriptive Statistics from Chiefs Survey

|	(1) Full sample	(2) Council markets	(3) Cooperative markets
Age	47.68	46.87	48.04
	(8.156)	(8.181)	(8.197)
Female	0.08	0.09	0.08
	(0.271)	(0.288)	(0.267)
Secondary school	0.57	0.74	0.49
	(0.499)	(0.449)	(0.505)
Chief has a business	0.87	0.83	0.89
	(0.340)	(0.388)	(0.320)
Business age	8.13	8.13	8.14
	(1.456)	(1.125)	(1.584)
Business is inside mkt	0.92	0.95	0.91
	(0.267)	(0.229)	(0.282)
Home within 20 minutes	0.53	0.39	0.58
	(0.503)	(0.499)	(0.497)
GSS Trust	0.09	0.04	0.11
	(0.291)	(0.209)	(0.320)
Chief start year	2016	2016	2016
	(3.131)	(3.350)	(3.065)
How often market meetings	3.97	4.09	3.92
	(1.119)	(1.125)	(1.124)
Mandate length	3.80	4.29	3.60
	(1.653)	(1.953)	(1.485)
Less than 24 hours to arbitrate	0.82	0.78	0.83
	(0.390)	(0.422)	(0.379)
Share of services responsible for	0.52	0.56	0.50
	(0.306)	(0.307)	(0.307)
Chief is elected	0.87	0.74	0.92
	(0.340)	(0.449)	(0.267)
N	76	23	53

The table shows descriptive statistics from the Chiefs Survey. "Lives within 20 minutes" shows the percentage of chiefs who live within a 20 minutes walk from the market. The variable "Share of services responsible for" is computed as follows: we listed 12 different services that are commonly offered in the market (e.g., fee-paying toilets, waste management) and we asked the chief which ones are present and whether s/he is directly responsible for each of them. The variable is the average share of services that the chief is responsible for. Standard errors in parentheses.

Table D.2: Indices Components

	Question Type	Questions	Source
Gender Bias Index	Inter-gender Dispute Scenarios	Chief uses subjective decision rules (scenarios 1 and 2)	Own
		Woman is unlikely to win dispute (from scenario 1)	Own
		Differences in the likelihood that the chief will catch a man's vs woman's perpetrator (from scenario 2)	Own
	Self-reported Attitudes	Women are better political leaders	WVS
		Women are better business executives	WVS
		Men have more right to a job when jobs are scarce	WVS
		Women (but not other groups) are at a disadvantage in dealing with police	WJP
		Women (but not other groups) are at a disadvantage in dealing with market leaders	WJP
Legal Competence Index	Inter-gender Dispute Scenarios	Number of times perpetrator is caught	Own
		Written records of disputes are kept	Own
		Number of pieces of evidence to decide	Own
		Time to reach a decision	Own
	Market characteristics	Chief part of disciplinary committee	Own
Managerial Duties Index	Market characteristics	Number of services in market	Own
		Share of services under chief's responsibility	Own
		Written records of disputes, behaviors, fees, entrants/leavers	Own
	Relationship with marketeers	How many ways to meet	Own
		Chief's phone number available to all	Own
		Frequency of market meetings	Own

Figure D.1: Cooperation by Chief's Characteristics: Women Inside vs Outside Markets
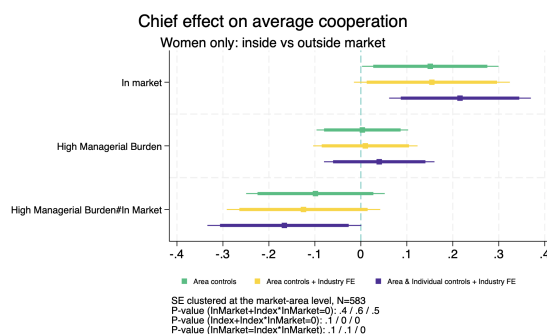
(a). Cooperation and Chief's Bias
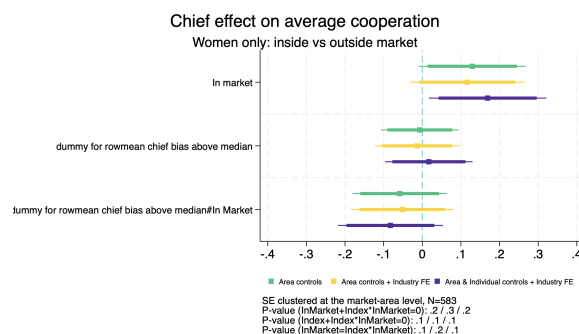
(b). Cooperation and Chief's Legal Competence
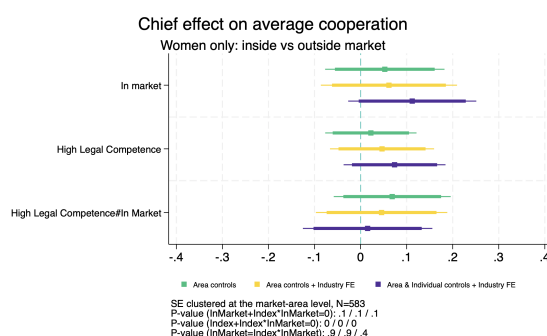
(c). Cooperation and Chief's Managerial Duties

Each figure in Panels (a), (b) and (c) show the coefficients of three regressions, which differ only in the controls used. All the specifications regress average cooperation on a dummy variable for being located inside a market ("In Market"), a dummy variable for having a chief with above average score on a certain index and the interaction between the two. In Panel (a), the dummy is for above-median score in the index of gender bias. In Panel (b), the variable is for above-median score in the index of legal competence and in Panel (c) for above-median managerial duties. The top coefficient (green) is for a regression with area density controls, the intermediate (yellow) coefficient is for a regression with density controls and industry fixed effects, and the bottom coefficient (purple) is for a regression with density and individual controls and industry fixed effects (Naics 3). Average cooperation is an index of cooperative behavior, calculated as the average of the four dummies "Joint Buy", "Lent", "Advice" and "Share Order" defined in the main tables. Error bars show 95% and 90% confidence intervals. The construction of the indexes on gender bias, legal competence and managerial duties is explained above.

D.2 Experiment 2: Chiefs survey

While collecting data from business owners located in markets in experiment 2, we also administered a Chief survey to the market chiefs of the 12 markets where we conducted the data collection.

Following the pre-registration of our experiment, we asked again the same questions that allows us to construct the variables that make up the "chief bias" index. However, to keep the survey length more limited, we skipped some questions and thus cannot construct the "managerial competence" and "legal competence" indices.

In addition to the chief bias index, to capture chief's behavior and attitudes related to gender, we also added the following questions:

- **Triad picture task**[59]:to assess whether respondents group individuals based on gender (Figure D.2). In this task, respondents are shown four different cards. Each card displays a woman or a man at the top and two individuals at the bottom—one male and one female—who differ or are similar compared to the person at the top along dimensions such as occupation or clothing. Respondents are asked to choose which of the two figures at the bottom is "the same type of person" as the one at the top. The task is designed to reveal which attributes respondents see as most salient when categorizing others. We construct a binary measure of gender salience, equal to 1 if the respondent selects the bottom figure who shares the same gender as the one on the top in more than two of the four cards, indicating a tendency to group based on gender (rather than on occupation or clothing). We get the same results if we consider more than three of the four cards, or all four cards. This measure is used in Figure 12a and called "Chief Gender Salience."

- **Hypothetical adjudication within the game**: We asked each chief how they would arbitrate a set of trust-game complaints brought by either a female or a male player. Chiefs responded to the same set of hypothetical complaints for both genders, allowing us to compare arbitration behavior by complainant gender. For each complaint, we recorded: (i) the maximum number of tokens sent back by the Trustee at which the chief would intervene (a proxy for the likelihood of action) and (ii) the number of tokens they would reallocate at this maximum threshold (a proxy for redistributive behavior). From the latter metric, we construct the implied percentage change in the return ratio implemented by the chief. These measures are used in Table C.22, Panel E.

Figure D.2: Triad Picture Task

---
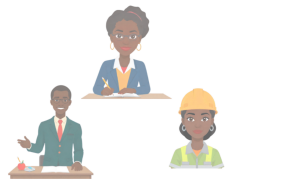
[59]Triad picture tasks have been used in the experimental psychology and cognitive neuroscience literature to study implicit categorization (Gentner and Brem, 1999; Lin and Murphy, 2001). Markman and Hutchinson (1984) first used them to examine how children group items based on taxonomic relations — grouping by shared category — versus thematic relations — grouping by functional association. More recently, Xu et al. (2022), use an adapted version of this task to show that categorization patterns are shaped by individuals' cultural environments. See Mirman et al. (2017) for a full review.

D.3 Experiment 2: Indices

Here is a more detailed explanation of the indices shown in the results of experiment 2. We divide them between survey measures and experimental measures:

Survey measures. The following indices were constructed from responses to questions in the pre-games survey and are used in Figure 12b:

- Index: "Means of social punishment"

  – This index is the market-level average for female business owners of three indicator variables for the woman reporting they have support in case of a dispute from someone in their family, someone in the police and another business owner. These questions were asked only to manufacturers in the market, and we impute with the market-gender mean for businesses in other industries.

- Index: "Historical Female Empowerment"

  – This index is the average for women in each market of 3 variables which relate to the historical gender norms of the ethnic group that a woman belongs to. Each of these three variables takes values -1 to 1. The first variable is equal to -1 if the woman's ethnicity is historically patrilineal, and 1 if it is matrilineal. The second variable is equal to -1 if the woman's ethnicity is historically patrilocal, and 1 if it is matrilocal. The third variable is equal to -1 if men used to participate "more" or "appreciably more" in agriculture, and 1 if women used to participate "more" or "appreciably more" in agriculture in the associated tribe. These three variables are constructed using Murdock's ethnographic Atlas.

Experimental measures. The following measures or indices are constructed from the trust game. These measures are used in Figures 12a, C.10 and Table C.22, Panel D.

- Preferences for having gender revealed to the chief

  – After knowing their random treatment assignment, participants are asked: "Which form were you hoping to get: the one where the chief would know you are a business man/woman or the one where he would not know? Did you have any preference?." We text code all the answers using two treatment-blind research assistance to distinguish between: i) Preference for blindness (blue form), ii) Preference for revelation (red form), iii) No preference.

- Incentivized beliefs on chief behavior

  – At the end of the experiment, we elicited incentivized beliefs about the chief's choices when s/he has to arbitrate complaints coming from men or women. The assignment of complaints brought by men or women was randomized between subjects, stratifying by the main strata and treatment assignment (as described in Appendix B.2). For each complaint, we recorded participants' beliefs on: (i) the maximum number of tokens sent back by the Trustee at which the chief would intervene (labeled "cutoff," which is a proxy for the likelihood of action) and (ii) the number of tokens they would reallocate at this threshold (a proxy for redistributive behavior). We incentivized guesses by matching responses with the chief's answers from their own survey. The measures used in Figure 12a

are the difference in beliefs for women's minus men's complaints at the market level. The index shown in Figure 12a aggregates these two measures of cutoff and return ratio by averaging them, and then standardizing the resulting average across markets.

Online Appendix - Survey Measures and Games Behavior

In this Appendix we show that our survey measures of trust are correlated with trustworthiness, as shown in previous research (Glaeser et al., 2000). Moreover, the survey cooperation measures are correlated with the number of tokens sent by the investors.

Table D.1: Trust Survey Measures and Behavior as Investor

| | Tokens Sent | | | |
	(1)	(2)	(3)	(4)
Lent	0.256	0.545**		
	(0.232)	(0.240)		
Gave Advice	1.124***	0.942***		
	(0.346)	(0.356)		
Joint buys	-0.0555	-0.0329		
	(0.214)	(0.230)		
Shared Order	-0.420*	-0.421*		
	(0.241)	(0.253)		
Average Cooperation			0.278	0.497
			(0.339)	(0.377)
Observations	474	474	474	474
Adjusted R^2	0.017	0.041	-0.001	0.023
Mean Dep Var	3.97	3.97	3.97	3.97
Industry FE		✓		✓
Market FE		✓		✓

Robust standard errors in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

This table shows correlations between survey measures of cooperation between businesses and the number of tokens that investors send to their partners. The dependent variable is the number of tokens sent by the Investor to the Trustee. The mean of the dependent variable is 3.97 and its standard deviation is 2.21. The variables "Joint Buys", "Lent", "Gave advice" and "Shared Order" are indicator variables that indicate whether a person ever engaged in the relevant activity. The variable "Average Cooperation" is an index of cooperative behavior, calculated as the average of the four dummies "Joint Buy", "Lent", "Advice" and "Shared Order". Robust standard errors in parentheses.

Table D.2: Trust Survey Measures and Behavior as Trustee

	Average Return Ratio	
	(1)	(2)
Trust in Strangers	0.0341**	0.0300*
	(0.0173)	(0.0180)
Trust in Neighbors	0.0182	0.0104
	(0.0119)	(0.0125)
Trust GSS	0.0207	0.0248
	(0.0142)	(0.0171)
Observations	476	476
Adjusted R^2	0.014	0.006
Mean Dep Var	.44	.44
Industry FE		✓
Market FE		✓

Robust standard errors in parentheses. * $p<0.10$, ** $p<0.05$, *** $p<0.01$.

This table shows correlations between survey measures of trust and the ratio of tokens sent by the Trustee to the number of tokens that the Trustee could have sent (average return ratio). The dependent variable is the number of tokens sent by the Investor to the Trustee. The mean of the dependent variable is 0.44 and its standard deviation is 0.14. The variables "Trust in Strangers" and "Trust in Neighbors" are measured on a scale from 1-4, and have been converted into dummy variables by combining low scores (1 and 2) and high scores (3 and 4). Low scores were given a value of 0, and high scores a value of 1. The variable "Trust GSS" is a dummy measuring whether most people can be trusted or not: a value of 1 indicates that "most people can be trusted", and a value of 0 indicates that "you cannot be too careful in dealing with people".