Title
Essays on Economic Development

Author
Leone, Samuel

Publication Date
2021

Peer reviewed|Thesis/dissertation

Essays on Economic Development

by

Samuel Leone

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Economics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Frederico Finan, Chair
Professor Edward Miguel
Professor Raul Sanchez de la Sierra

Summer 2021

Essays on Economic Development

Abstract

Essays on Economic Development

by

Samuel Leone

Doctor of Philosophy in Economics

University of California, Berkeley

Professor Frederico Finan, Chair

The mission of development economics is, fundamentally, to explain why some communities are wealthier than others. This dissertation is a collection of two essays that make progress towards fitting together key pieces of the puzzle. My hope is that the findings will inform, not only researchers, but policymakers working to accelerate economic development.

To that end, the essays make three thematic contributions. First, they leverage data and experiments from the Middle East and North Africa (MENA). While MENA's relevance to global markets is self-evident, to date economists have paid it less attention than other low- and middle-income regions. Second, the results advance the economic geography of development by exploring trade and forced migration. Third, the results advance the political economy of development by exploring corruption and the victims of conflict. These geographic and political factors are among the most important drivers of structural transformation, productivity growth, and poverty reduction.

In Chapter 1, joint with Nate Grubman and Jawaher Mbarek, I investigate corruption dynamics in international trade. Every year emerging markets import goods valued at more than $7 trillion, and in many countries shipments have to pass through corrupt customs administrations. Given these high stakes, policymakers require a deep understanding of both the causes and the effects of customs fraud. In addition, trade corruption can serve as a laboratory to study corruption writ large. One previously unexplored complexity is that bribe payers and bribe receivers often have repeated interactions. Given corruption's characteristic counterparty risks and information asymmetries (not to mention the impossibility of contract enforcement), these long-running relationships likely matter for a wide variety of outcomes across a wide variety of contexts.

To pursue these general learning objectives, we overcome the data and identification challenges inherent to investigating bribery: we build an original dataset on Tunisian customs transactions using an audit study to directly observe bribes, and we leverage a natural experiment in which a computer algorithm randomly assigns customs officials

to import shipments. There are three sets of results. First, we show that bribery and tax evasion are widespread, that bribery is collusive (not coercive), and that age (but not gender) predicts officials' corruptibility. Second, in line with a straightforward Nash Bargaining model, we show that the length of official/trader relationships increases tax evasion but decreases bribe amounts. Third, we zoom out to consider the larger macroeconomic implications and show that, in terms of lost tax revenue, bribery costs the Tunisian government 0.7% of GDP or $80 per citizen.

In Chapter 2, joint with Edward Miguel, Sandra V. Rozo, Emma Smith, and Sarah Stillman, I investigate affordable housing for forced migrants. There are 82 million forced migrants in the world today, and the vast majority reside in "host communities:" they live, work, and attend school in the same neighborhoods as the citizens of their host country. As both a cause and an effect of this arrangement, many governments, IGOs, and NGOs support these refugees' shelter through either direct provision or rental assistance. However, the evidence on the welfare implications of housing subsidies is inconclusive for refugees in particular and for high-poverty populations in general. On the one hand, studies have shown that shelter assistance has the potential to provide, not just a roof overhead, but a number of downstream benefits. On the other hand, some studies have found null or even negative results, and it is often unclear whether the positive results are due to direct housing-quality effects or indirect neighborhood effects.

To fill this knowledge gap, we ran a randomized controlled trial (RCT) testing a program that provides housing support to Jordan's Syrian refugees. The subsidies are large (approximately $190 per month) and are not vouchers; households can only use them for their current housing, allowing us to be among the first to identify housing-quality effects absent confounding neighborhood effects. There are three sets of results. First, we estimate the positive impacts of the program. Recipients experience improvements in their living situations, as well as better education and credit-market outcomes. Second, we observe that, while the program does not cause movement *of* households, it does cause movement *between* households. Recipients welcome new individuals into their homes, likely so that family, friends, and neighbors can take advantage of the upgraded housing. Third, we estimate the negative impacts of the program and argue that these unintended consequences likely follow from the endogenous household formation. For example, recipients are less food secure, consistent with having more mouths to feed.

Together, these studies have generalizable, actionable insights. Knowing that dynamics matter for corruption can help reform-minded governments decide which levers to pull as they attempt to enhance transparency. And the large economic cost of trade corruption implies a high return on investment for the governments that succeed. Our RCT reveals both the upside potential and the downside risk of large-scale affordable-housing interventions. By refining their design, implementation, and cost-effectiveness, policymakers could improve these programs and the lives of the vulnerable communities they serve.

To my parents,
Jo-Marie and Mark,
for their unconditional love.

Contents

List of Figures

List of Tables

Acknowledgments

My dissertation has been a team effort. I am extremely grateful to each and every member of the team.

First and foremost, I am humbled to have had such an esteemed committee of advisers. Frederico Finan has been the strongest advocate for my research, and he has invested more time into my studies than I have deserved. If not for his vote of confidence when I pitched my crazy-sounding idea to hire undercover bribe payers, and if not for his constant support from then to now, my corruption project would never have succeeded. Edward Miguel has been one of my role models since I was an undergraduate, and it has been a pleasure to work together. I have learned so much from his example, especially how to be an effective manager and how to compromise on neither academic rigor nor policy relevance. Raul Sanchez de la Sierra has been a mentor for seven years. Our very first conversation, about trying to answer the big questions in development economics, is what inspired me to attend Berkeley; and my experience as his research assistant in D.R. Congo strengthened and disciplined my intellectual creativity. Beyond my committee, numerous other academics have gone above-and-beyond to support my professional development. Before Berkeley, Daniele Paserman, Ty Turley, Todd Hare, and Ernst Fehr gave me the opportunities that led to my PhD. At Berkeley, Cecile Gaubert and Lucas Davis invited me to teach courses with them, and I ended up learning even more than the students did. Finally, Benjamin Faber inconvenienced himself on more than one occasion to help with the logistics of my grant and IRB applications.

I owe a special debt of gratitude to those who gave their labor and capital to my projects. Nate Grubman, Jawaher Mbarek, Sandra Rozo, Emma Smith and Sarah Stillman were the best coauthors I could have asked for. Their knowledge, skills, and insights complemented my own, and I quite literally could not have written this dissertation without them. Nate, Emma, and Sarah have been, not just coauthors, but friends; they made the work fun. Soufiene Alayet, Abir Cherif, Joaquin Fuenzalida, Simon Greenhill, Bailey Palmer, Abdulrazzak Tamim, Andy Theocharous, Saida Yengui, and Kais Zhioua were superlative research assistants who challenged me in all the best ways. I will never forget the dedication of my implementing partners, the Tunisian Institute for Strategic Studies, the Norwegian Refugee Council, the United Nations High Commissioner for Refugees, ELKA Consulting, and Mindset Social & Marketing Research. Nor will I ever forget the generosity of my financial backers, the Abdul Latif Jameel Poverty Action Lab, the Center for Effective Global Action, Innovations for Poverty Action, Private Enterprise Development in Low-Income Countries (especially its scientific coordinator Christopher Woodruff), the National Science Foundation Graduate Research Fellowship, Stefano DellaVigna, and Emmanuel Saez.

Above all, I thank my friends and family. My Berkeley classmates, especially Christina Brown, Isabelle Cohen, Marina Dias, Junyi Hou, Peter McCrory, Todd Messer, Preston Mui, and Harrison Wheeler, were the best teachers I have ever had. I will treasure our memories of solving problem sets together late into the night and of brainstorming new study ideas around campus. My parents, Jo-Marie and Mark Leone, made me the person I am today. They taught me the values of working hard, working smart, and having integrity. Most importantly, they have shown me unconditional love. My partner, Kathleen Kirsch, inspires me every day to be the best version of myself. She is my booster, my editor, and my therapist. She is my own personal comedian, the angel on my shoulder, and my best friend.

Chapter 1

Corruption Dynamics in International Trade

Joint with Nate Grubman & Jawaher Mbarek

1.1 Introduction

Every year emerging markets import goods valued at more than $7 trillion (UN, 2021). In some countries the process runs smoothly, but in many others shipments of everything from airplanes to bananas have to pass through corrupt customs administrations. This corruption likely has first-order welfare implications. International trade benefits economic growth by diffusing technological innovations and rewarding productive firms; it generates a nontrivial share of states' tax revenues; and it allows the global poor to access necessities produced overseas - as the scarcity of personal protective equipment (PPE) and vaccines during the COVID-19 pandemic has made all too clear (Acemoglu, 2009). Given these high stakes, policymakers require a deep understanding of both the causes and the effects of customs fraud.

In addition, trade corruption can serve as a laboratory to study corruption writ large. One previously unexplored complexity is that bribe payers and bribe receivers often have repeated interactions: just as traders at airports and seaports face the same customs officials on shipment after shipment, many truck drivers find the same police officers always assigned to their routes, and many restaurant owners find the same health inspectors always assigned to their neighborhoods. Given corruption's characteristic counterparty risks and information asymmetries (not to mention the impossibility of contract enforcement), these long-running relationships likely matter for a wide variety of outcomes across a wide variety of contexts.

To pursue these general learning objectives, we focus on three specific research questions: Are bribery relationships "coercive" (where officials extort traders) or "collusive" (where officials and traders partner to evade taxes), and which officials are more likely to participate in this abuse of power? How do officials and traders bargain over the bribe and tax-evasion amounts, and how does their bargaining change with an increase in the number of times that they have worked together in the past? What are the macroeconomic implications of this corruption for equilibrium trade volumes and tax revenues?

Answering these questions requires overcoming the data and identification challenges inherent to investigating bribery. One problem is the trouble with acquiring data on illegal and secretive behavior. To solve this problem, we build an original dataset on Tunisian customs transactions using an audit study to directly observe bribery. Audit studies allow researchers to measure real-time, fine-grained details on clandestine actions. Economists have used audit studies in contexts ranging from D.R. Congo to Indonesia, but to our knowledge ours is the largest to date in the international-trade sector and the first in the Middle East and North Africa. Overall, our dataset contains trade, taxation, and bribery information on thousands of transactions across hundreds of shipments. In addition, since Tunisia is in many ways a typical middle-income country (whose income *per capita* is halfway between Nigeria's and Brazil's), we believe the patterns in our data are likely to be representative beyond Tunisia's borders.

Another problem is that the factors giving rise to bribery are unlikely to be randomly assigned. This means that, in most contexts, it is hard to disentangle whether various characteristics are actually causing different corrupt behavior or merely correlated with different corrupt behavior. For example, if higher-ranked officials happen to take more bribes, then it would not be clear whether having a higher rank makes officials more corrupt intrinsically or if higher-ranked officials are just more likely to interact with more corrupt traders. To solve this problem, we leverage a natural experiment in which a computer algorithm randomly assigns customs inspectors to import shipments. The randomization allows us to more credibly estimate the models we use to answer some of our main research questions, as we describe in detail below.

There are three sets of results. First, we characterize the nature of seaport bribery and tax evasion. We show that bribery and tax evasion are widespread: 57% of transactions involve illicit money changing hands. We show that the dollar values are nontrivial: the average bribe amount is $719, and the average amount of taxes evaded is $1,202. And we show that bribery is collusive, not coercive: traders have the choice to pay less than what they owe in taxes by participating in bribery. In addition, we determine which types of officials tend to abuse their power. Using our natural experiment, which amounts to a judge-leniency design (Kling, 2006), we find that on the one hand younger officials are less likely to accept bribes, but on the other hand they evade even more taxes. We also find, contrary to previous evidence, that female officials are no more or less corrupt than their male colleagues.

Second, we analyze negotiations over bribery and tax evasion within the context of dynamic official/trader relationships. We build a game theory model to formalize our thinking and develop testable implications. The model is a straightforward Nash Bargaining model that, with only one additional set of assumptions, generates counterintuitive hypotheses consistent with our data. As the length of official/trader relationships increases, trust increases, risk decreases, and they engage in more tax evasion. However, the effect on bribes is ambiguous, since what matters is not risk *per se* but the change in the risk differential between the two players. Using this model and our natural experiment, we find that the length of official/trader relationship histories increases tax evasion but decreases bribes. We also use measures of trust and risk to provide evidence on mechanisms.

Third, we zoom out to consider the larger macroeconomic implications of bribery and tax evasion in international trade. Our data shows that on average each shipment sees 13% of taxes evaded, so one might conclude that in the aggregate the Tunisian government is losing 13% of its port tax revenue. But such a conclusion would not account for traders' demand response: if bribery affects the effective tax rate, then it likely affects overall trade volume too. To account for this equilibrium response, we use calibrated import-demand elasticities to perform a back-of-the-envelope calculation of bribery's effect on state fiscal capacity. Our preferred estimate is that bribery costs the Tunisian government 0.7% of GDP or $80 per citizen.

These three sets of results together make a number of contributions. Most directly, we contribute to the small literatures connected to trade corruption (Chalendard et al., 2020; Rijkers et al., 2014, 2017; Sequeira and Djankov, 2014) and corruption dynamics (Amodio et al., 2018; Campante et al., 2009; Dal Bo and Rossi, 2011; Macchiavello and Morjaria, 2015; Olken and Barron, 2009; Sanchez de la Sierra, 2020). We also speak to other corruption sub-literatures farther removed from our particular environment. On theory, we complement a body of work on the methods of bargaining between bribe payers and bribe receivers (Khan et al., 2016). On empirics, we complement a body of work showing that bribery is a net cost for emerging markets (Amodio et al., 2018; Colonnelli and Prem, 2021; Fismana and Svensson, 2007). And on methodology, we break ground by running an audit study in a new context (Olken and Barron, 2009; Sanchez de la Sierra and Titeca, 2017; Sequeira and Djankov, 2014). In international trade, most past studies of bribery have used indirect metrics, such as estimation-by-subtraction techniques that compare the value of imports (which face a tax-evasion incentive) to the value of exports (which do not). Ours is only the second study we are aware of that directly observes bribery in customs transactions.

We also speak to certain other policy-relevant fields of economics. By focusing on how bribery affects trade volume, we contribute to trade economics (Kee et al., 2008; Tokarick, 2010). To our knowledge, our study is the first to marry empirical estimates of bribery and tax evasion at seaports with trade elasticities to credibly estimate corruption's equilibrium effect on international trade. By focusing on how bribery affects tax revenue, we contribute to the public finance of development (Basri et al., 2019) and to "institutions" growth macroeconomics (Acemoglu, 2009). These studies tend to show that higher state fiscal capacity is a strong predictor of faster development, which underlines the welfare implications of our findings. In addition, our results on which officials are corruptible contribute to the personnel economics of the state (Finan et al., 2015).

The paper proceeds as follows. Section 2 describes our Tunisian context. Section 3 describes our data. Section 4 analyzes seaport bribery and tax evasion, including the coercive/collusive margin and officials' corruptibility. Section 5 analyzes repeated bargaining. Section 6 analyzes the macroeconomic implications. Section 7 concludes.

1.2 Context

Tunisian Customs Transactions

There are three facts essential for understanding Tunisian customs transactions, all of which have parallels to other ports around the world. First, a number of actors touch each shipment, either directly or indirectly. On one end is the firm that bought the shipment on the international marketplace. On the other end are the customs officers who clear the shipment and who fall into two categories: "liquidateurs" have primary responsibility for clearing shipments, while "reviseurs" review their liquidateur colleagues' work. (Think of the distinction as like that between the immigration official who stamps your passport and the one who checks for your stamped passport at the exit.) In the middle is the trader,[1] whose client is the firm and whose job is to intermediate between the firm and the customs officials; traders' business models vary, but in general they are responsible for managing all interactions with customs, with the shipping companies, and with other port employees. Finally, there are stevedores who move the goods from the ships to different areas around the port and eventually to the traders' trucks.[2] Figure 1.1 visualizes the connections these actors have to the shipment and to each other.

Second, the clearance process has a number of steps. Shipments arrive by ship and wait to be cleared by customs, usually within one to a few weeks. During the waiting process, sometimes shipments stay at the port, and sometimes they move to warehouses owned by the traders' clients but guarded by customs officials. The clearance itself involves inspecting the goods, inspecting the paperwork, and/or collecting the taxes; the intensity of these steps depends on the "lane" to which the goods are assigned, with the red, orange, and green lanes running from more to less intensive.

Third, the Tunisian customs administration uses a computer algorithm to randomly assign officials to shipments. We have done numerous focus groups with traders and retired officials to learn more about how this computer algorithm works. To our knowledge, when a shipment arrives at the port, it is assigned to one of all available officials, with equal probability and unconditional on any descriptive characteristics.

Tunisia's Political Economy

In Tunisia, bribery has long been quotidian (Yerkes and Muasher, 2017). One of the motivations for the 2010-2011 protests that toppled dictator Zine Abidine Ben Ali was a concern with the corruption that had come to characterize his regime. Investigations conducted following Ben Ali's 2011 removal have revealed the customs administration to be one of the main sites in which this corruption took place. Reports published by Tunisian state institutions (National Commission for Investigations of Bribery and Corruption, 2011; Truth and Dignity Commission, 2019) and academic researchers (Rijkers et al., 2017) have

[1]Tunisians call traders "transitaires."

[2]Tunisians call stevedores "STAM" workers.

detailed the ways that associates of Ben Ali manipulated the declaratory system to evade taxes or secure other advantages.

The political transition precipitated by Ben Ali's removal has allowed for some anti-corruption reforms. Authorities have periodically either arrested or forced the retirement of officials suspected of bribery.[3] In 2018, the customs administration began to employ cameras at a number of its facilities. In 2019, it began to require customs officials to wear photo identification. In 2019, the customs administration announced a five-year plan dedicated to the "modernization" of the administration; the strategic plan in part focused on establishing integrity and transparency. Also in 2019, the customs administration received an award by the anti-corruption commission for its progress in combating corruption.

But many Tunisians believe that the customs administration remains extremely corrupt, with many raising the prospect that it has in fact worsened during the transition. A December 2020 corruptions perception survey fielded by INLUCC asked respondents to name the institutions in the country most afflicted by corruption; more respondents (53%) named the customs administration than named any other institution outside of the security sector. Asked to specifically evaluate the level of corruption in the customs administration, approximately 90% estimated it to be a serious problem (GIZ and INLUCC, 2020). Some Tunisian elites have hypothesized a "democratization of corruption" after the revolution, meaning a spread beyond the groups that were close to Ben Ali. Academic research on tax evasion before and after the ouster of Ben Ali suggests that corruption remains serious - and has perhaps worsened. Indeed, Rijkers et al. (2017) find tariff evasion to have increased overall, even if its practice by those who were politically connected to Ben Ali has not.

Finally, anecdotally, bribery is consequential to the economy. The customs administration claims to contribute one quarter of all tax revenue collected by the Tunisian state, but Rijkers et al. (2014) and Rijkers et al. (2017) show that the corrupt practices of politically connected firms during the Ben Ali regime harmed market competition and tax collection. The qualitative work we did at the start of this project is consistent with this literature. For example, we spoke with a Tunisian businessman who imports pool chemicals and used to import mopeds. He went out of business as a moped-importer because his competitors were selling their goods at a markdown - a markdown that they could only afford by engaging in bribery-facilitated tax evasion. He decided to pivot to the market with the highest-margin tariff-free goods so that he would be protected against corrupt competitors, and that market ended up being pool chemicals. Similarly, ICG (2017) includes an interview with a smuggler who was dissuaded from engaging in the legal vegetable trade with Algeria by a customs official who preferred that the two collude in the illegal import of donkeys.

[3] In 2017, then-prime minister Youssef Chahed announced the beginning of a "war on corruption" starting with the arrest of customs official Colonel Ridha Ayari, along with three businessmen. In May 2020, the government announced the forced retirement of 21 customs officials suspected of engagement in corruption. In June 2020, the national anti-corruption commission (INLUCC) referred a captain in the customs administration for prosecution when it discovered that he had acquired a number of real estate holdings. In December 2020, the minister of the environment and a number of others, including a high-level customs official, were arrested for their alleged involvement in a *quid pro quo* allowing the illegal importation of 282 containers filled with municipal waste from Italy.

1.3 Data

Audit-Study Data

We build an original dataset on Tunisian customs transactions using an audit study to directly observe bribery, to our knowledge the largest to date in the international-trade sector and the first in the Middle East and North Africa.

The benefit of audit studies is that they allow researchers to measure real-time, fine-grained details on individual clandestine actions. For example, Olken and Barron (2009) hire Indonesian truck drivers to measure bribes at military checkpoints; Sanchez de la Sierra and Titeca (2017) hire Congolese bus riders to measure bribes to traffic police; and Sequeira and Djankov (2014) hire Southern African traders to measure bribes in customs transactions. In our case, our dataset contains 1,968 shipments processed between November 2019 and June 2021,[4] with information on the shipment itself, the trader, the trader's client, the customs officials who interact with the trader, legal characteristics of the transaction (e.g. taxation), and illegal characteristics of the transaction (e.g. bribery).

To build such a dataset, we recruited 18 traders to complete surveys on their shipments immediately after clearance.[5,6] The traders received financial compensation for the surveys they completed. Also, in order to mitigate their potential legal liability, we partnered with the Tunisian Institute for Strategic Studies (ITES), a thinktank embedded in the office of the president.[7] Because breach of confidentiality was the greatest risk to participants, we avoided collecting any personally identifiable information and created a firewall between participant- and data-facing members of the research team;[8] we also provided them with code names in the form of car brands, as seen in Figure 1.2. Although we initially allowed participants to fill out the survey using paper or digital questionnaires, the COVID-19 pandemic required that we adopt digital and contactless processes for all data-collection efforts.

To recruit our traders, we used a combination of random and snowball sampling. We obtained a list of all registered traders in Tunisia from the customs offices' public record, randomly selected names from the list, and called them asking them to participate. This method was fruitful, but imperfect: many of the names and much of the contact information was out of date, and the majority of traders in Tunisia are unregistered/informal. (Many even illegally "rent" credentials from retired or deceased traders.) To complement this sample,

[4]The data we focus on in this paper comes from Q1-Q2 2021. That is the period after which we made updates to the survey that allowed us to observe the random assignment of officials to shipments and which allowed us to solve a problem rated to the accidental over-reporting of certain product values and tax payments. See A.9 for selected results using our whole dataset.

[5]The vast majority of our data comes from ten of these 18 traders.

[6]In consultation with IRB, for the majority of the study we limited each trader to completing ten surveys per every two-week time period. Without this limit, IRB was concerned that participating in the study could become a meaningful revenue stream, thus potentially distorting participants' behavior.

[7]ITES issued a letter to each participant offering *de facto* legal protections.

[8]Research team members involved in obtaining consent, training participants, or liaising with them throughout the study did not have access to the data. Research team members involved in analyzing the data did not have access to participants.

we also asked well-established traders to make an approach on our behalf to some of their colleagues. Overall, while our sample of traders might or might not be representative, we view our recruitment strategy as strictly dominant over the social-science practice of persuading only one firm (versus, in our case, 18) to share proprietary data.

Administrative Data

To complement our survey data, we also acquired two sources of administrative data. First, a subset of traders shared documentation called a "Declaration en Detail des Merchandises" (DDM). Traders share relevant information about their shipments with the customs administration using a web interface, and the DDM is a one-page readout of that information. It contains the declared goods, the declared value, the taxes collected, and a number of other variables. Because of concerns for the confidentiality of the traders and their clients, we avoided collecting DDMs completed during the time period of our audit study, but several shared DDMs completed beforehand and for which we therefore do not have any corresponding information about criminal activities.

Second, we downloaded open-source data on Tunisian international trade published by the Tunisian government and certain IGOs like the OECD, the WTO, and the World Bank. These data contain important variables such as the country's total annual trade volume and trade tax revenues collected through tariffs and VATs.

Data Quality

A shortcoming of audit studies is that it is difficult to ensure data quality. One might be concerned, for example, that traders would feel uncomfortable revealing whether they had engaged in bribery, leading to an underestimate of its propensity. Reassuringly, Figure 1.2 shows that most of the participants in the study did reveal engaging in bribery at least once in a while. (It seems likely that traders would be more likely to lie about *whether* they pay bribes than about the *amount* of the bribes they pay.)

To further confirm the validity of the numbers in our surveys, we use the DDMs described above; although these declarations are not a random sample, they nonetheless give us benchmarks. We therefore adopt three criteria for dropping data of dubious quality. We drop observations for which either the product value or the tax owed was more than ten times the maximum among DDMs. We drop observations for which the tax owed is more than three times that of the product value, i.e. for which the reported tax rate is more than 300%. And we drop observations for which the reported tax paid is more than two times the tax owed.[9]

On the flip side, there are some patterns in the data that speak in favor of its quality, as we can see in Table 1.1. The median shipment in the dataset has a reported product

[9] In our 1968-N dataset, 185 observations exceeded at least one of these thresholds, most commonly due to an unusually large product value. These observations were concentrated in the early periods of the study. Only six observations included in the most recent period of the study exhibited questionable quality, according to these criteria.

value of approximately $32,062 and a reported tax liability of $7,027. This amounts to a tax rate of approximately 19%, which exactly matches the value-added tax rate of 19% applied to most goods imported by Tunisia (International Trade Administration, 2020). 18% of the officials in our sample are female, which is very similar to the 20% reported by the customs administration, and there is a strong correlation between officials' reported ages and their reported ranks.

In addition, the majority of shipments (53.6%) originated in the European Union, with large proportions also coming from Turkey (14.7%), China (11.2%), other parts of Europe (9.3%), or other parts of Asia (6.7%). By value, the proportion of the shipments originating in each region is largely reflective of Tunisian imports, although it slightly overrepresents shipments originating in Turkey and Asia and underrepresents those originating in the EU or elsewhere in the Arab world.[10]

1.4 Seaport Bribery & Tax Evasion

Bribery & Tax Evasion

Table 1.2 shows further descriptive statistics on Tunisian customs transactions.

Tax evasion is widespread. The mean shipment owed $8,853 in taxes but paid only $7,651 in taxes. This amounts to $1,202 stolen per shipment, or an evasion rate of 13% (1 minus the 87% in the table). One shipment stole as much as $14,400.[11]

Bribery is widespread too. 57% of transactions saw a bribe change hands, and the median bribe amount in those transactions was $540. One bribe was as much as $30,394.[12]

Coercion versus Collusion

Frequency and magnitude are only two of the three most important ways to categorize bribery. It is also important to know whether the bribery is coercive or collusive. Under coercion, the official uses their power to extort the trader into paying a bribe over and above the tax owed. The official says something like, "I see that the tax owed is $100. But pay me $120 or else I'll find a pretext to delay your clearance or confiscate your goods." The following equation generalizes this example.

$$\text{Tax Paid} = \text{Tax Owed} + \text{Official's Bribe}$$

Under collusion, the official and the trader work together to find a mutually beneficial way to evade taxes. The official says something like, "I see that the tax owed is $100. Let's

[10] The under-representation of shipments from the Arab world should not be surprising since our data collection happened at seaports whereas many goods from Algeria and Libya enter through Tunisia's land borders with Algeria and Libya.

[11] Our survey asks separately about the tax *owed* and the tax *paid*.

[12] Since the maximum bribe amount is greater than the maximum tax-evasion amount, the bribe must have been coercive. See below.

forge new paperwork, only pay \$60, and split the \$40 profit 50/50." The following equation generalizes this example.

$$\text{Tax Owed} = \text{Tax Paid} + \underbrace{\text{Tax Evaded}}_{=\text{Trader's Profit}+\text{Official's Bribe}}$$

In both of these cases, the bribe is the same amount, but the welfare implications are very different. Compared to the legal baseline, coercive bribery is beneficial to the official but costly to the trader; collusive bribery is beneficial to the official and the trader, but costly to the traders' competitors and to the government collecting the taxes.[13]

With this framework, the evidence shows that bribery in Tunisian customs transactions is decidedly collusive. Table 1.3 is a pivot table showing the percentage breakdown of transactions by when they have a bribe and when they evade taxes. We can see that approximately 42% neither bribe nor evade taxes,[14] and approximately 53% both bribe and evade taxes; in other words, a total of approximately 95% are characterized by discretionary collusion, whereby the risk-tolerant officials and traders can evade taxes while the risk-averse can opt out. Only approximately 4% involved bribery without tax evasion (coercion), and less than 1% featured tax evasion without bribery (crime).

We can also visualize this relationship graphically. Figure 1.3 is a scatter plot with bribery amounts on the x-axis and tax-evasion amounts on the y-axis (where a positive number means over-paying and a negative numbers means under-paying).[15] We can see that there is some clustering at the origin, indicating the legal transactions without any bribes or tax evasion. We can also see that venturing into positive bribe territory implies venturing into positive tax-evasion territory. Furthermore, that the implied slope is greater than one confirms that the bribe is the official's share of the tax evasion. Coercive bribery would have implied mass on the horizontal axis, with positive bribery but zero tax evasion.

Similarly Figure 1.4 is a kernel density plot showing taxes paid and owed both with bribery and without bribery. We can see that the distribution of taxes paid and owed without bribery are on top of each other, whereas the distributions with bribery are shifted relative to each other thus revealing clear tax evasion.

That bribery is more commonly an instrument of tax evasion than of official extortion is also apparent from answers to a direct question in the survey. Figure 1.5 shows that, when asked why they paid a bribe, more than 60% reported that the official would let them evade some sort of tax or regulation. Meanwhile, less than 5% reported that they had paid a bribe because they feared the official would punish them otherwise. Among those who reported paying a bribe to avoid taxation or regulation, large percentages reported doing so by misreporting the shipment's weight or value.

Our collusion finding is consistent with previous scholarship on bribery in the Tunisian customs administration. Hibou (2011) argues that allowing widespread tax evasion was

[13] For both types, there are also likely implications for firms deciding to enter or exit the market.

[14] We do not observe large differences in clearing times between shipments with and without bribes, suggesting that officials do not use the threat of delays to extort coercive bribes. See Figure A.1.

[15] See also Figures A.2-A.3.

a mechanism for the Ben Ali regime to discipline potential opponents by making them complicit. Similarly, the investigation of the Truth and Dignity Commission points to under- or mis-reporting of the contents of shipping containers as the primary means by which firms engaged in corruption. Finally, Rijkers et al. (2017) calculate "evasion gaps" at Tunisia's ports. Although each of these studies focuses on the period before the Arab Spring, our research suggests that collusive bribery remains far more common than coercive bribery.

Officials' Corruptibility

Summary

In addition to characterizing *how* bribery works, we can also characterize *who* gets bribed. The purpose of this section is to determine whether easily observable demographic characteristics (ones that the customs-office management has access to) can predict which types of officials tend to abuse their power. We find that age does predict officials' corruptibility but that gender does not.

Econometrics

For each trader i, official j, time k, and shipment l, our two outcomes of interest are the probability of a bribe $1\{b \neq 0\}_{ijkl}$ and the tax paid t_{ijkl}. Let x_{jk} be one of two demographic characteristics of the official, either the decade of age a_{jk} or an indicator for being female f_j. Let λ_k be time fixed effects.[16] Let U_{jk} be a vector of other official-level controls (e.g. age, gender, and rank).[17] Let V_{ijkl} be a vector of shipment-level controls (e.g. tax owed) that our balance regression in Table A.1 says are, despite random assignment, predictive of demographic characteristics.

With these variable definitions, our regression specification is as follows, with the βs as our coefficients of interest.

$$
\begin{aligned}
1\{b \neq 0\}_{ijkl} &= \alpha^b + \beta^b x_{jk} + \gamma^b U_{jk} + \delta^b V_{ijkl} + \lambda_k^b + \epsilon_{ijtk}^b \\
ln(t_{ijkl}) &= \alpha^t + \beta^t x_{jk} + \gamma^t U_{jk} + \delta^t V_{ijkl} + \lambda_k^t + \epsilon_{ijtk}^t
\end{aligned}
$$

Because of our natural experiment, the βs have a causal interpretation, but it is worth making a subtle distinction. Suppose we had found that officials with blue eyes were more corrupt. On the one hand, since demographic characteristics like blue eyes are not themselves randomly assigned, there could be omitted variable bias so it would be incorrect to say that blue eyes cause corruptibility. (Maybe people with blue eyes are less honest.) On the other hand, we are not interested in the true "structural relationship" between iris pigment and port operations; we are interested in providing the type of information that government H.R. managers and anti-corruption auditors could use to tell which officials are more or less likely

[16] Our results are not robust to trader fixed effects. See Table A.4.
[17] Table A.3 shows additional controls.

to abuse their office. To that end, our specification is in fact capable of saying causally that putting an official with blue eyes in charge will lead to a more corrupt portfolio.

In other words, random assignment of inspectors is useful because it eliminates the possibility that certain characteristics are correlated with bribery and tax evasion merely because of how individuals happened to self-select into working together.

Results

Figures 1.7-1.8 and Table 1.4 show the results of these econometric tests. Figure 1.7 contains bar plots showing the probability of bribery for different types officials. Figure 1.8 contains kernel density plots showing the distributions of bribe amounts for different types of officials. Table 1.4 shows regression estimates using the econometric specification described above.

First, we can see that each additional decade of age increases the probability of accepting a bribe by 1.6 percentage points. This finding is consistent with many mechanisms. It could be selection, in which less corrupt officials quit or get fired for not being team players. It could be treatment, in which over time corruption gets normalized or officials learn how to be more corrupt. It could also be a cohort effect, with younger generations coming of age since the revolution being more committed to good government. Interestingly, though on the extensive margin older officials are more likely to steal, when they do they steal less. We can see that each additional decade of age increases taxes paid by 1.1 percentage points.

Second, we can see that female officials are no more or less corrupt than their male colleagues in terms of both whether they steal and how much they steal. This finding is inconsistent with a stylized fact from other studies saying that females are less corrupt than males. This could be due to selection into becoming a customs officer, where less corrupt females are screened out.

1.5 Repeated Bargaining

Game Theory Model

Summary

The purpose of this model is twofold. First, it formalizes our thinking about bribery negotiations as fundamentally about bargaining over the surplus of tax evasion, mediated by risk and trust. Second, the model generates testable implications and structures the reduced-form econometrics we describe below.

Overall, we build off of Nash (1950) and Khan et al. (2016) to structure a model that, while static, is still able to generate insights on dynamic bargaining. With only one additional set of assumptions, our model generates counterintuitive hypotheses confirmed by the data: repeated interactions increase trust, decrease risk, increase tax evasion, and may *either increase or decrease* bribery. In other words, it is possible for bribes to go down even as tax evasion goes up.

Setup

There is a trader and an official who use bribery to bargain over tax evasion. The taxes owed $\tilde{t}$ is an exogenous parameter, while the taxes paid t and the bribe b are endogenous variables that the two players optimize. Their utilities $U_T(t, b; \tilde{t})$ and $U_O(t, b; \tilde{t})$ depend on whether they do engage in tax evasion $1\{t \neq \tilde{t}\}$ or do not engage in tax evasion $1\{t = \tilde{t}\}$.

$$
\begin{aligned}
U_T(t, b; \tilde{t}) &= U_T^0(t, b; \tilde{t}) \times 1\{t = \tilde{t}\} + U_T^1(t, b; \tilde{t}) \times 1\{t \neq \tilde{t}\} \\
U_O(t, b; \tilde{t}) &= U_O^0(t, b; \tilde{t}) \times 1\{t = \tilde{t}\} + U_O^1(t, b; \tilde{t}) \times 1\{t \neq \tilde{t}\}
\end{aligned}
$$

If they do not engage in tax evasion $1\{t = \tilde{t}\}$, then the trader's utility $U_T^0(t, b; \tilde{t})$ is decreasing in taxes owed $\tilde{t}$; the official's utility $U_O^0(t, b; \tilde{t})$ is zero.

$$
\begin{aligned}
U_T^0(t, b; \tilde{t}) &= -\tilde{t} \\
U_O^0(t, b; \tilde{t}) &= 0
\end{aligned}
$$

If they do engage in tax evasion $1\{t \neq \tilde{t}\}$, then the trader's utility $U_T^1(t, b; \tilde{t})$ is decreasing in taxes t, decreasing in the risk of tax evasion (a quadratic cost-of-risk function of the tax owed $\tilde{t}$ less the tax paid t mediated by the parameter $\alpha > 0$), and decreasing in the bribe b; the official's utility $U_O^1(t, b; \tilde{t})$ is decreasing in the risk of tax evasion (a quadratic cost-of-risk function of the tax owed $\tilde{t}$ less the tax paid t mediated by the parameter $\beta > 0$), and increasing in the bribe b.

$$
\begin{aligned}
U_T^1(t, b; \tilde{t}) &= -t - \alpha(\tilde{t} - t)^2 - b \\
U_O^1(t, b; \tilde{t}) &= -\beta(\tilde{t} - t)^2 + b
\end{aligned}
$$

The two play a static Nash bargaining game where they calculate, one, the tax-evasion amount that maximizes their joint surplus and, two, the bribe that optimally splits the joint surplus. Specifically, the equilibrium tax paid t^* (which, since $\tilde{t}$ is exogenous, automatically determines the tax-evasion amount) is just what maximizes the joint surplus. And the equilibrium bribe b^* is what maximizes the "Nash product" (the product of the differences in the two players' utilities between the two states of the world) as a function of t^*.

$$
\begin{aligned}
S(t;\tilde{t}) &= (\tilde{t}-t) - \alpha(\tilde{t}-t)^2 - \beta(\tilde{t}-t)^2 \\
\Rightarrow \frac{\partial S(t^*;\tilde{t})}{\partial t} &= 0 \\
\Rightarrow t^* &= \tilde{t} - \frac{1}{2}\frac{1}{\alpha+\beta}
\end{aligned}
$$

$$
\begin{aligned}
b^* &= \text{argmax}[U_T^1(t^*,b;\tilde{t}) - U_T^0(t^*,b;\tilde{t})][U_O^1(t^*,b;\tilde{t}) - U_O^0(t^*,b;\tilde{t})] \\
\Rightarrow b^* &= \frac{\alpha+3\beta}{8(\alpha+\beta)^2}
\end{aligned}
$$

Intuitively, using the Nash product is equivalent to choosing a bribe such that the two players' utilities with tax evasion are the same as their utilities without tax evasion (their "outside options") plus one half of the joint surplus.[18,19]

$$
\begin{aligned}
U_T^1(t^*,b^*;\tilde{t}) &= U_T^0(t^*,b^*;\tilde{t}) + \gamma S(t^*;\tilde{t}) \\
U_O^1(t^*,b^*;\tilde{t}) &= U_O^0(t^*,b^*;\tilde{t}) + (1-\gamma) S(t^*;\tilde{t}) \\
\text{s.t. } \gamma &\equiv 1-\gamma \equiv \frac{1}{2}
\end{aligned}
$$

One implication is that taxes paid are increasing in the sum of α and β; intuitively, tax evasion is decreasing in both players' aggregate risk. Second, bribes are decreasing in α, since the higher the risk to the trader, the more the official has to compensate her with a lower bribe. Third, bribes are either increasing or decreasing in β, depending on a numerical threshold. On the one hand, the higher is β, the higher is the risk to the official and the more the trader has to compensate her with a higher bribe. On the other hand, the higher is β, the lower the tax-evasion amount and the less money there is to go into the bribe.

$$
\begin{aligned}
\frac{\partial t^*}{\partial \alpha} &= \frac{1}{2(\alpha+\beta)^2} > 0 \\
\frac{\partial t^*}{\partial \beta} &= \frac{1}{2(\alpha+\beta)^2} > 0 \\
\frac{\partial b^*}{\partial \alpha} &= \frac{-\alpha-5\beta}{8(\alpha+\beta)^3} < 0 \\
\frac{\partial b^*}{\partial \beta} &= \frac{\alpha-3\beta}{8(\alpha+\beta)^3} > 0 \Leftrightarrow \alpha > 3\beta \\
\frac{\partial b^*}{\partial \beta} &= \frac{\alpha-3\beta}{8(\alpha+\beta)^3} < 0 \Leftrightarrow \alpha < 3\beta
\end{aligned}
$$

[18] See the Appendix for a proof.

[19] Note that the joint surplus is denominated in utility, not money, so the official receiving one half of the joint surplus does *not* necessarily imply that the official's bribe is one half of the tax-evasion amount.

So far the setup is the same as Khan et al. (2016), except we have quadratic cost instead of linear cost. Making the function differentiable allows for a continuous set of tax evasion amounts, which is a pattern we see in our data.

Matches, Trust & Risk

Within this setup, we make only one additional nested assumption: risk α and β are decreasing functions of trust u, which is itself an increasing function of matches m.

$$\begin{aligned}
\frac{\partial u(m)}{\partial m} &> 0 \\
\frac{\partial \alpha(u(m))}{\partial u(m)} &< 0 \\
\frac{\partial \beta(u(m))}{\partial u(m)} &< 0 \\
\Rightarrow \frac{\partial \alpha(u(m))}{\partial m} &= \alpha' < 0 \\
\Rightarrow \frac{\partial \beta(u(m))}{\partial m} &= \beta' < 0
\end{aligned}$$

This assumption is motivated by both our natural experiment and our many focus groups with traders and retired officials. First, these individuals say that trust and risk are central to their interactions with each other. For example, one trader said, "The more transactions [the trader and the official] have together, the stronger the relationship becomes. So, there is trust between them."[20] Second, the traders conceptualize matches as exogenous, backward-looking variables, not endogenous, forward-looking variables. This is why the model treats matches as a state variable that the two players take as given in each independent stage game.[21]

It follows automatically from this assumption that taxes paid are decreasing in matches and therefore tax evasion is increasing in matches.

$$\frac{\partial t^*(\alpha(u(m)), \beta(u(m)))}{\partial m} = \frac{(\alpha' + \beta')}{2(\alpha + \beta)^2} < 0$$

One might think that an increase in tax evasion automatically implies an increase in the bribe. If the pie gets bigger, then it seems natural for the pieces of the pie to get bigger too. In fact, the model has an ambiguous prediction about the relationship between bribes and taxes.[22]

[20] One alternative would have been to model trust as affecting utility, not indirectly through risk, but directly through social preferences.

[21] One alternative would have been to model an extensive-form game in which the players have subjective beliefs about future matches and use backward induction to arrive at a subgame perfect equilibrium. Such a model would be a less parsimonious representation of our context and would add unnecessary complexity.

[22] See the Appendix for a proof.

$$
\begin{aligned}
\frac{\partial b^*(\alpha(u(m)),\beta(u(m)))}{\partial m} &= \frac{1}{8(\alpha+\beta)^3}[\alpha'(-\alpha-5\beta)+\beta'(-\alpha+3\beta)] \\
\Rightarrow \frac{\partial b^*(\alpha(u(m)),\beta(u(m)))}{\partial m} &> 0 \Leftrightarrow \alpha < 3\beta \cup \frac{|\beta'|}{|\alpha'|} < \frac{\alpha+5\beta}{\alpha-3\beta} \\
\Rightarrow \frac{\partial b^*(\alpha(u(m)),\beta(u(m)))}{\partial m} &< 0 \Leftrightarrow \alpha > 3\beta \cap \frac{|\beta'|}{|\alpha'|} > \frac{\alpha+5\beta}{\alpha-3\beta}
\end{aligned}
$$

Intuitively, this two-part condition says that two factors matter for whether bribes are increasing or decreasing in matches: (1) whether bribes are increasing or decreasing in the β and (2) whether the effect of β on bribes or the effect of α on bribes is stronger.

First, recall that bribes are either increasing or decreasing in β. On the one hand, the official's risk puts upward pressure on the bribe because of risk compensation. On the other hand, the official's risk puts downward pressure on the bribe because risk decreases the joint surplus from tax evasion. $\alpha - 3\beta$ is the threshold that governs this tradeoff, so it makes sense that it would reappear here. Since β is decreasing in matches, if bribes are increasing in β, then matches put downward pressure on the bribe, but if bribes are decreasing in β, then matches put upward pressure on the bribe.

Second, in the latter case, there is a "competition" between the rate of change in the risk differential between the two players, $|\beta'|/|\alpha'|$. On the one hand, matches decrease β, which puts downward pressure on the bribe. On the other hand, matches decrease α, which puts upward pressure on the bribe. $(\alpha+5\beta)/(\alpha-3\beta)$ is the threshold that governs this tradeoff.

Consider an example of this competition. Suppose an official and a trader match for the first time on a shipment that owes \$100 in taxes. And suppose their risk parameters are such that they decide to pay only \$80, evade the other \$20, and split the surplus 50/50, generating a bribe of \$10. Suppose they match together for a second time, now with higher trust and therefore lower risk. Their new risk parameters are such that they pay only \$60 and evade the other \$40, but they do not automatically agree on another 50/50 split. If the risk to the official has gone down more than the risk to the trader, then the official will have to compensate the trader for that risk. The new split could be 40/60, resulting in a higher bribe, 25/75, resulting in the same bribe, or 20/80, resulting in a lower bribe. All three scenarios, and many more, are possible, since the official prefers a lower bribe amidst higher tax evasion to the outside option of getting nothing at all.[23]

Figures 1.9-1.11 visualize examples with a 50/50 split, uneven splits, and a split that is so uneven that tax evasion and bribery are negatively correlated.

Testable Implications

We take the following three testable implications to the data.

[23]Note that, whether or not the trader and the official split the *tax-evasion amount* 50/50, they split the *joint surplus* 50/50. This is because the tax-evasion amount is denominated in money, and the joint surplus is denominated in utility.

$$\frac{\partial u}{\partial m} > 0 \tag{1.1}$$
$$\frac{\partial t^*}{\partial m} < 0 \tag{1.2}$$
$$\frac{\partial b^*}{\partial m} \neq 0 \tag{1.3}$$

Econometrics

For each trader i, official j, time k, and shipment l, our three outcomes of interest are the relationship quality between the official and the trader q_{ijkl} (which is a question in our questionnaire that asks yes/no whether the trader has an especially friendly relationship with the official), the expected value of a bribe conditional on a bribe changing hands $E[b|b \neq 0]_{ijkl}$, and the tax paid t_{ijkl}.[24] Let m_{ijkl} be the number of times in the past that the official and the trader have been matched by random assignment. Let r_{jk} be an indicator for whether the official is a reviseur (versus a liquidateur). Let λ_i and λ_k be trader and time fixed effects. Let U_{jk} be a vector of official-level controls (e.g. age, gender, and rank). Let V_{ijkl} be a vector of shipment-level controls (e.g. tax owed) that our balance regression in Table A.1 says are, despite random assignment, predictive of demographic characteristics. (Note that tax owed is not such a variable, but we include it anyway in our vector of controls out of an abundance of caution.[25])

With these variable definitions, our regression specification is as follows, with the β_{m1}s and β_{m2}s as our coefficients of interest, referring to the match effect on the liquidatuers and reviseurs respectively.

$$\begin{aligned}
q_{ijkl} &= \alpha^q + \beta^q_{m1} m_{ijkl} + \beta^q_{m2}(m_{ijkl} \times r_{jk}) + \zeta^q r_{jk} + \gamma^q U_{jk} + \delta^q V_{ijkl} + \lambda^q_i + \lambda^q_k + \epsilon^q_{ijkl} \\
ln(b_{ijkl}) &= \alpha^b + \beta^b_{m1} m_{ijkl} + \beta^b_{m2}(m_{ijkl} \times r_{jk}) + \zeta^b r_{jk} + \gamma^b U_{jk} + \delta^b V_{ijkl} + \lambda^b_i + \lambda^b_k + \epsilon^b_{ijkl} \\
ln(t_{ijkl}) &= \alpha^t + \beta^t_{m1} m_{ijkl} + \beta^t_{m2}(m_{ijkl} \times r_{jk}) + \zeta^t r_{jk} + \gamma^t U_{jk} + \delta^t V_{ijkl} + \lambda^t_i + \lambda^t_k + \epsilon^t_{ijkl}
\end{aligned}$$

We use logs because our model says that bribes and taxes are nonlinear in matches,[26] and logs are the easiest function form to interpret: one match increases or decreases the outcome by β percent.

We have addressed a number of issues to ensure that this model is well-identified. First, if matches were endogenous (if, say, officials and traders could select into working together), then we would not know if the βs were the true effects or suffering from omitted variable bias. Luckily, our natural experiment means that matches are randomly assigned. Second,

[24] See Table A.5 for the probability of a bribe as an additional outcome.

[25] Our results are not robust to including *log* tax owed in our vector of controls, which we interpret as benign; it tells us merely that the "true" structural relationship is one in which matches change the percent of taxes paid, not the percent of the share of taxes owed paid. See Table A.6.

[26] Taking second-order conditions yields $\frac{\partial^2 t^*}{\partial m^2} \neq 0$ and $\frac{\partial^2 b^*}{\partial m^2} \neq 0$.

there is the possibility of noncompliance with random assignment. We therefore show the intention-to-treat (ITT) specifications in which our main regressor is the number of matches with the randomly assigned official, whether or not the trader actually ended up working with them. The first stage is very strong,[27] so the results do not change much between the ITT and the treatment-on-the-treated (TOT) specifications.[28,29] Third, there is also the possibility that the random assignment has been corrupted (as would be the case if, say, the computer code had been compromised). Chalendard et al. (2020) do in fact have evidence of this happening at a port in Madagascar, so we have taken this concern seriously. However, we are reassured that this is not the case for three reasons: Tunisia has much higher state capacity than Madagascar; our conversations with "reformed" former customs officials and the traders in our sample, who are very forthcoming about corruption in general, say this does not happen; and our balance regressions in Table A.2 show that few characteristics are predictive of the number of matches.

Fourth, under the best of circumstances, random assignment is not enough to identify the βs: officials and traders with more matches *overall*, because they have worked at the port longer or are more productive and manage more shipments, will mechanically have more matches with each other. This is why we control for trader fixed effects and inspector characteristics, including tenure.[30,31] Relatedly, the Law of Large Numbers says that in the limit traders should end up working with all officials the same number of times. We therefore estimate our effects using random sampling variation within each trader's relatively small sample of matches. The Ministry of Finance cycles customs inspectors between ports every 2-3 years, such that each clearing agent has a relatively short history with each customs official, and hence the median number of matches is 13.

Overall, this is a very conservative specification that leaves little room for selection on unobservables.

Results

Figures 1.13-1.12 and Table 1.5 show the results of these econometric tests of the implications of our game theory model.

Testable Implication 1.1 says that as the match history goes up, trust goes up as well. Table 1.5 Column 1 shows the regression above taking relations as the outcome. (Recall that the outcome is the answer to a question in our questionnaire that asks yes/no whether the trader has an especially friendly relationship with the official.) Each additional match

[27] See Table A.7.

[28] See Table A.8 for effects using the TOT specifications

[29] In Q1-Q2 2021 we can observe the number of matches with the randomly assigned official ("Matches Original"), whereas beforehand we can only observe the number of matches with the official with whom the trader actually ended up working ("Matches Terminal"). Hence, our I.V. specifications use a subset of our data, while our non-I.V. specifications (see Table A.9) use all of our data but make a conditional independence assumption.

[30] See Table A.10 for specifications with alternative controls.

[31] We do not have the data for official fixed effects.

increases the probability that the trader describes the relationship positively by almost 1%. The result is consistent with our model-based hypothesis.

Testable Implication 1.2 says that as matches go up, trust goes up, and risk goes down; consequently, taxes paid go down. Figure 1.13 is a bin scatter plot that visualizes the raw relationship (without controls) between matches and taxes paid for liquidateurs, and we can see a clear negative relationship. More formally, Table 1.5 Column 3 shows our main regression with taxes paid as the outcome. Each additional match for the liquidateur lowers taxes paid (which is mechanically negatively correlated with tax evasion) by almost 1%. The result is consistent with our model-driven hypothesis.

Testable Implication 1.3 says that, even amidst matches increasing tax evasion, the effect on bribes is ambiguous and depends on the rate of change of the relative risk between officials and traders. Hence the specific parameter values we estimate will determine which version of our hypothesis is consistent with the data. As a start, Figure 1.12 is a bin scatter plot that visualizes the raw relationship (without controls) between matches and bribe amount for liquidateurs, and we can see a clear negative relationship. More formally, Table 1.5 Column 2 shows our main regression with taxes paid as an outcome. Each additional match for the liquidateur lowers bribes by approximately 1% and is strongly statistically significant. Interestingly, the effect is different for the different types of officials. While for liquidateurs the coefficient is -0.8%, for reviseurs the coefficient is +1%. Intuitively, this means that, relative to the trader's, the risk to the liquidateur is decreasing *slower* but the risk to the reviseur is decreasing *faster*.

Finally, we can test for heterogeneity to further explore mechanisms. Table 1.6 takes our main regression specification of bribes on matches and controls but interacts matches with an indicator for whether the goods were assigned to the red lane. Either result would be consistent with our model, depending on whether the match effect is complementary to or substitutable with the overall riskiness of the environment. If being in the red lane weakens the match effect, then we could conclude that matches matter less in an environment where the procedures guarantee a certain high degree of risk. If being in the red lane strengthens the match effect, then we could conclude that matches are especially important in a risky environment. We can see that being in the red lane weakens the match effect for liquidateurs, consistent with the former hypothesis.

1.6 Macroeconomic Implications

Calculations

Our objective is to estimate the fiscal costs of bribery by comparing the amount of revenue the government currently collects to a counterfactual scenario without bribery.

We know that taxes collected today are just the number of shipments times the taxes owed per shipment times the share of taxes *not* evaded.[32] Taxes collected under the counterfactual

[32]We also know that taxes collected today are GDP times the tax-to-GDP ratio times the trade-taxes-to-total-taxes ratio (OECD et al., 2020; World Bank, 2021). In practice, we use this alternative calculation

scenario are more complicated. The off-equilibrium approach would be to assume the change is driven entirely by the share of taxes paid going up to 100%. But this fails to account for a demand response. The on-equilibrium approach says that both the taxes paid goes up to 100% and that traders change their behavior as a result of facing higher taxes. Without this latter behavioral response, the off-equilibrium numbers are an overestimate of bribery's fiscal cost.

In reality, the response to tax increases is governed by the import-demand elasticity ϵ, which is similar to the negative slope of a demand curve in any consumer-theory context. It says that that there is some percent change in trade volume for every percent change in trade costs. We know the trade cost, which is just the change in taxes, bribes, and other fees observed in our data. Hence, with calibrated estimates of ϵ from the trade literature, we can estimate the change in trade volumes.

$T(b)$ is the government's tax revenue. t is the tax revenue per shipment. $S(b)$ is the share of taxes *not* evaded. $V(b)$ is the trade volume in shipments. $C(b)$ is the trade cost per shipment. $T(b)$, $S(b)$, $V(b)$, and $C(b)$ are all functions of either the bribery equilibrium $b \neq 0$ or the no-bribery equilibrium $b = 0$. ϵ is the import-demand elasticity. Y is Tunisia's GDP, and L is Tunisia's population.

With these variable definitions, our calculation specification is as follows.

$$\begin{aligned}
\underbrace{T(b \neq 0)}_{\text{Factual}} &= t \times S(b \neq 0) \times V(b \neq 0) \\
\underbrace{T(b = 0)}_{\text{Counterfactual}} &= t \times S(b \neq 0)[1 + \%\Delta S] \times V(b \neq 0)[1 + \%\Delta V] \\
&= t \times S(b \neq 0)[1 + \%\Delta S] \times V(b \neq 0)[1 + (\epsilon \times \%\Delta C)]
\end{aligned}$$

$$\begin{aligned}
\Rightarrow \text{Anti-Corruption Dividend (\% GDP)} &= \frac{T(b = 0) - T(b > 0)}{Y} \\
\Rightarrow \text{Anti-Corruption Dividend (\$ \textit{Per Capita})} &= \frac{T(b = 0) - T(b > 0)}{L}
\end{aligned}$$

Importantly, this calculation is only focused on fiscal costs. Bribery likely has many other economic costs, such as generating uncertainty, disrupting price signals, and rewarding risk-tolerant, politically connected, or antisocial firms over potentially more productive ones. In addition, as the political-science literature cited above shows, bribery can also erode democratic legitimacy. In other words, while we argue that failing to account for bribery and tax evasions' potential positive trade effects leaves us with a overestimate of its fiscal costs, the fiscal cost is itself likely an underestimate of bribery's aggregate costs.

because our measure of taxes owed per shipment is noisy.

Results

Using our methodology, we can calculate the tax revenue effects of bribery as a function of the import-demand elasticity. We report our results in Table 1.7 and Figure 1.14 (which has a Laffer curve interpretation). We report the tax-revenue effects as a percent of GDP and in *per capita* terms. We also report our effects assuming different elasticities, with various numbers from the trade literature. Our preferred estimate, -1.75, is approximately the midpoint of the studies of which we are aware that estimate a trade elasticity for Tunisia specifically (Kee et al., 2008; Tokarick, 2010).

Note that our trade elasticity is less than some numbers cited in the trade literature because it measures the response to a *highly aggregated* change in trade costs. Some trade elasticities measure the response of the trade in one good to a change in the price of another good, similar to a cross-price elasticity in consumer theory; our trade elasticity measures the response of the trade in an average good to a change in the average effective tax rate. Necessarily, the latter is lower than the former, since it limits between-good substitution.

Overall, Table 1.7 shows that bribery's fiscal costs are meaningfully large. Our preferred estimate is that an annual anti-corruption dividend could send every Tunisian citizen an $80 check or allow the Tunisian government to spend an additional 0.7% of GDP on education, health, antiterrorism, or any other valuable use. $80 is enough to provide bread (the staple food for low-income Tunisians) to a family of four for nine months, and 0.7% of GDP is the approximate cost of the economic damage from climate change in the United States (Hsiang et al., 2017).

1.7 Conclusion

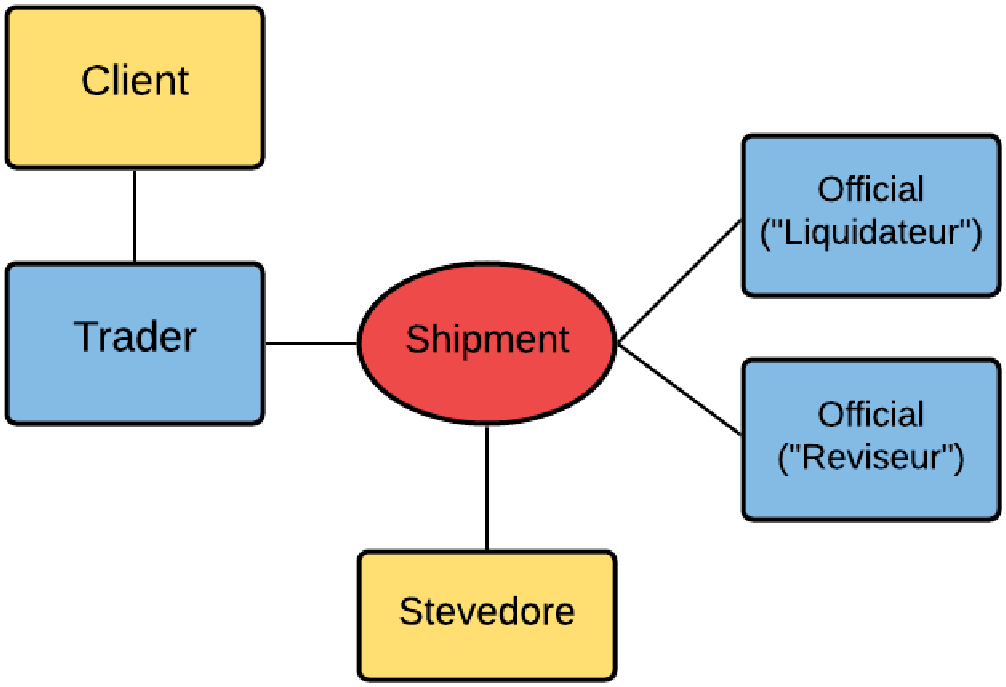
We overcome the data and identification challenges inherent to investigating bribery: we build an original dataset on Tunisian customs transactions using an audit study to directly observe bribes, and we leverage a natural experiment in which a computer algorithm randomly assigns customs officials to import shipments. There are three sets of results. First, we show that bribery and tax evasion are widespread, that bribery is collusive (not coercive), and that age (but not gender) predicts officials' corruptibility. Second, in line with a straightforward Nash Bargaining model, we show that the length of official/trader relationships increases tax evasion but decreases bribe amounts. Third, we zoom out to consider the larger macroeconomic implications and show that, in terms of lost tax revenue, bribery costs the Tunisian government 0.7% of GDP or $80 per citizen.

There are many opportunities for future work building off of these findings. First, we show that audit studies are a useful (if difficult) way of collecting data on bribery, so researchers should continue to use this methodology in a wide variety of contexts. Second, researchers should test our model, or an even richer version of it, to see if it is useful in general for explaining corruption dynamics. Third, given the difficulty in observing bribery and tax evasion in customs transactions, researchers should investigate whether firms and households have well-anchored beliefs about the costs of participating in international trade.

Our findings also have important and interesting policy implications. First, there are perhaps anti-corruption returns to hiring younger officials but not to hiring female officials, though there are of course many other reasons to get closer to gender parity in the Tunisian administration. Second, there is merit to the "regulatory capture" fear that motivates officials' random assignment and their frequent rotation to different postings around their country. Governments should continue these policies and also consider more aggressive monitoring of officials and traders known to have a high number of repeated interactions or other proxies for trustworthy relationships. Third, bribery is hugely consequential to the Tunisian government and economy and is therefore worth prioritizing. Overall, knowing that dynamics matter for corruption can help reform-minded governments decide which levers to pull as they attempt to enhance transparency. And the large economic cost of trade corruption implies a high return on investment for the governments that succeed.

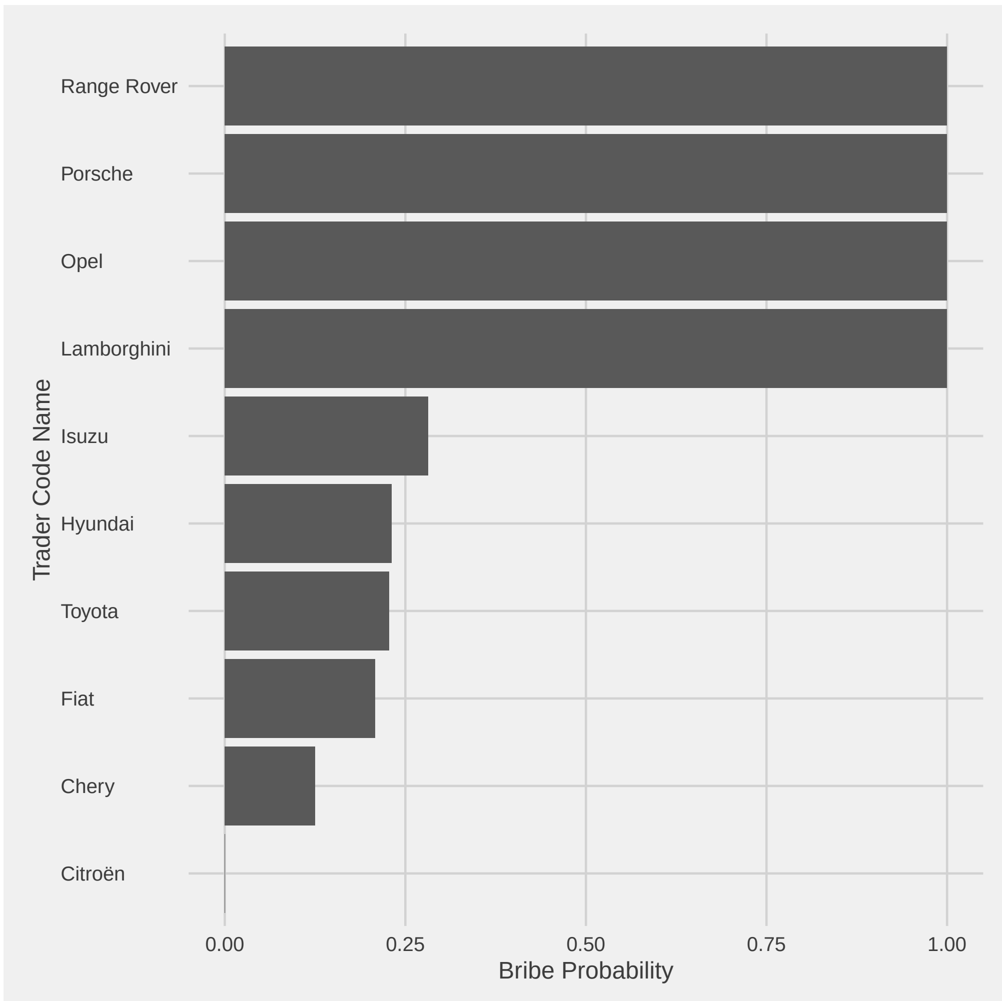
1.8 Figures

Figure 1.1: Actors Involved in Tunisian Customs Transactions



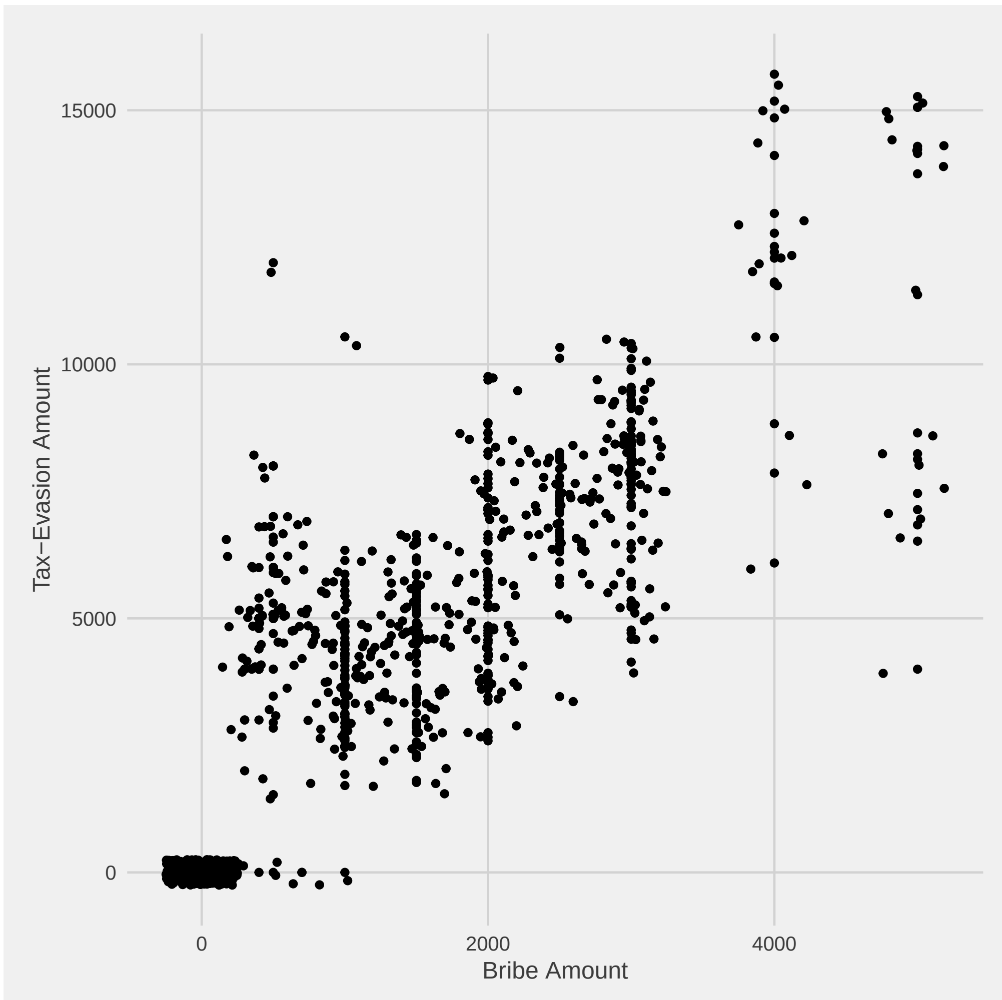
Note:
The figure is a flow chart visualizing the actors involved in Tunisian customs transactions, in relation to each import shipment and to each other. Actors who feature prominently in our study are in blue. Other actors are in yellow.

Figure 1.2: Traders' Willingness to Reveal Paying Bribes



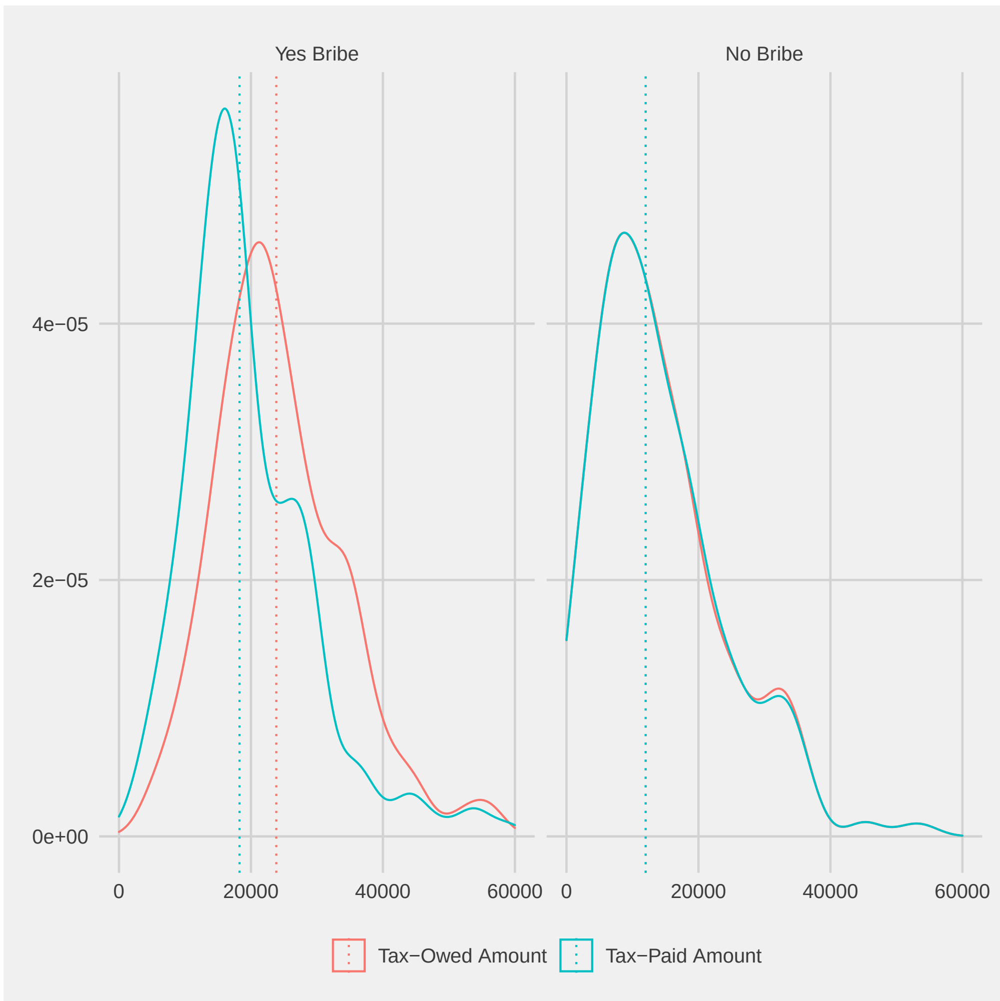
Note:
The figure is a bar chart visualizing the proportion of surveys for which the trader reveals paying bribes. The sample used is all cleaned data from the latest (2021) version of our survey, for which these ten traders who contributed the vast majority of the surveys (90%).

Figure 1.3: Tax Evasion as Function of Bribery



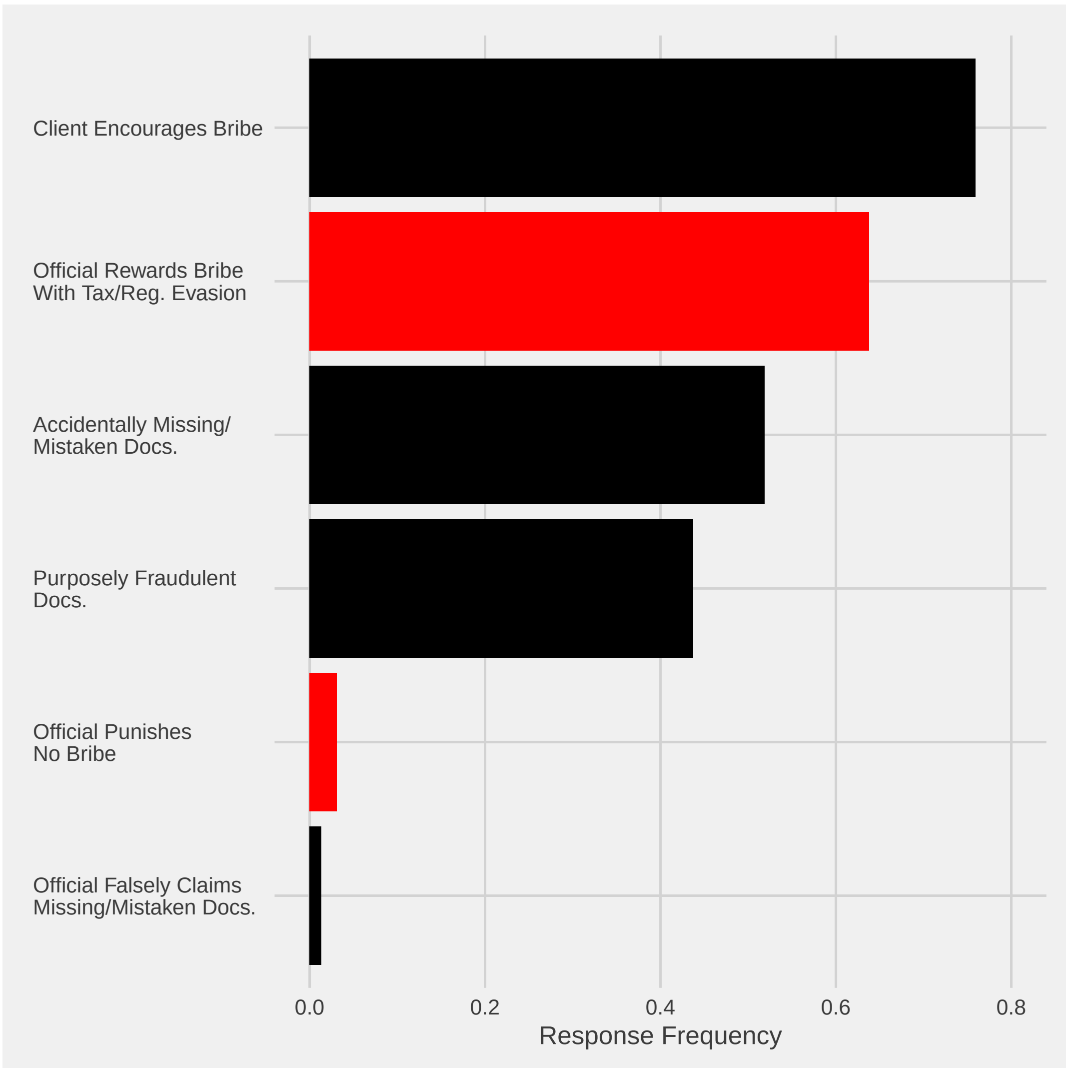
Note:
The figure is a scatter chart visualizing tax evasion as a function of bribery. The x-axis represents the bribe amount paid by the trader to the liquidateur official. The y-axis represents the tax-evasion amount, defined as the tax owed less the tax paid. Monetary values are in USD. The figure plots data jittered by up to $250 in each direction and is trimmed at the top/bottom 1%. The sample used is all cleaned data from the latest (2021) version of our survey.

Figure 1.4: Taxes Owed and Taxes Paid by Bribery



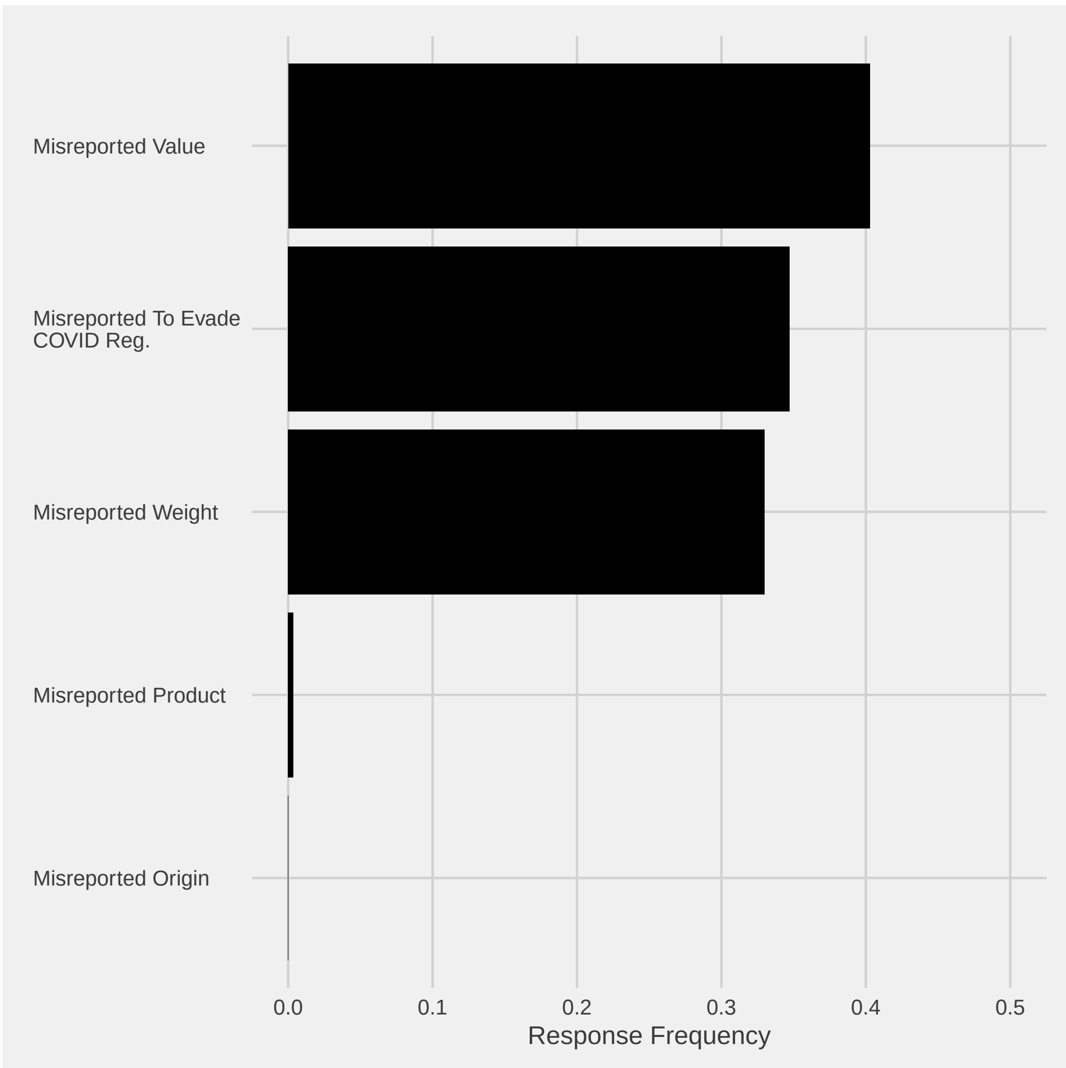
Note:
The figure is a kernel-density chart visualizing taxes owed and taxes paid broken down by bribery. The figure uses the kernel-density algorithm to plot nonparametric estimates of the probability density functions. In the left-hand panel, the two lines represent taxes owed and taxes paid for transactions in which the trader paid a bribe. In the right-hand panel, the two lines represent taxes owed and taxes paid for transactions in which the trader paid a bribe. Monetary values are in USD. The respective medians are represented by a dotted vertical line. Monetary values are in USD. The sample used is all cleaned data from the latest (2021) version of our survey.

Figure 1.5: Why Traders Pay Bribes



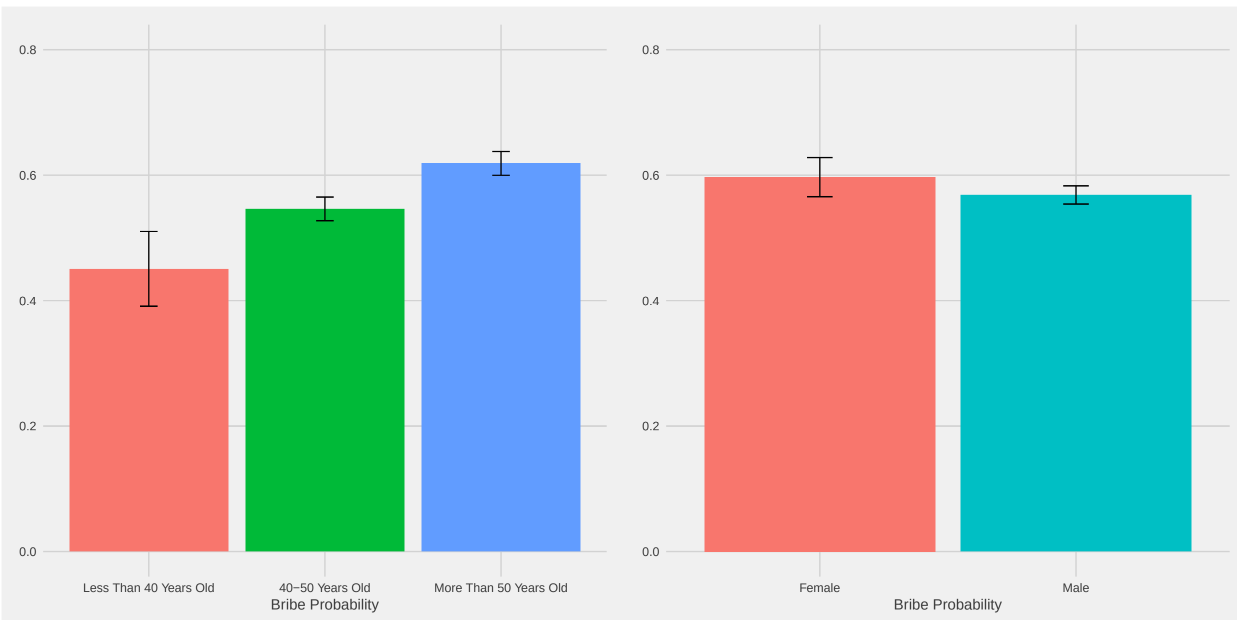
Note:
The figure is a bar chart visualizing how traders respond when asked why they paid a bribe. The the y-axis represents the responses, which were answers to a multiple-choice question to which the traders could select all the applied. The x-axis represents the frequency with which each answer was selected. The two red bars highlight the cases with clear collusion and clear coercion. The sample used is all cleaned data from the latest (2021) version of our survey.

Figure 1.6: How Traders Evade Taxes



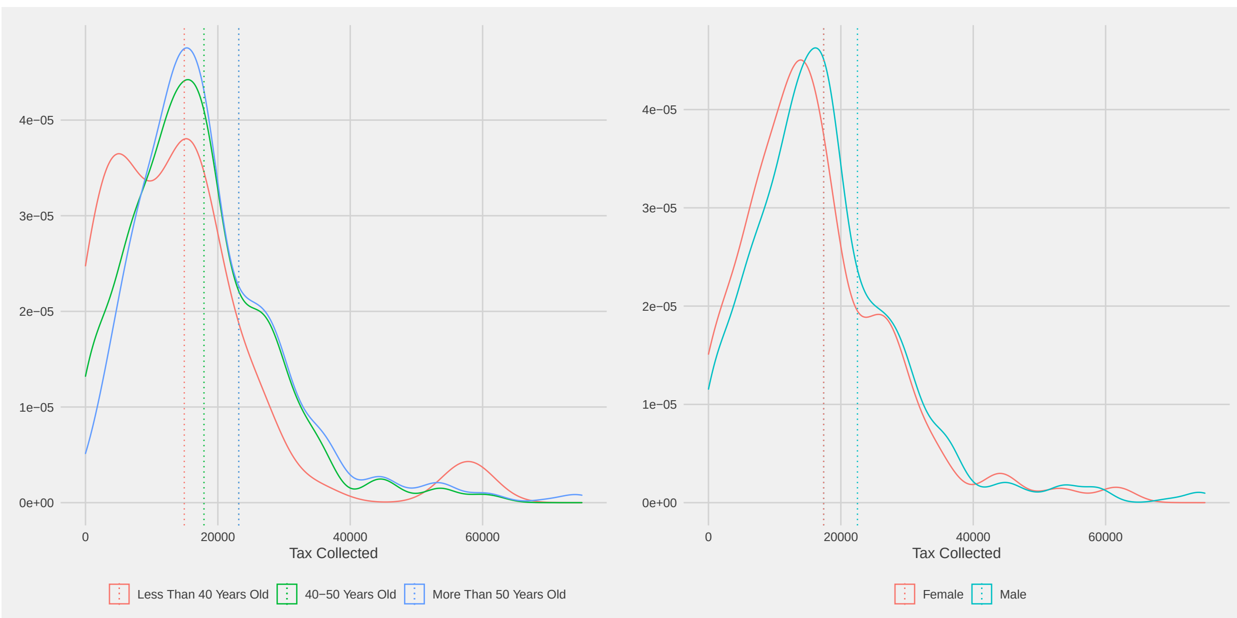
Note:
The figure is a bar chart visualizing how traders respond when asked how they evade taxes. The the y-axis represents the responses, which were answers to a multiple-choice question to which the traders could select all the applied. The x-axis represents the relative frequency with which each answer was selected. The sample used is all cleaned data from the latest (2021) version of our survey.

Figure 1.7: Bribe Probability by Officials' Demographics



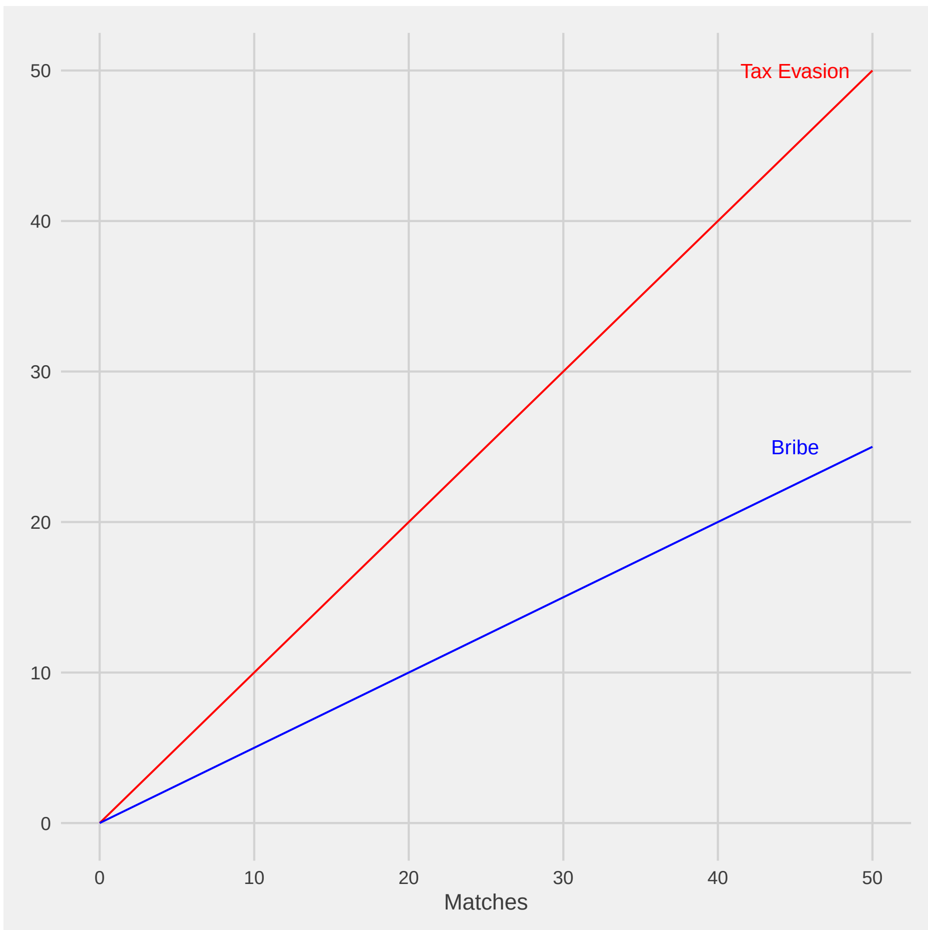
Note:
The figure is two bar charts visualizing bribe probability broken down by officials' demographics. In the left-hand panel, the three bars represent official age. In the right-hand panel, the two bars represent official gender. The y-axis represents bribe probability. Standard-error-bars cover the mean plus/minus two standard deviations. The sample used is all cleaned data from the latest (2021) version of our survey.

Figure 1.8: Tax Collected by Officials' Demographics



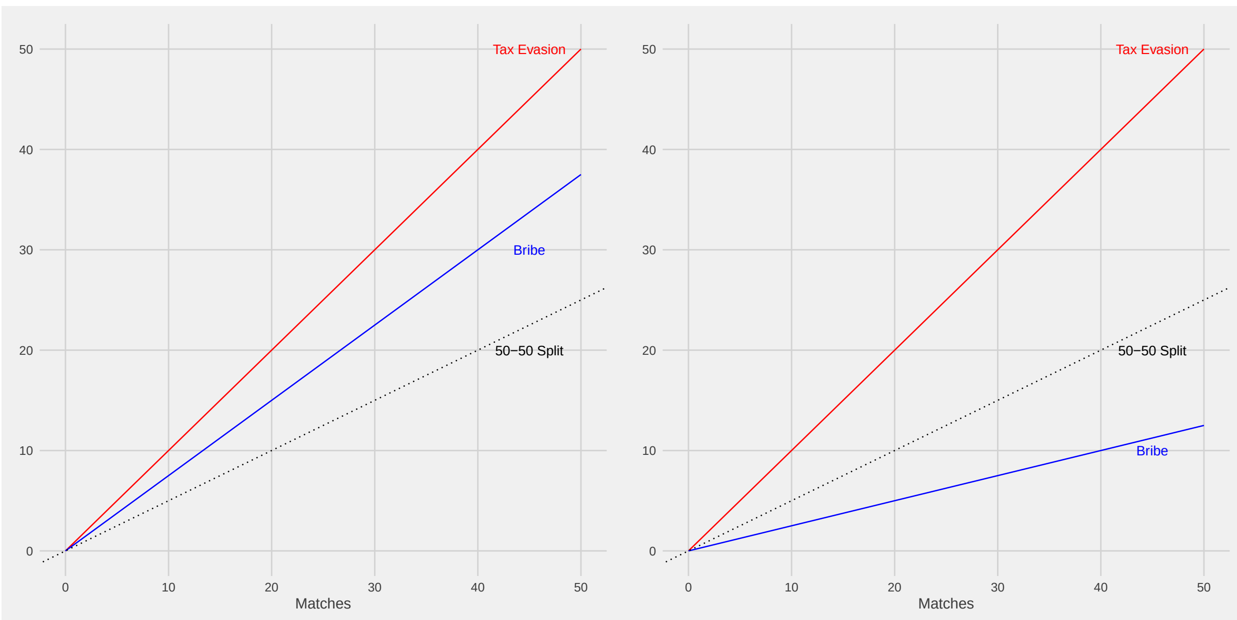
Note:
The figure is two kernel-density charts visualizing tax collected broken down by officials' demographics. The figure uses the kernel-density algorithm to plot nonparametric estimates of the probability density functions. In the left-hand panel, the three lines represent official age. In the right-hand panel, the two lines represent official gender. Monetary values are in USD. The respective medians are represented by a dotted vertical line. The sample used is all cleaned data from the latest (2021) version of our survey.

Figure 1.9: Game Theory Model Parameterization with 50/50 Tax-Evasion Split



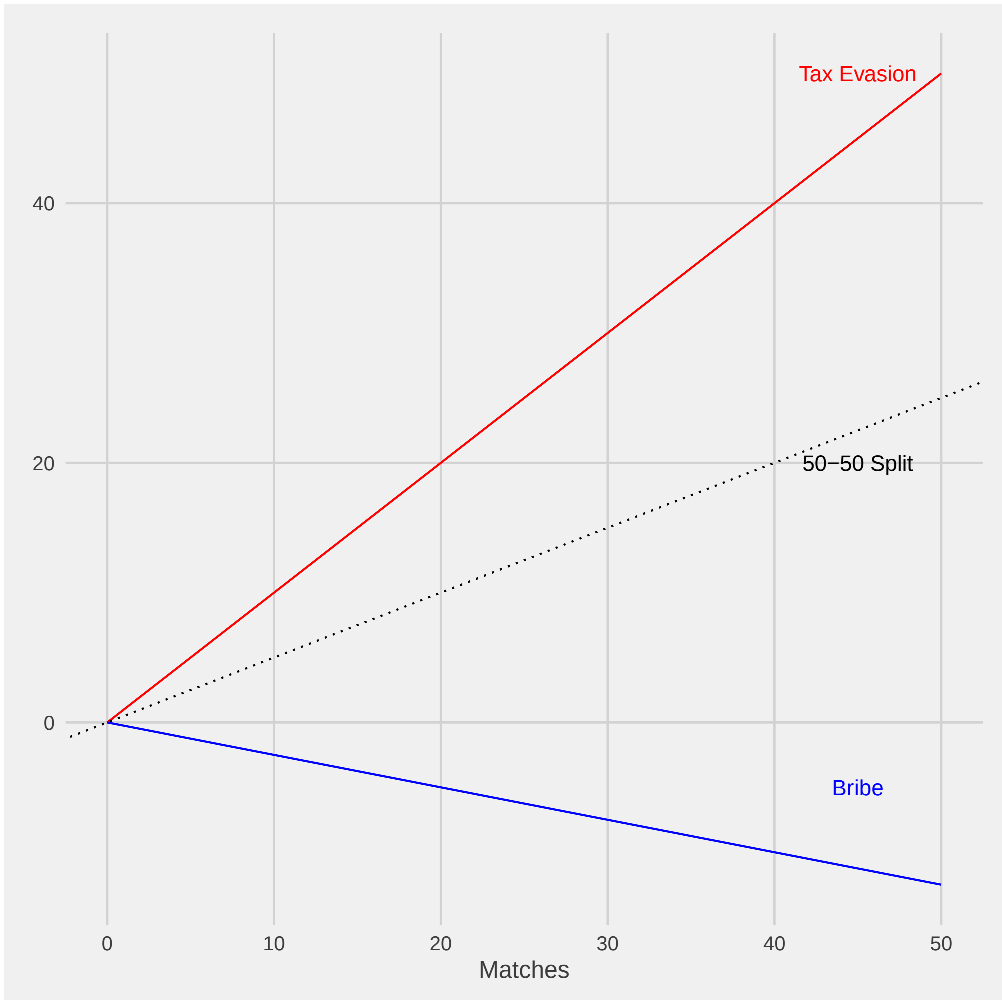
Note:
The figure is a simulation of our Nash Bargaining model. The x-axis represents the number of matches between the trader and the official from some baseline. The y-axis represents monetary amounts in excess of those at the baseline, where the red line is the tax-evasion amount and the blue line is the bribery amount. The paramterization is one with a 50/50 tax-evasion split. Mathematically: $\frac{\partial b^*}{\partial m} = \frac{1}{2}\frac{\partial(\tilde{t}-t^*)}{\partial m} = -\frac{1}{2}\frac{\partial t^*}{\partial m}$.

Figure 1.10: Game Theory Model Parameterization with Uneven Tax-Evasion Split



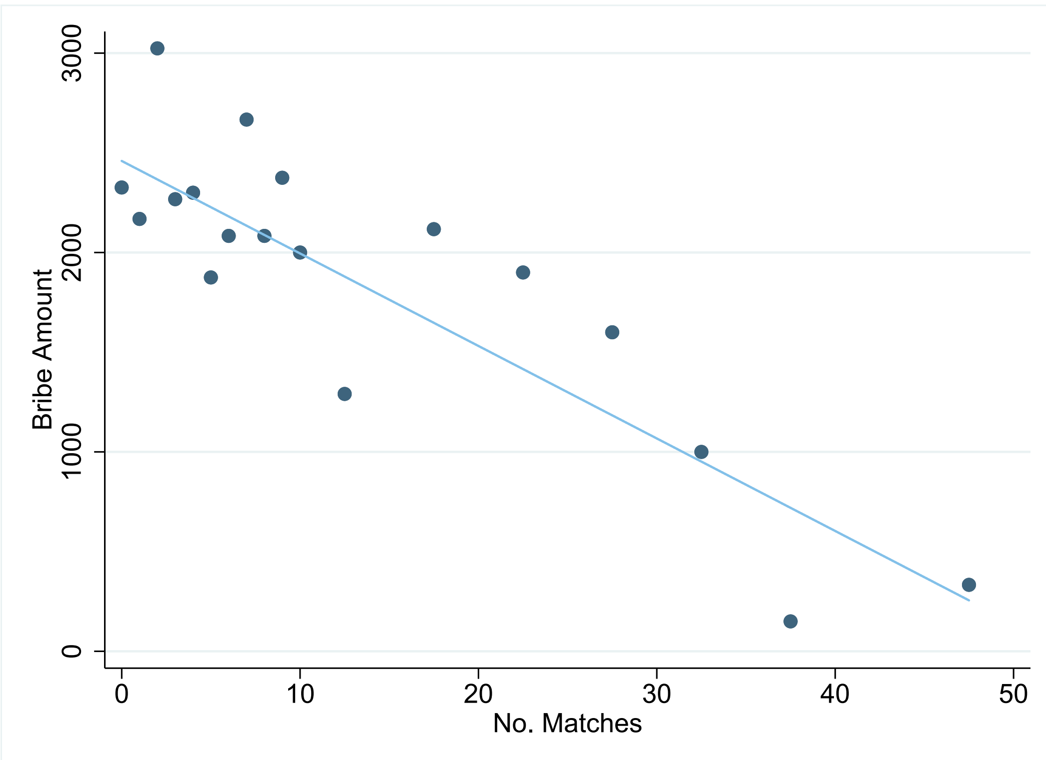
Note:
The figure is a simulation of our Nash Bargaining model. The x-axis represents the number of matches between the trader and the official from some baseline. The y-axis represents monetary amounts in excess of those at the baseline, where the red line is the tax-evasion amount and the blue line is the bribery amount. The black dotted line represents a 50/50 split of the tax-evasion amount between the trader and the official (whose share is the bribe). In the plot on the left, the blue line is above the black line. This parameterization is one for which matches increase the official's share of the pie. In the plot on the right, the blue line is below the black line. This parameterization is one for which matches decrease the official's share of the bribe. Mathematically: $\frac{\partial b^*}{\partial m} \neq \frac{1}{2}\frac{\partial(\tilde{t}-t^*)}{\partial m} = -\frac{1}{2}\frac{\partial t^*}{\partial m}$.

Figure 1.11: Game Theory Model Parameterization with Bribery/Tax-Evasion Divergence



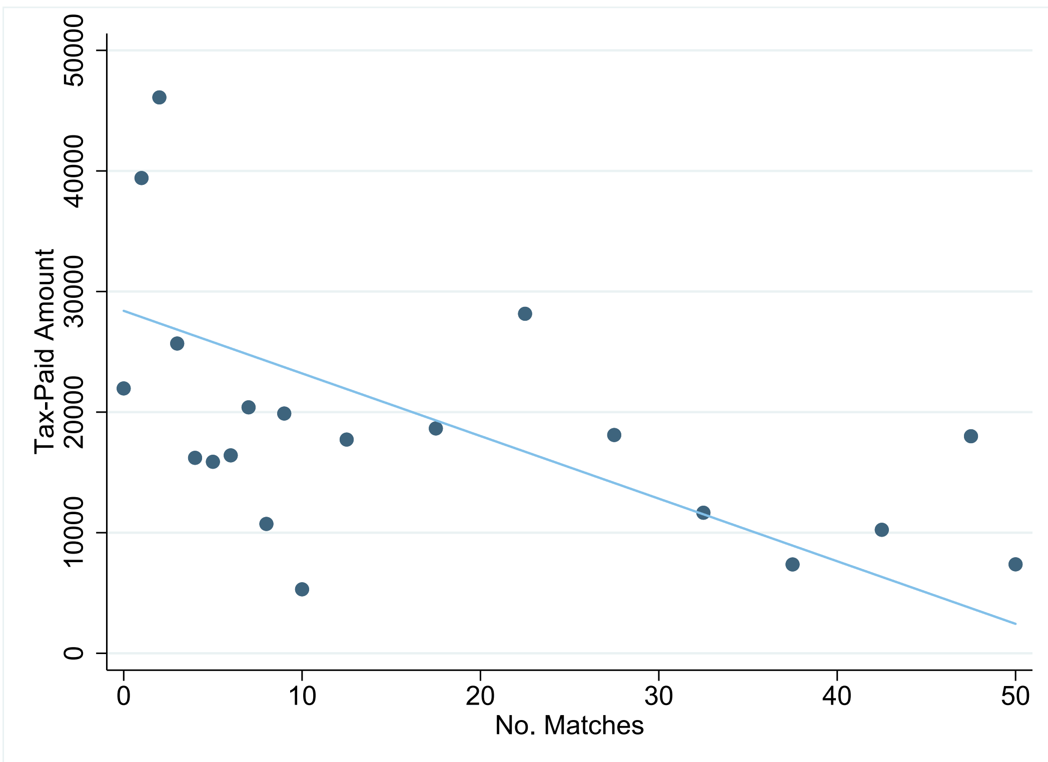
Note:
The figure is a simulation of our Nash Bargaining model. The x-axis represents the number of matches between the trader and the official from some baseline. The y-axis represents monetary amounts in excess of those at the baseline, where the red line is the tax-evasion amount and the blue line is the bribery amount. The black dotted line represents a 50/50 split of the tax-evasion amount between the trader and the official (whose share is the bribe). Since the blue line is below the x-axis, this parameterization is one for which matches decrease the official's share of the bribe so much that, even though the pie gets bigger, her amount of the pie gets smaller. Mathematically: $\frac{\partial b^*}{\partial m} < 0$, $\frac{\partial(\tilde{t}-t^*)}{\partial m} > 0$ & $\frac{\partial t^*}{\partial m} < 0$.

Figure 1.12: Effect of No. Matches on Bribery



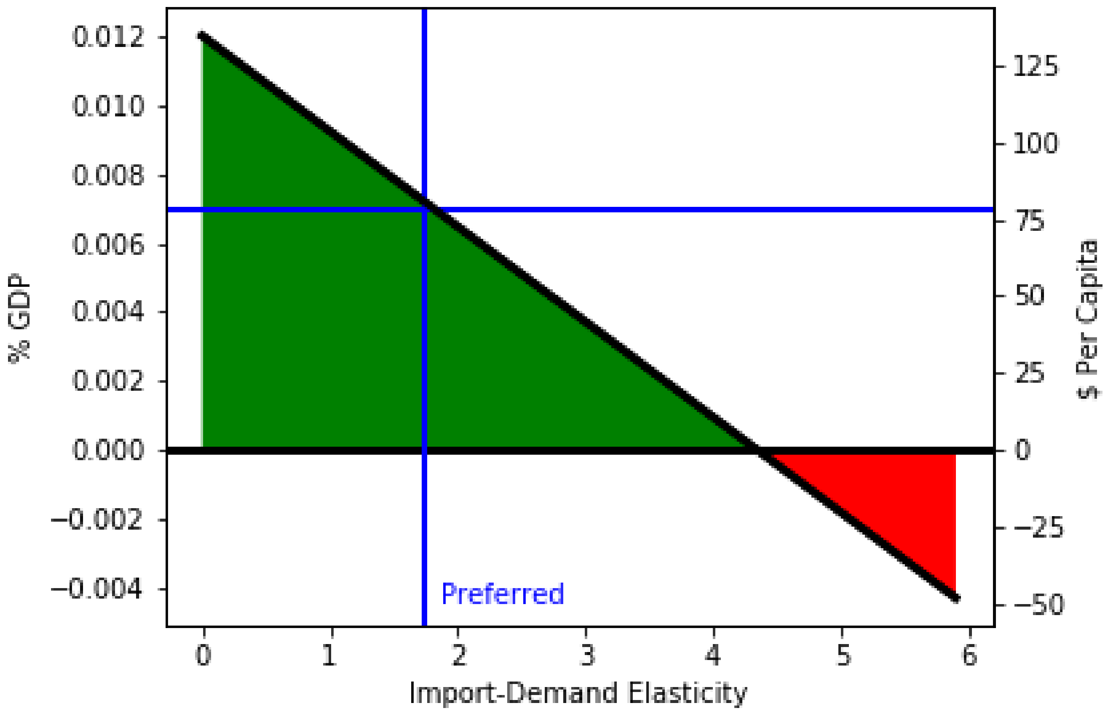
Note:
The figure is a binscatter chart visualizing the effect of the number of matches on bribery. The x-axis represents the number of matches between the trader and the official originally assigned as the trader. The y-axis represents the bribe amount. Monetary values are in USD. The figure sorts the data into equally sized bins and plots the average x and y values in each bin. The line is a line of best-fit calculated using Stata's *binscatter* package. The sample used is all cleaned data from the latest (2021) version of our survey.

Figure 1.13: Effect of No. Matches on Tax Evasion



Note:
The figure is a binscatter chart visualizing the effect of the number of matches on tax evasion. The x-axis represents the number of matches between the trader and the official originally assigned as the trader. The y-axis represents the tax-paid amount. Monetary values are in USD. The figure sorts the data into equally sized bins and plots the average x and y values in each bin. The line is a line of best-fit calculated using Stata's *binscatter* package. The sample used is all cleaned data from the latest (2021) version of our survey.

Figure 1.14: Anti-Corruption Dividend as Function of Import-Demand Elasticity



Note:
The figure visualizes the "anti-corruption dividend" as a function of the import-demand elasticity. The x-axis represents alternative assumed import-demand elasticities. The left-hand y-axis expresses the added tax revenue from eradicating bribery as a percent of GDP. The right-hand y-axis expresses the added tax revenue from eradicating bribery as U.S. dollars per citizen. The green area signifies positive returns to eradicating bribery. The red area signifies negative returns to eradicating bribery. The vertical blue line represents our preferred trade elasticity. The horizontal blue line represents the added tax revenue corresponding to our preferred trade elasticity.

1.9 Tables

Table 1.1: Shipment and Official Descriptives

	Mean	Median	Min	Max	SD	N
Shipment KGs	72,163	7,525	1	45,000,000	1,617,590	774
Shipment Value	36,265	32,062	53	354,992	34,715	774
Tax Owed	8,853	7,027	4	90,000	9,405	774
Other Trade Costs	2,526	324	21	73,962	7,435	774
Red Lane	0.88	1	0	1	0.32	774
Client Shipments/Y	60	55	5	100	16.96	737
Client Employees	56	55	0	105	20.41	738
Age Official	49	45	35	65	6.03	1423
Female Official	0.18	0	0	1	0.38	1423

Note:
The table shows descriptive statistics on shipment and official variables. Observations are at the shipment-level, except for the official characteristics, which are at the official-level. Red Lane is defined as an indicator variable for whether the goods passed through the high-enforcement "red lane." Monetary values are in USD. The sample used is all cleaned data from the latest (2021) version of our survey.

Table 1.2: Bribery and Tax Descriptives

	Mean	Median	Min	Max	SD	N
Bribe Yes/No	0.57	1	0	1	0.49	1423
Bribe Amount	719	540	10.8	30,394	1,242	815
Tax Paid	7,651	5,762	3.6	90,000	8984.57	774
Tax Owed	8,853	7,027	3.6	90,000	9404.76	774
Owed - Paid	1,202	949	-2,045	14,400	1,553	774
1 - (Paid / Owed)	0.13	0.14	-0.5	0.8	0.14	774
Paid / Owed	0.87	0.86	0.20	1.50	0.14	774
Tax Rate	0.30	0.19	0	2.76	0.26	774
No. Matches	14	13	0	50	11.94	1423

Note:
The table shows descriptive statistics on bribery and tax variables. Observations are at the official-level, except for the taxes, which are at the shipment-level. Monetary values are in USD. The sample used is all cleaned data from the latest (2021) version of our survey.

Table 1.3: Proportion of Bribery and Tax Evasion "Regimes"

		Bribery		
		No	Yes	Total
Tax Evasion	No	42.45	4.08	46.52
	Yes	0.21	53.27	53.48
	Total	42.66	57.34	100.00

Note:
The table shows our results on the proportion of bribery and tax evasion "regimes." Observations are at the official-level. Tax Evasion is defined as Yes if Tax Owed > Tax Paid and No otherwise. The sample used is all cleaned data from the latest (2021) version of our survey.

Table 1.4: Effect of Officials' Demographics on Bribery and Tax Evasion

	(1) Bribe Yes/No	(2) Tax Paid (Log)
Age Official	0.016***	0.011***
	(0.002)	(0.003)
Female Official	0.007	0.011
	(0.027)	(0.045)
Outcome Mean	.57	9.52
Observations	1,358	1,358
Month-by-Year FE	YES	YES
Controls	YES	YES

Note:
The table shows our regression results for the effect of officials' demographics on bribery and tax evasion. Observations are at the official-level. In Column 1, the dependent variable is an indicator for whether a bribe was exchanged. In Column 2, the dependent variable is the log of the amount of tax paid. The independent variables of interest are official age and gender, coded as ten-year bins and a female indicator respectively. Monetary values are in USD. The regressions also have month-by-year fixed effects, official controls (rank, tenure, and liquidateur/reviseur status), and shipment controls (weight, value, tax owed, other trade costs, red-lane assignment, client employees, and client trade volume). In the parentheses are robust standard errors clustered at the shipment level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) The sample used is all cleaned data from the latest (2021) version of our survey.

Table 1.5: Effect of No. Matches on Bribery and Tax Evasion

	(1) Esp. Friendly Yes/No	(2) Bribe (Log)	(3) Tax Paid (Log)
No. Matches (Liquidateur)	0.009***	-0.008**	-0.008**
	(0.003)	(0.004)	(0.004)
No. Matches (Reviseur)	0.001	0.010**	0.001
	(0.002)	(0.004)	(0.003)
Outcome Mean	.33	7.15	9.52
Observations	1,358	769	1,358
R^2	0.504	0.794	0.755
Trader FE	YES	YES	YES
Month-by-Year FE	YES	YES	YES
Controls	YES	YES	YES

Note:
The table shows our regression results for the effect of the number of matches on bribery and tax evasion. Observations are at the official-level. In Column 1, the dependent variable is the relationship quality between the official and the trader (which is a question in our questionnaire that asks yes/no whether the trader has an especially friendly relationship with the official). In Column 2, the dependent variable is the log of the bribe amount conditional on a bribe being exchanged. In Column 3, the dependent variable is log of the amount of the tax paid. The independent variables of interest are the number of matches between the trader and the official originally assigned to the trader, for both the liquidateur and the reviseur. Monetary values are in USD. The regressions also have trader fixed effects, month-by-year fixed effects, official controls (age, gender, rank, tenure, and liquidateur/reviseur status), and shipment controls (weight, value, tax owed, other trade costs, red-lane assignment, client employees, and client trade volume). In the parentheses are robust standard errors clustered at the shipment level. (*** p<0.01, ** p<0.05, * p<0.1.) The sample used is all cleaned data from the latest (2021) version of our survey.

Table 1.6: Heterogeneous Effects by Risk Environment

	Bribe (Log)
No. Matches (Liquidateur)	-0.037***
	(0.014)
No. Matches (Reviseur)	0.006
	(0.006)
No. Matches (Liquidateur)×Red Lane	0.029**
	(0.014)
No. Matches (Reviseur)×Red Lane	0.005
	(0.005)
Outcome Mean	7
Observations	769
R^2	0.794
Trader FE	YES
Month-by-Year FE	YES
Controls	YES

Note:
The table shows our regression results for the heterogeneous effects of the number of matches on bribery broken down by risk environment. The dependent variable is the log of the bribe amount conditional on a bribe being exchanged. The independent variables of interest are the number of matches between the trader and the official originally assigned to the trader, for both the liquidateur and the reviseur, and their interactions with a red-lane indicator. Monetary values are in USD. The regressions also have trader fixed effects, month-by-year fixed effects, official controls (age, gender, rank, tenure, and liquidateur/reviseur status), and shipment controls (weight, value, tax owed, other trade costs, red-lane assignment, client employees, and client trade volume). In the parentheses are robust standard errors clustered at the shipment level. (*** p<0.01, ** p<0.05, * p<0.1.) The sample used is all cleaned data from the latest (2021) version of our survey.

Table 1.7: Anti-Corruption Dividend as Function of Import-Demand Elasticity

Import-Demand Elasticity	-1.75 (Preferred)	0	-1	-2	-3	-4	-5
% GDP	0.7	1.2	0.9	0.6	0.4	0.1	-0.2
$ *Per Capita*	80	134	103	72	41	10	-21

Note:
The tables shows our results on the "anti-corruption dividend" as a function of the import-demand elasticity. The x-axis represents alternative assumed import-demand elasticities, with -1.75 being our preferred calibration because it is the midpoint of all the studies of which we are aware that estimate a trade elasticity for Tunisia specifically. The y-axis represents different ways of reporting the numbers, namely either as a percent of GDP or as U.S. dollars per citizen. The sample used is all cleaned data from the latest (2021) version of our survey.

Chapter 2

Affordable Housing for Forced Migrants

Joint with Edward Miguel, Sandra V. Rozo, Emma Smith & Sarah Stillman

2.1 Introduction

There are 82 million forced migrants in the world today, more than at any time in history and more than the population of France (UNHCR, 2021). Many of these individuals reside in the camps and informal settlements commonly represented in popular media, but in fact the vast majority reside instead in "host communities:" they live, work, and attend school in the same neighborhoods as the citizens of their host country, and on many dimensions their day-to-day experiences match those of low-income incumbents. As both a cause and an effect of this arrangement, many governments, IGOs, and NGOs support these refugees' shelter through either direct provision or rental assistance. Given the large (and growing) global refugee crisis and the scale of the fiscal resources in play, policymakers demand better information on the impacts of these affordable-housing programs (Ginn, 2020).

However, the evidence on the welfare implications of housing subsidies is inconclusive for both refugees in particular and high-poverty populations in general. On the one hand, studies have shown that shelter assistance has the potential to provide, not just a roof overhead, but downstream benefits on financial well-being, health, education, crime, and social cohesion (Aubry et al., 2015; Barnhardt et al., 2017; Bergman et al., 2020; Chetty et al., 2016; Clampet-Lundquist and Massey, 2008; Katz et al., 2001; Kling et al., 2007; Kumar, 2021; Ludwig et al., 2013). On the other hand, some studies have found null or even negative results, and it is often unclear whether the positive results are due to direct housing-quality effects or indirect neighborhood effects (Barnhardt et al., 2017; Chetty et al., 2016; Jacob et al., 2015; Jacob and Ludwig, 2012; Ludwig et al., 2013; Mills et al., 2006; Picarelli, 2019; van Dijk, 2019). In addition, if subsidizing demand fails to trigger a supply response, then it may cause housing-market inflation that burdens low-income tenants (Rozo and Sviatschi, 2021). Finally, though housing subsidies tend to be very expensive, there is relatively little work comparing their cost-effectiveness to other alternative interventions

(Bergman et al., 2020; Chetty et al., 2016).

To fill this knowledge gap, we ran a randomized controlled trial (RCT) testing a program that provides housing support to Jordan's Syrian refugees. Specifically, we partnered with the Norwegian Refugee Council (NRC) to run an impact evaluation of their Urban Shelter Program. The Urban Shelter Program offers subsidies to vulnerable refugees in the country's northern region. The program's subsidies are large (∼\$190/month or ∼\$3,100 total for ∼16 months) and go toward both rental assistance and structural improvements. Because of longstanding over-subscription, we were able to randomize receipt of the funding in order to causally identify its efficacy. Importantly, the subsidies are not vouchers; NRC designed the transfers so that households can only use them for their current housing, allowing us to be among the first to identify housing-quality effects absent confounding neighborhood effects.

We also partnered with the Syrian Refugee Life Study (S-RLS) for our data collection. S-RLS is a multi-year research initiative to better understand the refugee experience of both Syrians and other forcibly displaced populations. The S-RLS survey instrument covers all conceivable topics, from finances to education and health, and researchers are using it to build a representative longitudinal dataset. By leveraging an abridged version of this instrument (adapted into a phone survey during the COVID-19 pandemic), our data not only covers a wide variety of outcomes but will also be directly comparable to all future studies under the S-RLS umbrella.

There are three sets of results. First, we estimate the positive impacts of the program. Consistent with the structure of the transfers, recipients experience improvements in their housing quality and their housing-related finances. For example, the treatment group is more likely than the control group to have electricity, locks on their doors and windows, and clean drinking water. They also have better education and credit-market outcomes. For example, the treatment group is more likely than the control group to have children participating in learning activities and less likely to take on new debt. We find null results for most health and labor-market outcomes, but we argue that the COVID-19 pandemic limited effects on these fronts.

Second, we observe that, while the program did not cause movement *of* households, it did cause movement *between* households. Recipients welcome new individuals into their homes, likely so that family, friends, and neighbors can take advantage of the upgraded housing. Breaking down the movers demographically, teenage boys are the subpopulation more likely in the treatment group than the control group to join new households. We consider the possibility that this is because teenage boys are the minors least likely to attend school and therefore the most mobile.

Third, we estimate the negative impacts of the program and argue that these unintended consequences likely follow from the endogenous household formation. Recipients are less food secure, even though the treatment amounts to a large capital injection. We reason that this is consistent with having more mouths to feed, as each person's share of the pie gets smaller. (We also consider, and cannot eliminate, the possibility that our transfers make households ineligible for nutritional assistance.) Similarly, recipients are more likely to have COVID-19 symptoms and less likely to socially distance. We reason that this decrease in COVID-19 security might follow from more socialization taking place in upgraded housing, especially between the households that send and receive teenage boys. It might also follow from the epidemiological fact that with communicable diseases household risk increases with household size, even if behavior remains constant.

These results make a number of contributions to the literature. First and foremost, our unique context is one that policymakers are especially eager to understand. To our knowledge our study is the first to explicitly investigate refugee housing in host communities, as opposed to camps and camp-like informal settlements. In this way, our results speak to the long-running debate among humanitarian-aid organizations about the socially optimal way to house displaced persons (Ginn, 2020). Relatedly, COVID-19 directly attacked housing security by confining people to their homes and causing financial distress, and our study is among the first to investigate housing assistance during an acute public-health emergency. In this way, our results speak to how the international community can prepare for future crises (Egger et al., 2021).

At the same time, our study has generalizable implications for the way economists model and theorize about affordable-housing subsidies writ large. In particular, we separately estimate direct housing-quality effects and indirect neighborhood effects. Most earlier papers have studied housing vouchers, which allow households to move to new communities, as opposed to housing subsidies that support households in their current residence (Aubry et al., 2015; Barnhardt et al., 2017; Bergman et al., 2020; Chetty et al., 2016; Clampet-Lundquist and Massey, 2008; Jacob et al., 2015; Jacob and Ludwig, 2012; Katz et al., 2001; Kling et al., 2007; Kumar, 2021; Ludwig et al., 2013; Mills et al., 2006; Picarelli, 2019; van Dijk, 2019). We evaluate the latter type of program and therefore isolate the effect of housing *per se*, not the effect of housing bundled with access to desirable locations. Disentangling direct housing-quality effects and indirect neighborhood effects could be an important insight for policy design.

The paper proceeds as follows. Section 2 provides contextual information on both the housing situation for Jordan's refugee housing and its COVID-19 lockdown. Section 3 details our experiment, Section 4 details our data, and Section 5 details our econometric strategy linking the two. Section 6 presents positive impacts, Section 7 presents our results on endogenous household formation, and Section 8 presents negative impacts. Section 9 concludes.

2.2 Context

Jordan's Refugee Housing

Secure affordable housing is scarce for Jordan's Syrian refugees. At the time of a recent study, one in ten Syrian refugees was under immediate threat of eviction, and one in four who applied for rental assistance reported moving at least three times per year (NRC, 2015). Housing insecurity is such that refugees use untargeted transfers primarily to cover rental costs yet still these transfers are insufficient (Hagen-Zanker et al., 2018; Lehmann and

Masterson, 2014).[1] Furthermore, frequent movement due to an inability to pay rent also makes it more difficult for refugees to retain documentation and access local public services (NRC, 2015).

Jordan's COVID-19 Lockdown

To suppress COVID-19 cases, hospitalizations, and deaths, the Jordanian government enforced a country-wide lockdown from March 17 to May 3, 2020. The lockdown varied in its intensity over the two months, but it was very strict by global standards. On March 17, the government shut down all business (including groceries and pharmacies) and banned people from leaving their homes (even to exercise outdoors). On March 21, because plans to use the military to deliver food and medicine had failed, the government allowed people to buy staples from local stores, but it still banned travel for nonessential workers and isolated neighborhoods with confirmed cases. On April 30, the government decided to ease the lockdown by allowing most sectors to re-open gradually, but it kept schools, universities, gyms, churches, and mosques closed, and it issued fines for breaking mask and social-distancing mandates.

2.3 Experiment

The Urban Shelter Program

The Norwegian Refugee Council's Urban Shelter Program is the largest housing subsidy program for refugees in Jordan. It focuses on Syrian refugees in the country's northern governorates of Irbid and Mafraq, which are adjacent to the Syrian border and therefore contain a plurality of the country's displaced persons. Its funding comes from a variety of international actors, including certain OECD governments, and it works in partnership with the U.N. High Commissioner for Refugees (UNHCR).

The Urban Shelter Program is composed of several different shelter modalities, including ones focused on water and sanitation, eviction protection, and inclusion kits for individuals with disabilities. We study the Flexible Shelter ("FLEX") modality. Under FLEX, Jordanian landlords agree to provide Syrian refugee households with rent-free housing for a predetermined time period (between 9 and 18 months). In exchange, NRC provides rental assistance and structural improvements, such as lock installation, roof and wall construction, and mold removal.[2] (The program therefore benefits not just the Syrian tenants but their Jordanian landlords, as well as the overall Jordanian housing market by investing in the

[1] Syrians refugees' housing insecurity may have spillovers onto Jordanian citizens' housing insecurity. Researchers have found that an influx of displaced persons caused a rise in rental rates in areas adjacent to refugee camps (Rozo and Sviatschi, 2021)

[2] In the process, Jordanian landlords and Syrian refugees sign standard tenancy agreements in line with Jordanian law, and NRC works to support both landlords' and refugees' understandings of their rights and responsibilities.

housing stock.) Together the rental assistance and the structural improvements are on average valued at $3,100.

To test the Urban Shelter Program, our RCT included 2,126 households, with 1,099 (52%) randomly assigned to treatment and 1,027 (48%) randomly assigned to control. Our selection process followed a two-step procedure, as described in the following subsections and in Figures 2.1-2.2.

Community Randomization

In the first step, the 334 communities in Irbid and Mafraq were randomized into treatment or control. By community, we mean either (1) the Jordanian government's defined "localities" outside of major city centers or (2) neighborhoods inside of the major city centers of Irbid and Mafraq. One third of communities (114) were randomly assigned to treatment, while the remaining two thirds (220) were randomly assigned to control.

We randomized at the community level for three reasons: (1) to streamline implementation given NRC staff's required frequent site visits, (2) to prevent conflict between neighbors where, say, two similarly situated neighbors apply for the program but only one receives it, and (3) to allow for potential future analysis of treatment spillovers onto community housing rents/prices and social cohesion between refugees and host communities.[3]

Household Selection

Households applied to received Urban Shelter Program assistance, and NRC flagged them as eligible or not based on its standardized vulnerability assessment.[4]

To ensure that NRC could continue to uphold their humanitarian mandate, we developed a selection protocol that prioritizes the most vulnerable while still allowing for randomization. Every two weeks, we used the vulnerability assessment to sort applicants into the following groups: the least vulnerable 10% *did not* receive treatment and were excluded from the study; the most vulnerable 10% *did* receive treatment and were also excluded from the study; the middle 80% were included in the study and sorted based on their community's treatment/control status. All eligible applicants living in treatment communities were assigned treatment, and an equal number of randomly assigned households in the control communities were selected to be our control subjects. Thus our sample consisted of an equal number of program-eligible households in treatment and control, to maximize our statistical power.

NRC then implemented the program for those selected by entering into a negotiation with the landlords. These negotiations were one source of variation in how the money got spent;

[3]We originally intended to use our RCT to measure these community-level outcomes in addition to the household-level outcomes we discuss in this paper, but funding and statistical-power constraints led us to focus only on the latter.

[4]NRC previously used UNHCR's Vulnerability Assessment Framework, but they switched to a new in-house assessment shortly before we launched our RCT.

for example, some landlords agreed to waive rent for 12 months, whiles others agreed only to six months. More fundamentally, differences in baseline housing quality affected housing investments; for example, there was no need to invest in fixing roofs that were sound to begin with.

2.4 Data

Survey

We initially planned to deploy the S-RLS survey instrument six to 12 months after NRC completed implementation of the program, but the COVID-19 pandemic forced us to adapt. First, we decided to start data collection early. This was because NRC shut down implementation of its program early and also because we decided that it was scientifically interesting to surveil how housing subsidies were affecting refugees' resilience during an especially difficult time. Second, we decided to move from a full face-to-face survey to an abridged phone survey to protect both our enumerators and the participants.

The new survey included a subset of outcomes from seven of the pre-specified categories in our pre-analysis plan (see below): (1) dwelling characteristics and housing expenditures; (2) food security; (3) financial participation; (4) earnings and labor; (5) migration; (6) physical and mental health; and (7) child learning. We also included a module on specific COVID-19 outcomes. We reached 86% of randomized households over the phone, and we randomly selected whether to survey a male or a female respondent in each household.

It is worth noting that a large fraction (80%) of respondents had not received the full course of the housing subsidies prior to completing our survey; hence the effects we detect in this paper might be considered a lower bound on the true effects.

Retention, Compliance & Balance

Table 2.1 shows results on sample attrition and treatment compliance. Column 1 shows that sample attrition is uncorrelated with treatment assignment (the point estimate is close to zero and not statistically significant), so the treatment effect is unlikely to be correlated with sample attrition. Column 2 shows a strong first stage. Even though the compliance rate is 35%, the t-statistic is high. In addition to complications arising during the COVID-19 pandemic, the compliance rate is lower than it otherwise would have been because of program-implementation frictions, such as landlords being uncooperative, the baseline housing-quality being too high (e.g. very little required fixing), and the baseline housing quality being too low (e.g. not even $3,100 could make the home safe for living in).

Table 2.2 shows that we achieved balance between the treatment and control groups. (Since we were unable to run a baseline survey before treatment implementation, these data come from the NRC vulnerability assessment that we used to determine program eligibility.) Among the 18 covariates that we measure, spanning respondent, household, and housing characteristics, we find only one statistically significant difference, about as many as we would expect by pure chance. The joint significance of these covariates is also insignificant at traditional levels, per the reported Chi-squared test.

Descriptives

We can use both NRC's vulnerability assessment at baseline (Table 2.2) and the responses of our control group to our phone survey at endline (Table 2.3) as descriptive statistics. The average housing unit houses more than one family, only one in three households has access to piped water, and only one in ten households has a fully constructed roof. In addition, only one in four households has children that completed any learning activity (like reading) in the past 24 hours, which may be because schools were closed during surveying. Annualized labor incomes before the lockdown were $2,878, during the lockdown were $742 (much lower), and after the lockdown were $2,441 (close to the pre-lockdown level).

2.5 Econometrics

Specifications

For each household i in community c surveyed by enumerator e at time t, y_{icet} is some outcome of interest, A_{ic} is an indicator for treatment *assignment*, and T_{ic} is an indicator for treatment *implementation*. The vector X_c contains community-level controls, including the stratification variables used during randomization: whether community c is in Irbid or Mafraq and its population quartile. The vector Z_{ic} contains household-level controls: household i's vulnerability-assessment quartile, its number of children, its number of children plus adults, respondent gender, and respondent age. λ_e and λ_t are enumerator and month-by-year fixed effects, respectively.

With these variable definitions, our regression specification is as follows. Per our pre-analysis plan (see below), we focus on treatment-on-the-treated (TOT) specifications, where the first stage predicts treatment implementation using treatment assignment and the second stage regresses the outcomes of interest on the predicted values. (We prefer treatment-on-the-treated to intention-to-treat specifications because, from the perspective of an NGO like NRC, what matters is the effect of money spent, not the effect of money almost spent.) Standard errors are clustered at the community level, the level of randomization. β^y is our coefficient of interest.

$$T_{icet} = \alpha^T + \beta^T A_{ic} + X_c'\Lambda^T + Z_{ic}'\Gamma^T + \lambda_e^T + \lambda_t^T + \epsilon_{icet}^T \tag{2.1}$$

$$y_{icet} = \alpha^y + \beta^y \widehat{T}_{icet} + X_c'\Lambda^y + Z_{ic}'\Gamma^y + \lambda_e^y + \lambda_t^y + \epsilon_{icet}^y \tag{2.2}$$

Pre-Analysis Plan

To increase transparency and limit the temptations of specification search (so called "p-hacking"), before deploying our survey we filed a pre-analysis plan (PAP).[5] Our econometric specification respects the PAP with only one change. We added enumerator fixed effects because we observed more between-enumerator heterogeneity than expected. Each of our regressions also reports q-values, a form of multiple-inference adjustment of p-values.

We also added three outcomes. First, we added *per capita* food assistance to shed light on our food insecurity results. Second, we broke down our endogenous-household-formation effects by age and gender. Third, we added a measure of whether a dwelling can be locked or secured, since this is an important dimension of housing safety that we failed to include at the outset.

Finally, we de-emphasize the distinction between "primary" and "secondary" outcomes; since our phone survey is an abridged version of the full S-RLS survey, we lack data on two of the five primary outcomes and many of the secondary outcomes.

2.6 Positive Impacts

Housing Quality & Housing-Related Finances

Table 2.4 (Figure 2.3) shows our effects on housing quality and housing-related finances. The first takeaway is that the intervention is successful at improving the quality of refugees' housing. For example, households who receive the subsidy are 1 percentage point (1%) more likely to have electricity, 14 percentage points (18%) more likely to have locks on their doors and windows, and 20 percentage points (85%) more likely to have filtered drinking water. The households' living spaces also have one half of an additional room and improved building materials (e.g. dwelling walls made of concrete or brick versus tin or temporary materials).

The second takeaway is that the treatment is successful at improving refugees' finances by lessening their rent burdens. Specifically, households who receive the subsidy pay an additional \$19.60 in rent (\$3.80 *per capita* times an average household size of 5.16), *inclusive of their rental assistance.* Given that there are no statistically significant differences between the treatment and control groups in terms of their rent owed (see the Appendix), we take this increased housing expenditure as evidence of a decreased rental deficit (i.e. a decreased flow of new rental debt).[6]

Education, Health & General Finances

Table 2.5 (Figure 2.4) shows the program's effects on children's education. As shown in Table 2.3 above, the state of children's education among Jordan's Syrian refugees during the COVID-19 pandemic was bleak: fewer than one in four respondents reported that any

[5] The Pre-Analysis Plan, which was registered on July 13, 2020, can be found here: AEARCTR-0006141.

[6] We recognize that this metric is imperfect, but our abridged-phone-survey dataset limits our options.

children in their household had taken part in learning activities in the past 24 hours. In this context, the treatment dramatically improves refugee welfare. The housing subsidies raise the share of children taking part in learning activities by 16 percentage points (64%). This could be because the improved housing quality or housing-related finances reduce stress on parents and therefore create an environment more conducive to children's learning. Alternatively, it could be because it frees up time for parents to help with their children's schooling, although Table 2.5 shows that the treatment lowers hours spent on childcare.

Table 2.5 (Figure 2.4) also shows the program's effects on physical and mental health, namely depression (as measured using psychologists' validated CES-D scale) and a subjective self-report of overall health. Interestingly, we do not find any statistically significant treatment-control differences. These null results could follow from having collected data very soon after program implementation, since health likely changes with a longer lag than education. Null results are still clarifying in that improved health does not appear to be the mechanism by which the housing subsidies improve education.

Table 2.6 (Figure 2.5) shows the program's beneficial effects on credit- and labor-market outcomes, broken down by the periods before, during, and after Jordan's COVID-19 lockdown. Regarding loans, in all three periods the treatment households are between 21 and 33 percentage points (34-77%) less likely than control households to take on new debt. This result is consistent with the result we discuss above, whereby the housing subsidies decrease rent burdens. There are no statistically significant treatment-control differences in hours worked in any period, but there is a statistically significant and *negative* treatment-control difference on labor income during the lockdown. Given the very low hours worked by all groups during the lockdown, it may be that treatment households have the financial security to avoid risky income-generating activities.[7]

Overall, it is noteworthy that the treatment had few labor-market impacts. On the one hand, the housing subsidy amounts to a large increase in household wealth, which implies that there might be meaningful wealth effects. On the other hand, the intervention is unlikely to have substitution effects that change the payoff differential between labor and leisure.[8] In addition, unlike other similar housing interventions, ours does not involve households moving to new labor markets with new employment opportunities. Finally, the COVID-19 pandemic amounted to a large negative labor-demand shock.

That being said, even though there are few aggregate effects, there might be disaggregated effects; it is worth exploring whether the treatment had any credit- or labor-market impacts that vary by subpopulation. We therefore perform heterogeneity analysis on three dimensions directly relevant to program targeting: household-head gender (male or female), community density (urban or rural), and household vulnerability as measured by NRC's assessment.[9] We do not find many statistically significant results by community density (Appendix Table B.1)

[7]However, we are unsure what these alternative income-generating activities might be. Note that there is no treatment-control difference in asset liquidation.

[8]It is unlikely but not impossible. Upgraded housing could raise the utility of leisure relative to labor if it is more fun to "hang out" at home.

[9]74% of households are male-headed, 85% are urban (defined as localities with >5,000 inhabitants), and 49% are "vulnerable."

or household vulnerability (Appendix Table B.2), but we do by head-of-household gender (Table 2.7). We can see that the treatment tends to lower female-headed households' loan uptake and labor income. Overall, these results are encouraging from a gender-equality perspective because they suggest that the housing subsidies might lower female-headed households' debt burdens and offer them the freedom to choose to exit the labor market.[10]

2.7 Endogenous Household Formation

Per Table 2.8 (Figure 2.6), we see that the program does not cause households to move to different units or different neighborhoods. This is as expected, since the intervention does provide housing vouchers: the housing subsidies are only for the refugees' preexisting housing, not for moving to better housing in a better neighborhood.

However, we do see that the program causes a different type of movement. The number of children in treatment households is higher at a rate of approximately one new child for every third treatment household. Hypothetically this could be because of an increase in babies through an increase in fertility, but as indicated in the second panel it is instead because of teenage boys moving into the household. While the data do not directly speak to the mechanisms driving this inter-household movement, our interpretation is that treatment households take advantage of their upgraded housing to care for the children of their friends, families, or neighbors; and that the children are generally teenage boys because they are the least likely to attend school and therefore the most mobile (see Figure 2.7, Day and Cervero (2010), and GAGE (2020)).

One concern might be whether these movers into treatment households are coming from control households. If this were the case, then it could lead to a Stable Unit Treatment Value Assumption (SUTVA) violation that threatens our causal identification. We do not believe this to be the case. First, the average treatment density (defined as the number of treatment households divided by the total number of Syrian refugee households in each treatment community) is approximately 0.5%; and recall that one third of communities are in the treatment group and two thirds are in the control group. This means that the probability the movers are coming from a control household is only about 1 in 300. Second, a comparison of baseline and endline household size among the control group finds a small *increase*, which cuts against the idea that individuals are moving from control to treatment.[11]

[10]Jordan has one of the world's lowest female labor force participation (FLFP) rates.

[11]We attribute this to differences in how the baseline vulnerability assessment and the endline phone survey measure household size. While this confounds the measurement of out-migration from control households, any out-migration is bounded from above by the measurement difference between the two instruments.

2.8 Negative Impacts

Food Insecurity

In addition to the positive impacts described above, we also detect a number of negative impacts. Table 2.9 (Figure 2.8) shows results on food insecurity. As measured by a self-reported binary indicator, food consumption is lower for the treatment group than for the control group, whether we are focused on the respondent themself, other adults in the household, or children in the household.[12] Why are households who receive a capital injection of $3,100 more food insecure than statistically identical households who do not receive the money? Our leading hypothesis is that this may be driven by the teenage boys who move into the house. An increase in food insecurity is with consistent having another mouth to feed, as each person's share of the pie gets smaller.

An alternative hypothesis is that changes in food assistance may drive the increase in food insecurity, since not only is *per capita* food assistance lower, the *total* amount of food assistance is lower. This could imply that receipt of our assistance makes treatment households ineligible for other assistance, especially if the newest household members are not yet reflected in IGO/NGO records. Conversations with various humanitarian-aid organizations present mixed evidence on this hypothesis. On the one hand, an internal document from the Basic Needs Working Group (comprised of 30+ organizations) indicates that during the COVID-19 pandemic practitioners expanded coverage but attempted to exclude "any cases currently receiving basic needs assistance from any partner under an existing programme" (i.e. our subjects) (BNWG, 2020). On the other hand, despite these attempts, we find that 53% of our treatment group and 56% of our control group were included in this program, a statistically insignificant difference. Since this is only one program, and since other qualitative evidence suggests that community groups also attempted to account for our transfers in determining the most deserving households, we cannot reject this hypothesis.

Further complicating the findings is that, though overall food insecurity is higher in treatment households, control households report food consumption decreasing over the month preceding the survey, while treatment households' food consumption stayed about the same. Because this only reflects changes over a single month, it is difficult to know whether the difference in trends is idiosyncratic or points toward eventual convergence.

COVID-19

Table 2.10 (Figure 2.9) shows results on COVID-19-related health outcomes. The index we use to measure COVID-19 physical symptoms is higher in the treatment group. The treatment group is also less likely to socially distance and more likely to attend mosque.

[12] One might consider running two additional specifications: one that interacts treatment with household composition and one that uses treatment as an instrument to predict household composition. Unfortunately, these specifications are invalid: in the former case, household composition is a "bad control," and in the latter case, it violates the exclusion restriction.

Why does risky behavior go up? This might follow from more socialization taking place in upgraded housing, especially between the households that sent and received teenage boys. It might also follow from the epidemiological fact that with communicable diseases household risk increases with household size, even if behavior remains constant.

2.9 Conclusion

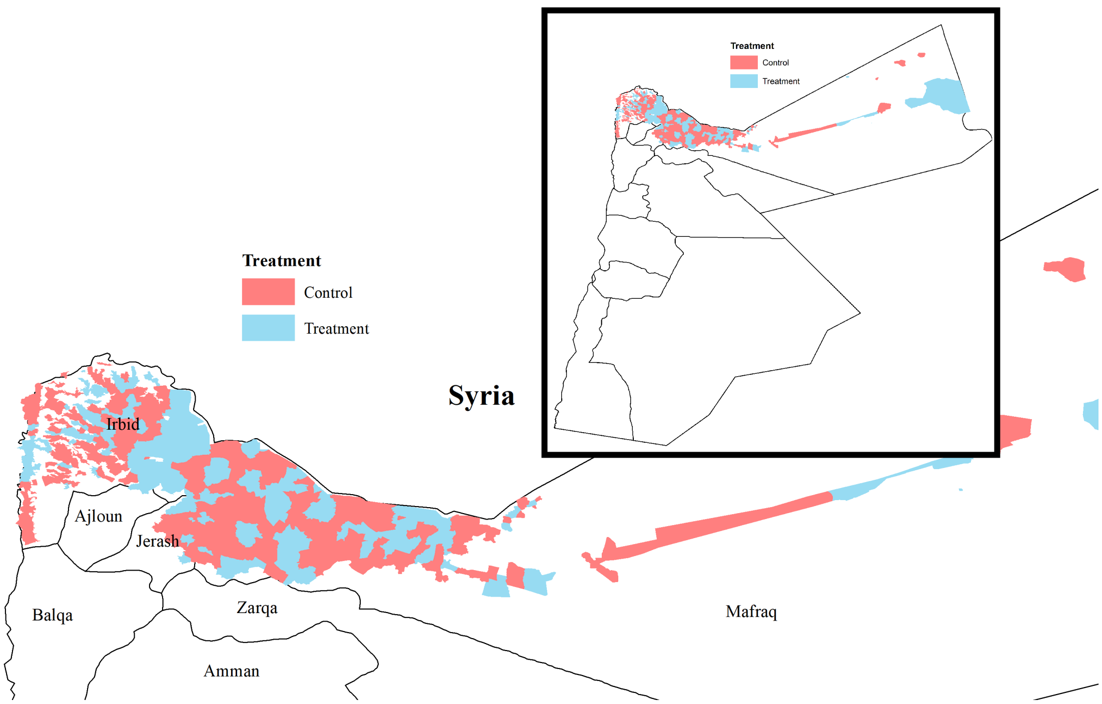
We ran a randomized controlled trial (RCT) testing a program that provides housing support to Jordan's Syrian refugees. The subsidies are large (approximately $190 per month) and are not vouchers; households can only use them for their current housing, allowing us to be among the first to identify housing-quality effects absent confounding neighborhood effects. There are three sets of results. First, we estimate the positive impacts of the program. Recipients experience improvements in their living situations, as well as better education and credit-market outcomes. Second, we observe that, while the program does not cause movement *of* households, it does cause movement *between* households. Recipients welcome new individuals into their homes, likely so that family, friends, and neighbors can take advantage of the upgraded housing. Third, we estimate the negative impacts of the program and argue that these unintended consequences likely follow from the endogenous household formation. For example, recipients are less food secure, consistent with having more mouths to feed.

Moving forward, there are many opportunities for future work. Because of the COVID-19 pandemic, we collected data using an abridged survey (to accommodate having to speak with respondents over the phone) and an accelerated timeline (to accommodate our implementing partner's revised program operations). Hence, it would be valuable to return to the field with a full face-to-face survey and test robustness, persistence, and mechanisms. For example, it would be interesting to explore whether, despite our initial null results, new data reveals health or labor-market impacts. It would also be interesting to explore what might explain our surprising food-insecurity and COVID-19 impacts.

To inform policymakers, this future work should perform a full cost-benefit or cost-effectiveness analysis of programs like ours that provide affordable housing to forced migrants. On the one hand, our program has many benefits on housing, finances, and education. On the other hand, it also appears to have (unintended) costs on food security and public health; and, at an implementation cost of $3,100 per recipient, it is expensive compared to many other development and humanitarian interventions. Overall, our RCT reveals both the upside potential and the downside risk of large-scale affordable-housing interventions. By refining their design, implementation, and cost-effectiveness, policymakers could improve these programs and the lives of the vulnerable communities they serve.

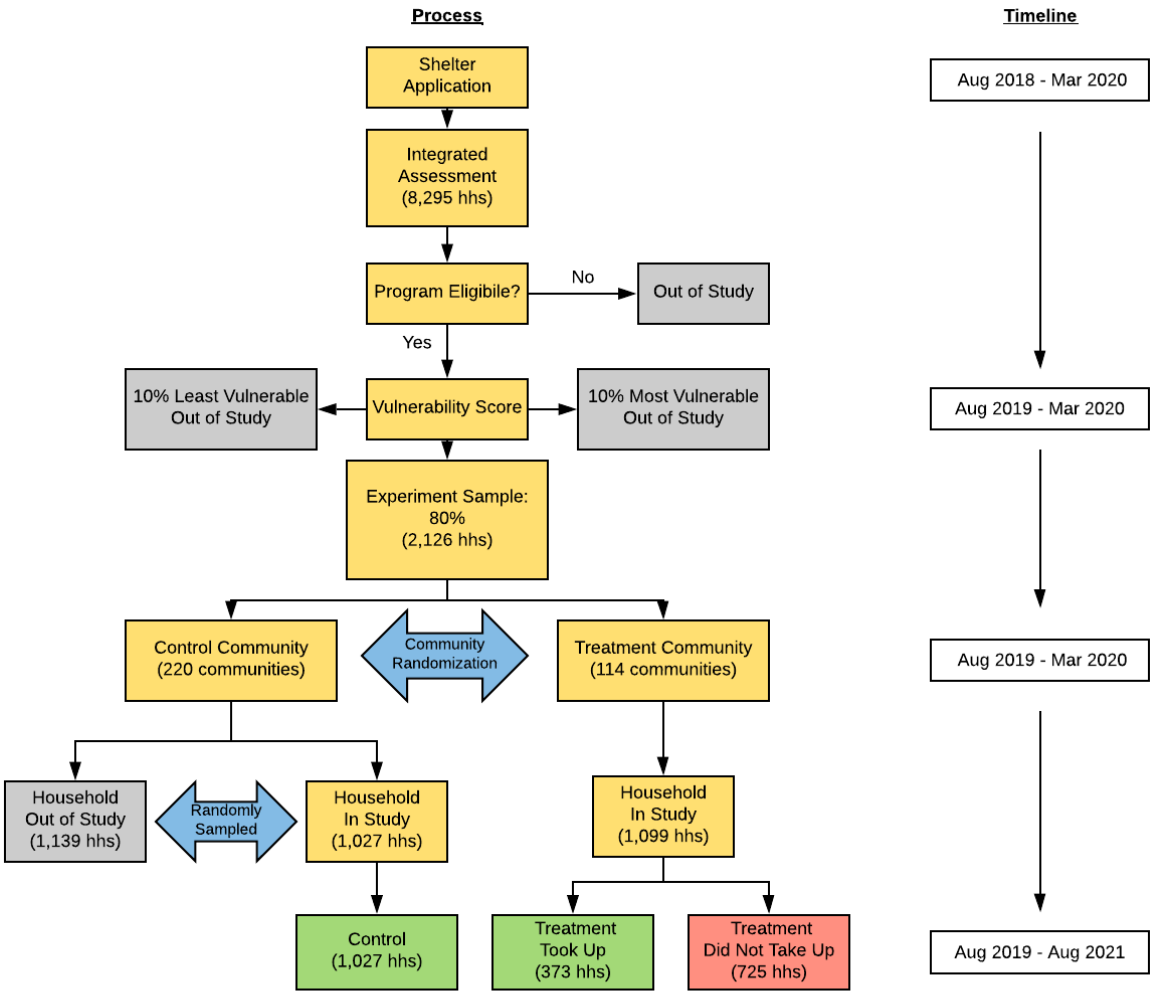
2.10 Figures

Figure 2.1: Community Randomization



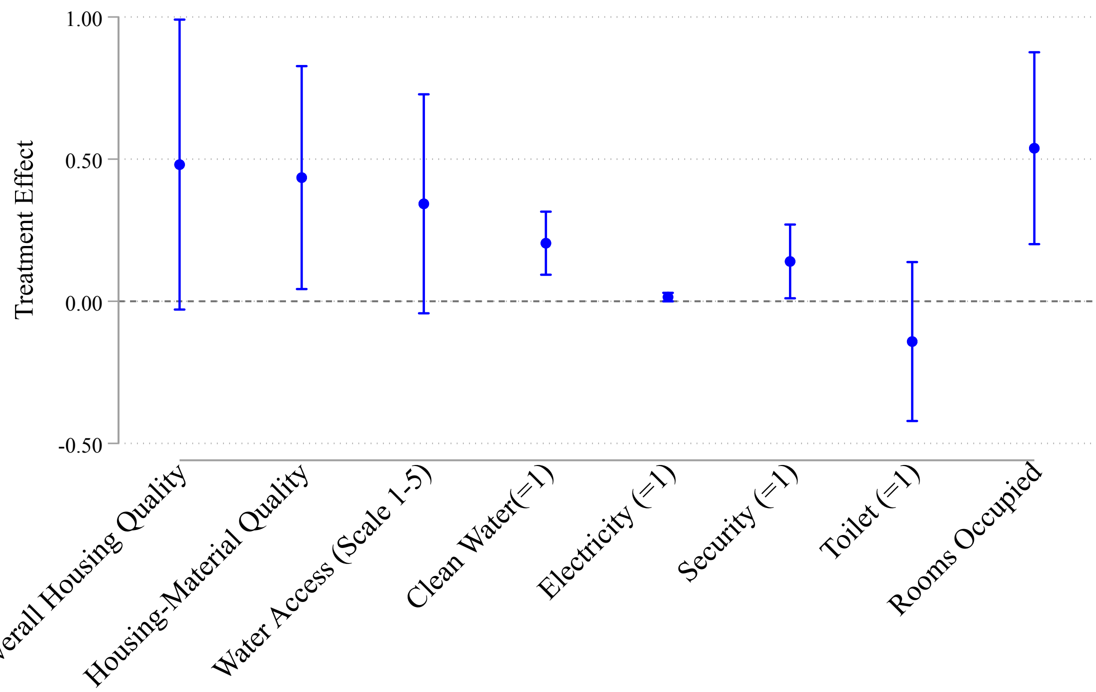
Note:
The figure maps our treatment and control communities. The communities are "localities" (as defined by the Government of Jordan) in the northern governorates of Irbid and Mafraq, which are adjacent to the Syrian border. Control communities are in red. Treatment communities are in blue. White areas within Irbid and Mafraq represent uninhabited territory.

Figure 2.2: Household Selection



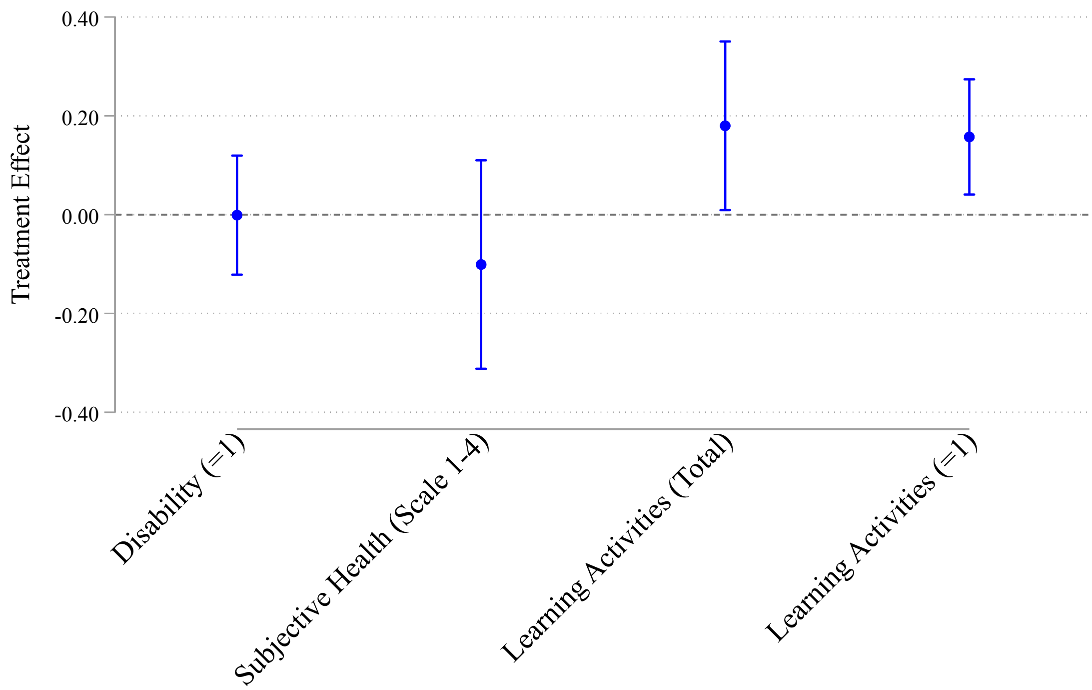
Note:
The figure visualizes our household-selection methodology. The right-hand column is the timeline, and the left-hand column is the process itself. Yellow boxes represent steps. Blue arrows represent randomizations. Green boxes represent households in the study. Gray boxes represent households out of the study due to assignment, while red boxes represent households out of the study due to noncompliance. "hhs" stands for households.

Figure 2.3: Effects on Housing Quality and Housing-Related Finance



Note:
The figure visualizes the regression in Table 2.4. The x-axis represents the dependent variables. On the y-axis, the dots represent the point estimates and the bars represent the 95% confidence intervals.

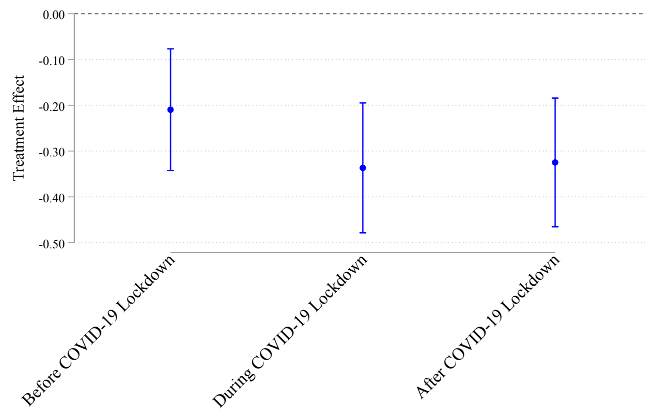
Figure 2.4: Effects on Education and Health



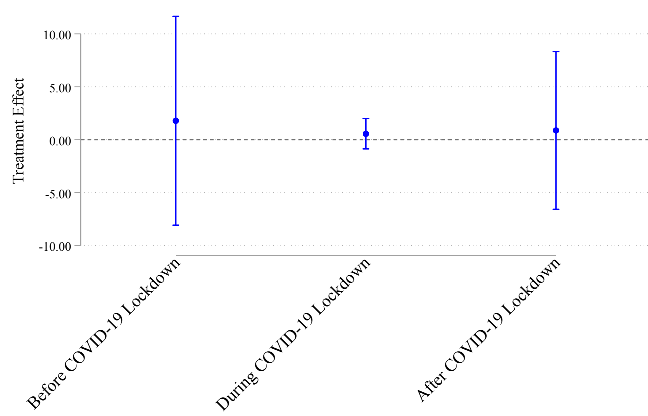
Note:
The figure visualizes the regression in Table 2.5. The x-axis represents the dependent variables. On the y-axis, the dots represent the point estimates and the bars represent the 95% confidence intervals.

Figure 2.5: Effects on General Finances

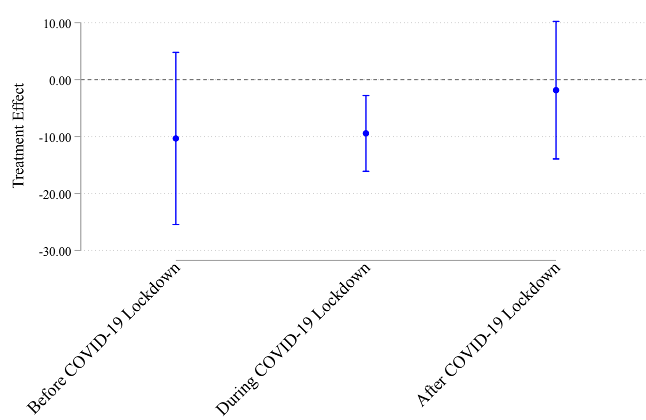
(a) Loan Uptake (=1)



(b) Hours Worked (Per Week)

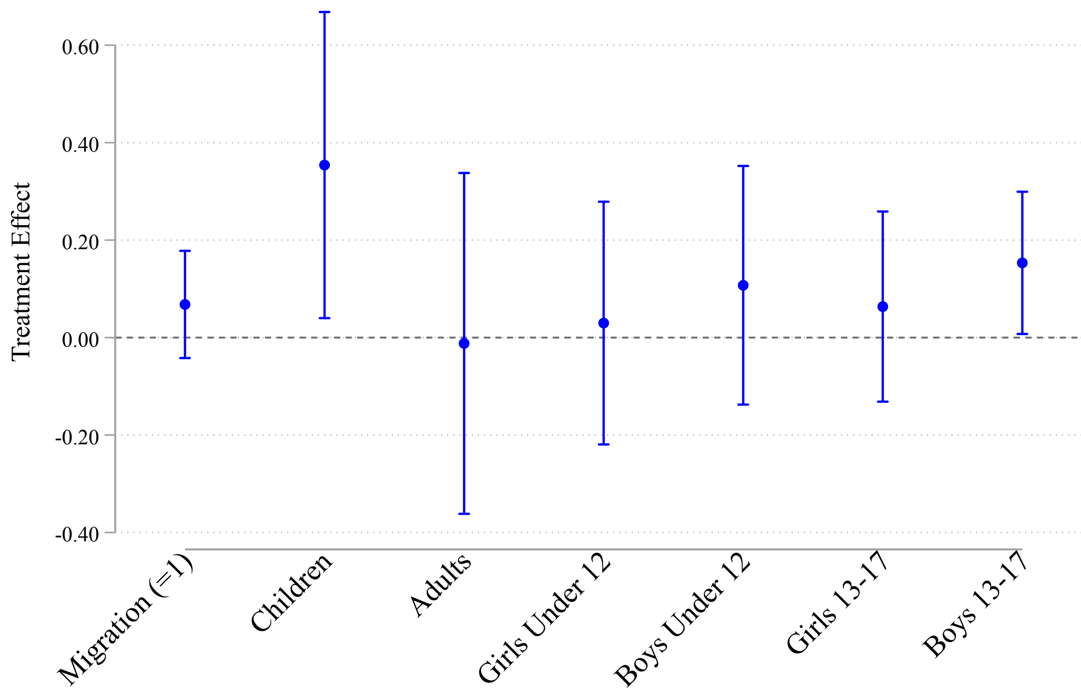


(c) Labor Income (USD PPP Per Week)



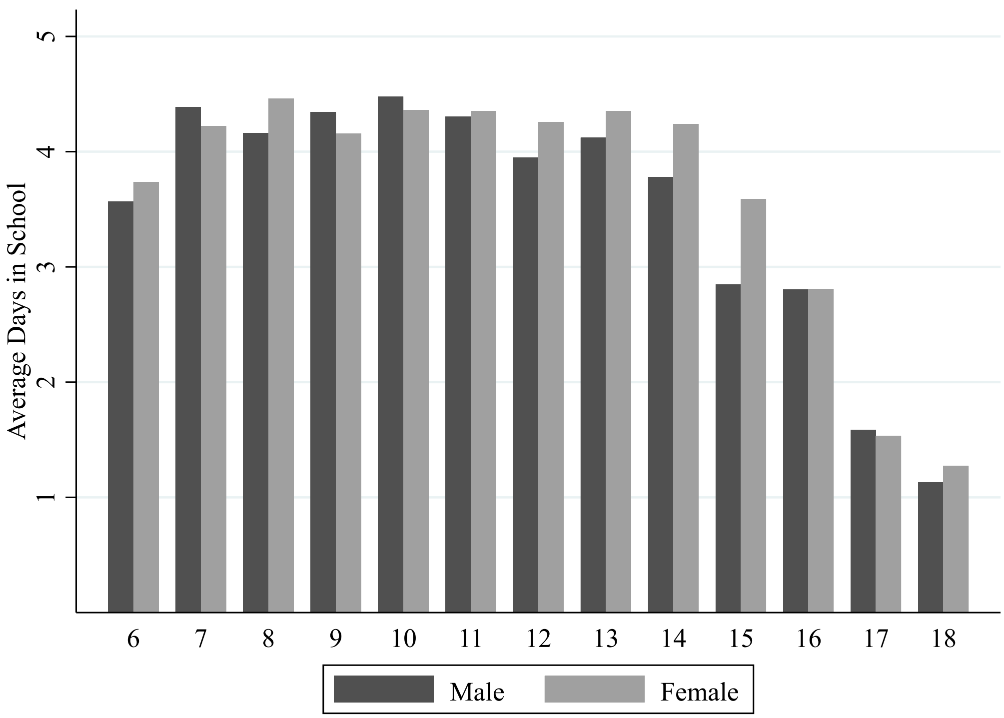
Note:
The figure visualizes the regression in Table 2.7. The x-axis represents the dependent variables. On the y-axis, the dots represent the point estimates and the bars represent the 95% confidence intervals.

Figure 2.6: Effects on Movement of and between Households



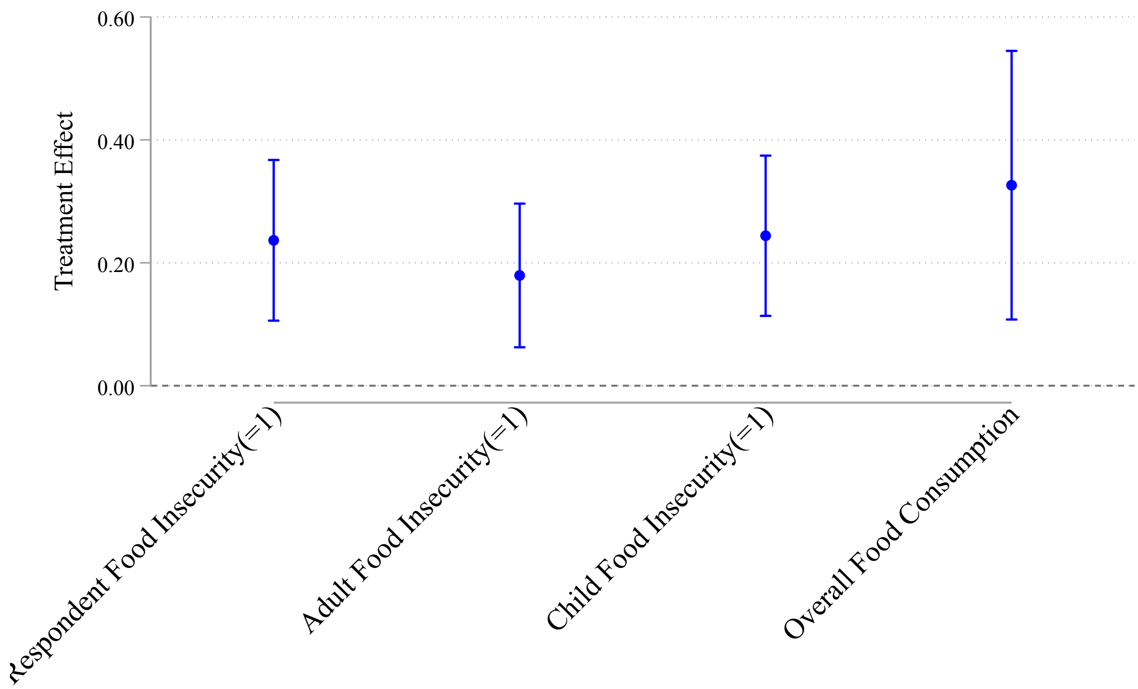
Note:
The figure visualizes the regression in Table 2.8. The x-axis represents the dependent variables. On the y-axis, the dots represent the point estimates and the bars represent the 95% confidence intervals.

Figure 2.7: School Attendance by Age and Gender



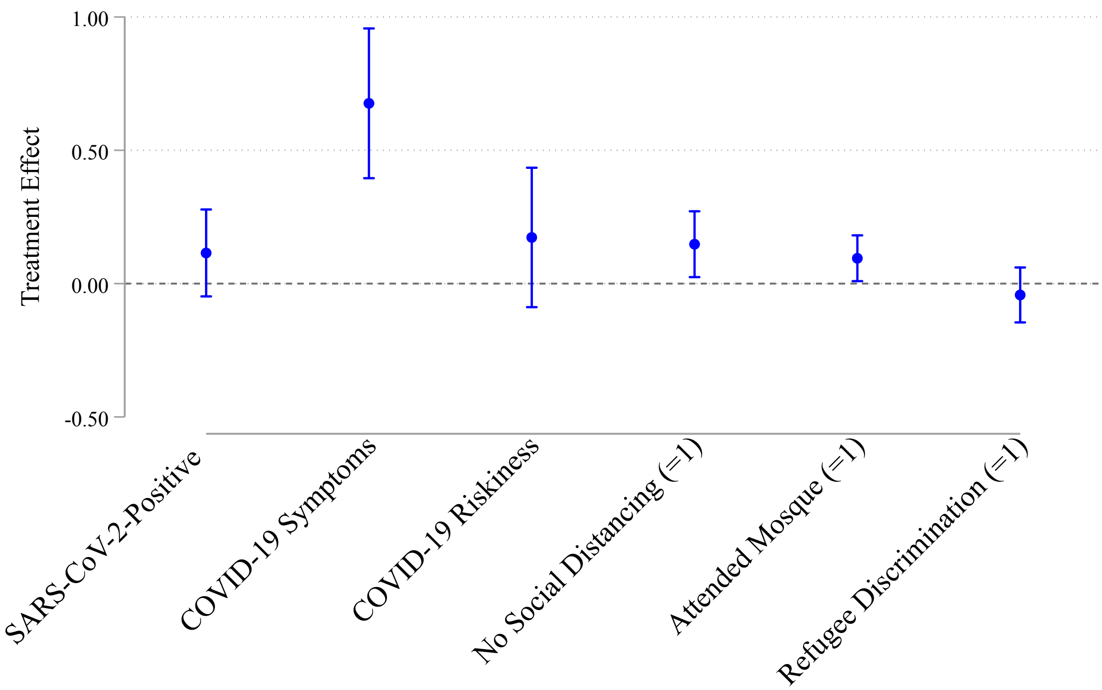
Note:
This figure visualizes a household's children's average days in school in the last week, broken down by age and gender.

Figure 2.8: Effects on Food Insecurity



Note:
The figure visualizes the regression in Table 2.9. The x-axis represents the dependent variables. On the y-axis, the dots represent the point estimates and the bars represent the 95% confidence intervals.

Figure 2.9: Effects on COVID-19 Outcomes



Note:
The figure visualizes the regression in Table 2.10. The x-axis represents the dependent variables. On the y-axis, the dots represent the point estimates and the bars represent the 95% confidence intervals.

2.11 Tables

Table 2.1: Retention and Compliance

	Retention	Treatment Implementation	
Treatment Assignment	-0.02	0.34***	0.35***
	(0.04)	(0.02)	(0.02)
Control Mean	0.77	0	
N	2,126	1,614	1,614
Controls		✓	✓
Enumerator FE			✓

Note:
The table shows our regression retention/attrition and compliance. The dependent variables across columns are retention (defined as an indicator for the household completing our phone survey after treatment implementation) and compliance (defined as an indicator for treatment implementation). The independent variable of interest is treatment assignment. The regressions also have enumerator fixed effects, month-by-year fixed effects, community-level controls (Irbid/Mafraq governorate and population quartile), and household-level controls (vulnerability-assessment quartile, number of children, number of children plus adults, respondent gender, and respondent age). In the parentheses are robust standard errors clustered at the locality level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.)

Table 2.2: Descriptives from Baseline Data & Balance

		Treatment (T)	Control (C)	$T-C$	N
Panel A: Respondent					
(1)	Female (=1)	0.46	0.52	-0.05*	1,619
				(0.03)	
(2)	Age (10-Years Bins)	33.99	34.19	-0.21	1,619
				(0.61)	
(3)	Married (=1)	0.84	0.84	0.00	1,616
				(0.02)	
(4)	Disability (=1)	0.26	0.25	0.02	1,616
				(0.03)	
Panel B: Household					
(5)	Dependency Ratio	1.31	1.27	0.04	1,616
				(0.07)	
(6)	Household Size	5.2	5.14	0.06	1,616
				(0.15)	
(7)	Families in Same House	1.3	1.3	-0.00	1,616
				(0.04)	
(8)	Children	2.96	2.89	0.07	1,619
				(0.12)	
Panel C: Housing					
(9)	Piped water (=1)	0.66	0.67	-0.01	1,616
				(0.07)	
(10)	Fully Constructed Roof (=1)	0.12	0.08	0.04**	1,616
				(0.02)	
(11)	Functional Windows (=1)	0.22	0.22	-0.01	1,616
				(0.03)	
(12)	Completed Floor (=1)	0.44	0.4	0.04	1,616
				(0.06)	
(13)	Toilet (=1)	0.93	0.92	0.01	1,616
				(0.04)	
(14)	Plan to Stay in Shelter (=1)	0.92	0.92	0.00	1,537
				(0.02)	
(15)	Monthly Rent (USD PPP)	155.93	159.42	-3.48	1,431
				(8.31)	
(16)	Lease Contract (=1)	0.77	0.78	-0.01	1,431
				(0.04)	
(17)	Times Changed Shelter	0.47	0.44	0.03	1,616
				(0.08)	
(18)	Permanent Shelter (=1)	0.91	0.86	0.05	1,616
				(0.06)	
Chi-squared test for joint orthogonality of all covariates (p-value): 0.113					

Note:
The table shows baseline descriptive statistics and baseline balance. Panels A, B, and C represent average respondent, household, and housing characteristics. The dependent variables across rows are baseline characteristics (except for gender and age, which we observe at endline). Disability is defined as the Washington Group Short Set on Disability (which classifies respondents by seeing, hearing, walking, cognition, self-care, and communication). Rent is winsorized at the top 1%. The independent variable of interest is the treatment assignment. Monetary values are in USD PPP. In the parentheses are robust standard errors clustered at the locality level. (*** p<0.01, ** p<0.05, * p<0.1.)

Table 2.3: Descriptives from Endline Data

	Mean	Median	Min	Max	SD	N
Annualized Income						
Before COVID-19 Lockdown	2878	2238	0	13953	2757	791
During COVID-19 Lockdown	742	0	0	8208	1642	794
After COVID-19 Lockdown	2441	2089	0	11864	2474	792
Annualized Rent	955	861	0	3616	1004	735
Depression (=1)	.56	1	0	1	.5	791
Learning Activities (Total)	.22	0	0	4	.56	670

Note:
The table shows descriptive statistics (for the control group) from our phone survey. Depression is defined as an indicator for the psychologist-validated Center for Epidemiological Studies Depression (CES-D) scale being greater than or equal to 10. Monetary values are in USD PPP.

Table 2.4: Effects on Housing Quality and Housing-Related Finance

	Outcome	Treatment	Q-Value	Control Mean	N
(1)	Overall Housing Quality	0.47* (0.26)	0.175	0	1,614
(2)	Housing-Material Quality	0.40** (0.19)	0.16	1.45	1,556
(3)	Water Access (Scale 1-5)	0.34* (0.20)	0.175	3.42	1,614
(4)	Clean Water (=1)	0.20*** (0.06)	0.009	0.17	1,614
(5)	Electricity (=1)	0.01* (0.01)	0.162	0.99	1,614
(6)	Security (=1)	0.14** (0.07)	0.16	0.79	1,614
(7)	Toilet (=1)	-0.14 (0.14)	0.257	0.23	1,614
(8)	Rooms Occupied	0.54*** (0.17)	0.024	2.85	1,614
(9)	Housing Expenditures	3.80* (1.96)	0.162	14.11	1,549

Note:
The table shows our regression results on housing quality and housing-related finances. Each row is its own dependent variable. Overall Housing Quality is defined as a normalized housing quality index that includes indicators for quality floors and roofs, indicators for access to grid electricity and piped water, and the number of people per room. Housing-Material Quality is defined as the summation of two indicators for high-quality floors and roofs. Water Access is defined on a scale from 1-5 where 1 is very inaccessible while 5 is very accessible. Clean Water is defined as an indicator for households having treated drinking water (such as by a filter). Security is defined as an indicator for whether the dwelling can be locked or secured. Housing Expenditures is defined as the total of rent and mortgage payments (inclusive of rental assistance) divided by household size. The independent variable of interest is the TOT treatment indicator, which is the predicted value from a first-stage regression of treatment implementation on treatment assignment. Monetary values are in USD PPP. The regressions also have enumerator fixed effects, month-by-year fixed effects, community-level controls (Irbid/Mafraq governorate and population quartile), and household-level controls (vulnerability-assessment quartile, shelter program, number of children, number of children plus adults, respondent gender, and respondent age). In the parentheses are robust standard errors clustered at the locality level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) Q-values are calculated per Anderson (2008) and correspond to Family 2 in our pre-analysis plan.

Table 2.5: Effects on Education and Health

	Outcome	Treatment	Q-Value	Control Mean	N
(1)	Disability (=1)	-0.00	0.7	0.26	1,614
		(0.06)			
(2)	Subjective Health (Scale 1-4)	-0.10	0.43	3.34	1,614
		(0.11)			
(3)	Learning Activities (Total)	0.18**	0.116	0.22	1,388
		(0.09)			
(4)	Learning Activities (=1)	0.16***	0.031	0.17	1,388
		(0.06)			
(5)	Childcare Hours	-6.18**	0.162	29.91	1,511
		(3.12)			

Note:
The table shows our regression results on education and health. Each row is its own dependent variable. Disability is defined as the Washington Group Short Set on Disability (which classifies respondents by seeing, hearing, walking, cognition, self-care, and communication). Learning Activities variables are defined as total activities or as an indicator done in the last 24 hours (including homework, e-learning, reading, and educational programs and videos). The independent variable of interest is the TOT treatment indicator, which is the predicted value from a first-stage regression of treatment implementation on treatment assignment. The regressions also have enumerator fixed effects, month-by-year fixed effects, community-level controls (Irbid/Mafraq governorate and population quartile), and household-level controls (vulnerability-assessment quartile, shelter program, number of children, number of children plus adults, respondent gender, and respondent age). In the parentheses are robust standard errors clustered at the locality level. (*** p<0.01, ** p<0.05, * p<0.1.) Q-values are calculated per Anderson (2008) and correspond to Family 6 (outcomes 1 and 2) and Family 14 (outcomes 3 and 4) in our pre-analysis plan.

Table 2.6: Effects on General Finances

	Outcome	Treatment	Q-Value	Control Mean	N
Panel A: Before COVID-19 Lockdown					
(1)	Loan Uptake (=1)	-0.21***	0.011	0.61	1,612
		(0.07)			
(2)	Hours Worked (Per Week)	1.80	0.746	30.03	1,608
		(5.03)			
(3)	Labor Income (USD PPP Per Week)	-10.33	0.414	55.42	1,610
		(7.71)			
Panel B: During COVID-19 Lockdown					
(4)	Loan Uptake (=1)	-0.34***	0.001	0.78	1,612
		(0.07)			
(5)	Hours Worked (Per Week)	0.57	0.738	1.34	1,614
		(0.73)			
(6)	Labor Income (USD PPP Per Week)	-9.44***	0.023	14.31	1,613
		(3.39)			
Panel C: After COVID-19 Lockdown					
(7)	Loan Uptake (=1)	-0.32***	0.001	0.43	1,611
		(0.07)			
(8)	Hours Worked (Per Week)	0.88	0.752	24.13	1,610
		(3.80)			
(9)	Labor Income (USD PPP Per Week)	-1.86	0.752	46.97	1,609
		(6.16)			

Note:
The table shows our regression results on general finances. Each row is its own dependent variable. Panels A, B, and C represent the three time periods of before, during, and after Jordan's COVID-19 lockdown. Hours Worked and Labor Income are winsorized at the top 1%. The independent variable of interest is the TOT treatment indicator, which is the predicted value from a first-stage regression of treatment implementation on treatment assignment. Monetary values are in USD PPP. The regressions also have enumerator fixed effects, month-by-year fixed effects, community-level controls (Irbid/Mafraq governorate and population quartile), and household-level controls (vulnerability-assessment quartile, shelter program, number of children, number of children plus adults, respondent gender, and respondent age). In the parentheses are robust standard errors clustered at the locality level. (*** p<0.01, ** p<0.05, * p<0.1.) Q-values are calculated per Anderson (2008) and correspond to Family 14 in our pre-analysis plan.

Table 2.7: Heterogeneous Effects on General Finances by HH-Head Gender

	Treat	Treat×Female Head	Female Head	Control Mean	N
Panel A: Before COVID-19 Lockdown					
(1) Loan Uptake (=1)	-0.12	-0.36**	0.07*	0.61	1,612
	(0.07)	(0.17)	(0.04)		
(2) Hours Worked (Per Week)	6.27	-16.54	-4.38	30.03	1,608
	(5.20)	(11.34)	(3.08)		
(3) Labor Income (USD PPP Per Week)	1.84	-45.63***	-5.09	55.42	1,610
	(7.55)	(17.62)	(4.99)		
Panel B: During COVID-19 Lockdown					
(4) Loan Uptake (=1)	-0.23***	-0.41***	0.10***	0.78	1,612
	(0.07)	(0.13)	(0.03)		
(5) Hours Worked (Per Week)	1.06	-1.85	0.10	1.34	1,614
	(0.96)	(2.15)	(0.63)		
(6) Labor Income (USD PPP Per Week)	-3.80	-21.65**	4.14*	14.31	1,613
	(3.92)	(8.59)	(2.38)		
Panel C: After COVID-19 Lockdown					
(7) Loan Uptake (=1)	-0.26***	-0.24	0.07*	0.43	1,611
	(0.07)	(0.20)	(0.04)		
(8) Hours Worked (Per Week)	4.98	-15.18*	-5.49**	24.13	1,610
	(4.72)	(9.13)	(2.68)		
(9) Labor Income (USD PPP Per Week)	7.72	-35.33**	-8.43*	46.97	1,609
	(6.81)	(14.06)	(4.33)		

Note:
The table shows our regression results on general finances, heterogeneous by head-of-household gender. Each row is its own dependent variable. Panels A, B, and C represent the three time periods of before, during, and after Jordan's COVID-19 lockdown. Hours Worked and Labor Income are winsorized at the top 1%. The independent variables of interest are the TOT treatment indicator, which is the predicted value from a first-stage regression of treatment implementation on treatment assignment, and its interaction with an indicator for the head of household being female. Monetary values are in USD PPP. The regressions also have enumerator fixed effects, month-by-year fixed effects, community-level controls (Irbid/Mafraq governorate and population quartile), and household-level controls (vulnerability-assessment quartile, shelter program, number of children, number of children plus adults, respondent gender, and respondent age). In the parentheses are robust standard errors clustered at the locality level. (*** p<0.01, ** p<0.05, * p<0.1.)

Table 2.8: Effects on Movement of and between Households

	Outcome	Treatment	Q-Value	Control Mean	N
(1)	Migration (=1)	0.07	0.292	0.1	1,613
		(0.06)			
(2)	Children	0.35**	0.073	3.27	1,614
		(0.16)			
(3)	Adults	-0.01	0.776	2.69	1,614
		(0.18)			
(4)	Girls Under 12	0.03	0.688	1.21	1,614
		(0.13)			
(5)	Boys Under 12	0.11	0.387	1.27	1,614
		(0.12)			
(6)	Girls 13-17	0.06	0.409	0.39	1,614
		(0.10)			
(7)	Boys 13-17	0.15**	0.086	0.4	1,614
		(0.07)			

Note:
The table shows our regression results on migration and inter-household movement. Each row is its own dependent variable. Migration is defined as an indicator for if the respondent has moved since October 2019. The independent variable of interest is the TOT treatment indicator, which is the predicted value from a first-stage regression of treatment implementation on treatment assignment. The regressions also have enumerator fixed effects, month-by-year fixed effects, community-level controls (Irbid/Mafraq governorate and population quartile), and household-level controls (vulnerability-assessment quartile, shelter program, number of children, number of children plus adults, respondent gender, and respondent age). In the parentheses are robust standard errors clustered at the locality level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) Q-values are calculated following Anderson (2008) and correspond to Family 5 (outcome 1) and Family 8 (outcomes 2-7) in the pre-analysis plan.

Table 2.9: Effects on Food Insecurity

	Outcome	Treatment	Q-Value	Control Mean	N
(1)	Respondent Food Insecurity (=1)	0.24***	0.005	0.35	1,608
		(0.07)			
(2)	Adult Food Insecurity (=1)	0.18***	0.006	0.38	1,614
		(0.06)			
(3)	Child Food Insecurity (=1)	0.24***	0.004	0.22	1,472
		(0.07)			
(4)	Total Food Assistance (USD PPP)	-24.07**	0.027	103.34	1,573
		(10.81)			
(5)	*Per Capita* Food Assistance (USD PPP)	-4.59***	0.015	16.98	1,573
		(1.63)			
(6)	Overall Food Consumption	0.33***	0.016	-0.3	1,610
		(0.11)			

Note:
The table shows our regression results on food insecurity. Each row is its own dependent variable. Respondent/Adult/Child Food Insecurity is defined as an indicator for if Respondent/Adult/Child indicated going to bed hungry because they did not have enough food on at least one day in the last seven days. Overall Food Consumption is defined as an ordinal variable that compares current consumption to before Jordan's COVID-19 lockdown (decrease (-1), same (0), and increase (1)). Food Assistance is winsorized at the top 1%. The independent variable of interest is the TOT treatment indicator, which is the predicted value from a first-stage regression of treatment implementation on treatment assignment. Monetary values are in USD PPP. The regressions also have enumerator fixed effects, month-by-year fixed effects, community-level controls (Irbid/Mafraq governorate and population quartile), and household-level controls (vulnerability-assessment quartile, shelter program, number of children, number of children plus adults, respondent gender, and respondent age). In the parentheses are robust standard errors clustered at the locality level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) Q-values are calculated following Anderson (2008) and correspond to Family 2 (outcome 1), Family 6 (outcomes 2 and 3), Family 8 (outcome 4), and Family 14 (outcome 5 and 6) in the pre-analysis plan.

Table 2.10: Effects on COVID-19 Outcomes

	Outcome	Treatment	Q-Value	Control Mean	N
(1)	SARS-CoV-2-Positive	0.11	0.414	0.04	105
		(0.08)			
(2)	COVID-19 Symptoms	0.68***	0.001	0	1,614
		(0.14)			
(3)	COVID-19 Riskiness	0.17	0.414	0	1,613
		(0.13)			
(4)	No Social Distancing (=1)	0.15**	0.058	0.51	1,613
		(0.06)			
(5)	Attended Mosque (=1)	0.10**	0.084	0.25	1,613
		(0.04)			
(6)	Foregone Health Visits	-1.28	0.437	2.66	1,610
		(1.09)			
(7)	Refugee Discrimination (=1)	-0.04	0.738	0.16	1,613
		(0.05)			

Note:
The table shows our regression results on COVID-19 outcomes. Each row is its own dependent variable. COVID-19 Symptoms is defined as a normalized index of an indicator for any household member testing positive for SARS-CoV-2 and the summation of total symptoms experienced by the respondent. COVID-19 Riskiness is defined as a normalized index of indicators that increase COVID-19 risk (not social distancing, leaving the house, leaving the community, going to a grocery store or market, attending church, attending mosque, and attending social gatherings). Refugee Discrimination is defined as an indicator for if the respondent felt discriminated against because of refugee status amidst the pandemic. The independent variable of interest is the TOT treatment indicator, which is the predicted value from a first-stage regression of treatment implementation on treatment assignment. The regressions also have enumerator fixed effects, month-by-year fixed effects, community-level controls (Irbid/Mafraq governorate and population quartile), and household-level controls (vulnerability-assessment quartile, shelter program, number of children, number of children plus adults, respondent gender, and respondent age). In the parentheses are robust standard errors clustered at the locality level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) Q-values are calculated following Anderson (2008) and correspond to Family 14 in the pre-analysis plan.

Bibliography

Acemoglu, D. (2009). *Introduction to Modern Economic Growth*. Princeton University Press.

Amodio, F., J. Choi, G. D. Giorgi, and A. Rahman (2018). Bribes vs. taxes: Market structure and incentives. IZA DP No. 11668, IZA – Institute of Labor Economics.

Aubry, T., G. Nelson, and S. Tsemberis (2015). Housing first for people with severe mental illness who are homeless: A review of the research and findings from the at home-chez soi demonstration project. *Canadian Journal of Psychiatry 60*(11).

Barnhardt, S., E. Field, and R. Pande (2017). Moving to opportunity or isolation? Network effects of a randomized housing lottery in urban India. *American Economic Journal: Applied Economics 9*(1), 1–32.

Basri, M. C., M. Felix, R. Hanna, and B. A. Olken (2019). Tax administration vs. tax rates: Evidence from Corporate taxation in Indonesia. Working Paper 26150, National Bureau of Economic Research.

Bergman, P., R. Chetty, S. DeLuca, N. Hendren, L. F. Katz, and C. Palmer (2020). Creating Moves to Opportunity: Experimental evidence on barriers to neighborhood choice. Working Paper 26164, National Bureau of Economic Research.

BNWG (2020). Standards for non-camp refugee response, Jordan. Internal document.

Campante, F. R., D. Chor, and Q.-A. Do (2009). Instability and the incentives for corruption. *Economics & Politics 21*(1), 42–92.

Chalendard, C., A. Duhaut, A. M. Fernandes, A. Mattoo, G. Raballand, and B. Rijkers (2020). Does better information curb customs fraud? Policy Research Working Paper No. 9254, World Bank.

Chetty, R., N. Hendren, and L. F. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *American Economic Review 106*(4), 855–902.

Clampet-Lundquist, S. and D. Massey (2008, 07). Neighborhood effects on economic self-sufficiency: A reconsideration of the moving to opportunity experiment. *American Journal of Sociology 114*, 107–143.

Colonnelli, E. and M. Prem (2021). Corruption and firms. *Review of Economic Studies 0*, 1–38.

Dal Bo, E. and M. Rossi (2011). Term length and the effort of politicians. *Review of Economic Studies 78*(14), 1237–1263.

Day, J. and R. Cervero (2010). Effects of residential relocation on household and commuting expenditures in Shanghai, China. *International Journal of Urban and Regional Research 34*(4), 762–788.

Egger, D., E. Miguel, S. S. Warren, A. Shenoy, E. Collins, D. Karlan, D. Parkerson, A. M. Mobarak, G. Fink, C. Udry, M. Walker, J. Haushofer, M. Larreboure, S. Athey, P. Lopez-Pena, S. Benhachmi, M. Humphreys, L. Lowe, N. F. Meriggi, A. Wabwire, C. A. Davis, U. J. Pape, T. Graff, M. Voors, C. Nekesa, and C. Vernot (2021). Falling living standards during the COVID-19 crisis: Quantitative evidence from nine developing countries. *Science Advances 7*(6).

Finan, F., B. Olken, and R. Pande (2015). The personnel economics of the state. Working Paper 21825, National Bureau of Economic Research.

Fismana, R. and J. Svensson (2007). Are corruption and taxation really harmful to growth? Firm level evidence. *Journal of Development Economics 83*(1), 63–75.

GAGE (2020). Exploring the impacts of COVID-19 on adolescents in Jordan's refugee camps and host communities. Brief, Gender And Adolescence: Global Evidence.

Ginn, T. C. (2020). Prison or Sanctuary? An evaluation of camps for Syrian refugees. Accessed: 2021-1-06.

GIZ and INLUCC (2020). Perception de la corruption en tunisie. Etude quantitative, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH.

Hagen-Zanker, J., M. Ulrichs, and R. Holmes (2018). What are the effects of cash transfers for refugees in the context of protracted displacement? Findings from Jordan. *International Social Security Review 71*(2), 57–77.

Hibou, B. (2011). *The Force of Obedience: The Political Economy of Repression in Tunisia*. Boston, Massachusetts: Polity Press.

Hsiang, S., R. Kopp, A. Jina, J. Rising, M. Delgado, S. Mohan, D. J. Rasmussen, R. Muir-Wood, P. Wilson, M. Oppenheimer, K. Larsen, and T. Houser (2017). Estimating economic damage from climate change in the United States. *Science 356*(6345), 1362–1369.

ICG (2017). La transition bloquée : corruption et régionalisme en tunisie. Rapport moyen-orient et afrique du nord n°177, International Crisis Group.

International Trade Administration (2020). Tunisia - Country Commercial Guide. https://www.trade.gov/country-commercial-guides/tunisia-import-tariffs. Accessed: 2020-12-06.

Jacob, B. A., M. Kapustin, and J. Ludwig (2015). The impact of housing assistance on child outcomes: Evidence from a randomized housing lottery. *The Quarterly Journal of Economics 130*(1), 465–506.

Jacob, B. A. and J. Ludwig (2012). The effects of housing assistance on labor supply: Evidence from a voucher lottery. *American Economic Review 102*(1), 272–304.

Katz, L. F., J. R. Kling, and J. B. Liebman (2001). Moving to Opportunity in Boston: Early results of a randomized mobility experiment. *The Quarterly Journal of Economics 116*(2), 607–654.

Kee, H. L., A. Nicita, and M. Olarreaga (2008). Import demand elasticities and trade distortions. *Review of Economics and Statistics 90*(4), 666–682.

Khan, A. Q., A. I. Khwaja, and B. A. Olken (2016). Tax farming redux: Experimental evidence on performance pay for tax collectors. *Quarterly Journal of Economics 131*(1), 219–271.

Kling, J. R. (2006). Incarceration length, employment, and earnings. *American Economic Review 96*(3), 863–876.

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

Kumar, T. (2021). Home-price subsidies increase local-level political participation in urban India. *Journal of Politics*.

Lehmann, C. and D. Masterson (2014). Emergency economies: The impact of cash assistance in Lebanon. Report, International Rescue Committee.

Ludwig, J., G. J. Duncan, L. A. Gennetian, L. F. Katz, R. C. Kessler, J. R. Kling, and L. Sanbonmatsu (2013). Long-Term neighborhood effects on low-income families: Evidence from Moving to Opportunity. *American Economic Review 103*(3), 226–31.

Macchiavello, R. and A. Morjaria (2015). The value of relationships: Evidence from a supply shock to Kenyan rose exports. *American Economic Review 105*(9), 2911–45.

Mills, G., D. Gubits, L. Orr, D. Long, J. Feins, B. Kaul, M. Wood, A. J. . Associates, C. Consulting, and the QED Group (2006). The effects of housing vouchers on welfare families. Report, U.S. Department of Housing and Urban Development, Office of Policy Development and Research.

Nash, J. F. (1950). The bargaining problem. *Econometrica 18*(2), 155–162.

National Commission for Investigations of Bribery and Corruption (2011). *Report of the National Commission for Discovering Facts Surrounding Corruption and Bribery [Taqrir al-lajna al-wataniyya li taqassi al-haqa'iq hawl al-fisad wa al-rashwa]*. Tunis, Tunisia: National Commission for Investigations of Bribery and Corruption.

NRC (2015). In search of a home: Access to adequate housing in Jordan. Report, Norwegian Refugee Council.

OECD, A. U. Commission, and A. T. A. Forum (2020). *Revenue Statistics in Africa*.

Olken, B. A. and P. Barron (2009). The simple economics of extortion: Evidence from trucking in Aceh. *Journal of Political Economy 117*(3), 417–452.

Picarelli, N. (2019). There is no free house. *Journal of Urban Economics 111*, 35–52.

Rijkers, B., L. Baghdadi, and G. Raballand (2017). Political connections and tariff evasion evidence from Tunisia. *World Bank Economic Review 31*(2), 459–482.

Rijkers, B., C. Freund, and A. Nucifora (2014). All in the family: State capture in Tunisia. Policy Research Working Paper No. 6810, World Bank.

Rozo, S. V. and M. Sviatschi (2021). Is a refugee crisis a housing crisis? Only if housing supply is unresponsive. *Journal of Development Economics 148*, 102563.

Sanchez de la Sierra, R. (2020). On the origins of the state: Stationary bandits and taxation in eastern congo. *Journal of Political Economy 128*(1).

Sanchez de la Sierra, R. and K. Titeca (2017). The state as organized crime: Industrial organization of the traffic police in the Democratic Republic of the Congo.

Sequeira, S. and S. Djankov (2014). Corruption and firm behavior: Evidence from African ports. *Journal of International Economics 94*(2), 277–294.

Tokarick, S. (2010). A method for calculating export supply and import demand elasticities. Working Paper WP/10/180, International Monetary Fund.

Truth and Dignity Commission (2019). The final comprehensive report. Report, Truth and Dignity Commission.

UN (2021). *Key Statistics and Trends in International Trade 2020: Trade Trends Under the Covid-19 Pandemic*. United Nations Publications.

UNHCR (2021). Global Trends in Forced Displacement. Report, United Nations High Commissioner for Refugees.

van Dijk, W. (2019). The socioeconomic consequences of housing assistance. Job market paper.

World Bank (2021). World Bank Open Data. https://data.worldbank.org.

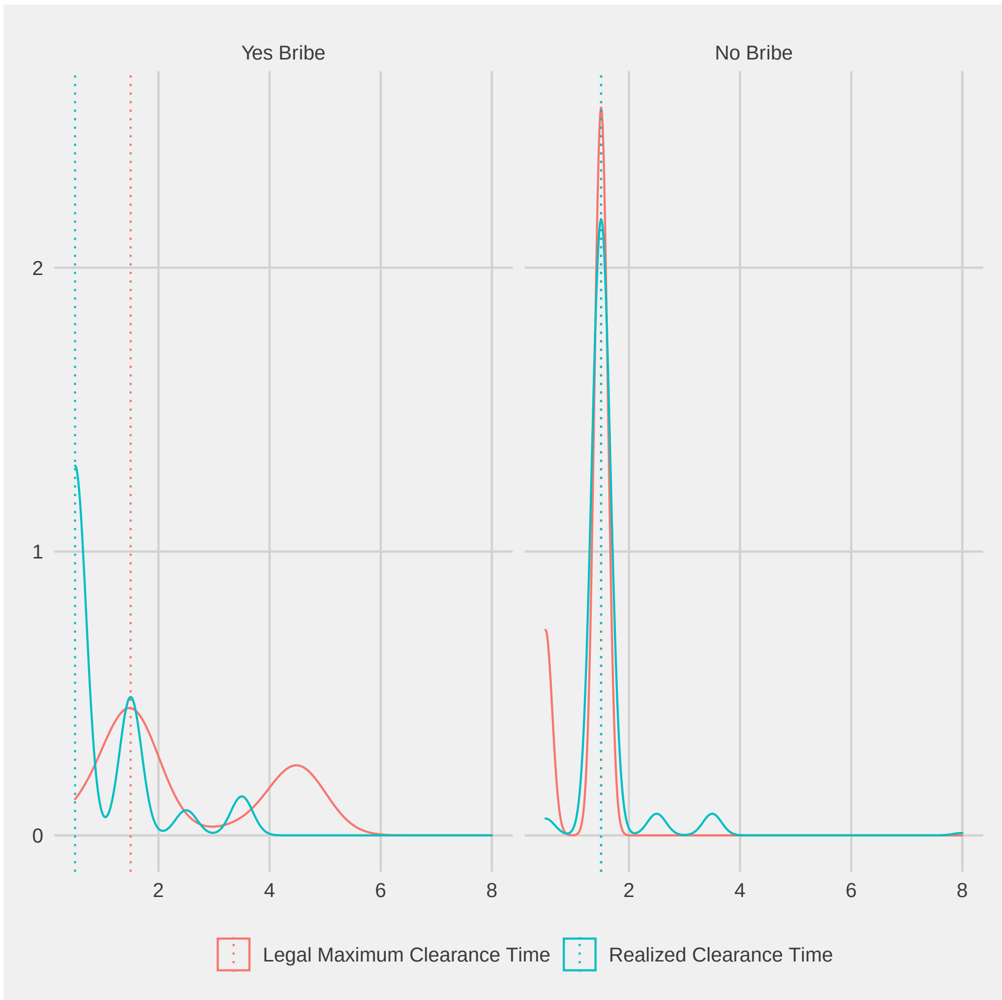
Yerkes, S. and M. Muasher (2017). Tunisia's corruption contagion: A transition at risk.

Appendix A

Corruption Dynamics in International Trade

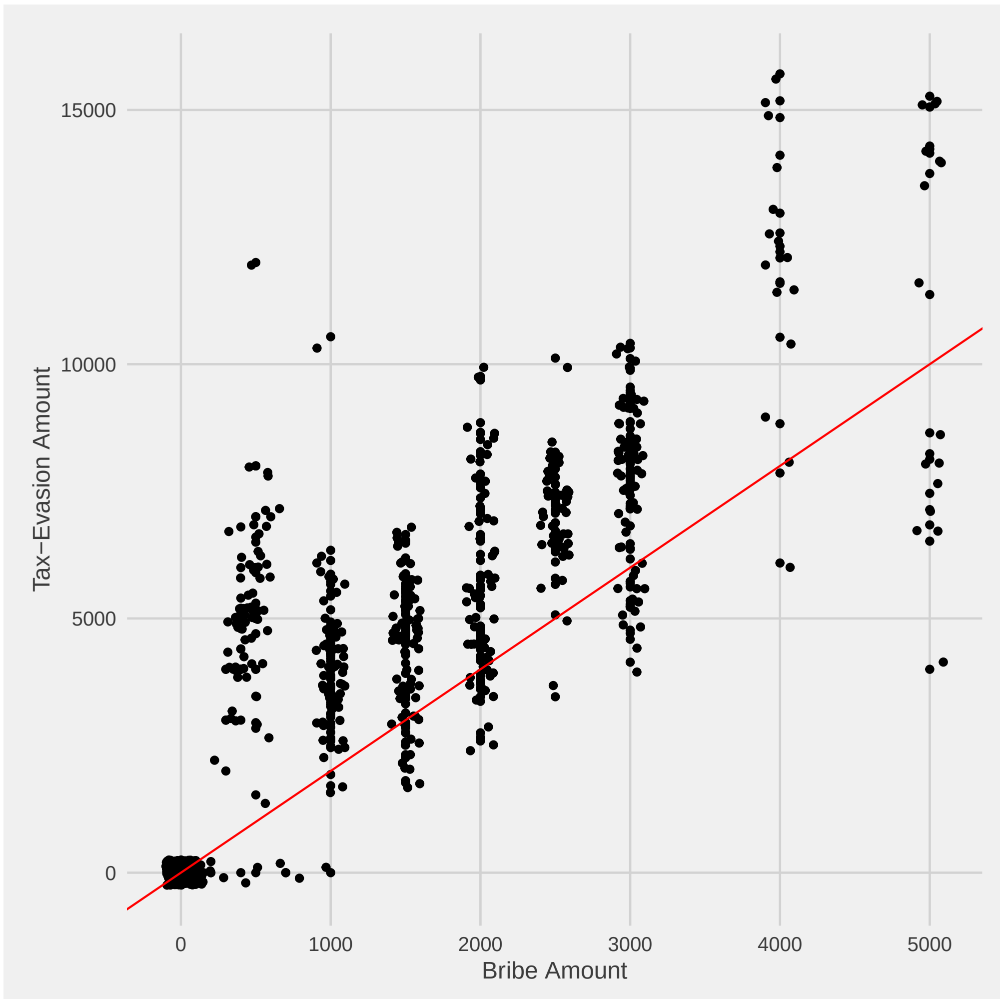
A.1 Figures

Figure A.1: Clearance Time by Bribery



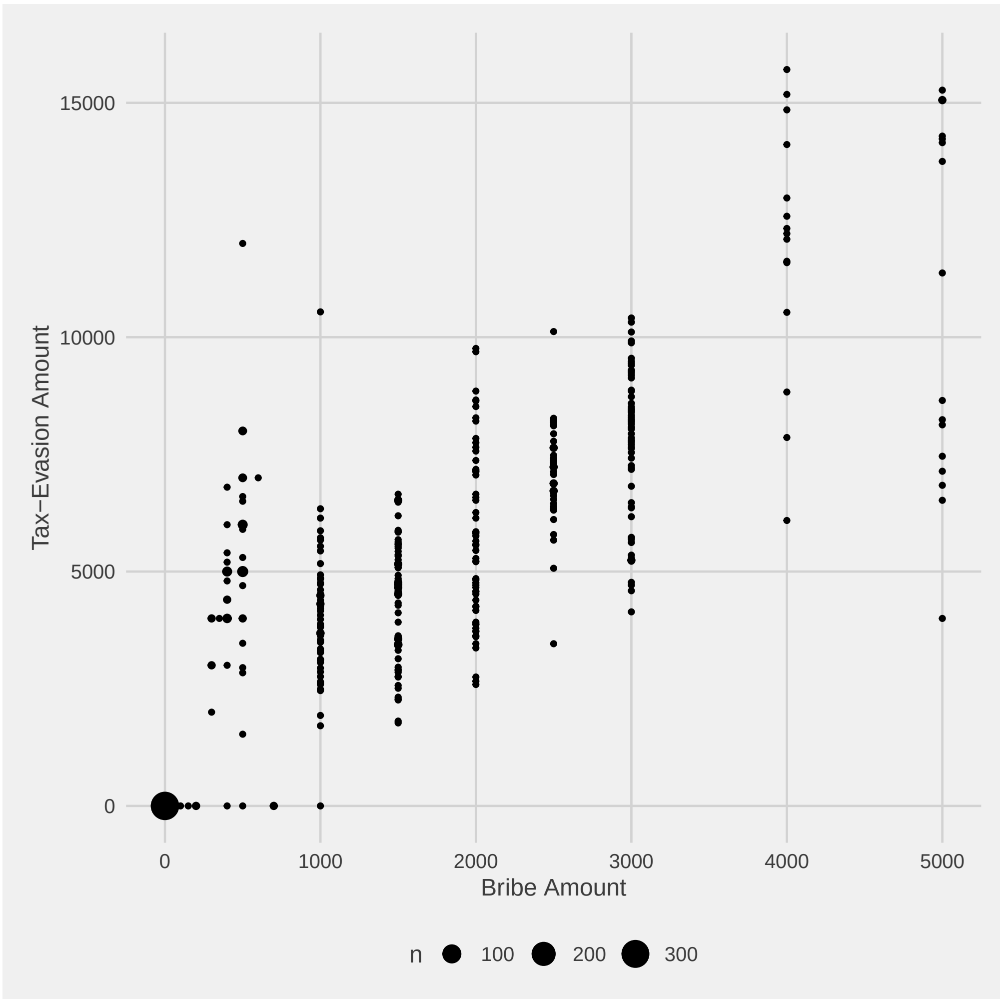
Note:
The figure is a kernel-density chart visualizing clearance time by bribery. The figure uses the kernel-density algorithm to plot nonparametric estimates of the probability density functions. In the left-hand panel, the two lines represent the legal maximum clearance time and the realized clearance time for transactions in which the trader paid a bribe. In the right-hand panel, the two lines represent the legal maximum clearance time and the realized clearance time for transactions in which the trader did not pay a bribe. Monetary values are in USD. The respective medians are represented by a dotted vertical line. The sample used is all cleaned data from the latest (2021) version of our survey.

Figure A.2: Tax Evasion as Function of Bribery (Alt. #1)



Note:
The figure is a scatter chart visualizing tax evasion as a function of bribery. The x-axis represents the bribe amount paid by the trader to the liquidateur official. The y-axis represents the tax-evasion amount, defined as the tax owed less the tax paid. Monetary values are in USD. The figure plots data jittered by up to $250 in each direction and trimmed at the top/bottom 1%. The red line represents a 50/50 split of the tax-evasion amount between the official and the trader; data above the line means the trader has a higher share; data below the line means the official has a higher share. The sample used is all cleaned data from the latest (2021) version of our survey. The figure mirrors Figure 1.3, but with different formatting.

Figure A.3: Tax Evasion as Function of Bribery (Alt. #2)



Note:
The figure is a scatter chart visualizing tax evasion as a function of bribery. The x-axis represents the bribe amount paid by the trader to the liquidateur official. The y-axis represents the tax-evasion amount, defined as the tax owed less the tax paid. Monetary values are in USD. The figure plots data The figure plots data trimmed at the top/bottom 1%. Different sized dots represent more/fewer observations at that coordinate. The sample used is all cleaned data from the latest (2021) version of our survey. The figure mirrors Figure 1.3, but with different formatting.

A.2 Tables

Table A.1: Balance on Officials' Demographics

	(1) Age Official	(2) Female Official
Tax Owed	0.000	-0.000
	(0.000)	(0.000)
Shipment KGs	-0.000***	0.000***
	(0.000)	(0.000)
Shipment Value	-0.000*	-0.000
	(0.000)	(0.000)
Other Trade Costs	0.000***	0.000
	(0.000)	(0.000)
Red Lane	0.740***	-0.065**
	(0.270)	(0.032)
Client Shipments/Y	-0.009	-0.001*
	(0.009)	(0.001)
Client Employees	0.009	0.001
	(0.007)	(0.001)
Origin=Central & South America.	-1.995***	0.056
	(0.768)	(0.168)
Origin=China.	-0.366	0.018
	(0.329)	(0.031)
Origin=North America.	2.161	-0.002
	(1.621)	(0.110)
Origin=Other Asia.	0.360	0.084*
	(0.522)	(0.046)
Origin=Other Europe.	-0.223	0.078**
	(0.342)	(0.035)
Origin=Other Middle East & North Africa.	0.921	-0.073**
	(0.699)	(0.032)
Origin=Sub-Saharan Africa.	-3.144***	-0.081*
	(0.596)	(0.048)
Origin=Turkey.	-0.036	0.011
	(0.294)	(0.027)
Client=European Union.	1.128	-0.017
	(1.730)	(0.098)
Client=North America.	3.304*	0.344***
	(1.700)	(0.118)
Client=Other Europe.	0.139	0.481***
	(0.548)	(0.171)
Client=Turkey.	0.846***	0.315***
	(0.282)	(0.033)
Outcome Mean	49	0.17
Observations	1,358	1,358
R^2	0.330	0.123
Month-by-Year FE	YES	YES
Official Controls	YES	YES

Note:
The table shows our "balance" test of the random assignment of officials to shipments. The dependent variables are official demographics, namely age and gender. The independent variables are the other official characteristics (age or gender, rank, tenure, and liquidateur/reviseur status), month-by-year fixed effects, and shipment controls (weight, value, tax owed, other trade costs, red-lane assignment, origin client employees, client trade volume, client nationality). Monetary values are in USD. In the parentheses are robust standard errors clustered at the shipment level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) The sample used is all cleaned data from the latest (2021) version of our survey.

Table A.2: Balance on No. Matches

	No. Matches
Tax Owed	0.000
	(0.000)
Shipment KGs	-0.000***
	(0.000)
Shipment Value	-0.000
	(0.000)
Other Trade Costs	-0.000
	(0.000)
Red Lane	1.020***
	(0.350)
Client Shipments/Y	0.044***
	(0.012)
Client Employees	-0.033***
	(0.011)
Origin=Central & South America.	0.474
	(1.206)
Origin=China.	-0.915*
	(0.509)
Origin=North America.	-1.552
	(0.944)
Origin=Other Asia.	0.089
	(0.628)
Origin=Other Europe.	-1.185**
	(0.577)
Origin=Other Middle East & North Africa.	-0.379
	(0.654)
Origin=Sub-Saharan Africa.	-6.895***
	(1.186)
Origin=Turkey.	0.028
	(0.454)
Client=European Union.	0.455
	(0.767)
Client=North America.	7.537***
	(1.384)
Client=Other Europe.	-1.277
	(0.867)
Client=Turkey.	1.114
	(1.745)
Outcome Mean	14
Observations	1,358
R^2	0.824
Trader FE	YES
Month-by-Year FE	YES
Official Controls	YES

Note:
The table shows our “balance” test of the random assignment of officials to shipments. The dependent variable is the number of matches between the trader and the official to whom they were ostensibly randomly assigned. The independent variables are trader fixed effects, month-by-year fixed effects, official controls (age, gender, rank, tenure, and liquidateur/reviseur status), and shipment controls (weight, value, tax owed, other trade costs, red-lane assignment, origin client employees, client trade volume, client nationality). Monetary values are in USD. In the parentheses are robust standard errors clustered at the shipment level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) The sample used is all cleaned data from the latest (2021) version of our survey.

Table A.3: Effect of Officials' Demographics on Bribery and Tax Evasion (Fewer Controls)

	(1)	(2)	(3)	(4)	(5)	(6)
		Bribe Yes/No			Tax-Paid (Log)	
Panel A: Age						
Age Official	0.010***	0.010***	0.016***	0.043***	0.025***	0.011***
	(0.002)	(0.002)	(0.002)	(0.005)	(0.005)	(0.003)
Panel B: Gender						
Female Official	0.032	0.030	0.007	-0.172*	-0.134*	0.011
	(0.032)	(0.030)	(0.027)	(0.088)	(0.074)	(0.045)
Outcome Mean		.57			9.5	
Observations	1,422	1,360	1,358	1,423	1,361	1,358
R^2	0.001	0.230	0.400	0.003	0.418	0.722
Trader FE	NO	NO	YES	NO	NO	YES
Official Controls	NO	NO	YES	NO	NO	YES
Month-by-Year FE	NO	YES	YES	NO	YES	YES
Shipment Controls	NO	YES	YES	NO	YES	YES

Note:
The table shows additional regression results for the effect of officials' demographics on bribery and tax evasion. Observations are at the official-level. In Columns 1-3, the dependent variable is an indicator for whether a bribe was exchanged. In Column 4-6, the dependent variable is the log of the amount of tax paid. In Panel A, which represents its own set of regressions, the independent variable of interest is official age, coded as ten-year bins. In Panel B, which represents its own set of regressions, the independent variable of interest is official gender, coded as a female indicator. Monetary values are in USD. The controls include trader fixed effects, month-by-year fixed effects, other official controls (age or gender, rank, tenure, and liquidateur/reviseur status), and shipment controls (weight, value, tax owed, other trade costs, red-lane assignment, client employees, and client trade volume). In the parentheses are robust standard errors clustered at the shipment level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) The sample used is all cleaned data from the latest (2021) version of our survey. The table mirrors Table 1.4, but with different sets of controls.

Table A.4: Effect of Officials' Demographics on Bribery and Tax Evasion (Trader FEs)

	(1) Bribe Yes/No	(2) Tax Paid (Log)
Age Official	0.001	0.002
	(0.001)	(0.003)
Female Official	0.017	0.003
	(0.020)	(0.042)
Outcome Mean	.57	9.52
Observations	1,358	1,358
R^2	0.714	0.757
Month-by-Year FE	YES	YES
Trader FE	YES	YES
Controls	YES	YES

Note:
The table shows additional regression results for the effect of officials' demographics on bribery and tax evasion. Observations are at the official-level. In Column 1, the dependent variable is an indicator for whether a bribe was exchanged. In Column 2, the dependent variable is the log of the amount of tax paid. The independent variables of interest are official age and gender, coded as ten-year bins and a female indicator respectively. Monetary values are in USD. The other controls include trader fixed effects, month-by-year fixed effects, other official controls (age or gender, rank, tenure, and liquidateur/reviseur status), and shipment controls (weight, value, tax owed, other trade costs, red-lane assignment, client employees, and client trade volume). In the parentheses are robust standard errors clustered at the shipment level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) The sample used is all cleaned data from the latest (2021) version of our survey. The table mirrors Table 1.4, but with trader fixed effects.

Table A.5: Effect of No. Matches on Bribe Probability

	Bribe Yes/No
No. Matches (Liquidateur)	0.000
	(0.001)
No. Matches (Reviseur)	-0.002*
	(0.001)
Outcome Mean	.57
Observations	1,358
R^2	0.714
Trader FE	YES
Month-by-Year FE	YES
Controls	YES

Note:
The table shows additional regression results for the effect of the number of matches on bribery and tax evasion. Observations are at the official-level. The dependent variable is an indicator for whether a bribe was exchanged. The independent variables of interest are the number of matches between the trader and the official originally assigned to the trader, for both the liquidateur and the reviseur. The controls include trader fixed effects, month-by-year fixed effects, official controls (age, gender, rank, tenure, and liquidateur/reviseur status), and shipment controls (weight, value, tax owed, other trade costs, red-lane assignment, client employees, and client trade volume). In the parentheses are robust standard errors clustered at the shipment level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) The sample used is all cleaned data from the latest (2021) version of our survey.

Table A.6: Effect of No. Matches on Bribery and Tax Evasion (Log Control)

	Tax Paid (Log)
No. Matches (Liquidateur)	0.001
	(0.001)
No. Matches (Reviseur)	0.000
	(0.000)
Tax Owed (Log)	0.998***
	(0.008)
Outcome Mean	9.52
Observations	1,358
R^2	0.986
Trader FE	YES
Month-by-Year FE	YES
Controls	YES

Note:
The table shows additional regression results for the effect of the number of matches on bribery and tax evasion. Observations are at the official-level. The independent variables of interest are the number of matches between the trader and the official originally assigned to the trader, for both the liquidateur and the reviseur. Monetary values are in USD. The other controls include trader fixed effects, month-by-year fixed effects, official controls (age, gender, rank, tenure, and liquidateur/reviseur status), and shipment controls (weight, value, tax owed, other trade costs, red-lane assignment, client employees, and client trade volume). In the parentheses are robust standard errors clustered at the shipment level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) The sample used is all cleaned data from the latest (2021) version of our survey. The table mirrors Table Table 1.5, but controls for tax owed in logs instead of linearly.

Table A.7: First-Stage Predicting Terminal Official with Original Official

	Terminal Matches
Original Matches	0.905***
	(0.009)
Constant	2.491***
	(0.210)
Observations	1,423
R^2	0.892

Note:
The table shows the first-stage regression of "Matches Original" (the number of matches with the randomly assigned official) on "Matches Terminal" (the number of matches with the official with whom the trader actually ended up working). Observations are at the official-level. In the parentheses are robust standard errors clustered at the shipment level. (*** p<0.01, ** p<0.05, * p<0.1.) The sample used is all cleaned data from the latest (2021) version of our survey.

Table A.8: Effect of No. Matches on Bribery and Tax Evasion (TOT)

	(1) Esp. Friendly Yes/No	(2) Bribe (Log)	(3) Tax Paid (Log)
No. Matches (Liquidateur)	0.012***	-0.011**	-0.011**
	(0.004)	(0.005)	(0.005)
No. Matches (Reviseur)	0.000	0.014**	0.002
	(0.003)	(0.006)	(0.003)
Outcome Mean	.33	7.15	9.52
Observations	1,358	769	1,358
R^2	0.423	0.793	0.755
Trader FE	YES	YES	YES
Month-by-Year FE	YES	YES	YES
Controls	YES	YES	YES

Note:
The table shows additional regression results for the effect of the number of matches on bribery and tax evasion. Observations are at the official-level. In Column 1, the dependent variable is the relationship quality between the official and the trader (which is a question in our questionnaire that asks yes/no whether the trader has an especially friendly relationship with the official). In Column 2, the dependent variable is the log of the bribe amount conditional on a bribe being exchanged. In Column 3, the dependent variable is the log of the amount of the tax paid. The independent variables of interest are the number of matches between the trader and the "Terminal" official who actually ended up working with the trader, instrumented by the "Original" official who was randomly assigned, for both the liquidateur and the reviseur. Monetary values are in USD. The controls include trader fixed effects, month-by-year fixed effects, official controls (age, gender, rank, tenure, and liquidateur/reviseur status), and shipment controls (weight, value, other trade costs, red-lane assignment, client employees, and client trade volume). In the parentheses are robust standard errors clustered at the shipment level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) The sample used is all cleaned data from the latest (2021) version of our survey. The table mirrors Table 1.5, but instead of showing ITT specifications it shows TOT specifications.

Table A.9: Effect of No. Matches on Bribery and Tax Evasion (100% Sample & No I.V.)

	(1) Esp. Friendly Yes/No	(2) Bribe (Log)	(3) Tax Paid (Log)
No. Matches (Liquidateur)	0.019***	-0.005	-0.009**
	(0.003)	(0.004)	(0.004)
No. Matches (Reviseur)	-0.011***	0.007	0.002
	(0.002)	(0.004)	(0.003)
Draw No.	0.118***	0.075*	0.063**
	(0.040)	(0.043)	(0.029)
Outcome Mean	0.33	7.02	9.48
Observations	1,358	769	1,358
R^2	0.434	0.794	0.756
Trader FE	YES	YES	YES
Month-by-Year FE	YES	YES	YES
Controls	YES	YES	YES

Note:
The table shows additional regression results for the effect of the number of matches on bribery and tax evasion. Observations are at the official-level. In Column 1, the dependent variable is the relationship quality between the official and the trader (which is a question in our questionnaire that asks yes/no whether the trader has an especially friendly relationship with the official). In Column 2, the dependent variable is the log of the bribe amount conditional on a bribe being exchanged. In Column 3, the dependent variable is the log of the amount of the tax paid. The independent variables of interest are the number of matches between the trader and the official originally assigned to the trader, for both the liquidateur and the reviseur. Monetary values are in USD. The other controls include trader fixed effects, month-by-year fixed effects, official controls (age, gender, rank, tenure, and liquidateur/reviseur status), and shipment controls (weight, value, other trade costs, red-lane assignment, client employees, and client trade volume), and draw number (an indicator for whether the official worked with the randomly assigned "Original" inspector or a different non-randomly assigned "Terminal" inspector). In the parentheses are robust standard errors clustered at the shipment level. (*** p<0.01, ** p<0.05, * p<0.1.) The table mirrors Table 1.5, but it (1) uses all cleaned data from all versions of our survey (2019-2021), and (2) instead of using instrumental-variables it controls for draw number.

Table A.10: Effect of No. Matches on Bribery and Tax Evasion (Fewer Controls)

	(1)	(2)	(3)	(4)	(5)	(6)
	Esp. Friendly Yes/No		Bribe (Log)		Tax-Paid (Log)	
No. Matches (Liquidateur)	0.009***	0.009***	-0.013***	-0.008**	-0.006	-0.008**
	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
No. Matches (Reviseur)	0.001	0.001	0.011**	0.010**	0.002	0.001
	(0.002)	(0.002)	(0.004)	(0.004)	(0.003)	(0.003)
Outcome Mean	0.33		7.2		9.5	
Observations	1,420	1,358	814	769	1,420	1,358
R^2	0.492	0.504	0.772	0.794	0.733	0.755
Trader FE	YES	YES	YES	YES	YES	YES
Month-by-Year FE	NO	YES	NO	YES	NO	YES
Official Controls	YES	YES	YES	YES	YES	YES
Shipment Controls	NO	YES	NO	YES	NO	YES

Note:
The table shows additional regression results for the effect of the number of matches on bribery and tax evasion. Observations are at the official-level. In Columns 1-2, the dependent variable is the relationship quality between the official and the trader (which is a question in our questionnaire that asks yes/no whether the trader has an especially friendly relationship with the official). In Column 3-4, the dependent variable is the log of the bribe amount conditional on a bribe being exchanged. In Column 5-6, the dependent variable is the log of the amount of the tax paid. The independent variables of interest are the number of matches between the trader and the official originally assigned to the trader, for both the liquidateur and the reviseur. Monetary values are in USD. The controls include trader fixed effects, month-by-year fixed effects, official controls (age, gender, rank, tenure, and liquidateur/reviseur status), and shipment controls (weight, value, tax owed, other trade costs, red-lane assignment, client employees, and client trade volume). In the parentheses are robust standard errors clustered at the shipment level. (*** $p<0.01$, ** $p<0.05$, * $p<0.1$.) The sample used is all cleaned data from the latest (2021) version of our survey. The table mirrors Table 1.5, but with different sets of controls.

A.3 Proofs

Nash Bargaining Equilibrium Equivalence

Recall that the trader's and the official's utilities, with and without tax evasion, and the joint surplus from tax evasion are defined as follows.

$$\begin{aligned}
U_T^0(t, b; \tilde{t}) &= -\tilde{t} \\
U_T^1(t, b; \tilde{t}) &= -t - \alpha(\tilde{t} - t)^2 - b \\
U_O^0(t, b; \tilde{t}) &= 0 \\
U_O^1(t, b; \tilde{t}) &= -\beta(\tilde{t} - t)^2 + b \\
S(t; \tilde{t}) &= (\tilde{t} - t) - \alpha(\tilde{t} - t)^2 - \beta(\tilde{t} - t)^2
\end{aligned}$$

And recall that equilibrium taxes paid are as follows.

$$t^* = \tilde{t} - \frac{1}{2}\frac{1}{\alpha + \beta}$$

Hence the utility and surplus terms evaluated at equilibrium taxes paid are as follows.

$$\begin{aligned}
U_T^0(t^*, b; \tilde{t}) &= -\tilde{t} \\
U_T^1(t^*, b; \tilde{t}) &= -(\tilde{t} - \frac{1}{2}\frac{1}{\alpha + \beta}) - \alpha(\tilde{t} - (\tilde{t} - \frac{1}{2}\frac{1}{\alpha + \beta}))^2 - b \\
U_O^0(t^*, b; \tilde{t}) &= 0 \\
U_O^1(t^*, b; \tilde{t}) &= -\beta(\tilde{t} - (\tilde{t} - \frac{1}{2}\frac{1}{\alpha + \beta}))^2 + b \\
S(t^*; \tilde{t}) &= (\tilde{t} - (\tilde{t} - \frac{1}{2}\frac{1}{\alpha + \beta})) - \alpha(\tilde{t} - (\tilde{t} - \frac{1}{2}\frac{1}{\alpha + \beta}))^2 - \beta(\tilde{t} - (\tilde{t} - \frac{1}{2}\frac{1}{\alpha + \beta}))^2
\end{aligned}$$

Let b_1^* be the bribe that solves the Nash Product optimization problem.

$$b_1^* = \text{argmax}[U_T^1(t^*, b; \tilde{t}) - U_T^0(t^*, b; \tilde{t})][U_O^1(t^*, b; \tilde{t}) - U_O^0(t^*, b; \tilde{t})]$$

Substituting the utility terms evaluated at equilibrium taxes paid, taking first-order conditions, and rearranging algebraically yields the following.

$$b_1^* = \frac{\alpha + 3\beta}{8(\alpha + \beta)^2}$$

Let b_2^* be the bribe such that the players' utilities with tax evasion are the same as their utilities without tax evasion (their "outside options") plus one half of the joint surplus.

$$\begin{aligned}
U_T^1(t^*, b_2^*; \tilde{t}) &= U_T^0(t^*, b_2^*; \tilde{t}) + \gamma S(t^*; \tilde{t}) \\
U_O^1(t^*, b_2^*; \tilde{t}) &= U_O^0(t^*, b_2^*; \tilde{t}) + (1-\gamma) S(t^*; \tilde{t}) \\
\text{s.t. } \gamma &\equiv 1 - \gamma \equiv \frac{1}{2}
\end{aligned}$$

Substituting the utility and surplus terms evaluated at equilibrium taxes paid and rearranging algebraically yields the following.

$$b_2^* = \frac{\alpha + 3\beta}{8(\alpha+\beta)^2}$$

$b_1^* = b_2^*$ and therefore the two Nash Bargaining equilibria are equivalent.

Bribe-Match Comparative Statics

Recall that the derivative of equilibrium bribes with respect to matches is composed of five terms, which we can label Z, A1, A2, B1, and B2.

$$\frac{\partial b^*(\alpha(u(m)), \beta(u(m)))}{\partial m} = \underbrace{\frac{1}{8(\alpha+\beta)^3}}_{\text{Z}} [\underbrace{\alpha'}_{\text{A1}} \underbrace{(-\alpha - 5\beta)}_{\text{A2}} + \underbrace{\beta'}_{\text{B1}} \underbrace{(-\alpha + 3\beta)}_{\text{B2}}]$$

Z is positive, A1 is negative, A2 is negative, B1 is negative, and B2 is either positive or negative. Hence the sign of the derivative depends entirely on the sign of B2.

$$\frac{\partial b^*(\alpha(u(m)), \beta(u(m)))}{\partial m} = \underbrace{\frac{1}{8(\alpha+\beta)^3}}_{+} [\underbrace{\alpha'}_{-} \underbrace{(-\alpha - 5\beta)}_{-} + \underbrace{\beta'}_{-} \underbrace{(-\alpha + 3\beta)}_{?}]$$

If B2 is negative, then B1×B2 is positive, and then the derivative is positive and bribes are increasing in matches. If B2 is positive, then B1×B2 is negative, and then the sign of the derivative depends on the following inequality.

$$\begin{aligned}
&\alpha'(-\alpha - 5\beta) + \beta'(-\alpha + 3\beta) < 0 \\
&\qquad \Rightarrow |\beta'(-\alpha + 3\beta)| > |\alpha'(-\alpha - 5\beta)| \\
&\qquad \Rightarrow \frac{|\beta'|}{|\alpha'|} > \frac{\alpha + 5\beta}{\alpha - 3\beta} \\
&\Rightarrow \frac{\partial b^*(\alpha(u(m)), \beta(u(m)))}{\partial m} < 0
\end{aligned}$$

Appendix B

Affordable Housing for Forced Migrants

B.1 Tables

Table B.1: Heterogeneous Effects on General Finances by Urban/Rural

		Treat	Treat× Urban	Urban	Control Mean	N
Panel A: Before COVID-19 Lockdown						
(1)	Loan Uptake (=1)	0.07	-0.33	0.08	0.61	1,612
		(0.21)	(0.24)	(0.06)		
(2)	Hours Worked (Per Week)	-0.17	3.20	-1.72	30.03	1,608
		(9.12)	(10.76)	(2.44)		
(3)	Labor Income (USD PPP Per Week)	-33.22*	28.10	-8.39*	55.42	1,610
		(19.06)	(21.24)	(4.59)		
Panel B: During COVID-19 Lockdown						
(4)	Loan Uptake (=1)	-0.25	-0.09	0.00	0.78	1,612
		(0.17)	(0.19)	(0.04)		
(5)	Hours Worked (Per Week)	-1.78	2.41	-1.10	1.34	1,614
		(2.33)	(2.47)	(0.74)		
(6)	Labor Income (USD PPP Per Week)	-24.35***	17.05*	-5.05*	14.31	1,613
		(8.48)	(9.54)	(2.69)		
Panel C: After COVID-19 Lockdown						
(7)	Loan Uptake (=1)	-0.17	-0.19	0.06	0.43	1,611
		(0.14)	(0.16)	(0.04)		
(8)	Hours Worked (Per Week)	9.76	-10.53	-1.31	24.13	1,610
		(8.66)	(10.04)	(2.19)		
(9)	Labor Income (USD PPP Per Week)	-3.17	0.67	-7.01	46.97	1,609
		(17.64)	(19.50)	(4.86)		

Note:
The table shows our regression results on general finances, heterogeneous by community urban/rural status. Each row is its own dependent variable. Panels A, B, and C represent the three time periods of before, during, and after Jordan's COVID-19 lockdown. Hours Worked and Labor Income are winsorized at the top 1%. The independent variables of interest are the TOT treatment indicator, which is the predicted value from a first-stage regression of treatment implementation on treatment assignment, and its interaction with an indicator for the community being urban. Monetary values are in USD PPP. The regressions also have enumerator fixed effects, month-by-year fixed effects, community-level controls (Irbid/Mafraq governorate and population quartile), and household-level controls (vulnerability-assessment quartile, shelter program, number of children, number of children plus adults, respondent gender, and respondent age). In the parentheses are robust standard errors clustered at the locality level. (*** p<0.01, ** p<0.05, * p<0.1.)

Table B.2: Heterogeneous Effects on General Finances by Vulnerability

	Treat	Treat× Vulnerable	Vulnerable	Control Mean	N
Panel A: Before COVID-19 Lockdown					
(1) Loan Uptake (=1)	-0.15*	-0.14	-0.02	0.61	1,612
	(0.08)	(0.15)	(0.05)		
(2) Hours Worked (Per Week)	8.00	-14.73*	-7.76**	30.03	1,608
	(6.65)	(8.82)	(3.08)		
(3) Labor Income (USD PPP Per Week)	-10.58	0.59	-16.90***	55.42	1,610
	(10.69)	(13.49)	(4.60)		
Panel B: During COVID-19 Lockdown					
(4) Loan Uptake (=1)	-0.23***	-0.25*	-0.01	0.78	1,612
	(0.08)	(0.14)	(0.04)		
(5) Hours Worked (Per Week)	1.05	-1.15	-0.93*	1.34	1,614
	(1.13)	(1.93)	(0.55)		
(6) Labor Income (USD PPP Per Week)	-7.83*	-3.83	-2.46	14.31	1,613
	(4.61)	(7.13)	(1.80)		
Panel C: After COVID-19 Lockdown					
(7) Loan Uptake (=1)	-0.28***	-0.10	-0.00	0.43	1,611
	(0.08)	(0.11)	(0.04)		
(8) Hours Worked (Per Week)	6.23	-12.72	-5.31**	24.13	1,610
	(5.97)	(8.71)	(2.55)		
(9) Labor Income (USD PPP Per Week)	0.75	-6.17	-9.53**	46.97	1,609
	(9.59)	(14.07)	(4.28)		

Note:
The table shows our regression results on general finances, heterogeneous by household vulnerability. Each row is its own dependent variable. Panels A, B, and C represent the three time periods of before, during, and after Jordan's COVID-19 lockdown. Hours Worked and Labor Income are winsorized at the top 1%. The independent variables of interest are the TOT treatment indicator, which is the predicted value from a first-stage regression of treatment implementation on treatment assignment, and its interaction with household vulnerability-assessment score. Monetary values are in USD PPP. The regressions also have enumerator fixed effects, month-by-year fixed effects, community-level controls (Irbid/Mafraq governorate and population quartile), and household-level controls (vulnerability-assessment quartile, shelter program, number of children, number of children plus adults, respondent gender, and respondent age). In the parentheses are robust standard errors clustered at the locality level. (*** p<0.01, ** p<0.05, * p<0.1.)