American Economic Review 2021, 111(10): 3093–3122
https://doi.org/10.1257/aer.20201062

Jobs for Sale: Corruption and Misallocation in Hiring[†]

By Jeffrey Weaver*

Corrupt government hiring is common in developing countries. This paper uses original data to document the operation and consequences of corrupt hiring in a health bureaucracy. Hires pay bribes averaging 17 months of salary, but contrary to conventional wisdom, their observable quality is comparable to counterfactual merit-based hires. Exploiting variation across jobs, I show that the consequences of corrupt allocations depend on the correlation between wealth and quality among applicants: service delivery outcomes are good for jobs where this was positive and poor when negative. In this setting, the correlation was typically positive, leading to relatively good performance of hires. (*JEL* D73, I11, J16, J24, J45, M51, O17)

The recruitment of public sector workers is an important determinant of state capacity and service delivery (Dal Bó, Finan, and Rossi 2013; Finan, Olken, and Pande 2017; Best, Hjort, and Szakonyi 2019). Previous studies have theorized that corruption in the allocation of government jobs, in which hiring decisions are made on the basis of bribes or connections, will have negative consequences for service delivery and may be one of the root causes of bureaucratic inefficiency (e.g., Wade 1982, Shleifer and Vishny 1993, Muralidharan 2015, Sukhtankar and Vaishnav 2015). Such corruption is thought to be widespread in the developing world: for example, Kristiansen and Ramli (2006) interviewed 60 Indonesian civil servants and found that all had paid a bribe to be hired, while former Russian Prime Minister Dmitry Medvedev has publicly acknowledged that most government jobs in Russia can be purchased (NewsRu 2008). However, the difficulty of collecting data on corruption in hiring has resulted in little empirical evidence on its effects.

Theoretical models provide ambiguous predictions on the impact of corruption. Some papers have argued that corruption may be economically efficient in allocating

* University of Southern California (email: jbweaver@usc.edu). Stefano DellaVigna was the coeditor for this article. I am grateful to the anonymous referees for their insightful comments. I also thank Chris Udry, Mark Rosenzweig, Nick Ryan, Ebonya Washington, Vittorio Bassi, Alex Cohen, Sabyasachi Das, Lauren Bergquist, Sharat Ganapati, Robert Gonzalez, Dean Karlan, Dan Keniston, Matt Lowe, Karthik Muralidharan, Paul Niehaus, Sam Norris, Rossa O'Keefe-O'Donovan, Paulina Oliva, Ben Olken, Matt Pecenco, Nancy Qian, Santiago Saavedra, Gabriella Santangelo, Meredith Startz, Sandip Sukhtankar, Edoardo Teso, Jaya Wen, and seminar audiences at Berkeley, Berkeley Haas, Center for Global Development, EconCon, EMCON, NBER Summer Institute, NEUDC, Ohio State, UC-San Diego, USC, and Yale for helpful comments, and am especially grateful to Scott Lee for sharing materials. All fieldwork and analysis was conducted while the author was supported by a NSF Graduate Research Fellowship. Funding for data collection was provided by a Yale Economic Growth Center/Sylff Foundation Research Grant. Yale IRB: 1408014406. All views expressed are those of the author and do not necessarily reflect the opinions of any of the funding organizations.
[†]Go to https://doi.org/10.1257/aer.20201062 to visit the article page for additional materials and author disclosure statement.

scarce goods to individuals who value them most highly (Leff 1964, Beck and Maher 1986, Lien 1986, Shimer 1999). On the other hand, corruption may allocate valuable resources such as jobs to the corrupt or wealthy rather than the socially efficient recipients (Krueger 1974, Esteban and Ray 2006). This ambiguity generates important empirical questions: can corruption lead to efficient outcomes, and if so, what factors determine whether corruption leads to efficient or inefficient outcomes?

In this paper, I collect original data from a corrupt government hiring process for jobs as supervisors of community health workers (CHWs). There are over five million CHWs globally, and their work has been shown to be an important determinant of public health outcomes (Deserranno 2019, Ashraf et al. 2020). The study area is in a rural portion of a large developing country and is served by over a thousand CHWs.[1] These CHWs had worked in their positions for an average of eight years prior to this hiring, but the supervisor position did not previously exist. For the hiring process, current CHWs were grouped into geographically based clusters of 15 to 25 CHWs, and one supervisor was hired for each cluster. Only the CHWs within a given cluster were eligible to apply for that cluster's supervisor position, a feature which will prove important for the empirical strategy. Hiring processes like this one, with a small number of desirable positions and many similarly credentialed applicants, are common in developing country bureaucracies.

The first part of the paper studies the hiring process and shows that bribes were a key determinant of hiring decisions. I collect and cross validate data on bribe payments from key informants involved in the hiring process, including hires and unsuccessful applicants. All hires made bribe payments, and these payments averaged 17 months of salary as a supervisor. Since becoming a supervisor results in a roughly 40 percent increase in salary and supervisors are rarely fired, this proved a sound investment. Consistent with the hiring mechanism resembling a winner-pay auction, most unsuccessful applicants were solicited for bribe offers, but only hires actually made payments.

Taking the bribe offers and characteristics of both hires and unsuccessful applicants, I estimate a discrete choice model to establish what factors were part of hiring decisions. Bribe size appears to be the primary determinant of who is hired, while political connections and the education of applicants play a smaller role. I examine what factors are related to bribe offers, and find that candidates offer larger bribes if they are wealthier, value the job more highly, or face a stronger pool of competitors. As expected when bribes influence hiring, wealth and willingness to pay are strongly related to being hired; applicants in wealthiest tenth of CHWs are five times as likely to be hired as those in the poorest tenth. As a result, the quality of hires will depend on what characteristics are correlated with wealth and willingness to pay.

The second section of the paper examines the observable quality of hires. Theoretical work suggests that the quality of nonmeritocratic hires will be poor (Shleifer and Vishny 1993, Esteban and Ray 2006), but there is virtually no empirical evidence on bribery-based hiring as compared to patronage-based hiring

[1] Due to human subjects concerns about respondents and data collectors being identified and retaliated against, the identity of the country was removed from the paper. Exposing corruption can be costly, with many cases of anticorruption activists losing their jobs or being murdered. This precaution is also taken by Cole and Tran (2011), who examine a firm's records of bribe payments, and de Janvry, McIntosh, and Sadoulet (2015) due to a nondisclosure agreement.

(Xu 2018; Colonnelli, Prem, and Teso 2020; Ornaghi 2020). Since only current CHWs were eligible to apply and I collect information on all CHWs, I can compare hires to the universe of individuals eligible to apply as well as those who actually applied. Hires are better than nonhires across a large set of observable quality measures, including cognitive ability and past performance as a CHW. To focus this comparison on the set of characteristics related to supervisor performance, I take a 20-month panel dataset on delivery of health services by CHWs and determine which supervisor characteristics are related to improvements by the CHWs they supervise. I then construct an index of these characteristics, labeled as the predicted "supervisor performance index" (SPI). The value of SPI among hires is significantly higher than both nonhired applicants and those who did not apply; 62 percent of hires have one of the three highest SPI values among candidates for their position.

The policy-relevant counterfactual is how the observed hires compare to those hired under purely merit-based systems, where the observed hires appear to have been selected based on a combination of bribes, connections, and quality. I predict who would have been hired under two counterfactual systems—a knowledge-based test and the rules that were supposed to be used for hiring in this context—and compare them to actual hires. Actual hires have significantly higher values of SPI than predicted hires under the knowledge-based test (p-value < 0.01), and are similar to the predicted hires under the intended hiring rules. However, there is heterogeneity in quality across actual hires, where a quarter had values of SPI in the bottom half of applicants for their position. This generates two puzzles: first, why does the average quality of corrupt hires seem to be relatively high; and second, what might cause variation in the quality of hires?

The final section of the paper addresses these questions by empirically demonstrating a mechanism that determines when corruption can lead to efficient outcomes. Public sector jobs are one of a large class of goods that bureaucrats may allocate corruptly. Some theoretical work has argued that such allocations will be efficient (Leff 1964, Beck and Maher 1986, Lien 1986), while others have argued that corrupt allocations will be inefficient (Shleifer and Vishny 1993, Esteban and Ray 2006). An alternative possibility is that the consequences of corruption are not uniformly bad or good, but vary depending on features of the institutional context. I test a model based on Banerjee, Hanna, and Mullainathan (2013) in which the efficiency consequences of corruption depend on two economic parameters: in the case of a corrupt allocation of jobs, these are the correlation of candidates' wealth and willingness to pay for the job with how they would perform in the job. This model predicts that misallocation will be lower if these correlations are more positive since the best candidates will be more likely to offer larger bribes and be hired.

This context offers an ideal setting to empirically test this model. Each supervisor position had a fixed and externally determined pool of eligible candidates: only current CHWs could apply, and they were restricted to applying to the one position that their work area falls under. This generates variation across supervisor positions in the correlation of quality with wealth and valuation within the candidate pool. Using a "leave-out" version of SPI to measure predicted quality as a supervisor for each candidate, I find substantial variation in the correlation of predicted quality and wealth among positions: it is positive and greater than 0.2 for 64 percent of positions and negative for 26 percent of positions. As a test of the model, I examine whether service delivery outcomes are better when this correlation is more positive.

Consistent with the hypothesis, hires are of a higher predicted quality and service delivery outcomes improve by more after the hiring of supervisors in clusters where the correlation is more positive. I also show that the correlation is unrelated to pre-existing cluster characteristics (e.g., average wealth or predicted quality of CHWs), pre-existing service delivery outcomes, or trends in service delivery outcomes. Instead, the relationship between service delivery outcomes and the wealth-quality correlation emerges right when supervisors are hired, consistent with these gains being attributable to the hiring of better supervisors in candidate pools with a more positive wealth-quality correlation. This explains both puzzles from the second section of the paper: the observable quality of hires is good on average due to the strong positive correlations between SPI, valuation, and wealth across all CHWs in the region. However, there is heterogeneity due to the variation in those correlations across candidate pools.

Government allocations of scarce resources such as jobs are characterized by multiple market failures. Bureaucrats typically lack incentives to make socially optimal allocations and may also lack information on how to optimally screen. This paper observes that allocation via bribes need not lead to poor outcomes, but the consequences depend on a simple set of economic parameters. These findings are relevant to other contexts in which governments allocate scarce goods and for considering where anticorruption efforts may have the largest impact.

This paper contributes to a number of distinct literatures. First, the paper is related to work on recruitment and hiring of public sector workers (e.g., Dal Bó, Finan, and Rossi 2013; Hanna and Wang 2017; Finan, Olken, and Pande 2017). Along with the contemporary work of Colonnelli, Prem, and Teso (2020); Ornaghi (2019); and Xu (2018), who study patronage-based rather than bribery-based hiring, this is one of the first papers to directly measure the consequences of nonmeritocratic hiring.[2] In contrast to those papers, other work on selection based on connections (e.g., Fisman et al. 2018), and most theoretical discussions of corruption-based hiring (Shleifer and Vishny 1993, Muralidharan 2015, Sukhtankar and Vaishnav 2015), I find that nonmeritocratic hiring does not necessarily cause an inefficient allocation of jobs.

Second, the paper speaks to broader questions of how corruption affects allocative efficiency. The concept of efficient corruption is well known theoretically (Leff 1964, Beck and Maher 1986, Lui 1985), but there is minimal empirical evidence in practice. The vast majority of the empirical literature on how corruption affects economic outcomes has found negative consequences from embezzlement (e.g., Olken 2007; Ferraz, Finan, and Moreira 2012; Muralidharan, Niehaus, and Sukhtankar 2017), extortion (e.g., Sequeira and Djankov 2014), exchange of favors (e.g., Bertrand et al. 2007, Duflo et al. 2013), and corruption in allocation of in-kind transfers (Niehaus et al. 2013).[3] This paper departs from existing work in examining

[2] There is also literature on the historical sale of offices in medieval Europe (Allen 1998), colonial governorships (Guardado 2018), and tax farming (e.g., Johnson and Koyama 2014, White 2004). Those are about the sale of offices by the state, whereas sale by unauthorized actors is more common in the contemporary context, with potentially more negative consequences.

[3] I know of only two papers finding evidence of efficient corruption, and both are due to different underlying mechanisms than this paper. Dreher and Gassebner (2013) find evidence of "greasing the wheels" (Lui 1985), while in Sukhtankar (2015), there is a reallocation of licenses to efficient firms after an initially corrupt allocation.

the heterogeneous effects of corruption and documenting the economic parameters that underlie this heterogeneity. This framework can be applied to many other types of public goods and is relevant in considering where cracking down on corruption can have the highest returns.

Third, this paper contributes to the literature documenting how corrupt markets work (Olken and Barron 2009, Burgess et al. 2012). Corruption takes many forms with different underlying economic structures, and understanding the nature of the market is relevant in the design of anticorruption policy. Existing documentation of corrupt markets comes primarily from settings such as the exchange of money for favors (e.g., Bertrand et al. 2007, Duflo et al. 2013), kickbacks in bidding for contracts (Cole and Tran 2011), and embezzlement (e.g., Olken 2007, Niehaus and Sukhtankar 2013). In the case of markets for jobs, nearly all empirical papers have focused solely on the role of patronage (Iyer and Mani 2011; Fafchamps and Labonne 2017; Colonnelli, Prem, and Teso 2020) aside from the more qualitative accounts of Wade (1982) and Wade (1985). This paper uses detailed data to understand the role of bribery in nonmeritocratic hiring, including how bribery and patronage concerns interact. I find that patronage is a secondary consideration to bribes in this context, highlighting the importance of studying factors other than patronage in understanding developing country bureaucracies.

I. Background and Data

A. *Context*

This paper studies corruption in hiring in the context of a CHW program. CHW programs extend health-care services among underserved populations by providing targeted services for which a highly skilled provider is not required, with over five million CHWs globally (Perry, Zulliger, and Rogers 2014). In contrast to doctors and nurses, CHWs are usually hired without having an extensive background in health and are trained on the tasks they will perform. The responsibilities of CHWs range from basic checkups and distribution of basic medicines to roles such as delivering babies or serving as tuberculosis treatment providers.

The data in this study are from a rural area of a large developing country. The study area contains slightly fewer than two million individuals and is served by around a thousand active CHWs. I study the hiring of approximately 70 supervisors, each of whom oversees 15–25 CHWs. Due to concerns about respondents and individuals who assisted in the data collection being identified and retaliated against, the identity of the country was removed from the paper.

In the study area, all of the CHWs are women hired from the communities that they serve and are trained by the government. The typical CHW is between 30 and 40 years of age, married (94 percent), has between 8 and 12 years of education, and has worked as a health worker for 8 years. Depending on the local geography, a CHW typically provides services to between 850 and 1,500 individuals. On average, CHWs earn slightly more than the median household income in the area and have substantial attachment to their work: 85 percent expect to remain working in the same health bureaucracy for the rest of their lives, and an average of only 1.1 percent of CHWs exit from the job annually either voluntarily or involuntarily.

This attachment is a function of their relatively high pay and near complete job security.

The primary responsibility of these health workers is services for pregnant women and children. The average CHW serves 11.6 pregnant or recently delivered women at any one time and visits them regularly. During the visits, they distribute iron supplements, provide basic antenatal care counseling and advice, and perform postnatal checkups on newborns. They often serve as frontline health workers, with over half dispensing medical advice or basic medicines at least once per week. CHWs also bring women to give birth at formal birthing facilities, and slightly fewer than half of CHWs had served as tuberculosis treatment (DOTS, i.e. Directly Observed Treatment, Short-course) providers in the previous six months. While the CHWs provide additional services, these are their most important functions.

The studied hiring process was part of a wave of hiring of health worker supervisors across the country.[4] Prior to this hiring, CHWs lacked supervisors specifically tasked with monitoring their work, and national policymakers felt that more monitoring could improve training and effort provision. Only current CHWs were eligible to apply to become a supervisor, and so by collecting information on all the CHWs, I observe the universe of potential applicants. All current CHWs were grouped into geographically based clusters of 10 to 30 CHWs, with 85 percent in clusters of 15 to 25 CHWs. One supervisor was hired to oversee the work of each cluster, and crucially, only current health workers within a given cluster were eligible to apply for its supervisor position. All current CHWs were informed about the nature of the position and pay, and slightly more than a third of CHWs applied for the job. In nearly all cases, there was substantial competition: only 1 position had only 1 applicant, while 23 percent of positions had 2 to 3 applicants, 32 percent had 4 to 5 applicants, and 44 percent had 6 or more applicants.

Hiring for jobs as supervisors was conducted by committees of local health bureaucrats. Each committee included the lead bureaucrat overseeing health programs in the region as well as more junior bureaucrats. Higher-level government officials attempted to standardize hiring by creating a system in which applicants earned points based on education, past work as a CHW, and an interview with the hiring committee. The applicant within a cluster who had the most points was supposed to be assigned the position, but Section ID shows that selection appears to have been based on bribery. Given the lack of oversight at higher levels, it was easy to claim that the hire had the most points regardless of the truth.

After the supervisors were hired, there was a six month lag before their training due to administrative delays in preparing the curriculum (see online Appendix Figure A.1 for a timeline). During this period, those selected as supervisors did not receive training or take on supervisory duties, but continued in the same role as CHWs. Their eventual training focused on the tasks that they would carry out as supervisors, including their responsibilities when visiting CHWs in their villages, how to restock CHW supplies, strategies for "supportive supervision," and how to

[4] The study area was selected due to relationships that enabled data collection, but corruption in hiring seems to have been ubiquitous. I encountered similar bribe-paying arrangements in other regions when piloting the survey instrument, and such corruption is common knowledge.

conduct monthly meetings with CHWs. After the training, they began working as supervisors.

The primary duties of supervisors are to visit the village of CHWs to monitor their work and provide feedback, rate whether CHWs' performance is satisfactory given centrally defined criteria, and provide training. Based on data collected from the CHWs, supervisors visit CHWs in their villages between once and twice a month for one to two hours. During these visits, they review the records and recent work of the CHW (96 percent of visits) and give feedback and advice (81 percent of visits). In around half of visits, they directly observe the CHW's work, such as counseling of pregnant women. CHWs generally appreciate their supervisors, with 71 percent stating that their supervisors are helpful or very helpful. Supervisors cannot fire CHWs and have little other sanctioning power, so their power over workers is based primarily on verbal reprimands and informal incentives. Although less powerful than formal incentives, monitoring without formal incentivization has been found to be effective in many developing country bureaucracies (Callen et al. 2018, Dal Bó et al. 2018, Muralidharan et al. 2021).

Supervisors are compensated on a salary basis and earn around 40 percent more than CHWs. This is the main financial incentive to become a supervisor, as there are few opportunities to extract corrupt rents. Supervisors do not control significant flows of government funds from which they can skim and have little sanctioning power over CHWs that they could use to extract bribes. As with CHWs, the job security of supervisors is high: 95 percent of supervisors expect to remain in the health bureaucracy for the rest of their lives, and after two years in the job, none had left or been fired. From discussions with current supervisors, their motivation to perform well comes from a combination of intrinsic motivation, fear of verbal reprimands, and hope that better performance will result in promotion.

B. *Data Collection*

The paper relies on two main sources of data: two rounds of survey data collected by the author and administrative government data on CHW performance (Weaver 2021). Online Appendix Figure A.1 provides a project timeline, where data collection efforts are denoted in green and other events are in yellow. The first round of survey data was collected after supervisors had been hired, but before they had started their new duties, while the second round was conducted six months after supervisors began their work. The first round focused on the work of CHWs over the preceding six months and hiring of supervisors. The second round was similar, but included questions about the performance of supervisors. During the surveys, I administered tests of health knowledge and general ability, psychometric instruments, and behavioral games measuring prosocial preferences and honesty. Respondents were compensated at 1.5 times the prevailing daily wage and could earn more based on performance in the behavioral games, but were not told that prior to arrival. All supervisors and nearly all CHWs were surveyed (see online Appendix B for details).

I also use 20 months of monthly CHW-level administrative data on delivery of important health services. In each month, I observe the delivery of ten health services by each CHW, such as number of institutional deliveries assisted, number of newborn checkups conducted, and whether they served as a tuberculosis DOTS treatment provider. The government aggregates these outcomes into a single performance measure between 0 and 100 that I focus on in the paper, but I also report results for individual services for robustness.

One concern is manipulation of the administrative data by supervisors, and so I run three tests for manipulation that I briefly describe here (see online Appendix B.2 for further details). First, I compare the administrative data to performance evaluations that I collected for each supervisor from individuals overseeing the CHW program, where supervisors could not manipulate the performance evalutions since they did not know they were being collected. Online Appendix Figure A.2 and column 1 of online Appendix Table B.2 show that these measures line up well with the administrative data. Second, I check how administrative data on institutional deliveries assisted compare to data on institutional deliveries from surveys of the CHWs and find that they match well. Third, I add an interaction term to these regressions to test whether the match between the administrative data and independent measures is weaker among supervisors whom the administrative data indicate are good quality, consistent with those supervisors manipulating the data to look like better performers. I again find no evidence of manipulation.

C. *Bribery Data*

The first object of interest is if hires paid bribes to get the job. Practically all papers on bribery rely on self-reporting (see Olken and Pande 2013 for a review), but respondents may not want to admit to illegal activity. To deal with potential underreporting, I combine information from multiple sources on bribes. First, hires were asked if they had to pay any money for the job, and if so, how much they had given.[5] Forty-seven percent of hires told surveyors how much they had paid, while 33 percent declined to answer and 18 percent claimed that they did not pay anything. Given that a refusal to answer likely indicates the respondent paid a bribe, this immediately suggests that at least 80 percent of hires paid a bribe.

As an auxiliary source of data, I collected data from other CHWs. Health workers and supervisors regularly socialize at trainings and health facilities, and after the hiring process, bribe payments were a popular topic of conversation. As a result, many CHWs had talked directly with their supervisors about the bribes that the supervisor had to pay. During the survey, the other CHWs were asked if they knew whether their supervisor had paid money for the job, how much their supervisor had paid (if anything), and where they had heard this information (e.g., supervisor, member of hiring committee). Seventy-seven percent of respondents stated that their supervisor had paid money, 19 percent said that they did not know, and only 3.8 percent claimed that their supervisor had not paid a bribe. For 97 percent of supervisors, either the supervisor herself admitted to paying a bribe, or there are at least three

[5] The exact wordings were “Many CHW supervisors told us that it was necessary to give money to become a supervisor. Did you have to give anything?” and “How much did you have to give?” In order to avoid making respondents uneasy, surveyors were instructed to not probe: if they did not give a response on the first query, the surveyor moved to the next question.

independent reports of the supervisor paying a bribe. Based on this, I conclude that bribe payments to become a supervisor were basically universal.

The second object of interest is the size of the payments. In cases where the supervisor was unwilling to tell the surveyor about her bribe payment, I estimate it based on the reports of other CHWs. Seventy percent of CHWs who said their supervisor had paid a bribe also stated that they knew how much their supervisor had paid. If one or more respondents heard about the bribe directly from their supervisor, I privilege those direct accounts and take their average as the estimated bribe for that supervisor. In the remaining cases, I take the average of all respondents who said they heard their information from a reliable source such as a member of the hiring committee. These reports are highly consistent: even for cases where the supervisor did not admit to paying a bribe, the correlation across reported bribe amounts for the same supervisor is 0.54. In all but one case where the supervisor did not tell surveyors about their bribe payment, it is possible to estimate the size of the bribe using secondary sources; in most of these cases, the secondary sources heard directly from the supervisor. As a validation, I take cases in which the supervisor reported their payment and regress this on the bribe amount estimated from the secondary sources. As seen in panel B of Figure 1, the fit is remarkably tight, with an R^2 value of 0.69, a slope of 0.83 $(\mathrm{SE} = 0.11)$, and little systematic bias in the errors.

Panel A of Figure 1 shows the distribution of bribe amounts paid by hires, where the average bribe is 17 months of supervisor salary. There are minimal opportunities for rent extraction in this context, and so hires are mostly paying for the salary increase, career advancement, and nonpecuniary benefits of the job; the roughly 40 percent salary increase compensates for the bribe within a few years at local interest rates.[6] It is notable that there is substantial dispersion in payments of hires, indicating that there is not a fixed "fee" that candidates must pay for selection.

While 17 months of salary is a substantial amount, this is significantly less than in some other contexts. Wade (1982) finds that entry-level irrigation engineers in South India pay 2 to 3 times their annual salary for jobs in less desirable locations, while more desirable postings may cost up to 14 times their annual salary. Kristiansen and Ramli (2006) observe an average payment of 2.5 years salary for jobs as Indonesian civil servants, and Jauregui (2014) finds that Indian police officers pay 3–5 years of annual salary for an entry level job.[7] Those prices reinforce that payments of the magnitude I observe are not an outlier. The relative scarcity of opportunities for rent extraction in the position I study may explain why the payments are smaller than for jobs such as police officers or irrigation engineers.

[6] Since job separation for supervisors is rare, with 95 percent of supervisors stating they anticipate remaining in the health bureaucracy for the rest of their lives, this is a secure investment. Taking the annual separation rate for CHWs (1.1 percent) as a conservative estimate of the hazard rate for separation for supervisors, the average supervisor would expect to work for 25.6 years if they work until the age of 65. All of the hired supervisors were below the age of 45 (expected tenure of 18 years at a 1.1 percent separation rate if they worked until age 65), consistent with purchasing this job being a profitable investment in expectation.

[7] Journalists have documented similarly high prices in other locations. Walsh (2014) interviews an Afghani border official who paid approximately three years of salary for their job, and Filkins (2009) finds that Afghani provincial police chiefs pay over ten years of salary for their jobs. Chinese train attendants pay around one year's salary for their jobs (Moore 2013).

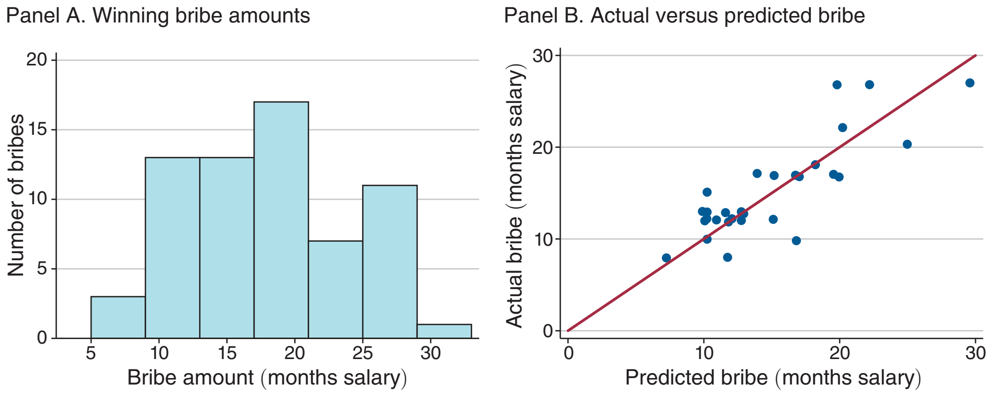

Figure 1. Bribe Payments

Notes: Panel A is the distribution of bribe payments that were made by those hired, where the payment unit is months of salary in the job that they are hired for. Since not all candidates were willing to report their bribes to the survey team, some bribe payments in panel A were estimated using reports from other sources. Panel B checks the accuracy of those secondary reports. In cases where the bribe is directly reported by the hire, the panel compares the direct reports (y-axis) to the estimated bribe for that hire from the secondary reports (x-axis). All points would fall on the superimposed 45-degree line if the secondary reports were exactly accurate, but the errors are small and only slightly biased.

Although only hires actually paid any money, I also surveyed unsuccessful applicants on whether they had been solicited for bribes and had made a bribe offer.[8] Eighty-two percent of unsuccessful applicants said that they had been solicited for a bribe offer, while 44 percent said they had made an offer for the job and how much they offered. Even for the same job, there is substantial variation in bribe offers (standard deviation of 4.74 salary months within a competition), indicating that bribes are not fixed job-specific fees that hires pay upon selection. The next section of the paper combines the bribe data to understand how these markets work.

D. *Understanding the Market for Jobs*

Due to the difficulty of collecting accurate data on illegal activities, there is only a small empirical literature on markets for jobs. Two of the only examples are Wade (1982) and Wade (1985), who qualitatively describe how Indian irrigation engineers pay bribes to get favorable postings, but those papers lack the data to more rigorously characterize the market. A number of papers have shown that political connections are related to public sector hiring decisions (e.g., Fafchamps and Labonne 2017; Jia, Kudamatsu, and Seim 2015; Colonnelli, Prem, and Teso 2020), but this section considers how patronage, bribes, and other considerations

[8] To elicit bribe offers, the question wording was “Many CHW supervisors told us that it was necessary to give some money to become a supervisor. Did you offer to give anything to become a supervisor?.” I lack secondary sources to validate this data, but a number of robustness checks suggest these data are high quality (see online Appendix C.1).

interact. Although it is a particular case, it has a wider relevance in demonstrating how markets for jobs can work.

Corrupt hiring systems can be conceptually split into three categories: (i) fully meritocratic, in which hiring decisions are made on the basis of merit and those selected are extorted for a "fee" in order to be hired; (ii) fully nonmeritocratic, in which hiring decisions are made solely based on bribes and connections; and (iii) partially meritocratic, in which hiring decisions are based on both meritocratic and nonmeritocratic elements. This section uses the data to determine which category the observed hiring process falls into.

There are a number of pieces of evidence that suggest the hiring was partially meritocratic. First, I surveyed nonapplicants on the reasons they did not apply. The two most frequently cited reasons were lacking money for a bribe (28.1 percent) and not having sufficient education (48.4 percent), suggesting those characteristics were relevant in hiring decisions. At least one nonapplicant mentioned that reason in 91 percent and 97 percent of clusters respectively, suggesting that both nonmeritocratic (bribes) and meritocratic (education) factors were relevant everywhere. A large fraction of applicants also reported having been asked for bribe offers, whereas under a fully meritocratic system, it would only be necessarily to solicit bribes from the preferred applicant.

The revealed preference of the hiring agent can be used to more rigorously test this classification. Applicants for the job differ in characteristics such as bribe offers and education. The hiring agent can be thought of as solving a discrete choice problem in which they face a pool of applicants and hire the one who maximizes their utility. More formally, the hiring agent can be modeled as having utility $u_{i,j}$ if they hire applicant i for a job j. For each job j, the hiring agent will hire the applicant who has the highest value of $u_{i,j}$ among applicants for that job. The hiring agent's utility may be based on the size of the bribe offer made by the candidate $(b_{i,j})$, political connections of the candidate $(c_{i,j})$, as well as a vector of other candidate characteristics $(\alpha_{i,j})$, where they may place different weights on each of these inputs. I model this utility as $u_{i,j} = \theta_1 b_{i,j} + \theta_2 c_{i,j} + \theta_3 \alpha_{i,j} + \epsilon_{i,j}$, where the error term $\epsilon_{i,j}$ represents unobserved or idiosyncratic preferences for applicant i. In a fully nonmeritocratic system, the hiring agent's utility depends on applicant bribe offers and connections, and so $\theta_1 \neq 0$, $\theta_2 \neq 0$, and $\theta_3 = \vec{0}$. In a fully meritocratic system, selection is based on personal characteristics related to merit, and so $\theta_1 = \theta_2 = 0$ but $\theta_3 \neq \vec{0}$. In a partially meritocratic system, all of the weights θ_1, θ_2, and θ_3 are nonzero. I estimate $\hat{\theta}_1$, $\hat{\theta}_2$, and $\hat{\theta}_3$ empirically, and categorize the hiring process accordingly as fully meritocratic, partially meritocratic, or fully nonmeritocratic.

I focus on bribe value, use of political connections, and education as potentially relevant predictors of hiring since those were the characteristics mentioned in the aforementioned survey of nonapplicants. Following a large discrete choice literature, I assume that the error term $\epsilon_{i,j}$ has a type-1 extreme value distribution (Train 2009), which implies that if these are the determinants of hiring, the probability of the hire i being the one hired in job j will be equal to

$$p_j = \frac{e^{\theta_1 bribe_{i,j} + \theta_2 education_{i,j} + \theta_3 connection_{i,j}}}{\sum_{i' \in I_j} e^{\theta_1 bribe_{i',j} + \theta_2 education_{i',j} + \theta_3 connection_{i',j}}},$$

Table 1—Hiring Decisions

	All (1)	Primary reports (2)	Supervisor imputed (2)	
Panel A. Discrete choice estimates of hiring determinants				
Bribe amount	1.00	0.95	0.66	
	(0.10)	(0.21)	(0.12)	
Political connections	4.68	3.72	4.29	
	(1.15)	(1.81)	(1.17)	
Education	1.64	0.60	1.50	
	(0.27)	(0.44)	(0.27)	
Observations	191	89	191	
	Bribe offer (1)	Bribe offer (2)	Hired (3)	Hired (4)
Panel B. Wealth and hiring				
Wealth	4.876	4.543	0.239	0.322
	(1.746)	(1.786)	(0.066)	(0.082)
Valuation proxy	9.464	11.212		
	(2.311)	(2.401)		
Average wealth bin	0.882			
	(0.383)			
Total CHWs	0.363			
	(0.120)			
Intercept			0.060	
			(0.038)	
Cluster fixed effects	No	Yes	No	Yes
Observations	191	179	339	338

Notes: This table evaluates factors related to hiring decisions. Panel A reports the discrete choice estimates of factors related to hiring decisions. The coefficient values measure the estimated increase in utility of the hiring agent (in salary months of bribe) from hiring a candidate with one unit higher value of the relevant independent variable. Column 1 includes all clusters. Column 2 restricts to clusters in which the supervisor told surveyors their bribe payment, while column 3 imputes bribe offers for supervisors who did not report their bribe offers to surveyors. The imputation is based on their observable characteristics and characteristics of their cluster, taking the statistically significant predictors of bribes from panel B. Coefficients are estimated via maximum likelihood. Panel B investigates the relationship between wealth and hiring decisions. Columns 1 and 2 examine the correlates of bribe offers. Wealth is the percentile rank of the individual in the wealth distribution of CHWs. Average wealth bin splits the clusters into five quintiles based on average wealth of CHWs in the cluster, so the coefficient corresponds to the average increase in bribe value moving up one quintile in the wealth distribution. Total CHWs is the total number of CHWs in the cluster. Columns 3 and 4 show that wealthier individuals are much more likely to be hired as supervisors. Columns 2 and 4 include cluster fixed effects. All standard errors in panel B are clustered at the supervisor cluster level.

where $i' \in I_j$ represents all of the applicants for job j, and the connections variable is equal to one if the applicant reported using any type of connection to get the job (zero otherwise). Since there is one hire selected from each cluster, the overall likelihood expression across all competitions is $\prod_j p_j$. I use maximum likelihood to estimate the coefficient values $\hat{\theta}_1$, $\hat{\theta}_2$, and $\hat{\theta}_3$ that maximize the overall likelihood expression, i.e., use the revealed preference of the hiring agent over applicants to infer the weights they place on bribe/education/connections.[9]

[9]The exact mechanism used to elicit bribe offers does not matter since for those making the hiring decisions, the relevant bribe is the highest offer that the applicant has made; in a sealed bid auction, this is submitted secretly, whereas in an ascending price auction, this is their last offer.

The first column of panel A of Table 1 displays the coefficient estimates, where the relative values indicate how the hiring agent valued each factor. Since the coefficient is 1.00 on bribes and 1.64 on education, this implies that selection agent derives the same utility from a candidate having an additional year of education as from that candidate offering an additional 1.64 salary months of bribe. This preference may explain why few hires had very low levels of education and why the most common reason for not applying was lack of education. Political connections are also highly valued, where a connection is valued as equivalent to 4.68 salary months of bribe. This estimates the average effect of a connection, but in practice, the effect will vary depending on the power and strength of that connection; a connection to a high-level politician presumably carries a value of more than 4.68 salary months, while a connection to a lower level politician may have less value. Given that both nonmeritocratic and meritocratic factors are statistically significant determinants of hiring, I conclude this was a partially meritocratic process. The point estimates should be viewed as suggestive since there may be other characteristics that are relevant in hiring. However, as discussed in more detail in online Appendix C.1, that would still not change the conclusion of a partially meritocratic process.[10]

In some clusters, the hire's bribe offer was based on reports from other CHWs rather than the hire. Column 2 of panel A reruns the estimation with only clusters where the hire directly reported her bribe. The coefficients on bribes and political connections are statistically significant and similar in size, but the coefficient on education is no longer statistically significant, likely due to the much smaller number of clusters. Column 3 instead imputes a bribe for supervisors who did not report their bribe based on observable characteristics such as wealth and valuation that are predictive of bribe offers (panel B). With this fuller sample, results are quite similar to column 1. To get a sense of which of bribes, education and connections is most important, I take the estimated coefficients, estimate the difference in implied utility between the hire and runner-up, and examine the fraction of the difference that is due to each factor. Differences in bribe offers account for two-thirds of the difference, while educational differences account for 23 percent and differences in connections for 11 percent. The applicant who offers the largest bribe is also selected 76 percent of the time. Online Appendix C.1 discusses further robustness checks related to missing data, omitted variables, or other factors that could explain hiring.

Since bribes are important for determining hires, the quality of hires will depend on the factors that determine the magnitude of bribe offers. Columns 1 and 2 of panel B of Table 1 examine how bribe offers are related to applicant and cluster-level characteristics. Wealth and valuation are strongly related to bribe offers, where wealth is percentile rank on a wealth index and valuation proxy is based on whether they would be willing to apply under different probabilities of being selected. If an individual is willing to apply at lower probabilities, this implies they place a higher value on the job; for ease of interpretation, the probabilities are flipped so valuation proxy is equal to one minus the lowest probability at which they are willing

[10] Even if there were an unobserved meritocratic characteristic on which hiring decisions are made and this were correlated with education, that would still be hiring based on a meritocratic factor. And even if bribe amounts were partially correlated with unobserved quality measures, the correlation would have to be implausibly high for fully meritocratic selection to hold.

to apply (see online Appendix C.1 for details). The wealthiest CHWs offer bribes averaging nearly five months of salary more than the poorest CHWs ($p = 0.007$). Although the coefficient on valuation proxy does not have as clear of an economic interpretation, bribe offers are clearly related to willingness to pay ($p < 0.001$). Applicants also respond to competitive pressures by offering larger bribes when they are in wealthier clusters ($p = 0.024$) and clusters with a larger number of CHWs ($p = 0.003$). These results are quite consistent with applicants using bribe offers to compete for the job. They are less consistent with alternative explanations such as hires being selected based solely on skill and then paying a fee for selection; in that case, the size of bribe payments should not depend on the wealth of the competition.[11]

Columns 3 and 4 of panel B regress whether an individual is hired as supervisor on their wealth and find that the wealthiest applicants are more than five times more likely to be hired than the poorest ($p < 0.001$ in both columns).[12] The degree to which bribes are related to wealth helps explain why wealthy individuals are much more likely to be hired as supervisors. This suggests that the quality of supervisors will depend heavily on what applicant characteristics are correlated with wealth—if wealthier individuals also tend to be better qualified for the job, this could lead to the selection of good quality hires.

II. Assessing the Quality of Observed Hires

Existing literature suggests that corruption in hiring processes will result in poor quality hires (Shleifer and Vishny 1993, Shleifer 2004, Muralidharan 2015, Sukhtankar and Vaishnav 2015). As a first pass at evaluating this claim, I compare the characteristics of observed hires, for whom bribes and connections appear to have been part of the hiring decision, to others who might have been hired. Since only current CHWs are eligible to apply, I observe the universe of potential applicants and so can compare not only to unsuccessful applicants but also to those who chose not to apply. That is potentially important in this case: high quality individuals might have chosen not to apply under a corrupt process, but may have applied under a merit-based process.

Table 2 tests for differences between hires, unsuccessful applicants, and nonapplicants on a range of variables. These are divided into three categories—hard skills (e.g., education, test scores), soft skills (e.g., honesty, prosociality), and past performance as a CHW. There are two main patterns. First, applicants are of a higher observable quality than nonapplicants across all three categories (joint p-value < 0.001 for hard skills and past performance; joint p-value $= 0.059$ for soft skills). Second, those hired are of a higher quality than unsuccessful applicants in both hard skills (joint p-value < 0.001) and past performance as a CHW (joint p-value $= 0.023$), while the difference is not statistically significant for soft skills

[11] Online Appendix C.1 further explores the correlates of bribe offers and provides robustness checks related to this analysis.

[12] Columns 3 and 4 of online Appendix Table A.1 rerun these specifications including nonapplicants. The coefficients are scaled down since the unconditional probability of selection is lower in the full sample. However, the ratio of the slope to the intercept implies an even more extreme selection on wealth, reflecting wealthier individuals being more likely to apply.

Table 2—Characteristics of Nonapplicants, Unsuccessful Applicants, and Successful Applicants

	Nonapplicant mean (1)	Unsuccessful applicant mean (2)	Successful applicant mean (3)	Col 1 vs. col 2 p-value (4)	Col 1 vs. col 3 p-value (5)	Col 2 vs. col 3 p-value (6)
Panel A. Hard skills						
Health knowledge	0.46	0.57	0.63	0.001	0.001	0.082
Short-term memory	4.32	4.56	4.88	0.000	0.000	0.023
Raven's score	2.49	3.66	4.61	0.000	0.000	0.185
Education	8.83	11.48	12.68	0.000	0.000	0.000
Reading skill	4.39	5.46	5.85	0.000	0.000	0.041
Writing skill	4.10	5.28	5.80	0.000	0.000	0.019
Joint p-value				0.000	0.000	0.000
Panel B. Soft skills						
Prosocial (donation)	0.23	0.22	0.24	0.289	0.504	0.497
Honesty (dice game)	0.49	0.46	0.46	0.169	0.092	0.382
Public service motivation	3.76	3.86	3.94	0.092	0.063	0.302
Intrinsic motivation	3.34	3.56	3.64	0.009	0.008	0.016
Extrinsic motivation	2.92	2.93	2.92	0.012	0.004	0.878
Extroversion	3.68	3.70	3.77	0.576	1.000	0.689
Joint p-value				0.059	0.039	0.840
Panel C. Past performance as a community health worker						
DOTS patients	1.14	1.53	1.79	0.026	0.111	0.780
CHW work hours (weekly)	13.66	15.83	19.09	0.008	0.002	0.162
Client visit tasks	3.75	4.14	4.27	0.002	0.008	0.488
Medical advice frequency	4.92	5.07	5.48	0.771	0.144	0.542
Contraception work	2.04	2.35	2.92	0.169	0.063	0.784
Immunization work	19.81	23.17	27.18	0.000	0.002	0.082
Deliveries assisted	3.30	3.51	3.78	0.576	0.672	0.174
Joint p-value				0.001	0.000	0.023

Notes: This table compares hires, the median nonapplicants and the median applicant from each competition. It uses a Wilcoxan signed-rank test to sign selection across the margins of applicants and being hired for a large set of variables that are plausibly correlated with performance in the job. The first three columns give the overall mean of the three groupings of candidates, while the last three give the p-value from a Wilcoxan signed rank test comparing the groups: column 4 compares nonapplicants to unsuccessful applicants, column 5 compares nonapplicants to hires, and column 6 compares unsuccessful applicants and hires. Health knowledge refers to the percent of questions answered correctly on a test of health knowledge administered during the survey. Short-term memory is the maximum number of digits successfully recalled during a digit-span memory test. Raven's score refers to the number of correct answers (out of 12) on the short-form Raven's matrices test. Education is years of education. Reading and writing skills are scores out of six on a test of reading and writing ability. Prosocial (donation) is the fraction donated in a dictator game to an orphanage, while honesty is the number of dice rolls other than fives and sixes reported on a modified Hanna and Wang (2017) honesty task. Public service motivation, intrinsic motivation, extrinsic motivation, and extroversion are on a five point scale, with five as the highest score. DOTS patients is the number of DOTS patients served by the CHW. CHW work hours are hours per week worked as a CHW in the baseline survey, while client visit tasks are the number of tasks carried out when visiting clients (e.g., providing information about pregnancies and conducting check-ups). Medical consultation frequency is how often they are consulted for medical advice, where higher values indicate more frequent consultations. Contraception and immunization work indicates the number of households assisted in the last month. Joint p-values are based on the joint test of Kling, Liebman, and Katz (2007) across all of the variables in the category.

(joint p-value $= 0.840$). The presence of corruption does not appear to preclude high quality candidates from applying or being hired.[13]

However, many of these characteristics may not actually be related to performance as a supervisor. I next consider comparisons based on the set of supervisor

[13] Online Appendix C.3 finds similar patterns using data collected prior to the hiring of supervisors and addresses concerns related to the timing of data collection.

Table 3—Post-LASSO Regression on CHW Performance

	Functionality score
Education	0.11
	(0.044)
Raven's score	0.063
	(0.026)
Writing skill	0.20
	(0.077)
Reading skill	0.43
	(0.202)
Extroversion	0.30
	(0.101)
Observations	917
Mean	0.058

Notes: This table takes supervisor characteristics selected by least absolute shrinkage and selection operator (LASSO) and regresses them on the average change in functionality score for the CHW over the 20 months of administrative data. Standard errors are clustered at the supervisor level.

characteristics related to performance. As a supervisor, the key performance metric is service delivery by the CHWs they oversee. I thus investigate which supervisor characteristics are related to improvement in delivery of health services by their CHWs. I use monthly administrative data on the delivery of ten health services by each CHW (e.g., institutional deliveries assisted, tuberculosis patients treated), where the first month of data is from after the supervisor has been in place for two months, and the panel covers the following 20 months. The government aggregates these ten services into a single "functionality score" between 0 (worst) and 100 (best), which I focus on as a comprehensive measure of performance. For each CHW i under a supervisor j, I calculate Δy_{ij}, the average monthly change of CHW i on the functionality score over the 20 months of data. A positive value of Δy_{ij} indicates that the CHW was improving, while a negative value indicates performance declines.

I then run a differences regression, regressing Δy_{ij} on characteristics v of each supervisor j in equation (1) in order to determine which supervisor characteristics are related to changes in delivery of health services.[14] Note, β_v is the average monthly change in the functionality score as a function of supervisor characteristic v: a positive value indicates that CHWs under a supervisor with a higher value of v improved over time relative to CHWs under a supervisor with a lower value of v. For example, if $\beta_{education}$ were equal to 0.05, this would imply that the functionality score increased by an average of 0.05 more points per month for CHWs under a supervisor with 12 years of education relative to CHWs under a supervisor with 11 years of education:

$$\Delta y_{ij} \;=\; \alpha + \sum_{v} \beta_v X_{vj} + \epsilon_{ij}. \tag{1}$$

[14] Since the data begin after the supervisors were in place, I am unable to run a differences-in-difference regression.

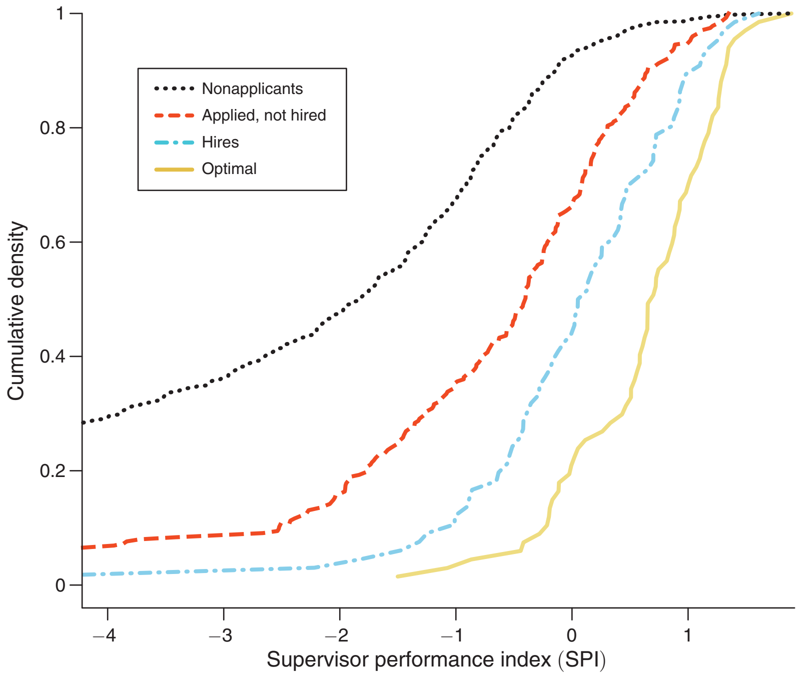

FIGURE 2. SELECTION ON PREDICTED SUPERVISOR PERFORMANCE INDEX

Notes: This figure plots the cumulative densities of predicted SPI for nonapplicants, unsuccessful applicants, those hired, and the predicted optimal set of hires. Comparisons between the groups show how the application and hiring decisions select on quality. The leftmost line (lowest values of SPI) is for individuals who did not apply for the job. The second leftmost line is for the set of individuals who applied for the job but were not hired. The third line is SPI for the set of individuals who were hired in the job. The final line takes the individual with the highest value of SPI from each of the competitions, i.e., the "predicted optimal" set of hires. All differences between groups are statistically significant at the 1 percent level in a Kolomogorov-Smirnov test for equality of distributions.

I use LASSO for variable selection (Belloni, Chernozhukov, and Hansen 2014) (see online Appendix Table B.1 for the full list of variables). The variables selected are (i) score on the Raven's Progressive Matrices, a test of problem-solving ability; (ii) years of education; (iii) scores on tests of writing and reading abilities administered by surveyors; and (iv) extroversion, as measured by the Big Five personality index. In order to scale the relative contributions of each factor, I use the estimated coefficients β_v from the post-LASSO regression in Table 3 to aggregate the variables into an index. This index is denoted as the predicted SPI for the remainder of the paper.

Figure 2 plots the cumulative density function of SPI for (i) nonapplicants, (ii) unsuccessful applicants, (iii) hires, and (iv) the predicted first best case, i.e., if the candidate with the highest value of SPI eligible for each job were selected. The index has been normalized so that higher values of SPI are positively associated with performance as a supervisor, and the mean and standard deviation of SPI among hires are equal to zero and one respectively. There is a strong positive selection on both the application and hiring margins. The median value of SPI for the nonapplicants is -1.83 and for unsuccessful applicants is -0.40, where the SPI distribution of unsuccessful applicants first-order stochastically dominates the distribution of nonapplicants (Kolmogorov-Smirnov p-value < 0.001). Despite the presence of bribes, hires are also of a much higher predicted quality than unsuccessful applicants, where the SPI distribution of hires first-order stochastically dominates that of unsuccessful applicants (Kolmogorov-Smirnov p-value $= 0.001$). As demonstrated by the lack of mass in the upper tail of the SPI distribution for nonapplicants, the presence of corruption did not deter individuals predicted to be high quality from applying.

The distribution of SPI among actual hires is relatively close to the distribution of SPI among predicted first-best hires, where one of the top three candidates was selected in 62 percent of cases.[15] However, although outcomes appear to have been relatively good on average, there are some important differences between the actual hires and predicted first best case. In over a quarter of cases, a candidate was selected whose predicted quality (SPI) is in the bottom half of applicants for their position, demonstrating the potential for corrupt hiring practices to have heterogeneous impacts.

A limitation of using SPI to assess quality is that it uses collected variables to predict performance as supervisor. This misses unmeasured characteristics that are uncorrelated with the collected predictors and may be related to performance as a supervisor. As a result, an individual may have observed characteristics that suggest they would be a poor supervisor (e.g., low level of education), but unmeasured characteristics that would cause them to perform well at the job (e.g., strong leadership skills). To get a sense of the significance of these unmeasured characteristics, online Appendix B.3 tests whether alternative measures of the quality of the hired supervisors (performance evaluations and process measures of performance) provide additional information on supervisor quality beyond that captured by SPI, as measured by effects on service delivery. I find that while these alternative measures are strongly related to service delivery outcomes, they have limited additional explanatory power beyond that provided by the SPI index. This suggests that even though there are certainly characteristics related to supervisor quality aside from that captured by SPI, SPI captures a meaningful fraction of the variation in supervisor quality.[16]

A. *Comparisons with Counterfactual Hires under Rule-Based Selection Policies*

The observed hiring process was based at least partially on corrupt factors, but it selected hires who appear to be of a good quality. However, other hiring systems might perform even better. This section compares actual hires to individuals

[15] The likelihood of such extreme selection on SPI happening by chance is extremely low. As a test, I rank each applicant within their cluster by their value of SPI and calculate the sum of the within-cluster ranks for hires. I then run 10,000 simulations in which I take one randomly selected applicant from each cluster and calculate the sum of SPI ranks for those "hires." None of the simulations have a sum of SPI ranks as low as observed in the data, meaning the simulated p-value is <0.001.

[16] Online Appendix B.3 tests for a number of other potential issues with using SPI, such as if the relationship between these characteristics and service delivery is due to how supervisors are matched to CHWs. Using data on deliveries from prior to the supervisors beginning in their role, I find no evidence of more positive pre-trends in clusters overseen by high SPI supervisors, suggesting that improvements reflect the supervisor rather than matching.

predicted to be hired under two alternative merit-based hiring systems. The first counterfactual hiring system is a test of health knowledge, as knowledge-based standardized testing is a common hiring tool. The second is the hiring procedure that was supposed to be used: the government created a formula that assigned points to applicants based primarily on their past performance as a CHW, and the applicant with the most points was supposed to be hired.

The predicted counterfactual hire under each system is the applicant who has the highest test score or number of points respectively. One complication with predicting hires is that the observed applicants may differ from applicants under merit-based systems. Fortunately, I observe the universe of potential applicants (as only current CHWs can apply), including those who did not apply in the actual hiring process but might have under alternative hiring procedures. Since counterfactual application behavior is not directly observable, I consider three different assumptions about counterfactual application behavior, predict the counterfactual hires under each assumption, and examine how predicted quality varies across assumptions.

The first assumption is that the set of counterfactual applicants is the exact same as the observed applicants. The second assumption is an opposing extreme assumption: all CHWs who want the job would apply for it.[17] The third assumption is that each candidate has a individual-specific cutoff probability where they will apply if their probability of being hired is above the cutoff and will not if it is lower than the cutoff. Online Appendix C.2 estimates an empirical industrial organization-style endogenous entry model to implement this: I use survey data to estimate individual-specific cutoffs, estimate the probability of being hired under each counterfactual, and combine those two pieces to determine which candidates would apply.[18] All three assumptions produce similar results.[19]

I compare predicted counterfactual hires to actual hires using SPI, an index of the observable characteristics related to performance as a supervisor. Since the scale of SPI is not intuitive, I translate into a more intuitive measure that I call the "percent of first-best SPI." For a given cluster, this is equal to the SPI of the hire minus the lowest SPI of a candidate in their cluster, divided by the difference between maximum and minimum SPI of potential candidates in that cluster. This equals one if the hire has the highest value SPI in their cluster and zero if the lowest SPI candidate is selected. Consistent with the graph of SPI in Figure 2, the actual hires are at 90.6 percent of the predicted first-best case. Section ID provides evidence that hiring decisions were

[17] This is based on survey questions: CHWs are considered to "want" the job if they would have chosen to apply if they were guaranteed to be hired without paying money.

[18] As an example of this method, consider the example of one particular CHW in the data, whom I denote CHW A. The survey indicates CHW A would be willing to apply as long as her probability of being hired upon applying is above 0.25. I observe that she scored a 96 percent on the health knowledge test and there are 15 other CHWs in her cluster eligible to apply. Based on her position in the overall distribution of test scores, CHW A has an estimated probability of 0.71 of being hired if she applies. Since that is above 0.25, she is predicted to apply for the job. I follow the same procedure for the other individuals in her cluster and determine that three others would apply (CHWs B, E, and G, with test scores of 83 percent, 85 percent, and 91 percent). Since CHW A's test score of 96 percent is higher than CHWs B, E, and G, she is the predicted counterfactual hire.

[19] This is because nearly all candidates who are likely to be hired under the counterfactuals (e.g., have high test scores) either do not want the job or highly value it. Those who do not want the job will never apply under any of the three assumptions. Those who have a high valuation are willing to apply even if their probability of selection is low. Therefore the predicted hires are similar regardless of whether I assume the same set of applicants, all who want the job apply, or account for the probability of selection in the endogenous entry model.

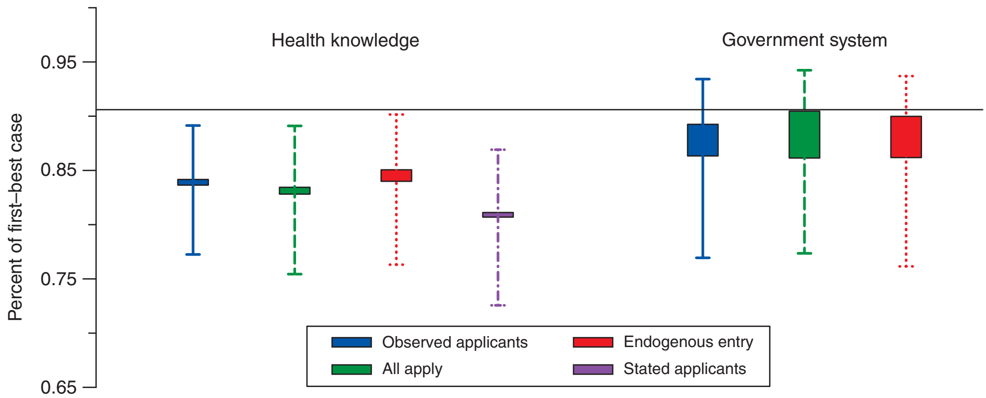

Figure 3. Comparison of SPI across Counterfactuals

Notes: This figure plots the predicted outcomes under two counterfactual hiring systems: a test of health knowledge and the system the government intended to be used for hiring. The percent of first-best SPI for a particular position is equal to the SPI of the hire under that system minus the lowest possible SPI of a potential candidate for that position (including those who did not apply), divided by the difference between maximum and minimum SPI of potential candidates. This figure plots the percent of first-best SPI for the predicted counterfactual hire under different assumptions about application behavior: (i) the set of applicants are the same as under the actual hiring system; (ii) all that want the job apply; (iii) counterfactual hires are predicted using the endogenous entry model in online Appendix C.2; and (iv) respondents' stated application behavior if hiring were based on a test of health knowledge. The figure also plots 95 percent confidence intervals, using bootstrapping to construct the confidence intervals.

based on a combination of bribes, connections, and education, so this value reflects selection on all of those factors as well as possibly other unmeasured factors.

The first counterfactual is a standardized test of health knowledge. Governments frequently use standardized testing to hire civil servants, but tests will only select good candidates to the extent that test scores are related to future job performance. This test was administered during the survey, and consisted of 30 questions taken from CHW training manuals. In cases where multiple candidates are tied for the top score in their cluster, I take the lowest and highest values of SPI among tied candidates to generate bounds on the value of SPI for predicted counterfactual hires in that cluster. The left side of Figure 3 plots these bounds for the health knowledge counterfactuals under each assumption about application behavior. Under all three assumptions, the lower/upper bounds for the counterfactual estimates fall between 82.3 percent and 84.2 percent of the first best case.[20] I also include bootstrapped 95 percent confidence intervals, where the top bar of the confidence intervals is the upper bound on the confidence interval for the upper bound estimate of SPI and the lower bar is the lower bound of the confidence interval for the lower bound. Even the upper bound of predicted hire quality is a statistically significant 6 percentage points lower than the value of SPI under the actual corrupt system (p-value < 0.01 under all assumptions).[21]

[20] I also directly surveyed CHWs on whether they would apply if selection were based on a health knowledge test (without any corruption). When these responses are used to determine the set of applicants, the resulting hires are at 80.7–81.1 percent of first-best SPI (fourth bar in Figure 3).

[21] p-values are generated using a cluster bootstrapped paired t-test of difference in means in order to account for SPI being a constructed variable (Efron and Tibshirani 1994). For the actual hires and counterfactual hires, I generate demeaned versions of each of the variables that are part of the SPI index by subtracting out the within-group

The second counterfactual examines who would have been hired under the intended government procedure discussed in Section IA, which is the closest to the true counterfactual. Although it was not possible to access the points as allegedly tallied by the hiring committee, I can approximate 65 percent of them using administrative data and survey responses since almost all of the points are based on performance as a CHW; the remaining 35 percent are from an interview with the hiring committee. Under all three application assumptions, the percent of first-best SPI under the intended government method is bounded between 86.1 percent and 90.5 percent, slightly lower than that of actual hires.[22] Note that although the *average* quality is similar between the intended government procedure and actual hiring procedure, only a third of predicted hires are the same as actual hires. While I can reject equivalence of the actual hires and the lower bound of hires under the government method (p-value < 0.032 in all cases), I cannot reject equivalence with the upper bound of government hires. This suggests that the observable quality of hires may not have been any better under the uncorrupted procedure.

To get a sense of how differences in SPI may be related to individual service delivery outcomes, online Appendix Table A.2 examines the cross-sectional relationship between supervisor SPI and delivery of four services (tuberculosis treatment, assisting institutional deliveries, newborn care, and nutritional counseling of pregnant women and new mothers). The dependent variable is the average monthly change in each of the services over the 20 months of administrative data for each CHW. Although these estimates are not causal, they give a sense of the economic magnitude of the relationships described using SPI. For example, these estimates imply that after 20 months, a health worker's probability of serving as a tuberculosis treatment provider was 3.5 percentage points higher if they were under a supervisor who was one standard deviation above the mean SPI as compared those under the mean supervisor. As a back-of-the-envelope calculation, if one were to interpret coefficients from table A.2 as causal, they would imply that relative to hires under the health knowledge test, the actual hires are associated with approximately 8 percent more tuberculosis patients being cared for and 2.5 percent more institutional deliveries.[23]

There are a number of limitations to this analysis. In this setting, candidates were already prescreened to be CHWs. Even though there are sizable differences between hires across the counterfactuals, the variance in quality of prospective hires may be lower than in settings with fresh candidates. The analysis also focuses on the static allocational effects of corruption, but dynamic effects could be important: if an individual expects to be able to purchase a job, this could

mean from each group (i.e., actual and counterfactual hires) and adding the overall mean. I redraw samples with replacement at the cluster level, re-estimate SPI in each sample using the new variables, and save the t-statistic from a paired t-test of the differences between actual and counterfactual hires in the sample with the re-estimated SPI. The p-value is based on a comparison to this bootstrapped t-statistic distribution.

[22] One concern is that education is one of the factors going into the points calculation but is also part of SPI, generating a mechanical positive relationship. As a robustness check, I recalculate the government points with education excluded and predict counterfactual hires under these points. This somewhat reduces the SPI of the predicted counterfactual hires, which is between 83.2 and 87.9 percent under any of the assumptions on application behavior.

[23] Online Appendix C.2 translates this into value of a statistical life and finds that the value of hiring better supervisors far exceeds the value of bribe payments, suggesting that allocation of positions is the most important criterion to examine in evaluating counterfactual systems.

reduce human capital investment. In the setting of this paper, the creation of the position of supervisor was a surprise, and so dynamic effects are not possible. To the best of my knowledge, there is no empirical evidence on dynamic effects in other settings, but there are reasons to think that they would be small. If public sector jobs are scarce, it would be risky to underinvest in human capital formation, and if there are complementarities between meritocratic and nonmeritocratic considerations (Jia, Kudamatsu, and Seim 2015), the incentive to invest in human capital is still strong.

There may also be other costs of corrupt hiring, such as decreasing confidence in institutions or degrading social norms. It is also unambiguously bad for less wealthy potential hires who are unable to compete based on bribes. Yet while bribery-based hiring is regressive, merit-based hiring systems also tend to be regressive; the predicted counterfactual hires under the intended government system are nearly as wealthy as the hires in the corrupt system. Furthermore, the degradation in norms from a corrupt hiring process is likely also true of other forms of corruption. The policy takeaway is thus not to ignore corruption, but rather to identify and direct resources to combating the most costly types given the difficulty of ending all corruption in the short run. The next section of the paper provides evidence for a framework to understand when corruption in hiring will have the most distortionary consequences.

III. Understanding Efficient Corruption

A broad question in economic development is how corruption affects economic outcomes. Although most empirical papers have found that corruption has harmful effects, the theoretical literature has highlighted ways that corruption can lead to efficient outcomes. For example, suppose a government is allocating a limited number of "slots" among applicants (e.g., licenses, procurement contracts, jobs), as in this paper. Some theories suggest that if the socially optimal beneficiary of the slot has the highest willingness to pay, then under a corrupt allocation system, they would offer the largest bribe and receive the good (Leff 1964, Beck and Maher 1986, Lien 1986). On the other hand, even if the socially optimal beneficiary has the highest willingness to pay, credit constraints may lead to an inefficient allocation towards the wealthy rather than to the optimal beneficiary (Esteban and Ray 2006). Finally, the consequences of corruption may be heterogeneous: corrupt allocations may be relatively efficient in cases where the socially optimal beneficiaries are wealthier and value the slot most highly, but may be inefficient in cases where individuals' wealth and valuation of the slot are negatively correlated with the extent to which they are the socially optimal beneficiary (Banerjee, Hanna, and Mullainathan 2013).

In this section, I empirically test how key economic parameters determine the extent of misallocation under corruption. These parameters can be characterized within a simple model in which applicants apply for jobs allocated by a corrupt selection agent. The key features of this model are based on the model in Banerjee, Hanna, and Mullainathan (2013) and are described visually in online Appendix Figure A.3. Applicants are characterized by their quality, wealth (w), and valuation of the job. There are two quality types—high quality (H) and low quality (L)—with measure one of each type. The wealth of low quality applicants is distributed uniformly between $[0,1]$, while that of high quality applicants is distributed uniformly over $[\underline{w}, 1+\underline{w}]$, where $-1 \leq \underline{w} \leq 1$. For simplicity, a potential applicant's valuation of the job takes on two values, high (h) and low (l), where $h > 1+\underline{w}$ and $l < \min(0, \underline{w})$; in other words, high valuation bidders are credit constrained, while low valuation bidders do not want the job. I also assume the set of applicants with high valuations is measure one, where a fraction p of high valuation applicants are high quality and $(1-p)$ are low quality. To highlight the relative importance of each factor separately, I will assume that valuation and wealth are independent, although this is easily relaxed.

Suppose that the number of jobs to be assigned is measure $S \ll 1$, and the agent uses a multiunit, second-price auction to assign jobs to the highest bidders; this simplifies the solution, but the allocation outcome will be the same under other standard winner-pay formats. High valuation candidates are credit constrained and cannot place bids greater than their wealth, meaning that their bids are equal to the maximum of zero or their wealth level. Candidates with low valuations can be ignored since they do not bid.

Under values of $\underline{w}$, S, and p such that at least some low quality and some high quality candidates are selected, the proportion of hires who are high quality is equal to $p + (p\underline{w}/S)(1-p)$.[24] The two parameters determining the relatively efficiency of outcomes are $\underline{w}$ and p, where both have one-to-one mappings to parameters that can be estimated in data: the correlation of valuation and quality is equal to $p - 0.5$, while the correlation of wealth and quality is equal to $\underline{w}/\left(\sqrt{(1/3)+\underline{w}^2}\right)$. Since p is bounded between zero and one, the model predicts that the efficiency of the allocation (i.e., proportion of hires who are high quality) increases as these correlations increase.

This model suggests that the effect of corruption is heterogeneous: corruption leads to relatively efficient outcomes in situations where applicant wealth and valuation are positively correlated with their quality (panel A of online Appendix Figure A.3), but leads to relatively inefficient outcomes in cases where these are poorly aligned (panel B of online Appendix Figure A.3). The institutional context will determine the nature of these correlations. For example, the wealth-quality correlation may be positive if wealthier individuals have higher levels of human capital that improve productivity, but may be negative if the wealthy are biased against poorer households in delivery of services. If those with a high public service motivation more highly value a job providing services to their community, then valuation may be positively correlated with quality. On the other hand, the valuation-quality correlation may be negative for jobs with higher scope for extracting corrupt rents, as more corrupt individuals can extract more (Shleifer and Vishny 1993). The next section applies data from this context to test the hypothesis that the correlations of wealth, valuation, and quality determine when corruption is allocatively efficient.[25]

[24] In equilibrium, S is equal to the sum of the high valuation and low quality candidates who are selected. Note that if x is the lowest wealth of a candidate who is selected, then $S = \int_x^{1+\underline{w}} p\,\partial y + \int_x^1 (1-p)\,\partial y$. Solving for x, the lowest wealth candidates selected for the job will have wealth equal to $(1 - S + p\underline{w})$. Note, $p(S + \underline{w} - p\underline{w})$ of hires will be high quality and $(1-p)(S - p\underline{w})$ will be low quality.

[25] This focuses on the allocational role of bribes rather than connections since bribes are most important in this context. Using a connection to try to get a job is also only weakly correlated with predicted quality (SPI), with a nonstatistically significant correlation of -0.04. Connections are thus unlikely to significantly affect hire quality here, but may in other cases.

A. *Testing the Hypothesis*

To empirically test the relationship between those correlations and the efficiency of allocations, I take advantage of the externally determined, independent pools of candidates for each supervisor position. Only current CHWs can apply to be supervisors, and they are restricted to applying to the one supervisor position that their work area falls under. This produces naturally occurring variation across supervisor positions in the correlations of wealth, valuation, and quality within the pool of potential applicants: e.g., for some positions, the best candidates tend to be wealthier than the average for that pool, whereas in others, they happen to be poorer. I measure these correlations, and test the model prediction that hires and service delivery outcomes are better in clusters where the correlations are more positive.

I focus on the correlation of predicted quality with wealth due to a lack of sufficiently granular data on valuation of the job.[26] Wealth is measured with survey data on assets and earnings,[27] while the predicted quality of a candidate is measured using SPI. To remove the mechanical relationship between SPI and service delivery due to how SPI is constructed, I use "leave-out SPI" rather than SPI: for each cluster j, I estimate the relationship between supervisor characteristics and performance using all of the other clusters other than j, create an index SPI_{-j} of those characteristics, and calculate the correlation of wealth and the leave-out index SPI_{-j} in cluster j among those who want the job.[28]

The correlation of wealth and predicted quality among potential candidates varies substantially across clusters: it is negative in 26 percent of clusters, and positive and greater than $+0.2$ in 64 percent of clusters. This is consistent with heterogeneity in the predicted quality of supervisors, where, for example, nearly a quarter of hires had values of predicted quality (SPI) that were in the bottom half of applicants for their cluster. Although I will test this empirically below, there is no reason that this correlation would be related to service delivery outcomes aside from how it affects the hiring of the supervisor: this is the *correlation* of wealth and SPI across CHWs in the cluster rather than the *level* of either variable. Since CHWs work independently, the correlation of characteristics across CHWs should not affect CHW productivity prior to the creation of the supervisor position; it becomes relevant at that point since the individual hired is a function of differences in wealth across CHWs in the cluster.

The first test is whether candidate pools with a higher correlation of wealth and "leave-out SPI" produce higher quality supervisors. Column 1 of Table 4 shows that there is a strong and statistically significant positive relationship between this

[26] Valuation is measured with survey questions on whether the CHW would apply for the job of supervisor under different probabilities of being selected, where those willing to apply under lower probabilities have higher valuations. This measure is too coarse to be used: nearly half of CHWs say they would apply at the lowest listed probability of selection.

[27] I construct an index using the first principal component of an index of household assets, earnings of the CHW and the earnings of other household members. There is substantial variation, primarily due to the earnings of their husbands, with a sixfold difference in total monthly household earnings between CHWs at the tenth and ninetieth percentiles.

[28] This correlation is among the set of individuals who would ever consider applying for the job rather than among those who applied since the application decision is endogenous. The correlation among those who do not want the job is irrelevant since they would never apply or be selected.

Table 4—Cluster-Level Wealth-Quality Correlations and Service Delivery Outcomes

		Change in CHW performance				
	Supervisor SPI (1)	Functionality score (2)	Institutional delivery (3)	Newborn checkups (4)	Nutritional counseling (5)	DOTS provider (6)
Wealth-quality correlation	0.917	0.741	0.0229	0.0115	0.00769	0.000909
	(0.41)	(0.36)	(0.011)	(0.0061)	(0.0048)	(0.0041)
Observations	66	930	930	930	930	930
Joint p-value						0.027

Notes: This table examines the relationship between the correlation of wealth and SPI in a cluster and later service delivery outcomes for that cluster. Column 1 tests whether this correlation is related to the quality of the selected supervisor, while the remaining columns test the relationship with the CHW-level average monthly change in service delivery outcomes in the administrative data. Since SPI is a constructed regressor, standard clustered standard errors will overestimate the true level of precision. To incorporate the additional uncertainty from construction of the index, I use cluster bootstrapped standard errors, where SPI is reconstructed in each bootstrap sample based on that sample of data. For the joint test, I use resampling to generate the distribution of the joint test statistic under the null hypothesis (permuting the value of wealth among CHWs in the same cluster on each random draw to eliminate any relationship between the correlation of wealth and SPI and performance changes). I again reconstruct SPI in each permutation sample to account for additional uncertainty from the construction of the index.

correlation and the SPI of the hired supervisor. Translated into the units from the counterfactual section, these estimates imply that going from a candidate pool at the tenth percentile of wealth-SPI correlation (a correlation of -0.32) to the ninetieth percentile (a correlation of 0.63) is associated with moving from a hire who captures 83.9 percent of the potential first-best SPI to one who captures 95.6 percent.[29]

Columns 2–6 of Table 4 estimate the differences regression in equation (2) to evaluate whether candidate pools with a higher correlation of wealth and quality experience larger increases in service delivery outcomes after the supervisor was hired. The dependent variable is the average monthly change in a service delivery outcome for CHW i in candidate pool j over the 20 months of administrative data. CHWs in pools with a stronger correlation of wealth and SPI experience statistically significant improvements in the index of services used by the government to measure overall CHW performance ("Functionality Score," column 2). This can also be seen when broken down into individual services such as assisting with institutional deliveries, newborn check-ups, and nutritional counseling. To account for multiple hypothesis testing, I conduct a test of joint significance across the five service delivery outcomes (joint p-value $= 0.027$). These relationships are also depicted visually in online Appendix Figure A.4. Formally,

$$\Delta y_{ij} = \alpha + \beta corr\_quality\_wealth_j + \epsilon_{ij}. \tag{2}$$

The key assumption for these tests is that the wealth-SPI correlation causes these improvements through the identity of the selected supervisor rather than pre-existing

[29] There is significant positive selection into applying for the job, as seen in Figure 2. As a result, even when the correlation of wealth and SPI is negative, hiring outcomes are still reasonably good since most of the worst quality candidates do not apply.

characteristics of the cluster. Since this is a differences regression, it is particularly important that there are not pre-existing differences in service delivery *trends* that are related to the correlation of wealth and quality in a cluster (e.g., clusters with positive wealth-SPI correlations were not already on an upwards trend relative to those with negative correlations).

Online Appendix Table A.4 tests whether the correlation of wealth-SPI in a cluster is related to pre-existing factors that may affect service delivery. These potential confounders include characteristics of CHWs in the cluster (e.g., wealth, SPI), characteristics of their work environment (e.g., population that they serve, number of CHWs in their cluster), and baseline levels of service delivery. I find no evidence of a relationship between the wealth-SPI correlation and these pre-existing factors (joint p-value $= 0.65$).

I also test whether there are pre-existing *trends* in service delivery outcomes related to the wealth-SPI correlation. This is a direct test of the identifying assumption, where it would be concerning if clusters with a more positive correlation of wealth and SPI were already on an upwards trajectory prior to the hiring of supervisors. Using CHW-level data on institutional deliveries over each of the six months prior to the start of the supervisors, online Appendix Figure A.5 shows no evidence of differential pre-trends related to the correlation of wealth and SPI. The figure also addresses a separate concern that the improvement in service delivery is not attributable to the correlation of SPI with wealth. The timing suggests that this is not the case, where the wealth-SPI correlation begins to be related to service delivery right when the supervisors are hired; it is hard to imagine what variable other than wealth would cause this relationship at this point given how hiring worked.

Given these pieces of evidence, I conclude that the relationship between the wealth-SPI correlation and service delivery improvements supports the model. This rationalizes why corrupt hires were typically of a high predicted quality, where SPI is strongly correlated with both valuation and wealth (correlations of $+0.47$ and $+0.27$ respectively). In clusters where the correlations were negative, a lower quality set of individuals tends to be hired and service delivery outcomes are worse. As a result, the effects of corruption are heterogeneous even though outcomes are relatively good on average in this setting.

IV. Conclusion

This paper examines how corruption may affect allocational efficiency. I collect data from a government hiring process and document the presence of substantial corruption. Despite that, few high quality candidates were deterred from applying and hires appear to be high quality, even as compared to counterfactual hires under merit-based systems. This is one of the first papers to empirically demonstrate how corruption can lead to relatively efficient outcomes aside from Sukhtankar (2015) and Dreher and Gassebner (2013), which have different underlying mechanisms. It also expands the existing corruption literature by empirically showing how corruption can have heterogeneous effects: in this case, the consequences of a corrupt allocation depend on how wealth and valuation are correlated with characteristics of the socially optimal beneficiary.

In this context, allocations were relatively good because the average correlations of quality with wealth and valuation were strongly positive. However, there are other types of jobs for which the sign of these correlations may differ. Consider promotions to supervisory positions for police officers in high corruption environments. The wealthiest applicants will be individuals who acted most corruptly as frontline police officers, and those with the highest valuation of the job will be individuals who anticipate extracting the most corrupt rents as supervisors. This generates a negative correlation between quality and wealth as well as quality and valuation; as seen in my sample, service delivery suffers when these correlations are negative. There is a long list of factors that will determine the sign of the wealth-quality and valuation-quality correlations, including ability to extract corrupt rents, expected tenure in the job, extent to which payment is based on performance, nonpecuniary motivations (e.g., status seeking), and intergenerational correlations of wealth. The extent to which each factor matters will depend on context, but the relatively positive outcomes found here will generalize better to jobs with fewer rent extraction opportunities (e.g., teachers, other health workers) and where wealth accumulation is a positive signal of quality (e.g., based on past job performance).

These results have a number of policy implications. First, it is notoriously difficult for organizations to identify talent. Here, bribes turned out to be a good signal for being a quality candidate, but with the unfortunate consequence of diverting resources from hires to corrupt superiors. Alternative hiring mechanisms that select on valuation or wealth without bribes, such as ordeal mechanisms, may have similarly good screening effects. However, the paper also points to the potential danger of such a mechanism: if valuation is negatively correlated with quality, it would be preferable not to screen on valuation as in an ordeal mechanism. Furthermore, these results emphasize the importance of screening on characteristics that are correlated with job-specific performance rather than generalized ability. As seen in the counterfactual exercises, screening mechanisms that seem ex ante similar, such as hiring based on health knowledge or past performance, can produce meaningfully different hires. Using data on bureaucratic performance to determine job-specific hiring criteria rather than hiring on generalized ability is a potentially low cost way to improve bureaucratic efficiency.

Second, the paper demonstrates empirically that corruption should not be thought of as categorically negative, but rather as having heterogeneous effects. Given the limited capacity of policymakers to fight corruption, governments should target anticorruption resources towards the sectors in which it leads to the biggest distortions. This paper offers a way to identify those sectors: for example, in the case of markets for jobs, there will be large distortions for jobs that have substantial ability to extract bribes as that produces a negative correlation between valuation and quality for the job. These results do not imply that corruption in hiring should be ignored, but that given the difficulty of eliminating corruption, it is better to first address its most costly forms.

REFERENCES

Allen, Douglas W. 1998. "Compatible Incentives and the Purchase of Military Commissions." *Journal of Legal Studies* 27 (1): 45–66.

Ashraf, Nava, Oriana Bandiera, Edward Davenport, and Scott S. Lee. 2020. "Losing Prosociality in the Quest for Talent? Sorting, Selection, and Productivity in the Delivery of Public Services." *American Economic Review* 110 (5): 1355–94.

Banerjee, Abhijit V., Rema Hanna, and Sendhil Mullainathan. 2013. "Corruption." In *Handbook of Organizational Economics*, edited by John Roberts and Robert S. Gibbons, 1109–47. Princeton: Princeton University Press.

Beck, Paul J., and Michael W. Maher. 1986. "A Comparison of Bribery and Bidding in Thin Markets." *Economics Letters* 20 (1): 1–5.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen. 2014. "Inference on Treatment Effects after Selection among High-Dimensional Controls." *Review of Economic Studies* 81 (2): 608–50.

Bertrand, Marianne, Simeon Djankov, Rema Hanna, and Sendhil Mullainathan. 2007. "Obtaining a Driver's License in India: An Experimental Approach to Studying Corruption." *Quarterly Journal of Economics* 122 (4): 1639–76.

Best, Michael Carlos, Jonas Hjort, and David Szakonyi. 2019. "Individuals and Organizations as Sources of State Effectiveness." Unpublished.

Burgess, Robin, Matthew Hansen, Benjamin A. Olken, Peter Potapov, and Stefanie Sieber. 2012. "The Political Economy of Deforestation in the Tropics." *Quarterly Journal of Economics* 127 (4): 1707–54.

Callen, Michael, Saad Gulzar, Ali Hasanain, Muhammad Yasir Khan, and Arman Rezaee. 2018. "Data and Policy Decisions: Experimental Evidence from Pakistan." Unpublished.

Cole, Shawn, and Anh Tran. 2011. "Evidence from the Firm: A New Approach to Understanding Corruption." In *International Handbook on the Economics of Corruption*. Vol. 2, edited by Susan Rose-Ackerman and Tina Søreide, 408–27. Cheltenham, UK: Edward Elgar Publishing.

Colonnelli, Emanuele, Mounu Prem, and Edoardo Teso. 2020. "Patronage and Selection in Public Sector Organization." *American Economic Review* 110 (10): 3071–99.

Dal Bó, Ernesto, Frederico Finan, Nicholas Y. Li, and Laura Schechter. 2018. "Government Decentralization Under Changing State Capacity: Experimental Evidence from Paraguay." NBER Working Paper 24879.

Dal Bó, Ernesto, Frederico Finan, and Martin A. Rossi. 2013. "Strengthening State Capabilities: The Role of Financial Incentives in the Call to Public Service." *Quarterly Journal of Economics* 128 (3): 1169–1218.

de Janvry, Alain, Craig McIntosh, and Elisabeth Sadoulet. 2015. "Fair Trade and Free Entry: Can a Disequilibrium Market Serve as a Development Tool?" *Review of Economics and Statistics* 97 (3): 567–73.

Deserranno, Erika. 2019. "Financial Incentives as Signals: Experimental Evidence from the Recruitment of Village Promoters in Uganda." *American Economic Journal: Applied Economics* 11 (1): 277–317.

Dreher, Axel, and Martin Gassebner. 2013. "Greasing the Wheels? The Impact of Regulations and Corruption on Firm Entry." *Public Choice* 155 (3–4): 413–32.

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan. 2013. "Truth-Telling by Third-Party Auditors and the Response of Polluting Firms: Experimental Evidence from India." *Quarterly Journal of Economics* 128 (4): 1499–1545.

Efron, Bradley, and Robert J. Tibshirani. 1994. *An Introduction to the Bootstrap*. New York: CRC Press.

Esteban, Joan, and Debraj Ray. 2006. "Inequality, Lobbying, and Resource Allocation." *American Economic Review* 96 (1): 257–79.

Fafchamps, Marcel, and Julien Labonne. 2017. "Do Politicians' Relatives Get Better Jobs? Evidence from Municipal Elections." *Journal of Law, Economics, and Organization* 33 (2): 268–300.

Ferraz, Claudio, Frederico Finan, and Diana B. Moreira. 2012. "Corrupting Learning: Evidence from Missing Federal Education Funds in Brazil." *Journal of Public Economics* 96 (9–10): 712–26.

Filkins, Dexter. 2009. "Afghan Corruption: Everything for Sale." *New York Times*, January 2.

Finan, Frederico, Benjamin Olken, and Rohini Pande. 2017. "The Personnel Economics of the State." In *Handbook of Economic Field Experiments*, edited by Abhijit Vinayak Banerjee and Esther Duflo, 467–514. Amsterdam: Elsevier.

Fisman, Raymond, Jing Shi, Yongxiang Wang, and Rong Xu. 2018. "Social Ties and Favoritism in Chinese Science." *Journal of Political Economy* 126 (3): 1134–71.
Guardado, Jenny. 2018. "Office-Selling, Corruption, and Long-Term Development in Peru." *American Political Science Review* 112 (4): 971–95.
Hanna, Rema, and Shing-Yi Wang. 2017. "Dishonesty and Selection into Public Service: Evidence from India." *American Economic Journal: Economic Policy* 9 (3): 262–90.
Iyer, Lakshmi, and Anandi Mani. 2012. "Traveling Agents: Political Change and Bureaucratic Turnover in India." *Review of Economics and Statistics* 94 (3): 723–39.
Jauregui, Beatrice. 2014. "Provisional Agency in India: Jugaad and Legitimation of Corruption." *American Ethnologist* 41 (1): 76–91.
Jia, Ruixue, Masayuki Kudamatsu, and David Seim. 2015. "Political Selection in China: The Complementary Roles of Connections and Performance." *Journal of the European Economic Association* 13 (4): 631–68.
Johnson, Noel D., and Mark Koyama. 2014. "Tax Farming and the Origins of State Capacity in England and France." *Explorations in Economic History* 51: 1–20.
Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz. 2007. "Experimental Analysis of Neighborhood Effects." *Econometrica* 75 (1): 83–119.
Kristiansen, Stein, and Muhid Ramli. 2006. "Buying an Income: The Market for Civil Service Positions in Indonesia." *Contemporary Southeast Asia* 28 (2): 207–33.
Krueger, Anne O. 1974. "The Political Economy of the Rent-Seeking Society." *American Economic Review* 64 (3): 291–303.
Lagarde, Mylene, and Duane Blaauw. 2014. "Pro-Social Preferences and Self-Selection into Jobs: Evidence from South African Nurses." *Journal of Economic Behavior and Organization* 107: 136–52.
Leff, Nathaniel H. 1964. "Economic Development Through Bureaucratic Corruption." *American Behavioral Scientist* 8 (3): 8–14.
Lien, Da-Hsiang Donald. 1986. "A Note on Competitive Bribery Games." *Economics Letters* 22 (4): 337–41.
Lui, Francis T. 1985. "An Equilibrium Queuing Model of Bribery." *Journal of Political Economy* 93 (4): 760–81.
Moore, Malcolm. 2013. "China Abolishes Powerful Railways Ministry in Battle against Corruption." *Telegraph*.
Muralidharan, Karthik. 2015. "A New Approach to Public Sector Hiring in India for Improved Service Delivery." Unpublished.
Muralidharan, Karthik, Paul Niehaus, and Sandip Sukhtankar. 2017. "General Equilibrium Effects of (Improving) Public Employment Programs: Experimental Evidence from India." NBER Working Paper 23838.
Muralidharan, Karthik, Paul Niehaus, Sandip Sukhtankar, and Jeffrey Weaver. 2021. "Improving Last-Mile Service Delivery Using Phone-Based Monitoring." *American Economic Journal: Applied Economics* 13 (2): 52–82.
NewsRu. 2008. "Medvedev Admitted That Government Posts Are Sold in Russia, and Zhironovsky Named the Price." *Novosti Rossii*.
Niehaus, Paul, Antonia Atanassova, Marianne Bertrand, and Sendhil Mullainathan. 2013. "Targeting with Agents." *American Economic Journal: Economic Policy* 5 (1): 206–38.
Niehaus, Paul, and Sandip Sukhtankar. 2013. "Corruption Dynamics: The Golden Goose Effect." *American Economic Journal: Economic Policy* 5 (4): 230–69.
Olken, Benjamin A. 2007. "Monitoring Corruption: Evidence from a Field Experiment in Indonesia." *Journal of Political Economy* 115 (2): 200–249.
Olken, Benjamin A., and Patrick Barron. 2009. "The Simple Economics of Extortion: Evidence from Trucking in Aceh." *Journal of Political Economy* 117 (3): 417–52.
Olken, Benjamin A., and Rohini Pande. 2013. "Governance Initiative Review Paper: JPAL Governance Initiative." Abdul Latif Jameel Poverty Action Lab.
Ornaghi, Arianna. 2019. "Civil Service Reforms: Evidence from US Police Departments." Unpublished.
Perry, Henry B., Rose Zulliger, and Michael M. Rogers. 2014. "Community Health Workers in Low-, Middle-, and High-Income Countries." *Annual Review of Public Health* 35: 399–421.
Sequeira, Sandra, and Simeon Djankov. 2014. "Corruption and Firm Behavior: Evidence from African Ports." *Journal of International Economics* 94 (2): 277–94.
Shimer, Robert. 1999. "Job Auctions." Unpublished.
Shleifer, Andrei. 2004. "Does Competition Destroy Ethical Behavior?" *American Economic Review* 94 (2): 414–18.

Shleifer, Andrei, and Robert W. Vishny. 1993. "Corruption." *Quarterly Journal of Economics* 108 (3): 599–617.

Sukhtankar, Sandip. 2015. "The Impact of Corruption on Consumer Markets: Evidence from the Allocation of Second-Generation Wireless Spectrum in India." *Journal of Law and Economics* 58 (1): 75–109.

Sukhtankar, Sandip, and Milan Vaishnav. 2015. "Corruption in India: Bridging Research Evidence and Policy Options." Brookings-NCAER India Policy Forum.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation*. Cambridge, UK: Cambridge University Press.

Tura, Gurmesa, Mesganaw Fantahun, and Alemayehu Worku. 2013. "The Effect of Health Facility Delivery on Neonatal Mortality: Systematic Review and Meta-Analysis." *BMC Pregnancy and Childbirth* 13 (1): 1–9.

Wade, Robert. 1982. "The System of Administrative and Political Corruption: Canal Irrigation in South India." *Journal of Development Studies* 18 (3): 287–328.

Wade, Robert. 1985. "The Market for Public Office: Why the Indian State Is Not Better at Development." *World Development* 13 (4): 467–97.

Walsh, Declan. 2014. "At Afghan Border, Graft Is Part of the Bargain." *New York Times*, November 11.

Weaver, Jeffrey. 2021. "Replication Data for: Jobs for Sale: Corruption and Misallocation in Hiring." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. http://doi.org/10.3886/E136061V1.

White, Eugene N. 2004. "From Privatized to Government-Administered Tax Collection: Tax Farming in Eighteenth-Century France." *Economic History Review* 57 (4): 636–63.

Xu, Guo. 2018. "The Costs of Patronage: Evidence from the British Empire." *American Economic Review* 108 (11): 3170–98.