Regulation by Reputation? Intermediaries, Labor Abuses, and International Migration

A. Nilesh Fernando[†]
Niharika Singh

February 13, 2023

Abstract

Migrants workers and employers rely on intermediaries to facilitate labor market placements. If information frictions obscure their reputation, however, intermediaries may under-invest in placement quality. Using data on over 1.5 million Sri Lankan migrants to the Gulf region, we examine the effects of an intermediary rating program that publicly revealed ratings two years after it was announced. Prior to the ratings being revealed, eligible under-performing agencies invest in the rating criteria and place migrants with less abusive employers who pay higher salaries. Our results suggest that the threat of quality revelation induced agencies to prospectively screen employers.

JEL Classification Numbers: O12, J63, O15
Keywords: Labor Intermediaries, Placement Quality, International Migration

[*]We thank Sam Bazzi, Gaurav Khanna, Caroline Theoharides, David McKenzie, Suresh Naidu, Ellora Derenoncourt, Simone Schaner, Mushfiq Mobarak, Dean Yang, Joe Kaboski, Renee Yaseen, and Taryn Dinkelman for comments on this paper. We thank seminar participants at Amherst, Boston University, Columbia, Harvard, Michigan, Notre Dame, Northwestern, UChicago, UCSD, UT-Austin, WB-LISER Conference, CEPR-MISUM-Site Conference, the Indian Statistical Institute, and the ND-PUC Luksburg Conference for comments. The Sri Lanka Bureau of Foreign Employment (SLBFE), the Ministry of Foreign Employment Promotion and Welfare, Verité Research, Azam Bakeer Markar, Tamra Ranasinghe, and the Hon. Min. Thalatha Atukorale were essential to facilitating this research partnership. We also thank L.K. Ruhunage for his expertise on the agency ratings program. We are grateful to the Ford Program in Human Development Studies and Solidarity and the Institute for Scholarship in the Liberal Arts at the University of Notre Dame for providing funding. Alison Lodermeier, Mika Inoue, Thashiya Nauki, Paul Shaloka, Mehak Siddiquei and Joe Tatarka provided excellent research assistance. The usual caveat applies.

[†]Fernando (corresponding author): University of Notre Dame, nilesh.fernando@nd.edu; Singh: Columbia University.

1 Introduction

Prospective migrants often rely on intermediaries to facilitate access to labor market opportunities abroad.[1] Though lucrative, seeking work in an unfamiliar context is risky and often leaves migrants vulnerable to abuse.[2] Our study encompasses over a million Sri Lankan migrants to the Gulf Region, and nearly one in ten migrants make a complaint about employer abuse to a *consulate*.[3] Descriptive evidence in this setting suggests that intermediaries greatly influence whether a migrant experiences abuse, and work by Naidu et al. (2016a) suggests employers exercise considerable monopsony power, limiting the ability of migrants to switch jobs and increasing the importance of their initial placement.[4] Migration intermediaries can influence placement quality by screening both employers and migrants, but their propensity to screen depends on the observability and returns to this costly investment (Autor, 2008). In particular, while an employer and a migrant may learn about investments in screening made by the intermediary that facilitated their match, information frictions may restrict whether these facts become more widely known. As a result, intermediaries may see a diminished return to screening and under-invest in placement quality.[5]

In this paper, we examine whether a program designed to rate intermediaries and publicly reveal their quality can relax information frictions and improve placement quality. Our study focuses on local migration intermediaries (hereafter, recruitment agencies) that recruit Sri Lankan workers for employment in the Gulf region. The (Sri Lankan) regulator first announced the criteria on which agencies would be rated and, two years later, revealed a star rating (zero to five stars) to the public. Using a difference-in-difference design, we find that the program successfully induced agencies eligible for the program—particularly, previously under-performing agencies—to screen foreign demand and place migrants with less abusive employers *prior* to the public revelation of the star ratings. To guide the interpretation of our empirical results, we develop a dynamic model of agency reputations and show how the ratings program improves the status quo information environment. It does so by providing the market with a credible signal of an agency's investment in placement quality. Future quality revelation induces agencies to invest in placement quality even before the ratings are made public, to influence an employer's belief about their type and, consequently, demand for their services in the future.

Migrant labor opportunities represent a Faustian bargain for Sri Lanka. Approximately 17% of the labor force is engaged in migrant work in the Gulf region, access to these employment opportunities can be transformative (Dinkelman and Mariotti, 2016; Theoharides, 2018), and remittances of foreign currency are crucial to macroeconomic stability. However, labor

[1] Across two of the world's largest migration corridors—the US Southern border and the Gulf region—estimates place the share of intermediated migration as high as 80%. See Demaret (2006) for the Gulf region, and Orrenius et al. (2001) regarding the US southern border.

[2] For example, see Clemens (2013), Gibson and McKenzie (2014), Hendricks and Schoellman (2018).

[3] By 'Gulf region' we refer to the countries belonging to the Gulf Cooperation Council (GCC).

[4] Fernando and Lodermeier (2022) find that the time-invariant characteristics of migration intermediaries systematically co-vary with the incidence of migrant abuse, suggesting they may have access to private information on the quality of foreign employers.

[5] Placement quality is a bundle of wage and non-wage amenities encompassing the safety of a workplace.

abuses are commonplace and employers routinely withhold pay, often renege on contractually agreed-upon amenities, and engage in physical abuse (HRW, 2010). These concerns have led the Sri Lankan government, among others, to institute sweeping restrictions on migration to the Gulf region at a great potential welfare cost.[6] However, there is limited research examining whether governments can instead use regulatory policy to induce local intermediaries to improve the quality of migrant placements.

We use a unique administrative dataset that allows us to link over 1.5 million Sri Lankan migrants over a decade (2005-2015) to recruitment agencies and foreign employers. To assess placement quality, we merge this dataset with data on contractual provisions and worker complaints—ranging from physical abuse to breach of contract—made to Sri Lankan consulates across the Gulf region. We complement these administrative datasets with original surveys of approximately 20% of active recruitment agencies in 2019.

The rating program created incentives for recruitment agencies to improve their placement quality on both an extensive (job types, sectors, and countries) and an intensive margin, by placing migrants with more reputable employers in existing sectors such as domestic work; placement quality was assessed using a government-designed metric: the *Quantitative Score*.[7] The program also encouraged eligible agencies to invest in recommended management practices that were subsequently audited by third-party inspectors—compliance with these recommendations was assessed using another metric: the *Qualitative Score*. The government combined the Quantitative and Qualitative Scores to produce a continuous score which, in turn, maps to a star rating. The government first announced the rating criteria in early 2010 and then publicly revealed a star rating in 2012.

Eligibility for the program is based on criteria that preceded the announcement of the program—agencies were eligible if they recruited at least 100 migrants in 2009. Consequently, agencies could not select into the program and, as per the program rules, they could not opt out of it. These facts motivate a difference-in-difference design that compares eligible agencies who recruited just above the 100 migrant threshold to a set of comparison agencies who recruited just below it.[8] Our identifying assumption is that eligibility is not systematically correlated with agency trends but, rather, reflect stochastic differences in recruitment; an assumption supported by F-tests of the lagged differences of our main outcomes.

We estimate the effects of the program for the period after the rating criteria were announced but prior to the revelation of ratings (i.e. 2010-2012)—which we refer to as the *Reputational Incentives* (RI) phase—and the period succeeding the public release of the star ratings, which we refer to as the *Quality Revelation* (QR) phase (i.e. > 2012). In addition, to estimate heterogeneous effects by (pre-program) agency quality, we use the ratings formula

[6] In recent years, the governments of Sri Lanka, Nepal, Indonesia and the Philippines have instituted policies that seek to ban groups of female domestic workers from seeking work in the Gulf region (Shivakoti et al., 2021).

[7] The Quantitative Score is based on total recruitment, the skill intensity of jobs matched, and an agency's performance in resolving complaints, see Appendix A.1.

[8] In order to select a set of comparable agencies, we use an algorithm developed by Calonico et al. (2014) that, in essence, assesses balance across a set of pre-specified pre-treatment characteristics. This algorithm yields an optimal sample that contains 207 agencies who recruited between 45 and 155 migrants in 2009.

to compute a *Predicted Rating* for both eligible and comparison agencies.[9] We denote agencies that have a below median Predicted Rating as pre-program low quality agencies. Importantly, outcomes during the RI-phase can *only* be influenced by agencies as the star ratings are yet to be made public. In contrast, the QR-phase combines the effects of revealing a rating—to which both employers and migrants can respond—and the effects of reputational incentives.

We have three main findings. *First, pre-program low quality agencies respond to reputational incentives, secure higher foreign employer demand, and are less likely to exit the market.* To assess investments in the rating criteria announced by the government, we use a summary measure developed by the government—the Quantitative Score—which aggregates outcomes across a set of criteria relating to the composition of recruitment and complaint management (see Appendix A.1). While eligible agencies have a higher Quantitative Score, this estimate is imprecise. In contrast, low quality agencies invest in the rating criteria and their Quantitative Score is 11% higher than that of comparison agencies prior to quality revelation. In addition, these agencies receive substantially more demand (65%) from foreign employers (i.e. job orders) relative to the comparison mean. Overall, while eligible agencies are approximately 44% less likely to exit the market, this result is driven by low quality agencies who are 75% less likely to exit the market relative to the comparison group[10]; in contrast, exit by high quality agencies is unaffected.

Second, the program induces eligible agencies to improve placement quality and this too is largely driven by pre-program low quality agencies. We find that migrants placed by eligible low quality agencies report monthly salaries that are $20 (12%) higher than the comparison mean. Further, these agencies are 29% more likely to place migrants with a *Good Employer*: i.e. one with a below median pre-program complaint rate. Using an index that aggregates multiple measures of contract quality, we find that low quality eligible agencies increase contract quality by 0.17 standard deviations.[11] Perhaps as a result of better placement quality, migrants placed by low quality eligible agencies are 19% more likely to renew their contract with an employer relative to the comparison mean.

Finally, improvements in placement quality occur prior to quality revelation, suggesting the program induced low quality agencies to prospectively screen foreign employers. Increases in average salaries and the share of job orders from Good Employers begin during the Reputational Incentives phase. As the star ratings are yet to be made public, these effects are *necessarily* a consequence of agency responses to the program announcement and preclude responses by foreign employers and migrants. During this period, we find little to distinguish migrants placed by eligible and comparison agencies on observable characteristics. In addition, eligible agencies are no more likely to place migrants with *new employers*: those we do not observe

[9] To compute the Predicted Rating we use the fact that the rating calculated by the government is, in part, a function of administrative data collected *prior* to the announcement of the program. As such, we use this data to predict the continuous rating and use the fitted values to construct a counterfactual Predicted Rating for comparison agencies.

[10] Exit here is defined as an agency that reports zero migrants recruited for at least 12 consecutive months.

[11] The index of contract quality is a normalized index incorporating whether or not an employer provides health insurance, accommodation, return airfare and the salary range paid to workers. Index components are weighted by the inverse of the covariance matrix as in Anderson (2008).

in our dataset prior to when the program began. Rather, we find that eligible agencies are more likely to place migrants with Good Employers that were previously known to them, and this shift begins during the RI-phase (see Figure 4). To verify the plausibility of our proposed explanation—that eligible agencies are induced to screen foreign employers—we collected detailed data on agency operations (n = 109) and find that agencies routinely screen employers: half of the surveyed agencies report rejecting a job order in the past year, with 52% citing safety concerns as a reason for doing so.

Collectively our results suggest that the introduction of the program created reputational incentives that induced eligible agencies to improve their placement quality prior to the public release of the star ratings. While we do not find that the program increased recruitment by eligible agencies, this may in part be due to spillover effects for comparison agencies that are in the immediate vicinity of an eligible agency. For these comparison agencies, we find that a nearby eligible agency induces positive spillovers, increasing migrant recruitment and job orders, though the latter estimate is imprecise. This result, at once, suggests that we may be underestimating the effect of the program and that eligible agencies may face large fixed costs in securing new demand. We explore the robustness of our results to using alternative sample thresholds and measures of pre-program agency quality and arrive at qualitatively similar conclusions.

We also find evidence to suggest that the program led foreign employers to reward highly rated agencies with more job orders. Using the exact ratings formula to construct a regression discontinuity design, we compare agencies who are just above and below star-specific thresholds. Pooling across the thresholds, we find that a marginal star leads to a 57% increase in the number of job orders. In contrast, we do not find that migrants responded to the star ratings: higher rated agencies do not recruit more migrants; migrants who had registered a complaint prior to the program and especially vulnerable groups of migrants were no more likely to pick a higher rated agency. This is unsurprising as the ratings were largely made available on the internet which, at the time, less that 12% of the Sri Lankan population had access to. Furthermore, surveys with migrants, suggest that just 14% had ever heard of the ratings program.

Even though low quality agencies drive many of the observed effects, they typically end up with relatively lower star ratings. Our results suggest that the program both induced these agencies to invest in placement quality prior to the QR-phase and had a certification effect that improved their visibility, regardless of the number of stars they ultimately received. Assessing whether the program improved overall market efficiency is complicated by the fact that the market experienced a steady decline over this period. However, we use a proxy for match surplus— Total Salaries: the product of migrants recruited and the salary they are contractually owed— and find that the program systematically increases this proxy. While (particularly, eligible low quality) agencies appear better off as a consequence of the program, we lack data on commissions and the entry and exit decisions of employers and are therefore unable to assess the effects of the program on their welfare.

Our paper is unique in its ability to capture the response of migration intermediaries

to changes in regulatory policy. The paper closest to our own, Bazzi et al. (2021), uses an experiment to estimate how migrants respond to information on intermediary quality. The policy we study, in contrast, influences the actions of employers and intermediaries. This is an especially relevant margin to examine, given that local intermediaries often contract out recruitment to `sub-intermediaries'—thereby obscuring and reducing the relevance of their own reputations—and wield considerable market power inhibiting the scope for migrant choice. The recent empirical literature has focused on product market intermediaries and how they address market failures and influence competition and prices. [12] Labor market intermediaries, in spite of their ubiquity, have received limited attention [13] and primarily act as `match-makers' rather than `market-makers' —in that they do not buy and sell a standardized product (Yavaş, 1994). Consequently, screening is an especially important function we are able to gauge by quantifying multiple dimensions of placement quality.

Second, we contribute to the literature on international migration. A number of studies have quantified the enormous wage gains from international migration (Gibson et al., 2017; Clemens, 2013; McKenzie et al., 2010; Hendricks and Schoellman, 2018), while others have estimated the effects of providing accurate information about the returns of migration (Shrestha, 2017; Shrestha and Yang, 2019; Batista and McKenzie, 2021). In spite of the recognition that migrant labor may constitute a `repugnant' transaction (Clemens, 2018), comparatively less attention has been paid to the quality of migrant placements—likely owing to the difficulty and political sensitivity of collecting reliable data. We use a combination of consular reports of abuse and contract characteristics (salaries and amenities) to assess migrant placement quality on multiple dimensions and estimate how it is influenced by regulatory policy in migrant-sending countries.

Finally, our paper contributes to a literature examining the effects of quality disclosure and reputational mechanisms in markets. In the development literature, research has focused on the importance of reputations and relational contracting in buyer-seller relationships where contract enforcement is costly (Macchiavello and Morjaria, 2015; 2021) and how quality disclosure can influence market-level outcomes (Andrabi et al., 2017). Outside the development literature, Jin and Leslie (2003) study the introduction of restaurant hygiene cards on customer demand, while Luca (2016) investigates the impact of Yelp reviews on similar outcomes (see Dranove and Jin (2010) for a review). In a similar study to our own, Benson et al. (2019) examine the value of employer ratings on an online labor market. They find that employers exogenously assigned a good reputation attract workers at nearly twice the rate and show that this is driven by workers screening employers. We contribute to this literature by showing how quality revelation in the high-stakes world of international migration serves as a reputational mechanism that induces intermediaries to improve placement quality.

The remainder of the paper is structured as follows: Section 2 describes the background

[12] See, for example Bardhan et al. (2013); Startz (2016); Mitra et al. (2018); Bergquist and Dinerstein (2020); Atkin and Donaldson (2021); Grant and Startz (2019).

[13] An exception is the work of Lee (2007) on public employment agencies in the US. He shows that the creation of public recruitment agencies in the U.S. was a feasible policy response to increase competition with private agencies and temper firm-side adverse selection.

of our study and the design of the agency ratings program. Section 3 outlines a conceptual framework with which to interpret our empirical results. Section 4 describes our data and Section 5 presents our empirical strategy. Section 6 presents and discusses our main results. Section 7 explores the robustness of our results, while Section 8 examines how the program in uences market ef ciency. Section 9 concludes.

2 Background: Migration Overview and Regulatory Framework

2.1 International Labor Migration in Sri Lanka

An estimated 1.5 million Sri Lankan workers, roughly 17% of the labor force, were employed abroad in the Gulf region in 2015.[14] Each year, roughly 200,000 migrants—evenly split between men and women—seek employment abroad primarily as construction workers, vehicle operators, and (female) domestic workers (see Appendix A. 12). Remittances from these migrants are an important source of foreign currency, accounting for 8.5% of GDP in 2015. However, there are persistent concerns about labor abuses experienced by migrants in the Gulf region, as documented by human rights groups (HRW, 2007; International, 2020) and ethnographic studies[15] that are consistent with a descriptive paper that uses our administrative data to systematically explore the extent of abuse (Fernando and Lodermeier, 2022). In late 2013, renewed concerns about migrant abuse led the government to impose restrictions on young women intending to migrate to the Gulf countries for domestic work (Peru, 2022). As a consequence, female domestic worker migration to the Gulf countries saw a large reduction after 2014.

2.2 Market Structure and Regulatory Framework

Concerns about abuse following the rapid expansion of migrant labor demand in the 1970s,[16] led to the creation of a regulatory body to oversee migration: the Sri Lanka Bureau of Foreign Employment (SLBFE) in 1985. The SLBFE was given legal authority to enter into agreements with foreign countries and regulate local recruitment agencies through the creation and enforcement of legal standards. All recruitment agencies are required to apply for a license with the SLBFE and migrant registration with the SLBFE has been compulsory since 1996. This is in large part enforced by monitoring at the airport; prospective migrants are denied boarding if they do not have a SLBFE registration stamp on their passport.

[14] This estimate was calculated using publicly available data from the International Labor Organization (ILO), which estimates the Sri Lankan labor force in 2020 at 8.6 million and a report from the Foreign Ministry of Sri Lanka which estimates the stock of migrants in the Gulf region at 1.5 million (Aryasinha, 2020).

[15] See (Abu-Habib, 1998; Agunias, 2011; Frantz, 2008; Jureidini and Moukarbel, 2004).

[16] See Alfano (2010); Somarathna (2011); Silva (2013); Somarathna (2015); Jayawardhane (2018) for examples of reporting on abuses faced by Sri Lankan migrants in local newspapers.

2.2.1 Consular Assistance and Complaints

While consular assistance is available to all Sri Lankans, only registered migrants have access to support from the SLBFE.[17] This support includes a dedicated labor section at consulates that registers and addresses complaints, and arranges repatriation and access to safe houses in response to severe cases of abuse (ILO, 2013). Migrant workers or their family members are able to register complaints at Sri Lankan consulates (see Appendix A.3 for an example of back and forth between the SLBFE and an agency regarding a complaint).

2.2.2 Current Recruitment Process

The process for recruiting a migrant worker involves obtaining a number of approvals from the SLBFE and can take anywhere from 2-5 months to complete (see Appendix A. 5 for details). In short, in order to recruit a Sri Lankan worker, a foreign employer must register a `job order' with the SLBFE. A job order lists the number of vacancies to be lled, the salary and bene ts that are contractually obliged and up to ve licensed agencies who would recruit these workers on an employer's behalf. The foreign employer rst reaches out to a Sri Lankan consulate and then completes the job order process on the SLBFE website. If the job order is approved, agencies recruit migrants, obtain passports and visas on their behalf and must typically satisfy a number of requirements (e.g. pre-departure training) stipulated by the SLBFE. Once these requirements are satis ed, the SLBFE stamps a migrant worker's passport.

2.2.3 Market Structure

Foreign employers pay recruitment agencies a xed commission per migrant recruited. While we do not observe commissions in our administrative data, a survey of agencies (n=109) we conducted suggests commissions are largely xed (see Section 4).[18] In our administrative data spanning 2005-2015, there are between 600-700 active recruitment agencies each year, that recruit approximately 80 migrants per year on average; the top 20 agencies account for 25% of migrants recruited during this period (see Section 4 for details).

The agency survey suggests that job orders are plentiful—50% of agencies report rejecting a job order in the prior year— but less than 1 in 5 agencies say they were able to ll all vacancies on a job order. Two-thirds of surveyed agencies state the main reason they were unable to ll a job order was because they couldn't nd workers. Though plentiful, the characteristics of job orders vary considerably and are consistent with prior studies demonstrating a high degree of employer monopsony power in this setting (Naidu et al., 2016b). We nd that the least abusive employers pay higher salaries and offer more amenities both within and across sectors (see Appendix A14). The lack of a compensating differential is suggestive of a

[17] Note, as Sri Lanka is an island, air travel is typically the only means to reach migrant destinations in the Gulf. This greatly reduces the possibility of 'irregular' migration ows through porous land borders.

[18] In the agency survey, 22 agencies specializing in the recruitment of domestic workers reported earning per worker commissions between $342-$382, in spite of placing these workers in very different countries.

dual market consisting of two types of employers that offer identical workers jobs with with very different characteristics (Bulow and Summers, 1986).

2.3 Agency Ratings Program

2.3.1 Motivation and Rationale

Due to domestic political pressures and the increased salience of migrant abuses, the SLBFE devised a program to rate local recruitment agencies towards improving professionalism in the industry and improve the quality of job opportunities available to Sri Lankan migrants. [19] The program's stated goals include assisting foreign employers to identify suitable local agencies, motivating agencies to improve their own standards, and providing information to migrants on reputable agencies. In the booklet provided to agencies (see Appendix A. 4 for a summarized version), the criteria listed for evaluation and the `Proposals and Suggestions' made by the government reveal two key concerns that the program intends to address:

- Management Practices: the program encourages the standardization of record-keeping practices towards minimizing fraud and maintaining orderly premises. [20]
- Placement Quality: the program encourages agencies to develop new markets and high-skill opportunities for Sri Lankan workers. The goal here in particular is to reward placements in sectors and destinations that may have less abusive conditions. In addition the program encourages agencies to resolve complaints made by migrants and be more vigilant about issues migrants face at their workplace. [21]

2.3.2 Eligibility for Program, Timing, & Program Phases

Agencies meeting the following criteria were eligible for the program: (i) recruited more than 100 migrants during 2009; (ii) held a valid license for more than a year. [22] The program consisted of two key phases:

Reputational Incentives (RI) Phase (2010-2012): This is the phase of the program during which eligible agencies are informed about the program and given the rating criteria.

[19] The case of Rizana Nafeek, a Sri Lankan domestic worker who was sentenced to death in 2009 and subsequently executed in Saudi Arabia was an pivotal moment in galvanizing reforms (Silva, 2013).

[20] For example, the government suggests that agencies should `Update the records and documents [including] data sheets of Sri Lankan employees and approved job orders' and '[Maintain] records of statutory refunds made by SLBFE, passport details, employers and commissions received.'

[21] For example, the booklet states that `agencies will receive minus points if there are complaints against agencies lodged by the migrant worker, relatives, family members, foreign agencies and foreign employers' and that agencies should `pay attention towards the bene ts and salary of the migrant worker as mentioned in the agreement throughout the contract period and persuade the foreign employer if necessary.'

[22] Technically, the eligibility also required that agencies had not previously been engaged in underage recruitment, were not convicted of submitting false documents, did not have more than 10 pending legal cases. These conditions are virtually guaranteed by having a valid license.

- January 2010: Agencies were made aware of the program in early 2010 through an online system managed by the SLBFE. Agencies were provided with a booklet that described the eligibility criteria for the program, the goals of the program, and the criteria used in the assessment (see Appendix A.4).

- July 2010–May 2011: Third-party audits to assess compliance with recommended management practices were conducted by the Sri Lanka Standards Institute (SLSI) and the Sri Lanka Institute of Marketing (SLIM). The visits were not pre-announced and took place in 2011; only eligible agencies were audited.

Quality Revelation (QR) Phase (2012–2015): This is the phase during which the star ratings are released to the public.

- February, 2012: The ratings were announced in a public award ceremony.

- June, 2012: A searchable database of the star ratings was released on the SLBFE website.

2.3.3 Ratings Formula

Agencies were awarded a star rating between zero to five stars based on thresholds corresponding to a continuous score (hereafter, the Combined Score) that aggregates the rating criteria.[23] The Combined Score is itself a weighted combination of two constituent scores: the Quantitative Score and the Qualitative Score. The Quantitative Score uses administrative data to assess placement quality and its criteria include the volume of recruitment, skill and market diversification, and the successful resolution of complaints. The Qualitative Score uses data collected through third-party audits to assess compliance with recommended management practices such as record-keeping, the maintenance of office premises, and the qualifications of agency staff (see Appendix A.1 for details on the criteria).

As the third-party audits occurred after the program announcement, agencies were able to influence their Qualitative Score. However, the government used pre-program data from 2009 to compute the Quantitative Score that was used for the star ratings. To reiterate, in spite of informing agencies that their star rating would be based on an assessment of their investments in the rating criteria during the RI-phase, part of the star rating was a function of data that, unknown to agencies, could not be influenced by their actions.[24]

In summary, agencies were made aware of their eligibility for the program, the criteria for the rating, and when they would be released to the public. Agencies are not told the precise formula which maps the criteria to a star rating, but were told that their rating would be determined by their quantitative and Qualitative Score during the RI-phase (2010–2012).

[23] Where c is the Combined Score used by the government to determine star ratings. c = F(quantitative score, qualitative score) and c is then mapped to a star rating as follows — 0 stars: 0 ≤ c ≤ 19.9, 1 star: 20 ≤ c ≤ 34.9, 2 stars: 35 ≤ c ≤ 49.9, 3 stars: 50 ≤ c ≤ 59.9, 4 stars: 60 ≤ c ≤ 79.9, and 5 stars: c ≥ 80.

[24] We compute that if the government instead used the 2011 Quantitative Score, 5/77 eligible agencies in our main sample would receive a different number of stars.

2.3.4 Distribution of Star Ratings

Based on the eligibility criteria, 306 out of 800 operational agencies were eligible for the program. Out of these eligible agencies, 34 agencies did not receive a star, 58 agencies received 1 star, 155 agencies received 2 stars, 52 agencies received 3 stars, and 8 agencies received 4 stars. No agency received 5 stars (see Appendix A.6). Consequently, while the star ratings certainly have a cardinal interpretation—a four star agency is certainly better than a two star agency—the more salient effect might have been a certification effect (i.e. whether or not they received a star) particularly for less well-known agencies.

3 Conceptual Framework

3.1 Migration Intermediaries and Information Frictions

Recruitment agencies perform two key functions: (1) they assist migrants and employers with the transaction costs—e.g. obtaining job order clearances, passports, and visas— and (2) they screen both employers and migrants in order to produce good matches. Fernando and Lodermeier (2022) show that agency and employer fixed effects matter five times as much as country, sector, and time fixed effects in explaining the incidence of migrant abuse in the Gulf region. Whether an agency obtains job order clearances and visas (i.e. transaction costs) is easily observable, but the effort an agency invests in screening is not. Screening may involve an agency scrutinizing an employer's history of abuse or a migrant's contractual history.[25] After a match is realized, the employer and migrant party to this match may learn about the screening effort invested by an agency, but other migrants and employers—the market at large— may not owing to information frictions that restrict the diffusion of this information.

These information frictions may result from agencies working with employers in different countries and, likewise, recruiting migrants from different parts of Sri Lanka and thereby limiting the ability of either party to share and coordinate on this information.[26] Consistent with the presence of such information frictions, we find little evidence of assortative matching between reputable employers—those with low pre-program complaint rates—and high quality agencies prior to the introduction of the ratings program (see Appendix A.13). Consequently, in the absence of a mechanism that aggregates this information on agency quality (e.g. a ratings program), investments in screening cannot be contracted upon and agencies may have little incentive to invest in the quality of their placements.

[25] For examples of how agencies may invest in screening employers, see Panel A of Appendix B.1.

[26] In the migrant survey, less than half the respondents surveyed could remember the name of the agency that recruited them. This is likely because a network of recruiters or `sub-agents' is a migrant's main point of contact and effectively obscures an agency's identity.

3.2 A Two-Period Bayesian Model of Agency Reputations

We adapt a framework developed by Cabral (2005) and model the strategic interaction between foreign employers and recruitment agencies, when information frictions obscure an agency's type (adverse selection) and their screening effort (moral hazard). The model has two periods that mimic the structure of the SLBFE agency ratings program: the first period is the Reputational Incentives (RI) phase, while the second period is the Quality Revelation (QR) phase. The key insight of the model is that the credible threat of quality revelation in the second period can induce agencies to invest in screening and improve placement quality in the first period.

Agencies invest in placement quality because it positively influences employers' beliefs about their type (i.e. their reputation) and raises demand for their services (i.e. job orders) in the future. The model focuses on agency responses during the RI-phase for two reasons. First, as the following sections will reveal, our key results arise prior to quality revelation implying that behavioral responses to the announcement of the program are critical to interpreting our estimates. Second, agencies are the only actors who can respond to the program during the RI-phase as the star ratings are yet to be publicly revealed.

3.3 Setup

Assume there are High (H) type and Low (L) type agencies that facilitate matches between employers and migrant workers in each period in order to fulfill a job order (i.e. recruit a fixed number of workers). Agencies differ in their ability to produce `good' matches, which we define as those where a contract is not prematurely terminated by either party. Further assume that H-type agencies always produce good matches (or placements) but the placement quality of L-type agencies is a function of their screening effort: e.

Screening an employer influences the placement quality of all migrants recruited under the job order. We, therefore, assume an agency's screening effort affects placement quality for the entire job order. This placement quality can either be `good' with probability $p(e)$ or 'bad' with probability $1-p(e)$, where $p(e)$ is concave in effort. Screening improves placement quality, but is costly: $c(e)$ is the cost of effort, which is convex and increasing in e.[27]

The market consists of many risk-neutral foreign employers, who live for one period and would like to enlist the services of agencies to recruit migrant workers.[28] In particular, employers would like to offer a job order to agencies with good placement quality, but avoid agencies with bad placement quality because migrant turnover is costly.[29] However, employers can't contract on an agency's type (H or L) or their effort (e) since this is private information. Instead, employers know the share of H-type agencies in the market q, where $q < 1$, and this informs their first period prior on the likelihood of good placement quality

[27] Specifically, we assume $c(0) = 0$, $c'(0) = 0$, $c''(e) > 0$, $p(0) = 0$, and $p''(e) < 0$.

[28] We make this assumption in order to make the model tractable. While repeated interactions would allow a specific employer to learn about an agency's type, other employers would not benefit from this information.

[29] For example, an employer may invest in a migrant's firm-specific human capital which is lost if a migrant decides to quit their job.

and, consequently, the expected number of job orders (q) offered to an agency.

In each period, an L-type observes the demand (q) for its services from foreign employers and then recruits migrants in order to ful ll these job orders. In both periods, the L-type agency must decide whether to invest effort (e) in screening to increase placement quality. However, as the game ends after two periods, L-type agencies will never invest effort to screen in the second period—rational employers realize this and would, therefore, like to avoid contracting the services of L-type agencies in the second period. Whether or not L-type agencies invest effort in the rst period is the key decision our model explores under two information environments: (1) where employers can't learn about rst period placement quality because of information frictions, and (2) when rst period placement quality is revealed in the second period. The equilibrium concept we use is a perfect Bayesian equilibrium, where the choice of rst period effort by an L-type is a mutual best response given employers beliefs about their effort.

3.4 Equilibrium without the Ratings Program

In the presence of information frictions, rst period placement quality is not revealed to the market. Consequently, an L-type agency will not invest effort (e) in the rst period because it will not in uence employers beliefs about their type nor second period demand. Employers realize this and offer an agency a level of demand ($\bar{q}$) equal to the probability of good placement quality, $Pr(G)$:

$$\bar{q} = Pr(G) = Pr(G|H)Pr(H) + Pr(G|L)Pr(L) = q + (1 - q)p(e) \tag{1}$$

Where $Pr(G|.)$ is the conditional probability of good placement quality given an agency-type and $p(e)$ is the probability of good placement quality for an L-type. Since there is no mechanism to reveal rst period placement quality, employer beliefs about good placement quality are unchanged in the second period. Consequently, employers offer an identical number of job orders ($\bar{q}$) in each period and equilibrium rst period screening $e = 0$.

3.5 Equilibrium with the Ratings Program

Now, the regulator commits to publicly revealing an agency's rst period placement quality in the second period. Moreover, agencies are informed about this credible threat in the rst period, as in the RI-phase of the ratings program. To derive the equilibrium under this information environment, we begin by assuming that $\bar{e}$ represents employers beliefs about an L-type agency's rst period effort. Consequently, an employer would offer q_1 job orders in the rst period, where $q_1 = q + (1 - q)p(\bar{e})$. In the second period, however, employers observe rst period placement quality and update their beliefs using Bayes' rule, where q is the proportion of H-types. :

$$\Pr(H|G) = \frac{\Pr(G|H)\Pr(H)}{\Pr(G)} = \frac{q}{q + (1 - q)\,p(\bar{e})} \tag{2}$$

As noted previously, an L-type agency will not exert effort in the second period, since it is the last period of the game. In the second period, employers condition job orders on revealed first period placement quality. Since H-type agencies never shirk on effort, employers would like to offer them a job order in the second period. If an employer observes bad placement quality [1-p(e)], the agency is revealed as an L-type and does not receive a job order. From an L-type agency's perspective, conditional on employers observing good match quality (p(e)) they believe this agency to be a H-type with $\Pr(H|G)$ as in equation 2. Where q_1 is the fixed level of first period employer demand we noted in the beginning, an L-type agency's expected payoff is:

$$p = q_1 + 1 \quad p(e) \quad \Pr(H|G) + 0 \quad [1 \quad p(e)] \quad c(e) = q_1 + p(e)\Pr(H|G) \quad c(e) \tag{3}$$

As such, an L-type agency picks first period screening effort, e, so as to equate its marginal benefit (increased second period demand) with its marginal cost, $c'(e)$, yielding the following first order condition:

$$\frac{q}{q + (1 - q)\,p(\bar{e})} = \frac{c'(e)}{p'(e)} \tag{4}$$

Given the assumptions on $c(e)$ and $p(e)$, the LHS of equation 4 is decreasing in $\bar{e}$ while the RHS is increasing in e. Consequently, there is a unique interior solution $e = \bar{e}$ such that an L-type agency's first period screening effort (e) is set equal to employers beliefs about its effort ($\bar{e}$).

3.6 Discussion

3.6.1 Mapping Theory to Empirical Results

Our framework demonstrates how easing information frictions can encourage L-type agencies to prospectively invest in placement quality to increase future employer demand. In the empirical results that follow, we find that agencies—especially, low quality agencies—likewise improve placement quality in the RI-phase. Moreover, we also find that agencies with higher ratings receive more job orders suggesting that there is a payoff to investing in their reputation (see section 6.4.2).

To simplify the model, we assume employers are homogeneous and agencies observe a level of demand and then decide whether to invest in screening. However, as we discuss in Section 2.2.3, foreign employers are characterized by a dual market with less abusive Good Employers also offering higher salaries and more amenities than more abusive Bad Employers. This suggests that an important dimension of screening is the type of employer at which a migrant is placed. An agency can improve placement quality by avoiding job orders from Bad

Employers or by filling fewer vacancies on an open job order (i.e. one with unfilled vacancies) from a previously known bad employer. Conversely, they can improve placement quality by increasing the number of migrants they recruit for open job orders from Good Employers, reach out to previously known Good Employers and solicit their business, or increase the likelihood of accepting a new job order from an employer they have screened and believe to be good.

3.6.2 Employer and Migrant Responses

Migrants are assumed to play a passive role as we believe their access to the ratings was extremely limited (see section 6.4.3 for further discussion).[30] We do not model employer responses here because our main empirical strategy allows us to identify agency responses and we do not observe commissions paid by employers to agencies. We note, however, that Good and Bad Employers may value screening differently. In particular, if Good Employers are more likely to provide amenities and invest in training their workers, they may also be especially wary of turnover. Consequently, if they view agency ratings as a credible signal of screening effort, they will be especially responsive to them.

Conversely, Bad Employers may care little about screening if they intend to renege on contractual obligations and expect high turnover. These employers may prefer agencies with lower ratings or agencies that were ineligible for the program. We think the former is unlikely, because most agencies that were eligible for the program received a star and therefore their cardinal interpretation is dominated by the 'certification' effect of whether or not they have a star; Section 8.2 discusses empirical evidence in support of this view. We also think it is unlikely that Bad Employers selected unrated agencies in response to quality revelation; their demand is likely more elastic and they could switch to recruiting migrants from other countries—nevertheless we explore the possibility of such spillovers in section 7.1.

4 Data

We employ data from four sources in our analysis: (1) publicly available data on the agency ratings program, (2) administrative data from the SLBFE covering the period 2005-2015, (3) a detailed survey of 100 recruitment agencies conducted in 2019, and (4) a detailed survey of 246 migrants conducted in 2009.

Administrative Data: The primary data we use are administrative data collected by the SLBFE on the universe of foreign job orders, local agencies, and migrants. We are able to match these data to each other and to the agency ratings data. That is, for every migrant departure through a local recruitment agency, we are able to link the migrant worker to the local agency

[30] The ratings were primarily made available on the SLBFE website and in the migrant survey we conducted, just 3% of surveyed migrants say they used the internet to learn about an agency; this is perhaps unsurprising given that just 12.1% of the Sri Lankan population had access to the internet in 2015—the final year of our dataset (Bank, 2021).

they used and to the speci c job order under which they were recruited. The speci c data we observe are as follows:

- Foreign Job Orders: We observe all job orders that are initiated by an employer and then approved by the SLBFE (data is missing for 2005 and 2014). These data include information on the number of workers requested by an employer (i.e. vacancies) and the salary, and amenities being offered.
- Migrant Trips: We observe the universe of migrant trips through 1510 unique licensed recruitment agencies (1,543,441 trips, 1,066,825 unique individuals) from Sri Lanka to foreign destinations. This dataset includes agency identi ers, migrant characteristics (age, sex, village, and marital status), departure information, the salary paid to migrant and the sector in which they work.
- Complaints: We observe complaints made by migrants or their family members to a Sri Lankan consulate. There are 127,238 complaints from 112,000 unique individuals. Complaints range from breach of contract to sexual harassment (see Section A2 for the composition).

Agency Survey : In June 2019, we conducted a survey of 109 recruitment agencies to obtain a richer understanding of how agencies solicit job orders and recruit migrants and their beliefs about the rating program. We used administrative data to sample an even split between agencies based in Colombo or Kurenegala—the two migration hubs—that recruited above and below 100 migrants.

Migrant Survey We also surveyed 246 prospective and returned migrants in June 2019 to understand worker decision-making regarding migration and awareness of the agency rating program. Our sample is by no means representative, we sampled workers at areas where they are likely to congregate (e.g. temples, schools) using a `right-hand rule' and snowball sampling across two major migrant-sending cities: Colombo and Kurunegala.

5 Empirical Strategy

5.1 Identifying Variation

As agencies were informed in early 2010 about the program, their eligibility is retrospective in that it depends on the number of migrants an agency recruited in 2009. Consequently, agencies can neither select into the program, nor select out of it since they cannot refuse a rating per the program rules. As such, a natural counterfactual are the set of agencies that have been licensed for at least a year but recruited fewer than 100 migrants in 2009. At rst glance, the eligibility threshold may suggest the appropriateness of a regression discontinuity design. However, a key concern is whether agencies that were narrowly eligible or ineligible were on different trends.

Our ideal counterfactual would be the set of agencies who, on average, follow similar trends in recruitment but recruited fewer migrants in 2009 due to the stochastic nature of recruitment. A difference-in-difference design allows us to transparently assess pre-trends for our outcomes of interest and control for theoretically important sources of agency-speci c, time-varying heterogeneity. In addition, given multiple phases of the program, the event study graphs also allow us to assess when program effects occur—a key issue for interpreting our empirical results as we discuss in the conceptual framework in Section 3.

To pick a non-arbitrary set of eligible and comparison agencies located about the 100 migrant threshold, we use an algorithm developed to conduct balance tests on pre-treatment variables by Calonico et al. (2014). In short, the algorithm conducts balance tests for pre-speci ed pre-treatment characteristics using a progressively larger window across the threshold until the test statistic suggests an imbalance.[31] We speci ed variables intended to capture the extent and diversity of agency operations—migrant recruitment, job orders, and complaint intensity between 2005-2008— and nd that the optimal sample contains agencies who recruited between 45 and 155 migrants in 2009. We explore the robustness of our results to the use of alternative samples in Section 7.3.

This sample restriction yields 207 agencies—32% of all active agencies—and Figure 2 plots average migrant recruitment by eligibility status for 6 quarters preceding the introduction of the program. As Figure 2 makes clear, recruitment is highly stochastic. The scatter plot for eligible (red) and comparison (blue) agencies overlap considerably lending credence to the assumption that our identifying variation results from the stochastic nature of recruitment. To wit, had the government instead used the 2008 migrant recruitment total to determine eligibility for the program, 35% of our eligible agencies would have been ineligible and, conversely, 26% of our comparison agencies would have been eligible.

5.2 Difference-in-Differences Speci cation

For agency i in month-year t, our main speci cation is:

$$Y_{it} = d_t + a_i + b_1 \text{Elig} \quad 1[2010 \quad t < 2012]_{it} + b_2 \text{Elig} \quad 1[t \quad 2012]_{it} + t X_{it} + e_{it} \tag{5}$$

In the speci cation above, d_t are month-year xed effects, a_i is an agency xed effect, Elig is an indicator variable for an eligible agency, $1[2010 \quad t < 2012]$ is an indicator variable for the period between January 2010 and May 2012, $1[t \quad 2012]$ is an indicator variable for the period after (and including) June, 2012, and $t X_{it}$ are two sets of time varying controls: the interaction of month-year xed effects with average migrant recruitment between 2005-08, and the interaction of month-year xed effects with the average Quantitative Score between

[31] Speci cally, the test constructs a sequence of nested windows around our 100-migrant eligibility cutoff into the program; we specify a minimum window of 20 migrants to ensure a large enough sample for inference. The algorithm successively increases the window size by 2.5 migrants, conducts binomial tests for our running variable and balance tests for our given covariates. The algorithm then chooses the largest window around the cutoff such that the p-value of the balance test is larger than a pre-speci ed level for all nested (smaller) windows, ensuring these nested windows also cannot reject the null. The p-values are calculated using randomization inference.

2005-08.[32] We cluster standard errors at the agency level. Under the assumption of parallel trends, b_1 is the causal effect of the program during the Reputational Incentives phase, i.e. the period preceding the public release of the star ratings, while b_2 is the casual effect of the program in the period after the Quality Revelation phase, which includes the effects of the RI-phase that precedes it.

Next, we estimate how the program effect varies by (pre-program) agency quality—the key source of differential agency response in our conceptual framework. Ideally, we would compare an eligible agency to an ineligible one of comparable quality that would have received a similar rating had they been eligible. To approximate this ideal, we compute a Predicted Rating for all agencies, using the fact that an eligible agency's 2012 star rating is partly determined by the pre-treatment (i.e. 2009) value of their Quantitative Score, which we are also able to construct for the comparison group (see Appendix A. 10).[33]

The Predicted Rating varies considerably within each group of agencies (see Appendix A.7), allowing us to assess whether the effects of the program vary by pre-program agency quality. To do so, we interact equation (5) with an indicator variable for whether an agency has an above (High) or below (Low) median Predicted Rating

$$Y_{it} = d_t + a_i + g_1 \text{High} \quad \text{Post}_{it} + b_1 \text{High} \quad \text{Elig} \quad \text{Post}_{it} + b_2 \text{Low} \quad \text{Elig} \quad \text{Post}_{it} + t\, X_{it} + e_{it} \tag{6}$$

The terms are as de ned above, however, Post_{it} refers to month-year observations after the introduction of the program. The High Post_{it} term captures the trend for pre-program high quality agencies in the comparison group. [34] In the model above, b_1 is the causal effect of the program for pre-program high quality agencies, while b_2 is the analogous effect for pre-program low quality agencies. We note that since eligible agencies are compared to comparison agencies of the same quality, our estimates cannot be driven by mean reversion (see Section 7.2). Finally, we estimate equation (6) for each program phase, to understand how agencies of varying quality respond to the program in each phase.

[32] Even armed with a set of comparable agencies, we are concerned that agencies of varying size and quality may be subject to different shocks. As such, our main empirical speci cation allows for trends that vary with average pre-program (2005-08) migrant recruitment and the analogous average pre-program Quantitative Score, which we use as a proxy of agency quality: see Appendix A. 10 for details of how the average Quantitative Score is computed. We show in Appendix C. 7 that our results are robust to excluding these trend controls; however, given that eligible agencies, on average, recruit more than comparison agencies we believe this is an important control to include.

[33] We regress an eligible agency's Combined Score—the continuous value that determines star ratings—on their 2009 Quantitative Score. We then use the estimated coef cients to construct a predicted value for the Combined Score for the comparison group which we call a Predicted Rating The Quantitative Score in 2009 is highly predictive of an agency's Combined Score, with a t-stat of 15.05.

[34] When we estimate heterogeneous effects by agency quality, we omit the interaction between month-year xed effects and the average Quantitative Score between 2005-08. This is both because the model already allows for differential trends by quality and so as to avoid issues with collinearity.

5.2.1 Parallel Trends

Our identifying assumption is that eligible agencies and comparison agencies would have followed a parallel trend in the absence of the program. To assess the plausibility of this assumption, we first plot the raw quarterly average of migrants recruited by eligible agencies and comparison agencies (Panel A in Appendix B.5) and see that the trends track each other closely prior to the introduction of the program (vertical blue line). Analogous charts plotting job orders (Panel B) and the complaint rate[35] (Panel C) yield a similar conclusion.

To investigate pre-trends more formally, we test whether we can reject that the lagged effects for a series of outcomes are jointly zero. In Appendix C.1 we find across eight key outcomes that we can't reject the joint null for the lagged effects with p-values ranging from 0.38 to 0.93—we provide plots of these lagged effects in Appendix C.2. Moreover, we also report analogous F-tests for the sample of low and high quality agencies for each outcome and we are invariably unable to reject the null; an exception is the test for Good Orders (i.e. job orders from employers with below median pre-program complaint rates) with low quality agencies. As Roth (2022) discusses, owing to power concerns and variation in outcomes, failure to reject the joint null need not preclude pretends nor would rejecting the null necessarily imply a violation of the assumption. A strength of our design that reduces concerns about power is that the F-tests we report are based on at least 15 lagged periods.

5.3 Summary Statistics

Table 1 provides descriptive statistics for our sample at the agency-month-year level using data from 2005 to 2015. The average firm in our sample recruits 8.23 migrants per month. A majority of these migrants are female, engage in domestic work, and nearly a quarter sought work in Saudi Arabia: the most common destination for Sri Lankan migrants. The median salary for a migrant worker during this period is $ 178 (2015 US dollars); see Section A12 for how salaries vary by major job categories. In terms of demand from foreign employers, an agency receives 0.48 job orders per month, each of which has 72 vacancies on average. We note here that the job order data is missing for 2005 and 2014, so estimates that use this data will have a different sample size. Over our 10-year sample period, we find that on average 8% of migrants placed by an agency made a formal complaint to a consulate. In Appendix A.2, we categorize these complaints: non-receipt of wages (17%), breach of employment contract (13%), and harassment (11%). Agencies resolve 88% of complaints filed, taking 4.6 months on average to reach a resolution.

6 Results

In this section, we report difference-in-difference estimates of the program. We first assess whether eligible agencies invested in the rating criteria and then estimate how the program

[35] The complaint rate is the number of complaints divided by the number of migrants recruited.

in uenced the extent of agency operations (Table 2). Finally, we estimate the effects of the program on placement quality, towards understanding the mechanisms underlying the prior extensive margin effects (Table 3).

6.1 Investment in Rating Criteria

To investigate whether agencies responded to the rating criteria, we construct a contemporaneous monthly version of the Quantitative Score that captures investment in a subset of the rating criteria that we observe for both eligible and comparison agencies (see Section 2.3.3). To recap, the Quantitative Score is a function of a number of criteria including the number of migrants recruited, the skill-level of jobs, and the complaint solve rate (see Appendix A. 10 for details of all criteria and how we constructed this measure).

In Panel A of Table 2, we nd that eligible agencies, on average, have a higher monthly Quantitative Score during the RI-phase (i.e. the coef cient on Elig [2010 12]), but this estimate is somewhat imprecise (t-stat = 1.64). However, in Panel B, we nd that pre-program low quality agencies (i.e. below median Predicted Rating), are signi cantly more likely to invest in the rating criteria: their Quantitative Score increases by 0.66 points, which represents an increase of 11% relative to the pre-program comparison mean. In Panel C, we estimate program effects by agency quality and program phase and nd that the increase in the Quantitative Score for low quality agencies (largely) occurs during the RI-phase.

In examining the components of the Quantitative Score, we do not nd that any speci c criterion drives the observed increase in the Quantitative Score during the RI-phase (See Panel A of Appendix B. 6). While low quality agencies are more likely to send skilled migrants abroad (Panel B), this effect arises after the RI-phase—suggesting it may be driven by employers and/or migrants responding to the ratings—and, therefore, cannot explain the higher monthly Quantitative Score during the RI-phase. One interpretation of these results is that in contrast to examining the effects on a speci c criterion, the monthly Quantitative Score yields a more powerful test of investments in the rating criteria. This is because the monthly Quantitative Score combines the criteria and weights them in a way that re ects the government's priorities.

In sum, our results suggest low quality eligible agencies invest in the rating criteria announced by the government during the RI-phase, consistent with the theoretical prediction for L-type agencies in Section 3. We next consider the consequences of this investment for the extent of agency operations.

6.2 Market Outcomes

In column 2 of Table 2, we nd that the program did not signi cantly affect migrant recruitment by eligible agencies on average (Panel A), nor by pre-program agency quality (Panels B & C); we note, however, that the estimates are generally positive. In column 3, we see that job orders for eligible agencies—a measure of demand from foreign employers—increase by

31% during the RI-phase. After quality revelation, job orders for eligible agencies are 42% higher relative to the pre-program comparison mean. As with investments in the rating criteria examined in the prior section, we again find that the effect on job orders is driven by low quality agencies. Job orders for these agencies increase by 65% relative to the comparison mean. Panels A and D of Figure 3 use event study plots to demonstrate these effects graphically and, additionally, show the absence of pre-trends which is considered in further detail in Panel D of Appendix C. 2.

Importantly, agencies are not passive recipients of job orders. While employers initiate job orders, they may do so as a consequence of finding an agency suited to their needs or because an agency has reached out to them to secure their business. Consistent with agencies taking proactive steps to secure job orders, in Panel C we find that the increase in job orders for low quality agencies begins prior to the QR-phase and persists thereafter. These point estimates are also consistent with the event study graph in Panel D of Figure 3, where the increase in job orders begins immediately after the government announces the rating criteria (right of the vertical blue line). These actions are necessarily a consequence of eligible agency responses, because prior to the the release of the star ratings, employers and migrants do not have any new information to respond to and are likely unaware of the program and (certainly) the rating criteria. [36]

In column 4 of Table 2, we estimate whether the program influenced the survival of eligible agencies in the market. To do so, we construct an `Agency Exit' variable that records whether an agency had 12 consecutive months of zero migrant recruitment. [37] We see that even prior to quality revelation, eligible agencies are 44% less likely to exit the market. [38] The estimated effect is even higher for pre-program low quality agencies, who are 75% less likely to exit the market relative to similar comparison agencies; low quality agencies are both less likely to exit prior to quality revelation and after it (column 4, Panel C). In contrast, we do not find that the program affects the likelihood of high quality agencies exiting the market.

The agency exit effects in the QR-phase are consistent with the theoretical predictions of the model in Section 3. We show that L-type agencies invest in the rating criteria so as to secure more job orders once quality is revealed. Though we lack cash flow data that would permit a thorough analysis, the increase in job orders for low quality eligible agencies during the QR-phase (column 3, Panel C), likely increases their revenues and likelihood of exit.

However, low quality eligible agencies are also less likely to exit before the star ratings are revealed. This immediate payoff may instead result from agencies investing in the qualitative score that captures compliance with recommended management practices (see Section 2.3.3); de Oliveira (2022) finds a similar result on the effect of management practices on firm survival. As agencies ineligible for the program were not audited by third-party inspectors, we are

[36] We note here that the authors were unable to learn about how the rating criteria mapped to the star ratings even after scrutinizing official SLBFE documents on the program. To learn about the latter, we had extensive discussions with a bureaucrat who designed the program.

[37] These results are robust to using a 6 month, 12 month or 24 month definition instead. See Appendix C. 3.

[38] Note, the comparison mean here is post-program as it is underfined in the period prior to the introduction of the program.

unable to test this hypothesis with a plausible counterfactual. However, we do find that having a higher Qualitative Score is correlated with being less likely to exit the market for low quality eligible agencies (t-stat = -2.69).[39]

Given that eligible low quality agencies invest in the rating criteria and have more job orders in both phases of the program, it is perhaps surprising that they do not increase the number of migrants recruited. One possibility here, is that low quality agencies are prioritizing the quality of their placements rather than the extent of their operations which we examine in the following section.

6.3 Placement Quality

In Table 3, we assess whether the program influenced a series of proxies for placement quality. Importantly, we adjust all outcomes for agency exit and, therefore, our sample sizes are smaller than in Table 2.[40] In column 1, we find that the program increased the average salary of migrants placed by eligible agencies in both phases of the program, but these estimates are imprecisely estimated (t-stat 1.3). Turning our attention to effect heterogeneity by agency quality (Panel B), we find that pre-program low quality agencies place migrants in jobs with significantly higher salaries ($ 20 or 12%). This increase in average salaries is driven by migrants placed by low quality agencies during the RI-phase (Panel C), consistent with event study graph in Panel E of Figure 3.

An alternative measure of placement quality is to assess whether migrants placed by agencies decide to renew their contracts with the same employer, as turnover is an important proxy of match quality. In column 2, we compute the program effect on `Contract Renewal' which reports the share of migrants placed by an agency who renew their contract with an employer. While we do not find that the program had an average effect on contract renewal (Panel A), we find that migrants placed by low quality eligible agencies are significantly more likely to return to the same employer after the program; the effect size is substantial and reflects an increase of 15% relative to the pre-program comparison mean (Panel B). Panel C shows that the contract renewal effect largely arises after quality revelation; this is likely due to the fact that contracts of migrants placed during the RI-phase would only be renewed two to three years later (i.e. the standard contract length), during the QR-phase.

In column 3, we turn our attention to the quality of foreign employers at which migrants were placed.[41] We use pre-program information on these employers to identify job orders originating from less abusive Good Employers (i.e. those with below median pre-program complaint rates) and find that the program resulted in a 31% increase in such job orders (Good Orders) for eligible agencies prior to quality revelation. While both high and low

[39] This regression includes all eligible agencies with below median pre-program agency quality. The regression clusters at the agency level and includes month-year fixed effects interacted with pre-program recruitment (2005-08) and average yearly Quantitative Score (2005-08).

[40] Adjusting the previously reported estimated effects on rating criteria, job orders, and migrant recruitment for exit reduces their precision but does not change our qualitative findings (see Appendix C.6).

[41] Our data do not allow us to distinguish between an individual employer (e.g. a home) and a foreign recruitment agency that places domestic workers in households; we refer to both of these groups as foreign employers.

quality eligible agencies have signi cantly more Good Orders (Panel B), the increase in Good Orders for low quality eligible agencies occurs during the RI-phase, but after it for their high quality counterparts (Panel C).

Finally, we use an index of contract quality that aggregates the provision of contractual amenities such as health insurance, accommodation, and return airfare to assess the quality of placements. We nd that the program increases contract quality by 0.17 standard deviations for low quality agencies (Panel B) and this increase occurs during the RI-phase (Panel C) as demonstrated graphically in Panel F of Figure 3.[42]

An alternative way to assess placement quality is to estimate the effect of the program on the complaint rate. In Appendix C. 5, we do not nd that the program in uenced the complaint rate in either phase of the program or by agency quality. Our design is likely under-powered to detect changes in the incidence of complaints owing to considerable variation even within an employer-agency match owing to idiosyncratic factors. [43]

Collectively, these results suggest the program induced eligible agencies—and, especially, pre-program low quality agencies—to improve their placements prior to quality revelation. We again note that this nding is consistent with predictions of our model for L-type agencies in Section 3. In the model, the mechanism by which L-type agencies improve placement quality is by screening employers. In the following section, we assess the evidence in support of this mechanism.

6.4 Discussion

6.4.1 Agency Screening during the Reputational Incentives Phase

Why doesn't the increase in job orders for eligible agencies result in additional migrant recruitment? Job orders contain multiple vacancies, many of which may remain un lled over time. Consequently, additional job orders need not imply an expansion in migrant recruitment. An agency intent on prioritizing placement quality may restrict the quantity of migrants placed by selectively lling vacancies for job orders from more reputable employers.

Figure 4 provides a visualization of how matches between agencies and Good Employers—those who have a below median pre-program complaint intensity—evolve with the introduction of the program. Each panel plots a series of distributions showing the density of Good Employers matching with agencies of varying quality, as measured by their Combined Score/2012 agency rating (x-axis). Although we plot all density functions with respect to the 2012 Combined Score, the agency ratings only became public after 2012, i.e. the green density functions in Panel C. Consequently, the density functions in Panel A and B visualize matches between Good Employers and agencies in an information environment where the ratings are

[42] Appendix B. 3 computes program effects for each of the constituent variables of the contract quality index. We again nd that low quality agencies drive the results and place migrants at employers who are more likely to provide them with higher salaries, accommodation, and health insurance.

[43] Using a migrant-year level dataset, agency x employer xed effects explain ve times as much variation in whether a migrant makes a complaint as country, sector, and time xed effects, but they jointly yield an adjusted r-squared of just 0.25 (Fernando and Lodermeier, 2022).

yet to be made public.

In Section 3, we argue that information frictions obscure the quality of agencies prior to quality revelation. The red line in Panel A of Figure 4 supports this assumption by showing that Good Employers do not systematically match with higher quality agencies prior to the introduction of the program. However, Panel B shows that the (darker) blue line corresponding to the second year of the RI-phase (i.e. matches in 2011) shifts its mass towards agencies with a higher SLBFE rating. Two points are of note here: first, the only actors in the market that can influence this shift are agencies as the star ratings are only revealed in the QR-phase (i.e. the green density functions). Second, this shift in the mass is driven by agencies who eventually ended up with more than 50 points on their Combined Score, corresponding to agencies that received 3 stars or better.

This graphical evidence is consistent with an average increase in Good Orders in column 3 in the period prior to quality revelation for low quality agencies (Panel C of Table 3). In Appendix B. 2, we examine the characteristics of employers linked to job orders. We find that low quality eligible agencies increase the number of unique employers at which they place migrants during the RI-phase (column 1). However, these employers were previously known to the agencies and not new employers with whom the agencies had no prior interaction (column 2). Consequently, it is within this intensive margin of previously known employers, that low quality agencies improve the placement quality of the migrants they recruit, resulting in an increase in the number and share of Good Orders during the RI-phase.[44] In contrast, we do not find evidence suggesting that agencies screened the supply-side (i.e. migrants). In Appendix B. 6 and Appendix B. 4, we find no evidence to suggest compositional changes in the migrants recruited by eligible agencies prior to quality revelation.

Collectively, these results suggest eligible agencies screen job orders to improve placement quality prior to the revelation of the star ratings. In support of the plausibility of this mechanism, the agency survey shows that approximately half (55) of the 109 agencies surveyed report rejecting a job order in the last year; of those agencies who report rejecting, 52% cite the reason for rejection as being related to safety concerns related to the reputation of the employer (see Panel B of Appendix B.1).

6.4.2 Employer Response to the Star Ratings

Foreign employers are the most likely to have access to the star ratings. The manner in which ratings were made available—through a searchable database on the SLBFE website (see Appendix A. 8)—makes them far more accessible to employers relative to migrants who largely lack access to the internet. As we discuss in Section 2.2.2, employers are required to fill out several forms and enter details on the SLBFE's online information management system, so they have good reason to access this website. In addition, interviews with consular officials suggest that foreign missions sometimes provided prospective employers with lists

[44] In contrast, there is no change in either the share or number of Bad Orders, viz. job orders from employers with below median pre-program complaint rates.

of recruitment agencies and their star ratings, though this was not a uniform practice.

In Panel C of Figure 4, we find that in the years after the star ratings are revealed (i.e. the green lines of varying intensity), there is a noticeable intensification of the likelihood of a match between a Good Employer and a highly rated agency; Appendix A. 15 shows that we do not find similar evidence that Bad Employers responded to the ratings. To more rigorously test if employers respond to the star ratings, we use a regression discontinuity design (RDD) that estimates the effect of an agency receiving a marginal star. In particular, we use discontinuities in the star-specific cut-offs that map the Combined Score—the continuous value that combines the Quantitative Score and the Qualitative Score—to a star rating. We compare agencies just above and below the cut-offs and pool the thresholds to arrive at an estimate of the causal effect of receiving a marginal star.[45].

We report the results in Table 4 and find that an additional star more than doubles the monthly number of job orders (significant at the 5% level). We interpret this result as providing evidence of employer response to the star ratings but, in addition, we note this result also supports the structural assumptions of the conceptual framework in Section 3, by showing that there is a payoff in terms of demand to investing in a higher rating. In the following section, we address the possibility that the star rating regression discontinuity design also capture migrant responses.

6.4.3 Migrant Response to Star Ratings

In theory, migrants may have responded to the star ratings and influenced placement quality. However, we believe this is unlikely because the primary manner in which the SLBFE chose to disseminate the ratings was through their website. In 2015 (the final year of our dataset), just 12.1% of the entire Sri Lankan population had access to the internet, making it very unlikely that prospective migrants had access to the ratings (Bank, 2021). Nevertheless, using the same RDD discussed in the previous section, we test whether migrants respond to higher ratings. In Table 4, we find that migrant recruitment (column 2) and the share of domestic workers (column 3) do not respond to an increase in the star rating of an agency. Migrants who made complaints prior to the ratings program may be the most sensitive to the ratings, but we find that agencies with higher ratings are no more likely to recruit such migrants. Finally, when shown pictures of the awards and an advertisement that included a rating, just 11-14% of respondents in the migrant survey recognized either ($n = 240$).

[45] We detail this specification in Appendix B. 7. In short, where $Treat_i$ is an indicator variable that switches on for the right-hand side of the window, using a uniform kernel with the MSE optimizing bandwidth we report b_2 for agency i in month m and year t:

$$Y_{imt} = d_t + b_1 Combined\ Score_i + b_2 Treat_i + b_3 Treat \times Combined\ Score_i + e_{imt}$$

7 Robustness

7.1 Spillover Effects

Agencies are often clustered in the same area in large cities as, by law, they are not allowed to have multiple physical premises. Panel A of Appendix A. 11 shows that many agencies are located in two large cities (Colombo and Kurunegala), while Panel B shows clusters of agencies in Colombo District. The results in the previous sections suggest that eligible agencies secure more job orders and are less likely to exit the market. When we estimate heterogeneous program effects by (pre-program) local market competition, we nd even more pronounced effects for eligible agencies in competitive markets (see Appendix B. 8). Consequently, these program effects may result from eligible agencies poaching the job orders and/or supply of migrants available to comparison agencies. To test this hypothesis, in Table 5, we estimate whether the presence of an eligible agency in a 100 foot radius of a comparison agency in uences outcomes. The speci cation includes a xed effect of the local market interacted with year xed effects and, in addition, we cluster standard errors at the local market level. [46]

We nd that the presence of an eligible agency results in a signi cant increase in migrant recruitment (column 1) for comparison agencies and also increases their job orders though this estimate is (marginally) imprecise (t-stat = 1.64). We also do not nd that comparison agencies in close proximity to eligible agencies are more likely to exit the market or less likely to obtain job orders from Good Employers. These results suggest that the positive market outcomes that eligible agencies experience as a consequence of the program do not come at the expense of comparison agencies. If anything, the positive spillovers—perhaps owing to high xed costs in securing new job orders which are then shared with neighboring agencies—on migrant recruitment and job orders suggest that we are underestimating the effects of the program.

7.2 Agency Quality & Mean Reversion

Measurement error in our classi cation of pre-program agency quality may imply that the heterogeneous effects we report re ect mean reversion rather than treatment effects. For example, it is not that the program improves outcomes for low quality agencies, but that these were simply high quality agencies that were improperly classi ed and who then trended back to their mean. The key argument against mean reversion is that high and low quality eligible agencies are always being compared to high and low quality comparison agencies (see Appendix A. 7). As such, any measurement error would apply similarly to both eligible and comparison agencies and the heterogeneous treatment effects we report, therefore, capture the differential effect net of any measurement error. We discuss this point in more detail in Appendix C. 8. In addition, we assess the robustness of our main results to two alternative

[46] We consider the 78 Divisional Secretariats that we observe in our data—the administrative unit below a district—as roughly corresponding to a local market. Roughly 40% of the 78 Divisional Secretariats do not have any eligible agencies. As such, the effects are likely being estimated off the set of local markets with variation in eligibility status where agencies are clustered as in Panel B of Appendix A. 11.

measures of pre-program agency quality: the Quantitative Score for 2009 and the average Quantitative Score for 2007-09). Overall, across all three measures of agency quality our estimates have a similar magnitude albeit with varying precision in the case of Good Orders and contract quality (see Appendix C. 8).

7.3 Sample Restrictions

For our main difference-in-difference speci cation, the algorithm we used led us to restrict our sample to agencies that recruited between 45-155 migrants in 2009. We chose this method of selecting the agencies both so to pick agencies that were comparable across their eligibility status and so as to `tie our hands' as to the speci c interval about 100 chosen to restrict the sample. In Appendix C. 9, we explore how our estimates vary when comparing the sample suggested by the algorithm (Panel A), to a sample of agencies that recruited between 25 and 175 migrants in 2009 (Panel B), and then consider all agencies (Panel C). In general, the estimated effects are qualitatively similar across all three samples but, unsurprisingly, the less restrictive samples yield substantially more precise estimates.

8 Discussion: Did the Program Improve Market Ef ciency?

In the presence of information frictions, adversely selected labor market intermediaries may exploit their informational advantage to the detriment of both workers and employers (Autor, 2008). A policy that relaxes these frictions may both provide intermediaries with incentives to improve their match quality and result in the exit of low quality intermediaries from the market as in Lee (2007). In assessing the effect of the program on market ef ciency, we rst focus on the quantity and quality of migrant placements. In particular, we analyze how the program affects the extensive margin (i.e. the extent of migration) and the intensive margin (i.e. placement quality) towards combining these effects to assess changes to match surplus. We then discuss how the program affects the entry and exit of agencies and employers.

8.1 Extent of Migration, Placement Quality, and Match Surplus

Inference about how the program in uences the extent of migration is complicated by the fact that the overall trend in the market is one of steady decline—see Panel A of Appendix B.5. This decline is in part due to the partial ban on migration discussed in Section 2.2. However, our estimates suggest that the program may have increased the amount of migrants recruited by eligible agencies, though these estimates are imprecise. If migrants are scarce, an increase in recruitment by eligible agencies may come at the expense of comparison agencies. Yet, when we explore this possibility in section 7.1 we nd that comparison agencies located near to eligible agencies increased the amount of migrants they recruited. Overall, our results suggest that the program may, in anything, have increased the volume of trade.

Though the effect of the program on the extent of migration is inconclusive, our ndings

suggest the program was successful in improving the intensive margin or placement quality. The results in Table 3 suggest eligible agencies place migrants with employers that pay higher salaries, provide more contractual amenities, and are less abusive. Unsurprisingly, this results in more contract renewals. Moreover, we do not find evidence to suggest that these improvements came at the cost of some migrants. Eligible agencies are not induced by the program to alter the composition of migrants placed abroad (Table B. 4) and, likewise, whether or not a comparison agency is exposed to competing eligible agency does not change their likelihood of placing migrants with Good Employers (Table 5).

In assessing the effect of the program on match surplus we must jointly consider how the program jointly affected both the extensive (extent of migration) and intensive margins (placement quality). To create a proxy for overall match surplus that combines these margins, we note again that a consequence of a dual-market characterized by monopsony power is that less abusive employers also tend to offer higher salaries and more amenities both within and across industries (see Section 2.2.3). Consequently, a proxy for the match surplus generated by an agency is sum of the salaries of migrants they have placed (total salaries). In Appendix C.4, we find that total salaries for eligible agencies increases by 69% prior to the QR-phase and stays 71% higher thereafter (column 1); the effect is larger for pre-program low quality agencies (109%) and positive though not significant for high-quality agencies. We also do not find evidence for negative spillover effects—to the contrary, total salaries are higher for comparison agencies in close proximity to eligible agencies, though this estimate is imprecise (t-stat = 1.26).

8.2 Agencies and Employers

Our results suggest the program left migrants better off with both higher quality and, likely, more placements. We weigh these gains in match surplus against entry and exit for agencies and employers, respectively. First, we find that eligible agencies are less likely to exit the market both prior to and after the QR-phase in Panel A of Table 2. Pre-program low quality eligible agencies drive this result, while their high quality counterparts are unaffected (Panel B). Even though eligible agencies are less likely to exit the market, this does not appear to come at the cost of comparison agencies. Whether or not a comparison agency is located near to an eligible agency does not influence their likelihood of exit (Panel B, column 2 of Table 5). While the program effects on agency exit are large, this reflects the declining trend in the market. By 2015, 23% of comparison agencies (both high and low quality) have exited the market, and this is equal to the percentage of high quality eligible agencies that exit the market (23%)—in contrast, just 6% of low quality eligible agencies exit the market.

Turning to employers, we first note that we neither observe their exit and entry decisions nor the commissions they pay agencies. Consequently, our ability to assess the effects of the program on employers is limited. However, we note that Good Employers are more likely to match with highly rated agencies after quality revelation (Panel C of Figure 4). While this result suggests the program may leave them better off, if competition intensifies for the services of highly rated agencies, they may also have to increase the commissions paid to

these agencies. In contrast, we think the elasticity of demand for Bad Employers is more inelastic; these employers likely value match quality less and may recruit non-Sri Lankan migrants in response to the agency ratings program. In support of this view, we find little evidence to suggest that Bad Employers responded to quality revelation in Appendix A.15. In sum, while our results suggest the program benefited migrants and agencies, we are unable to assess its effects on employers.

9 Conclusion

Labor abuse experienced by international migrants is a phenomenon of great human concern, yet it is understudied in the economics literature. This, in part, owes to the politically sensitive nature of the topic and the difficulty of collecting reliable data on migrant outcomes. We leverage a rich dataset that allows us to link individual migrants to local recruiters, foreign employers and, ultimately, whether or not they sought assistance from a Sri Lankan consulate for an employment-related dispute.

In the Quality Revelation phase, our results suggest that the ratings program acts similarly to a certification program. Rather than conveying a precise signal of quality, to an external observer or a prospective employer, agencies that carry any star rating—in contrast to a mass of agencies who do not—may appear more reputable. However, by inducing eligible agencies to care about their reputation, the program improved placement quality. Migrant sending countries typically have limited power to influence labor laws in destination countries and must usually resort to pre-departure training or banning migration to curb abuses experienced by their citizens. We show that a ratings system for local intermediaries is an important addition to the policy toolkit for these countries.

A limitation of the program, however, is that it does not influence the incidence of employer abuse so much as the burden on whom it falls. While placement quality for Sri Lankan migrants improves, this must come at the expense of non-Sri Lankan migrants placed at Bad Employers. In addition to understanding these reallocation effects, our results suggest a number of directions for further inquiry. Are migrant intermediaries able to extract surplus from migrants, limiting the benefits of large wage differentials they experience across borders? As labor intermediaries do not engage in arbitrage, the lack of prices through a supply chain limits the use of `pass-through' as a way to deduce market structure and estimate surplus extraction as in (Atkin and Donaldson, 2021). Future work may benefit from imposing structure on the process of matching so as to estimate how intermediary market power relates to surplus extraction.

References

Abu-Habib, L. (1998). The use and abuse of female domestic workers from sri lanka in lebanon. Gender & Development, 6(1):52–56.

Agunias, D. R. (2011). Running in circles: Progress and challenges in regulating recruitment of lipino and sri lankan labor migrants to jordan.

Alfano, S. (2010). Sri lankan maid accuses saudi bosses of hammering 24 nails into her body because she was complaining. Daily News.

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

Andrabi, T., Das, J., and Khwaja, A. I. (2017). Report cards: The impact of providing school and child test scores on educational markets. American Economic Review, 107(6):1535–63.

Aryasinha, R. (2020). The role of foreign employment in the economic revival of sri lanka; challenges and opportunities. Ministry of Foreign Relations.

Atkin, D. and Donaldson, D. (2021). Who's getting globalized? the size and implications of intra-national trade costs.

Autor, D. (2008). The economics of labor market intermediation: an analytic framework.

Bank, W. (2021). World development indicators 2021. The World Bank.

Bardhan, P., Mookherjee, D., and Tsumagari, M. (2013). Middlemen margins and globalization. American Economic Journal: Microeconomics, 5(4):81–119.

Batista, C. and McKenzie, D. J. (2021). Testing classic theories of migration in the lab.

Bazzi, S., Cameron, L., Schaner, S., and Witoelar, F. (2021). Information, intermediaries, and international migration.

Benson, A., Sojourner, A., and Umyarov, A. (2019). Can reputation discipline the gig economy?: Experimental evidence from an online labor market. Management Science.

Bergquist, L. F. and Dinerstein, M. (2020). Competition and entry in agricultural markets: Experimental evidence from kenya. American Economic Review, 110(12):3705–47.

Bulow, J. I. and Summers, L. H. (1986). A theory of dual labor markets with application to industrial policy, discrimination, and keynesian unemployment. Journal of labor Economics, 4(3, Part 1):376–414.

Cabral, L. M. (2005). The economics of trust and reputation: A primer. New York University and CEPR, 2005.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014). Robust nonparametric con dence intervals for regression-discontinuity designs. Econometrica, 82(6):2295–2326.

Clemens, M. A. (2013). Why do programmers earn more in houston than hyderabad? evidence from randomized processing of us visas. American Economic Review, 103(3):198–202.

Clemens, M. A. (2018). Testing for repugnance in economic transactions: Evidence from guest work in the gulf. The Journal of Legal Studies, 47(S1):S5–S44.

de Oliveira, P. (2022). Why businesses fail: Underadoption of improved practices by brazilian micro-enterprises. Mimeo.

Demaret, L. (2006). Private employment agencies: The challenges ahead from the workers' perspective. Merchants of Labour, page 159.

Dinkelman, T. and Mariotti, M. (2016). The long-run effects of labor migration on human capital formation in communities of origin. American Economic Journal: Applied Economics, 8(4):1–35.

Dranove, D. and Jin, G. Z. (2010). Quality disclosure and certification: Theory and practice. Journal of Economic Literature, 48(4):935–963.

Fernando, A. N. and Lodermeier, A. (2022). Understanding adverse outcomes in gulf migration: Evidence from administrative data from sri lanka. International Migration Review.

Frantz, E. (2008). Of maids and madams: Sri lankan domestic workers and their employers in jordan. Critical Asian Studies, 40(4):609–638.

Gibson, J. and McKenzie, D. (2014). The development impact of a best practice seasonal worker policy. Review of Economics and Statistics, 96(2):229–243.

Gibson, J., McKenzie, D., Rohorua, H., and Stillman, S. (2017). The long-term impacts of international migration: Evidence from a lottery. The World Bank Economic Review, 32(1):127–147.

Grant, M. and Startz, M. (2019). Cutting out the middleman: The structure of chains of intermediation.

Hendricks, L. and Schoellman, T. (2018). Human capital and development accounting: New evidence from wage gains at migration. The Quarterly Journal of Economics, 133(2):665–700.

HRW (2007). Exported and exposed: Abuses against sri lankan domestic workers in saudi arabia, kuwait, lebanon, and the united arab emirates.

HRW (2010). Rights on the line. Human Rights Watch.

ILO (2013). Operational manual for labour sections of sri lankan diplomatic missions in labour receiving countries.

International, A. (2020). Reality check 2020. Countdown to the 2022 world cup migrant workers' rights in qatar.

Jayawardhane, K. (2018). Saudi to repatriate sl housemaid's body following postmortem report. Daily Mirror .

Jin, G. Z. and Leslie, P. (2003). The effect of information on product quality: Evidence from restaurant hygiene grade cards. The Quarterly Journal of Economics, 118(2):409–451.

Jureidini, R. and Moukarbel, N. (2004). Female sri lankan domestic labour in lebanon: contractual, slavery-like practices and conditions. Journal of Ethnic and Migration Studies, 30(4):581–607.

Lee, W. (2007). Private deception and the rise of public employment of ces in the united states, 1890-1930. Technical report, National Bureau of Economic Research.

Luca, M. (2016). Reviews, reputation, and revenue: The case of yelp. com.

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

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

McKenzie, D., Stillman, S., and Gibson, J. (2010). How important is selection? experimental vs. non-experimental measures of the income gains from migration. Journal of the European Economic Association, 8(4):913–945.

Mitra, S., Mookherjee, D., Torero, M., and Visaria, S. (2018). Asymmetric information and middleman margins: An experiment with indian potato farmers. Review of Economics and Statistics, 100(1):1–13.

Naidu, S., Nyarko, Y., and Wang, S.-Y. (2016a). Monopsony power in migrant labor markets: evidence from the united arab emirates. Journal of Political Economy, 124(6):1735–1792.

Naidu, S., Nyarko, Y., and Wang, S.-Y. (2016b). Monopsony power in migrant labor markets: evidence from the united arab emirates. Journal of Political Economy, 124(6):1735–1792.

Orrenius, P. M. et al. (2001). Illegal immigration and enforcement along the us-mexico border: An overview. Economic and Financial Review-Federal Reserve Bank of Dallas, (4):2–11.

Peru, D. A. (2022). Women's labor market opportunities and fertility decisions: Evidence from sri lanka.

Roth, J. (2022). Pre-test with caution: Event-study estimates after testing for parallel trends. American Economic Review: Insights.

Shivakoti, R., Henderson, S., and Withers, M. (2021). The migration ban policy cycle: a comparative analysis of restrictions on the emigration of women domestic workers. Comparative Migration Studies, 9(1):1–18.

Shrestha, M. (2017). Get rich or die tryin': perceived earnings, perceived mortality rate and the value of a statistical life of potential work-migrants from nepal.

Shrestha, S. A. and Yang, D. (2019). Facilitating worker mobility: A randomized information intervention among migrant workers in singapore. Economic Development and Cultural Change, 68(1):63–91.

Silva, D. (2013). Rizana nafeek executed. Daily Mirror .

Somarathna, R. (2011). Alleged sexual abuse of workers. Daily News.

Somarathna, R. (2015). Lankan maid held captive for 16 years in saudi arabia returns. Daily News.

Startz, M. (2016). The value of face-to-face: Search and contracting problems in nigerian trade. Available at SSRN 3096685.

Theoharides, C. (2018). Manila to malaysia, quezon to qatar international migration and its effects on origin-country human capital. Journal of Human Resources, 53(4):1022–1049.

Yavaş, A. (1994). Middlemen in bilateral search markets. Journal of Labor Economics, 12(3):406–429.

Figure 1: Program Timeline

Panel A: Program Phases

- Reputational Incentives (RI) Phase (2010-2012):
  - January 2010: Program announced to agencies and booklet with grading criteria is sent through SLBFE management system
  - July 2010 - May 2011: Third-party inspections
- Quality Revelation (QR) Phase (2012-2015):
  - February, 2012: Ratings announced at public ceremony
  - June, 2012: Ratings made available on online database

Panel B: Timeline with Potential Effects

Notes: Panel A of Figure 1 shows a simple timeline (time on vertical axis) for the agency ratings program and lists the potential effects below it (time on horizontal axis). In Panel B, the timeline with potential effects illustrates the difference-in-difference design. We compare a set of eligible agencies to comparison agencies below the 100 migrant recruitment threshold in 2009. After the announcement, the rating criteria (signi ed by the notebook icon) are announced to all agencies, and third-party inspections take place shortly thereafter. The ratings are released to the public in February, 2012. We refer to the period between the program announcement and the public release of ratings as the 'Reputational Incentives' phase of the program; in theory all agencies can respond to these criteria but we expect they are especially binding on eligible agencies. We refer to the period after the public release of the ratings as the 'Quality Revelation' period. This period, of course, also includes the effects of RI. The blue and red shading denotes agencies of varying pre-program quality.

Figure 2: Quarterly Variation in Recruitment

Notes: This gure shows quarterly recruitment at the agency level for the period determining eligibility (four quarters in 2009) and two quarters prior to 2009. Attention is restricted here to agencies that recruited between 45 and 155 migrants in 2009, where eligible agencies recruited at least 100 migrants.

Figure 3: Event Study Graphs

Panel A: Job Orders

Panel B: Avg. Salary

Panel C: Contract Quality

Panel D: Job Orders - Low

Panel E: Avg. Salary - Low

Panel F: Contract Quality - Low

Notes: These gures show event study graphs for three outcomes: the number of job orders, the average salary of migrants placed and an index of contract quality. They plot the interaction coef cients—an indicator for an eligible rm interacted with a series of quarterly indicators for 2008-2015. Panels A-C consider all agencies, while Panels D-F restrict attention to pre-program low quality agencies. The capped dotted lines are the 95% con dence intervals. In each case the reference is the 6 month period preceding 2010. The blue vertical line is the point just before the program is announced, the green vertical shows the point just before the ratings are announced to the public. In each case, the vertical lines are shifted half a quarter to the left, so as not to obscure a point estimate.

Figure 4: The Distribution of Matches to Good Employers by Agency Quality

Notes: This gure plots a series of probability density functions that describe which recruitment agencies are matched to `Good Employers': de ned as employers who have a below median pre-program ( < 2010) complaint rate. A match here is de ned as whether an agency sends a migrant to a `Good Employer' in a speci c year. The domain is the continuous Combined Score/ 2012 agency rating which is only known to the public after 2012. Panel A plots the density function for matches occurring prior to the program (red line) and during the rst year of the program (2010), i.e. the Reputational Incentives (RI) phase (the blue line). Panel B adds 2011, the second year of RI. Panel C plots the data for the remaining years (2012-2015) of the Quality Revelation (QR) period (the green lines).

Table 1: Agency-Month Level Summary Statistics, 2005-2015

	Mean	SD	Median	N
# Migrants	8.23	20.8	4.00	27324
# Female Migrants	4.95	8.8	2.00	27324
# Domestic Worker Migrants	5.02	9.7	2.00	27324
# Migrants to Saudi Arabia	1.81	8.0	0.00	27324
Salary (USD), 2015 dollars	178.80	118.4	176.92	23777
# Job Orders	0.48	0.9	0.00	22356
# Vacancies	68.30	177.2	0.00	22356
# Complaints	0.66	1.6	0.00	27324
Complaint rate	0.08	0.1	0.00	19659
Complaint solved rate	0.88	0.3	1.00	6765
Complaint solved time	4.62	5.1	3.00	6765

Notes: This table reports summary statistics at the agency-month level. Salaries are converted to 2015 US dollars. Job Orders are the number of orders that an agency received in a speci c month from employers abroad (i.e. demand), while vacancies reports the average number of vacant positions contained in a job order. Note, data for 2005 and 2014 is missing for the job order data. Complaints reports the number of complaints made to Sri Lankan consulates by migrants who were placed by agencies in the relevant month. The complaint rate measures the average number of complaints by migrants placed by an agency in a given month conditional on the agency placing non-zero migrants. The complaint solved rate and time further condition on the set of agencies who had non-zero complaints in a given month.

Table 2: Investment in Rating Criteria and Market Outcomes

	(1) Quant Score	(2) Migrants	(3) Job Orders	(4) Agency Exit
	Panel A: Effects by Program Phase			
Elig*2010-12	0.3426	0.1991	0.1067*	-0.0701***
	(0.209)	(1.585)	(0.062)	(0.026)
Elig*2012-15	0.2537	3.3784	0.1451*	-0.0826*
	(0.195)	(3.266)	(0.078)	(0.044)
	Panel B: Effects by Agency Quality			
Elig*High*Post-2010	-0.3422	2.1827	-0.0105	0.0067
	(0.265)	(5.407)	(0.114)	(0.074)
Elig*Low*Post-2010	0.6574***	1.8596	0.2204***	-0.1256***
	(0.200)	(2.093)	(0.083)	(0.039)
	Panel C: Effects by Agency Quality and Program Phase			
Elig*High*2010-12	-0.0722	-4.4285	-0.0620	-0.0234
	(0.346)	(3.392)	(0.098)	(0.053)
Elig*High*2012-15	0.3984	5.9194	0.0288	0.0237
	(0.335)	(7.754)	(0.145)	(0.093)
Elig*Low*2010-12	0.6715***	2.3932	0.2151***	-0.0923***
	(0.240)	(2.009)	(0.078)	(0.031)
Elig*Low*2012-15	0.3577*	1.5580	0.2243**	-0.1444***
	(0.212)	(2.224)	(0.094)	(0.046)
N	27324	27324	22356	27324
Comparison Mean	5.80	7.61	0.34	0.16

Notes: This table reports difference-in-difference estimates on the overall effects of the program. Job order data for 2005 and 2014 are missing (i.e. column 3). `Comparison Mean' reports the pre-program (2005-2009) mean of the dependent variable for the comparison group. Panel A compares eligible agencies to comparison agencies during the period where agencies were exposed to Reputational Incentives (2010-2012) and the period after Quality Revelation (2012-2015). Panel B computes heterogeneous effects by pre-program agency quality; `High' denotes above median pre-program quality, while `Low' is below median pre-program quality, where quality is based on the `Predicted Rating' variable (See Appendix A. 10 for details). Panel C reports effects by program phase and by pre-program agency quality. In column 1, 'Quant Score' summarizes an agency's investment in the rating criteria introduced by the government (See Appendix A. 10). In column 2, `Migrants' reports the total number of migrants recruited by an agency in a given month. `Job Orders' are the number of orders that an agency received in a speci c month from employers abroad (i.e. demand). `Agency Exit' is a proxy for an agency closing down. It is coded as an indicator variable that switches on at the beginning of a 12 month sequence of zero recruitment after the program was introduced. All speci cations cluster at the agency level and include agency xed effects. Panels B & C contain month-year xed effects interacted with average migrant recruitment (2005-08). Panel A contains the same controls as Panel B and month-year xed effects interacted with the average yearly Quantitative Score (2005-08).

Table 3: Placement Quality

	(1) Avg. Salary (USD)	(2) Contract Renewal	(3) Good Orders	(4) Contract Quality
	Panel A: Effects by Program Phase			
Elig*2010-12	11.4884	0.0093	0.0526**	0.0499
	(8.858)	(0.011)	(0.025)	(0.069)
Elig*2012-15	13.4147	0.0138	0.0351	0.0684
	(10.329)	(0.011)	(0.025)	(0.087)
	Panel B: Effects by Agency Quality			
Elig*High*Post-2010	8.7314	0.0007	0.0764*	-0.0629
	(19.051)	(0.013)	(0.045)	(0.151)
Elig*Low*Post-2010	19.7544**	0.0232**	0.0498**	0.1669**
	(9.570)	(0.011)	(0.025)	(0.078)
	Panel C: Effects by Agency Quality and Program Phase			
Elig*High* 2010-12	-5.6466	0.0086	0.0614	-0.1401
	(18.393)	(0.024)	(0.044)	(0.133)
Elig*High* 2012-15	18.9260	-0.0006	0.0900*	0.0090
	(22.908)	(0.014)	(0.052)	(0.194)
Elig*Low* 2010-12	24.9239***	0.0131	0.0678**	0.1905***
	(9.123)	(0.013)	(0.029)	(0.073)
Elig*Low* 2012-15	16.6415	0.0280**	0.0348	0.1476
	(11.781)	(0.013)	(0.026)	(0.099)
N	21865	21865	18365	18365
Comparison Mean	164.90	0.15	0.17	0.31

Notes: Job order data for 2005 and 2014 is missing (column 3 & 4), in addition all outcomes have missing values if an agency exited the market (i.e. 12 consecutive months of zero migrant recruitment). This table reports difference-in-difference estimates on placement quality. `Comparison Mean' reports the pre-program (2005-2009) mean of the dependent variable for the comparison group. Panel A compares eligible agencies to comparison agencies during the (post-treatment) period preceding Quality Revelation during which agencies were exposed to Reputational Incentives (2010-2012) and the period after Quality Revelation (2012-2015). Panel B computes heterogeneous effects by pre-program agency quality; `High' denotes above median pre-program quality, while `Low' is below median pre-program quality, where quality is based on the `Predicted Rating' variable (See Appendix A. 10 for details). Panel C reports effects by program phase and by pre-program agency quality. `Average salary' is the average salary of migrants placed by an agency in a given month in 2015 USD. `Contract Renewal' is the share of migrants placed by an agency who returned to the same employer on a subsequent contract. `Good Orders' are job orders corresponding to employers with a below median pre-program complaint rate. `Contract Quality' is a normalized index incorporating whether or not a job order speci es that health insurance, accommodation, and return airfare are provided as well as the salary range paid to workers. Index components are weighted by the inverse of the covariance matrix as in Anderson (2008). All speci cations cluster at the agency level and include agency xed effects. Panels B & C contain month-year xed effects interacted with average migrant recruitment (2005-08); Panel A contains the preceding controls as Panels B and month-year xed effects interacted with the average yearly Quantitative Score (2005-08).

Table 4: The Effect of Receiving a Higher Rating

	(1) Job Orders	(2) Migrants	(3) Domestic Share	(4) Pre-Complaints
Treatment	0.860*	5.065	-0.018	0.113
	(0.464)	(5.144)	(0.108)	(0.228)
N	20714	28060	28058	28060
Depvar. Mean	1.50	25.43	0.25	0.56

Notes: This table estimates the effect of receiving an additional star on the government's star rating program using a regression discontinuity design. The sample consists of all agencies eligible for the rating program. Reported estimates pool across the star speci c thresholds (see Appendix B. 7 for details), use a uniform kernel, and use the MSE-optimal bandwidth using the rdrobust package. All speci cations include year xed effects and standard errors are clustered at the agency-level. Job Orders are the number of orders that an agency received in a speci c month from employers abroad (i.e. demand). Migrants are the number of migrants placed by an agency in a given month. Domestic Share is the share of migrants who are domestic workers placed by the agency in a given month. Pre-Complaints is a count variable for the number of migrants who made a complaint prior to the program and subsequently migrated again after the program.

Table 5: Spillover Effects

	(1) Migrants	(2) Agency Exit	(3) Job Orders	(4) Good Orders
Any Elig*Post-2010	42168*	0.0016	02418	00821
	(2.275)	(0.059)	(0.147)	(0.061)
N	16104	16104	13176	10650
Comparison Mean	5.93	0.20	0.32	0.17

Notes: This table estimates spillover effects of the program and restricts attention to the sample of comparison agencies. `Comparison Mean' reports the pre-program (2005-2009) mean of the dependent variable for comparison agencies that do not have an eligible agency within 100 ft. `Agency Exit' is a proxy for an agency closing down. It is coded as an indicator variable that switches on at the beginning of a 12 month sequence of zero recruitment after the program was introduced. `Job Orders' are the number of orders that an agency received in a speci c month from employers abroad (i.e. demand). Note, data for 2005 and 2014 is missing for the job order data. `Good Orders' are job orders corresponding to employers with a below median pre-program complaint rate. `Any Elig' is a dummy variable coded as 1 if there is an eligible agency within 100 ft of a comparison agency. The estimates are clustered at the local market level and the speci cation includes agency xed effects, month-year xed effects interacted with average 2005-2008 migrant recruitment, month-year xed effects interacted with average Quantitative Score 2005-2008 xed effects, and local market x month-year xed effects.

Appendices - Table of Contents

For Online Publication

A Program Information and Context 44

A.1 Criteria used for Agency Rating Program 44

A.2 Complaint Types 46

A.3 Example of Complaint Mechanism 47

A.4 Booklet Provided to Agencies 48

A.5 Job Order Process 51

A.6 Distribution of Ratings 52

A.7 Histogram of Predicted Rating 53

A.8 Agency Rating Online Database 54

A.9 Online Advertisements 55

A.10 Data Appendix 56

A.11 Agency Locations 58

A.12 Job Types 59

A.13 Pre-Program Employer to Agency Matches 60

A.14 Wages and Amenities by Employer Complaint Rates 61

A.15 The Distribution of Matches to Bad Employers by Agency Quality 62

B Additional Results 63

B.1 Results from Agency Survey 63

B.2 Characteristics of Job Orders 65

B.3 Components of Contract Quality 66

B.4 Composition of Migrants Recruited 67

B.5 Raw Trends 68

B.6 Components of Quantitative Score 70

B.7 The Effect of Receiving a Higher Rating 71

B.8 Heterogeneous Effects by Local Market Competition 72

C Robustness Tests 73

C.1 Pre-Trends 73

C.2 Plots of Lagged Effects 74

C.3 Agency Exit 75
C.4 Total Salaries 76
C.5 Complaint Rate 77
C.6 Robustness to Agency Exit 78
C.7 Robustness to Trend Controls 79
C.8 Mean Reversion and Alternative Measures of Agency Quality 80
C.9 Program Effects with Alternative Samples 82

Appendix A Program Information and Context

Appendix A. 1 Criteria used for Agency Rating Program

A1.1: Quantitative Criteria Used in Rating from SLBFE Data

- Total Numbers of Persons Recruited
- Skill Level of Jobs
  - Professional jobs
  - Middle-level Employment
  - Skilled workers
  - Semi-skilled workers
  - Unskilled workers
  - Domestic Workers
- Recruitment to New Destinations
- Dispute Settlement Rate
- Diversi cation of Job Market
- Cess payment

A1.2: Qualitative Criteria Used during Audits

- Timely submission of license renewal application
- Timely submission of bank guarantee for license application
- Book A: Register containing details of migrants
- Book B: Register of job offers
- Book C: Register on remittance reimbursed from the Bureau
- Book D: Register containing details of Foreign Employers or Commissions
- Maintaining Receipts
- Register containing details of passports
- Proper Maintenance of Of ce
  - Location of agency
  - Display of name boards
  - Attractive appearance
- Interior Layout
  - Division of of ce area

  - Process Management
  - Good equipment
- Sustainability of business (continuity of operation)
- Outlook of Of ce Staff
  - Uniform
  - ID Card
- Implementing Awareness Programs
  - Registry of bio-data of employers
  - Registry of bio-data of recruits and details of their family background
  - Maintenance of a notice board
- Implementing pre-migration awareness programmes which educate them on:
  - Destination country
  - Company or employer
  - Nature of employment
  - Material data on country
  - Common customs to be followed
  - Do away with social stigmas
  - Actions to be taken if problems crop up
- Assessment of Human Resources
  - Graduate and diploma holders
  - A-Level holders
  - O-Level holders
- Bonus Points
  - Recruiting persons without charging fees
  - Sending the employee direct to employer
  - License not temporarily discontinued during license year
  - Not receiving any letters of warnings over non-addressing of complaints
  - Not receiving any other letters of warning issued by the Bureau

Appendix A. 2 Complaint Types

Notes: The table above shows the decomposition of complaints over the span of the dataset 2005-2015.

Appendix A. 3 Example of Complaint Mechanism

Appendix A. 4 Booklet Provided to Agencies

A4.1: Importance of the grading methodology

- To bring all the agencies into a well-recognized standard.
- The objectives are to recognize, encourage and appreciate the agencies maintaining high standards.
- Establishing an impartial, justi able and scienti c methodology to grade agencies.
- The Act requires SLBFE to guide agencies towards their growth, development and standardizing their levels.

A4.2: Policies, qualifying/disqualifying and prospects

- Agencies will be quali ed for this grading system only if the agency holds a valid license for at least a year and have sent more than 100 people for employment during the assessment year.
- There are 11 indicators selected for this grading methodology which are listed out below:

    1. (i) The total number of recruitment (ii) Recruitment according to manpower groups (Professional, middle level, clerks and classical related jobs, skilled labour, semi-skilled labour, unskilled labour and house maids).
    2. Settlement of complaints is assessed in two ways: (i) Settlement rate of the conciliation division looks at the annual rate of the resolved complaints by an agency in that division
       (ii) Investigation and complaints settlement rate looks at the annual rate of resolved complaints in the special investigation division The nal rate is calculated by taking the average of the two divisions.
    3. Recruitment to new destination has two categories, (i) Entering into agreements with new countries (countries that have not received a SL employees through SLBFE) and should have at least sent one person for employment (ii) Entering into agreements with new agencies and companies in the countries where SL has already started business.
    4. Ratio difference between the number of vacancies approved and the number of vacancies ful lled (the number of people recruited).
    5. The payment of cess out the commission they receive from foreign countries for providing employees
    6. Evaluating human recourses- points if employees are granted with employee's provident fund (EPF) and points according to the educational quali cations such as graduate or a diploma holder, A/L and O/L quali ed.
    7. Prosecution- If complaints are forwarded to legal division as it has not been solved, law will be enforced. If 10 prosecutions are against an agency, then they will be disquali ed from the grading system.
    8. Agencies will receive minus points if there are complaints against agencies lodged by the migrant worker, relatives, family members, foreign agencies and foreign employers.

9. Minus points are given to agencies for receiving the approval of a job order but fail to provide employment within the valid period of approval.

A4.3: Disqualification: Any agency will not be considered if the following conditions can be applied:

- If the average complaint ratio is 100%
- Have sent less than 100 people
- If they are punished or warned by a court of law for producing forged documents, recruiting underage people, recruited disabled people and has been prosecuted by SLBFE for more than 10 court cases.

A4.4: Proposals and Suggestions

1. Updating the records and documents that are maintained at present such as data sheets of Sri Lankan employees and approved job orders.
2. Records of statutory refunds made by SLBFE, passport details, employers and commissions received.
3. Maintain a log book for government authorized officers to lodge their recommendations.
4. Classifying the letters received from SLBFE according to licensing division, conciliation division and general letters
5. Provide a valid letter of appointment for everyone serving in the agency, listing out the duties and mark attendance.
6. Maintain records of migrant workers who have returned before the contract period expired and sent back to Sri Lanka before the contract period and provide it to SLBFE which has to be authorised by bureau officers.
7. Make the agency staff thorough with laws, regulations, functions of the agency, the agency grading methodology and pay them well and on time.
8. Get the staff to be members of EPF, provide them with uniforms, ID cards and train and educate them on modern technology, the importance of public relations.
9. Submit the application for extending the license and bank guarantee on time.
10. Be vigilant of the authenticity of certificates forwarded by applicants.
11. Maintain a record of the family members of the migrant workers and take maximum action for the welfare of their family members.
12. Be vigilant of the workers being paid regularly and extend insurance if necessary.
13. Pay attention towards the benefits and salary of the migrant worker as mentioned in the agreement throughout the contract period and persuade the foreign employer if necessary.

A4.5: Benefits

Period of Supervision is from 1st of January to 31st of December every year

1. They will be invited to participate in foreign employment promotion organized by SLBFE.
2. Issue special identity card for the of cer in charge of the agency and provide special service counters.
3. Their names will be advertised in SLBFE publications.
4. Granting approval to open web page in the SLBFE website.
5. Discussing to provide all the bene ts of government, through the Ministry of Foreign Employment.
6. Standardization of agencies.

Appendix A. 5 Job Order Process

The gure below details a 12-step process from the point at which contact is initiated by a foreign employer to a consulate to to the nal approval for a migrant worker's departure. In the rst step, the foreign employer creates a job order that describes the number of workers required, their salary, and their bene ts. These bene ts include whether an employer provides food, accommodation, air travel, and/or health insurance. The Foreign Mission will authenticate the job order and then the local agency will receive the job order. The foreign employer may send the same job order to up to ve local agencies. The local agency must then complete the paperwork and submit the job order to the SLBFE for a First Approval.

If the job order is approved by the SLBFE, the local agency may advertise in the SLBFE job bank or through local media. Potential migrants are then interviewed and selected. Selected individuals must pass a medical examination, a background check by the Sri Lankan police, and, in some cases, undergo a basic training. The local agency sends prospective worker information to the foreign employer who then obtains visa clearance for the migrant and the local agency obtains visa and air tickets. If the prospective migrant clears these requirements, the job order is sent for nal approval to the SLBFE. In the nal step, the SLBFE grants nal approval and stamps the migrant worker's passport.

Appendix A. 6 Distribution of Ratings

(a) Panel A: All Eligible Agencies

(b) Panel B: Recruited 100-155 Migrants in 2009

Notes: This gure displays a frequency distribution of the Combined Score that was calculated for agencies in the agency grading program. The star cutoffs are as follows: 0 stars: 0-19.9, 1 star: 20-34.9, 2 stars: 35-49.9, 3 stars: 50-59.9, 4 stars: 60-79.9, 5 stars: 80-100. Panel A show the distribution for all eligible agencies while Panel B restricts the sample to those who recruited between 100 and 155 migrants in 2009.

Appendix A. 7 Histogram of Predicted Rating

Notes: The gure above plots the Predicted Rating variable for eligible agencies and comparison agencies.