

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

¹Across 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.

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

³By ‘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.

⁵Placement 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

⁶In 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).

⁷The 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.

⁸In 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 group¹⁰; 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

⁹To 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.

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

¹¹The 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 than 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 attention¹³ 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

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

¹³An 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 influences market efficiency. 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 studies¹⁵ 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 1970's,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.

¹⁴This 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).

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

¹⁶See 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 filled, the salary and benefits that are contractually obliged and up to five licensed agencies who would recruit these workers on an employer’s behalf. The foreign employer first 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 satisfied, the SLBFE stamps a migrant worker’s passport.

2.2.3 Market Structure

Foreign employers pay recruitment agencies a fixed 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 fixed (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 fill all vacancies on a job order. Two-thirds of surveyed agencies state the main reason they were unable to fill a job order was because they couldn’t find 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 find that the least abusive employers pay higher salaries and offer more amenities both within and across sectors (see Appendix A.14). The lack of a compensating differential is suggestive of a

¹⁷Note, 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 flows through porous land borders.

¹⁸In 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.

¹⁹The 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).

²⁰For 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.'

²¹For 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 benefits and salary of the migrant worker as mentioned in the agreement throughout the contract period and persuade the foreign employer if necessary.'

²²Technically, 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).

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

²⁴We 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.

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

²⁶In 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 θ , where $\theta < 1$, and this informs their first period prior on the likelihood of good placement quality

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

²⁸We 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.

²⁹For 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 fulfill 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 first period is the key decision our model explores under two information environments: (1) where employers can't learn about first period placement quality because of information frictions, and (2) when first period placement quality is revealed in the second period. The equilibrium concept we use is a perfect Bayesian equilibrium, where the choice of first 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, first period placement quality is not revealed to the market. Consequently, an L -type agency will not invest effort (e) in the first period because it will not influence 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, $Pr(G)$:

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

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 first 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 first period screening $e^* = 0$.

3.5 Equilibrium with the Ratings Program

Now, the regulator commits to publicly revealing an agency's first period placement quality in the second period. Moreover, agencies are informed about this credible threat in the first 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 first period effort. Consequently, an employer would offer q_1 job orders in the first period, where $q_1 = \theta + (1 - \theta)p(\bar{e})$. In the second period, however, employers observe first period placement quality and update their beliefs using Bayes' rule, where θ is the proportion of H -types. :

$$Pr(H|G) = \frac{Pr(G|H)Pr(H)}{Pr(G)} = \frac{\theta}{\theta + (1 - \theta)p(\bar{e})} \quad (2)$$

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:

$$\pi = q_1 + 1 * p(e) * Pr(H|G) + 0 * [1 - p(e)] - c(e) = q_1 + p(e)Pr(H|G) - c(e) \quad (3)$$

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:

$$\left[\frac{\theta}{\theta + (1 - \theta)p(\bar{e})} \right] = \frac{c'(e)}{p'(e)} \quad (4)$$

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

³⁰The 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 specific job order under which they were recruited. The specific 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 identifiers, 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 [A.2](#) 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 first 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-specific, 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-specified pre-treatment characteristics using a progressively larger window across the threshold until the test statistic suggests an imbalance.³¹ We specified variables intended to capture the extent and diversity of agency operations—migrant recruitment, job orders, and complaint intensity between 2005-2008—and find 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 Specification

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

$$Y_{it} = \delta_t + \alpha_i + \beta_1 \text{Elig} * \mathbb{1}[2010 \leq t < 2012]_{it} + \beta_2 \text{Elig} * \mathbb{1}[t \geq 2012]_{it} + \tau X_{it} + \epsilon_{it} \quad (5)$$

In the specification above, δ_t are month-year fixed effects, α_i is an agency fixed effect, Elig is an indicator variable for an eligible agency, $\mathbb{1}[2010 \leq t < 2012]$ is an indicator variable for the period between January 2010 and May 2012, $\mathbb{1}[t \geq 2012]$ is an indicator variable for the period after (and including) June, 2012, and τX_{it} are two sets of time varying controls: the interaction of month-year fixed effects with average migrant recruitment between 2005-08, and the interaction of month-year fixed effects with the average Quantitative Score between

³¹Specifically, 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-specified 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, β_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 β_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*.

$$Y_{it} = \delta_t + \alpha_i + \gamma_1 High * Post_{it} + \beta_1 High * Elig * Post_{it} + \beta_2 Low * Elig * Post_{it} + \tau X_{it} + \epsilon_{it} \quad (6)$$

The terms are as defined 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, β_1 is the causal effect of the program for pre-program high quality agencies, while β_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.

³²Even 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 specification 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.

³³We regress an eligible agency’s Combined Score—the continuous value that determines star ratings—on their 2009 Quantitative Score. We then use the estimated coefficients 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.

³⁴When we estimate heterogeneous effects by agency quality, we omit the interaction between month-year fixed 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 rate³⁵ (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 pretrends 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 Arabi: 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 A.12 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

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

influenced 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 find that eligible agencies, on average, have a higher monthly Quantitative Score during the RI-phase (i.e. the coefficient on $Elig * [2010 - 12]$), but this estimate is somewhat imprecise (t -stat = 1.64). However, in Panel B, we find that pre-program low quality agencies (i.e. below median *Predicted Rating*), are significantly 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 find 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 find that any specific 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 *specific* 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 reflects 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 find that the program did not significantly 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 cashflow 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

³⁶We 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.

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

³⁸Note, the comparison mean here is post-program as it is undefined 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 \sim 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

³⁹This 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).

⁴⁰Adjusting 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).

⁴¹Our 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 significantly 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 find 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 find that the program influenced the complaint rate in either phase of the program or by agency quality. Our design is likely underpowered 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 finding 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 unfilled 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 filling 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

⁴²Appendix B.3 computes program effects for each of the constituent variables of the contract quality index. We again find 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.

⁴³Using a migrant-year level dataset, agency x employer fixed effects explain five times as much variation in whether a migrant makes a complaint as country, sector, and time fixed 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

⁴⁴In 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).

⁴⁵We 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 β_2 for agency i in month m and year t :

$$Y_{imt} = \delta_t + \beta_1 \text{Combined Score}_i + \beta_2 \text{Treat}_i + \beta_3 \text{Treat} * \text{Combined Score}_i + \epsilon_{imt}$$

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 find 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 influences outcomes. The specification includes a fixed effect of the local market interacted with year fixed effects and, in addition, we cluster standard errors at the local market level.⁴⁶

We find that the presence of an eligible agency results in a significant *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 find 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 fixed 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 classification of pre-program agency quality may imply that the heterogeneous effects we report reflect 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 classified 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

⁴⁶We 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 specification, 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 specific 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 Efficiency?

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 efficiency, we first 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 influences 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 find 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 findings

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 filipino 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 confidence 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 offices 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*, (1):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.

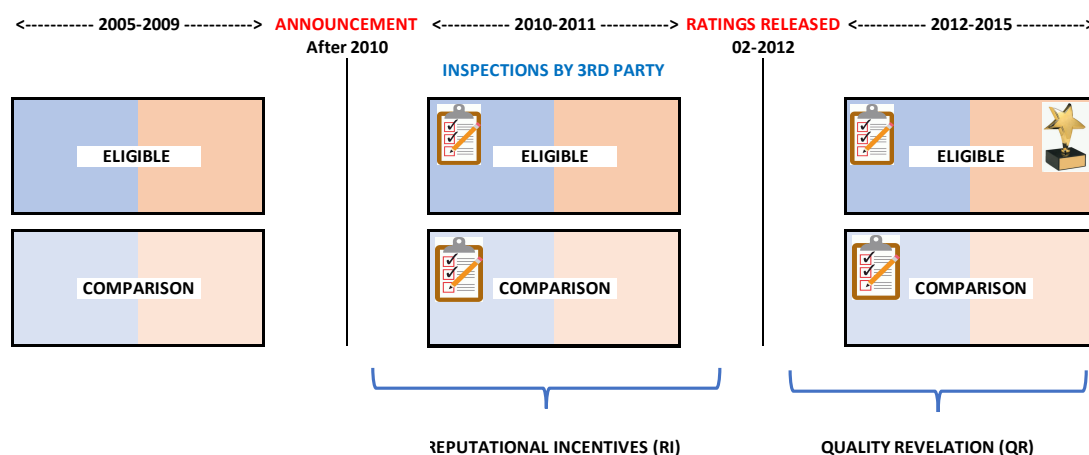
- Shrestha, S. A. and Yang, D. (2019). Facilitating worker mobility: A randomized information intervention among migrant workers in singapore. *Economic Development and Cultural Change*, 68(1):63–91.
- Silva, D. (2013). Rizana nafeek executed. *Daily Mirror*.
- Somarathna, R. (2011). Alleged sexual abuse 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

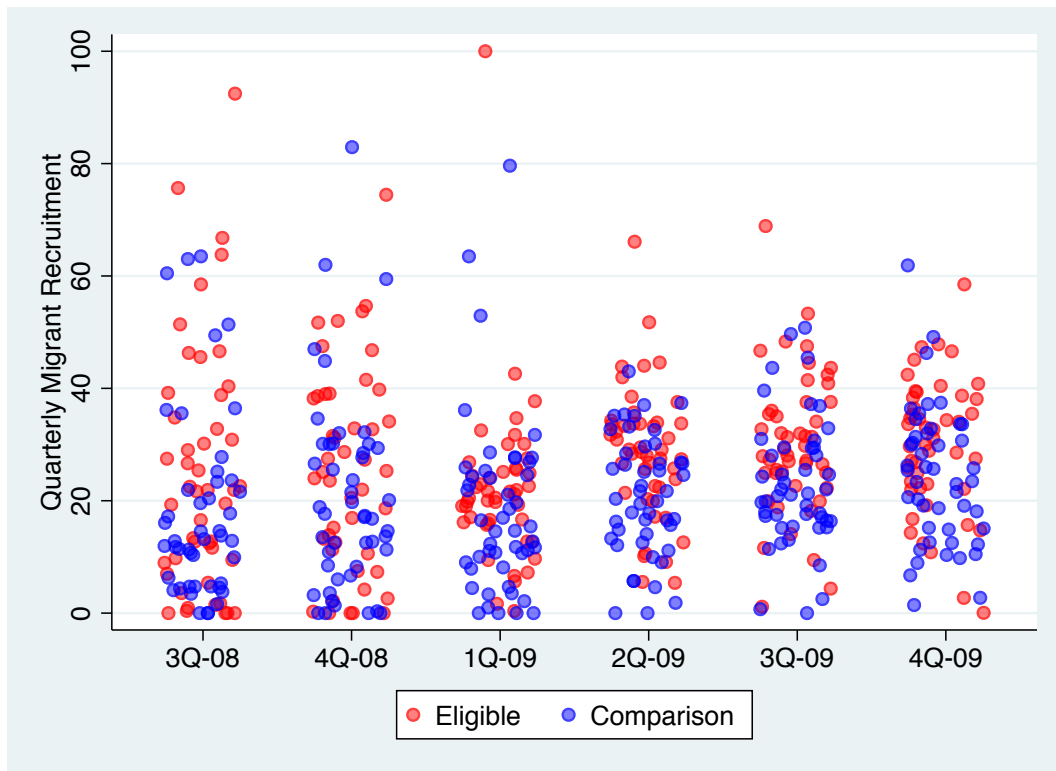
- **Reputational Incentives (RI) Phase (2010-2012):**
 - **January 2010:** Program announced to agencies and booklet with grading criteria is sent through SLBFE management system
 - **July 2010- May 2011:** Third-party inspections
- **Quality Revelation (QR) Phase (2012-2015):**
 - **February, 2012:**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

Panel B: Timeline with Potential Effects



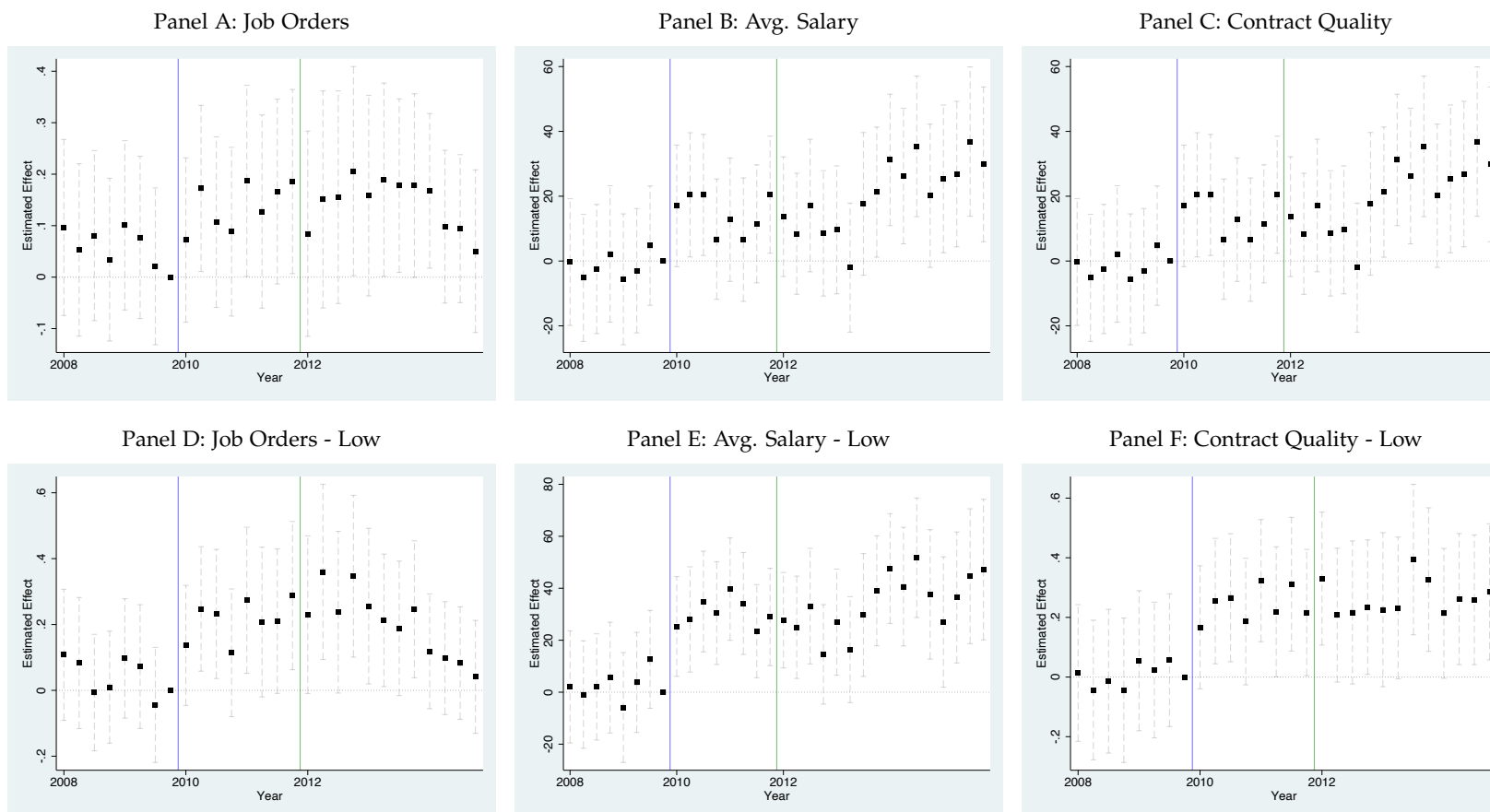
Notes: Panel A of Figure 1 shows a simple timeline (time on vertical axis) for the agency ratings program and lists the potential effects below it (time on horizontal axis). In Panel B, the timeline with potential effects illustrates the difference-in-difference design. We compare a set of eligible agencies to comparison agencies below the 100 migrant recruitment threshold in 2009. After the announcement, the rating criteria (signified 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 figure 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



35

Notes: These figures 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 coefficients—an indicator for an eligible firm 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% confidence 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 figure plots a series of probability density functions that describe which recruitment agencies are matched to ‘Good Employers’: defined as employers who have a below median pre-program (< 2010) complaint rate. A match here is defined as whether an agency sends a migrant to a ‘Good Employer’ in a specific 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 first 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 specific 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
<i>Panel A: Effects by Program Phase</i>				
Elig*2010-12	0.3426 (0.209)	0.1991 (1.585)	0.1067* (0.062)	-0.0701*** (0.026)
Elig*2012-15	0.2537 (0.195)	3.3784 (3.266)	0.1451* (0.078)	-0.0826* (0.044)
<i>Panel B: Effects by Agency Quality</i>				
Elig*High*Post-2010	-0.3422 (0.265)	2.1827 (5.407)	-0.0105 (0.114)	0.0067 (0.074)
Elig*Low*Post-2010	0.6574*** (0.200)	1.8596 (2.093)	0.2204*** (0.083)	-0.1256*** (0.039)
<i>Panel C: Effects by Agency Quality and Program Phase</i>				
Elig*High*2010-12	-0.0722 (0.346)	-4.4285 (3.392)	-0.0620 (0.098)	-0.0234 (0.053)
Elig*High*2012-15	0.3984 (0.335)	5.9194 (7.754)	0.0288 (0.145)	0.0237 (0.093)
Elig*Low*2010-12	0.6715*** (0.240)	2.3932 (2.009)	0.2151*** (0.078)	-0.0923*** (0.031)
Elig*Low*2012-15	0.3577* (0.212)	1.5580 (2.224)	0.2243** (0.094)	-0.1444*** (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 specific 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 specifications cluster at the agency level and include agency fixed effects. Panels B & C contain month-year fixed effects interacted with average migrant recruitment (2005-08). Panel A contains the same controls as Panel B and month-year fixed 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
<i>Panel A: Effects by Program Phase</i>				
Elig*2010-12	11.4884 (8.858)	0.0093 (0.011)	0.0526** (0.025)	0.0499 (0.069)
Elig*2012-15	13.4147 (10.329)	0.0138 (0.011)	0.0351 (0.025)	0.0684 (0.087)
<i>Panel B: Effects by Agency Quality</i>				
Elig*High*Post-2010	8.7314 (19.051)	0.0007 (0.013)	0.0764* (0.045)	-0.0629 (0.151)
Elig*Low*Post-2010	19.7544** (9.570)	0.0232** (0.011)	0.0498** (0.025)	0.1669** (0.078)
<i>Panel C: Effects by Agency Quality and Program Phase</i>				
Elig*High*2010-12	-5.6466 (18.393)	0.0086 (0.024)	0.0614 (0.044)	-0.1401 (0.133)
Elig*High*2012-15	18.9260 (22.908)	-0.0006 (0.014)	0.0900* (0.052)	0.0090 (0.194)
Elig*Low*2010-12	24.9239*** (9.123)	0.0131 (0.013)	0.0678** (0.029)	0.1905*** (0.073)
Elig*Low*2012-15	16.6415 (11.781)	0.0280** (0.013)	0.0348 (0.026)	0.1476 (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 specifies 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 specifications cluster at the agency level and include agency fixed effects. Panels B & C contain month-year fixed effects interacted with average migrant recruitment (2005-08); Panel A contains the preceding controls as Panels B and month-year fixed effects interacted with the average yearly Quantitative Score (2005-08).

Table 4: The Effect of Receiving a Higher Rating

	(1)	(2)	(3)	(4)
	Job Orders	Migrants	Domestic Share	Pre-Complaints
Treatment	0.860* (0.464)	5.065 (5.144)	-0.018 (0.108)	0.113 (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 specific thresholds (see Appendix B.7 for details), use a uniform kernel, and use the MSE-optimal bandwidth using the `rdr` package. All specifications include year fixed effects and standard errors are clustered at the agency-level. Job Orders are the number of orders that an agency received in a specific 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)	(2)	(3)	(4)
	Migrants	Agency Exit	Job Orders	Good Orders
Any Elig*Post-2010	4.2168*	0.0016	0.2418	0.0821
	(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 specific 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 specification includes agency fixed effects, month-year fixed effects interacted with average 2005-2008 migrant recruitment, month-year fixed effects interacted with average Quantitative Score 2005-2008 fixed effects, and local market x month-year fixed 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
- Diversification 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 Office
 - Location of agency
 - Display of name boards
 - Attractive appearance
- Interior Layout
 - Division of office area

- Process Management
 - Good equipment
- Sustainability of business (continuity of operation)
- Outlook of Office 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

	Complaints	Percentage of Total Complaints
Non Receipt of Wages	19741	16.62
Breach of Employment Contract	15201	12.80
Harassment	13407	11.29
Sickness	12522	10.54
Others	11775	9.92
Lack of Communication	11322	9.53
Not Sent Back After Completion of Contract Period	9509	8.01
Problem at Employees Home (Sri Lanka)	6207	5.23
No Correspondence to Complaints	3866	3.26
Being Stranded Without Employment	2849	2.40

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

Appendix A.3 Example of Complaint Mechanism

Licensee,
[REDACTED]

2018.10.30

Complaint no: [REDACTED]
Migrant worker: [REDACTED]
Passport no: [REDACTED] DEPARTURE DATE: 2016. [REDACTED]

The conciliation division received a complaint on the 2018.06.08 regarding the migrant worker mentioned above. This is regarding the complaint and the investigation on the 2018.06.20.

It was mentioned in her contract that her salary would be 1125 SR but she was actually paid 900 SR. Every month she was not paid 225 SR. Thus the total outstanding salary for 24 months is (9225x 24) for 24 months is 5400 SR. The letter dated 2018.06.20 states that the balance salary should be paid and make arrangements get her back to Sri Lanka. However, She came back on the 2018. 10. 15 but the balance salary was not paid.

Therefore, make sure that that the balance salary of 5400SR is paid before the 2018. 11.30 and handover the documents related to that to our office. We kindly inform you that if you fail to do pay the balance, we will take legal action.

Manager (Conciliation)
SLBFE

Licensee,
[REDACTED]

2018.12.24

Complaint no: [REDACTED]
Migrant worker: [REDACTED]
Passport no: [REDACTED] DEPARTURE DATE: 2016. [REDACTED]

The conciliation division received a complaint on the 2018.06.08 regarding the migrant worker mentioned above. This is regarding the complaint and the investigation on the 2018.06.20.

An order was given to you to settle the balance payment of the salary (225x24) 5400 SR. This was informed to you on 2018.06.20 and 2018.10.30. However, you have failed to settle this matter.

Therefore, we have forwarded your case to the legal division for legal action.

Manager (Conciliation)
SLBFE

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, justifiable and scientific 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 qualified 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 final 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 fulfilled (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 qualifications such as graduate or a diploma holder, A/L and O/L qualified.
 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 disqualified 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

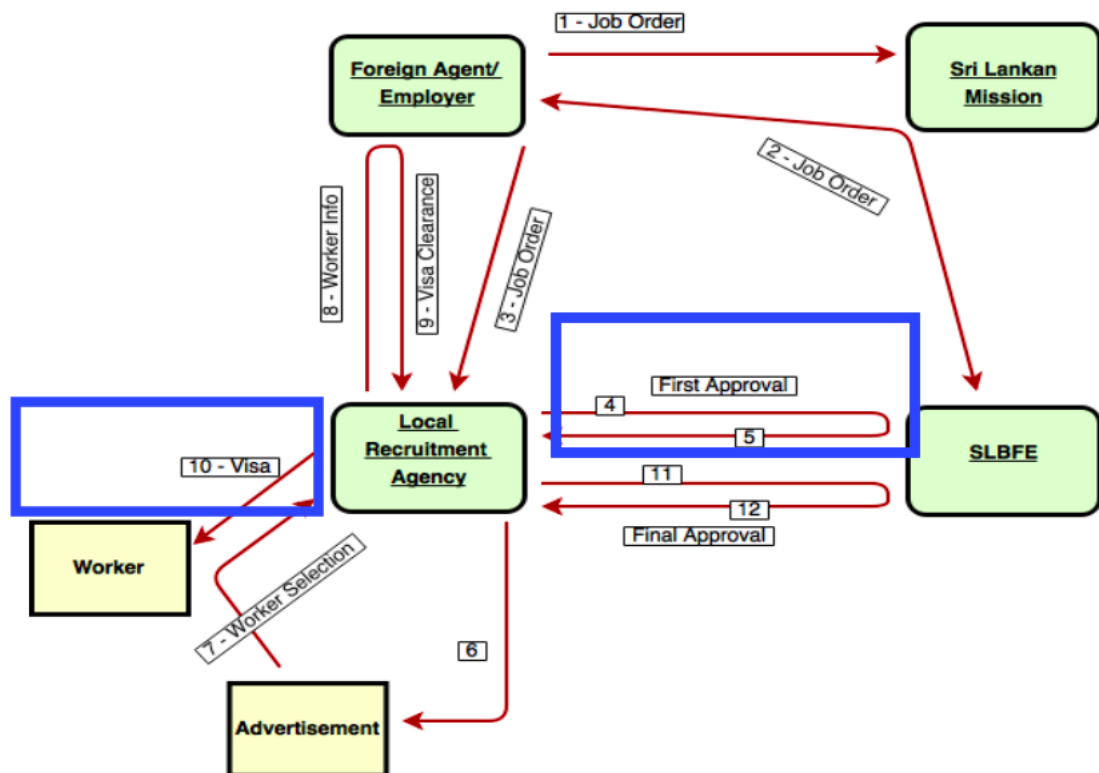
Period of Supervision is from 1st of January to 31st of December every year

1. They will be invited to participate in foreign employment promotion organized by SLBFE.
2. Issue special identity card for the officer 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 benefits of government, through the Ministry of Foreign Employment.
6. Standardization of agencies.

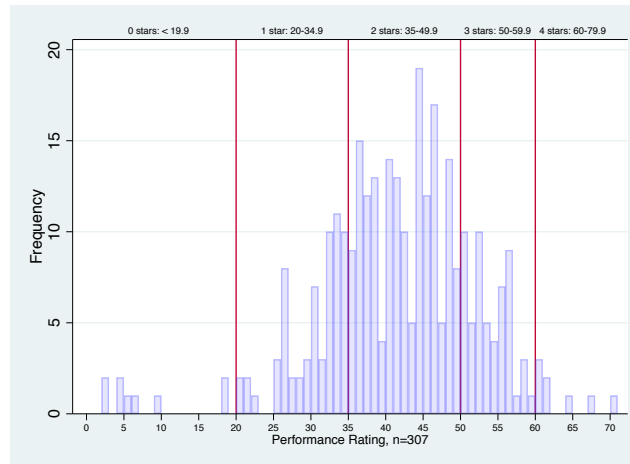
Appendix A.5 Job Order Process

The figure below details a 12-step process from the point at which contact is initiated by a foreign employer to a consulate to to the final approval for a migrant worker's departure. In the first step, the foreign employer creates a job order that describes the number of workers required, their salary, and their benefits. These benefits 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 five local agencies. The local agency must then complete the paperwork and submit the job order to the SLBFE for a First Approval.

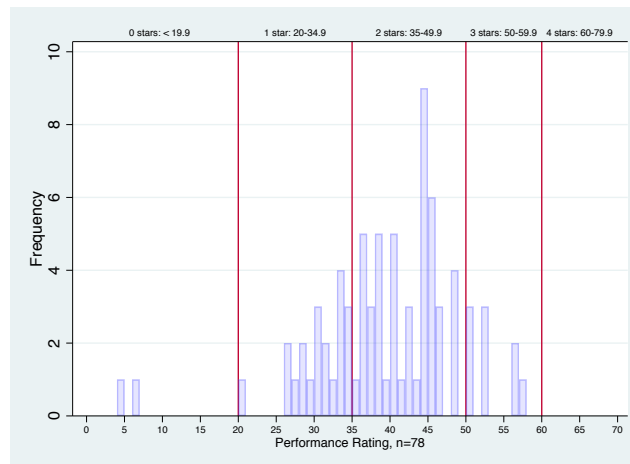
If the job order is approved by the SLBFE, the local agency may advertise in the SLBFE job bank or through local media. Potential migrants are then interviewed and selected. Selected individuals must pass a medical examination, a background check by the Sri Lankan police, and, in some cases, undergo 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 final approval to the SLBFE. In the final step, the SLBFE grants final approval and stamps the migrant worker's passport.



Appendix A.6 Distribution of Ratings



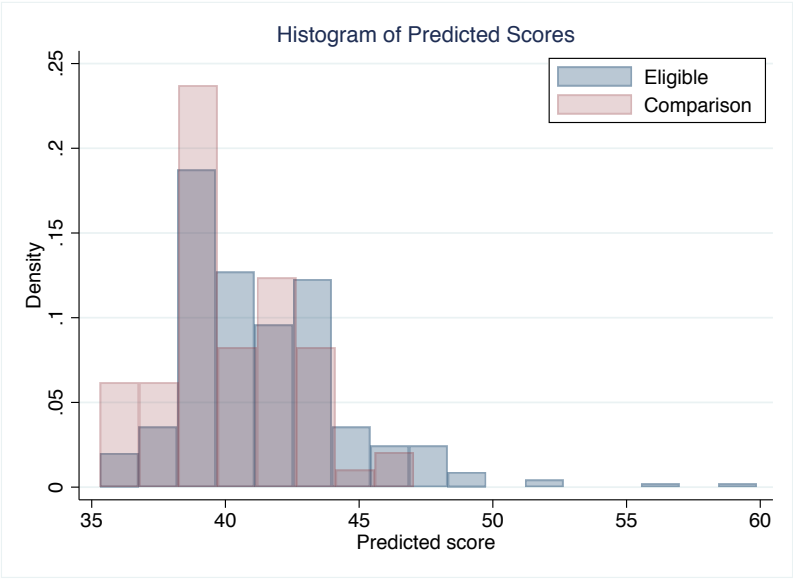
(a) Panel A: All Eligible Agencies



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

Notes: This figure 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 figure above plots the Predicted Rating variable for eligible agencies and comparison agencies.

Appendix A.8 Agency Rating Online Database

 **Sri Lanka Bureau of Foreign Employment**
ශ්‍රී ලංකා විදේශ සේවා හිඳුක්කි කාර්යාංශය இலங்கை வெளிநாட்டு வேலைவாய்ப்புப் பணியகம்

 [Home](#)

Agency Grading System

Excellence Awards for Licensed Foreign Employment Agencies - 2012



- Date of Award :28th February 2012
- Period of Assessment :Year 2009/2010
- [Criteria used for Assessment](#)

Ratings	Awards
4	
3	
2	
1	

Search by :

● Labour Licence Number :

● Name of Agency (any part) :

Search

Notes: The image above is a screenshot of the agency rating online database from the SLBFE website.

Appendix A.9 Online Advertisements

Vacancies in KUWAIT

Security Guards
Attractive Salary Package
300 Nos

IMMEDIATE DEPARTURE TO KUWAIT
One of the leading Security Company Requires Security Guards..

Minimum Requirements:
Age 24 - 37 | 42 inches of arched background | Able to speak and write English
Minimum education qualification | Minimum height 160cm (5'3")
Preferred with security guard experience

Jobs & Benefits:
30 days contract
Accommodation, Medical, Transport (Free/Shared)
Refreshment provided by company
Working hours: 08:00 am to 06:00 pm
Working days: 6 days per week
Other contract details at: www.alakeem.com

Final Interview with Foreign Principals
26th, 26th & 27th of Oct 2014
Register Today

Visit & Register at Al Akeem Agency Today

Send Your Cv's cv@alakeem.com

AL AKEEM ENTERPRISES (PVT) LTD
Manpower, Consulting and Recruiting Agency
0772 - 575 575 / 0771 - 415 415 / 0772 - 755855

(a) Panel A

Saudi Arabia Vacancies

Female Cleaners
Basic Salary SR 700

**2 year contract (Renewable),
Food allowance QR 200,
joining air ticket,
accommodation & medical
will be provided.**

**Apply now : cv@emeraldisle.lk
Call : 0777 809137 | 077 4372186**

Please visit us with CV, Originals of all appointment letters & service certificates, 2 photo copies of each. Two passport photos copies & 2 Recent size photos also required.

Our services, Your Future

Emerald Isle
Manpower & Travel Services
The Sri Lanka Recruitment Specialists

641, Mandirama Road, Colombo 19
Tel: 011 4 389762

487 2/1, Galle Road, Bandula Linnelle (Near Linnu Courts - Land Side)
Tel: 011 4 595488

186 1/1, Firdaus Building, Trinco Road, Trinco
Tel: 005 4 526803

Visit www.emeraldisle.lk for latest foreign employment

SL.BFF.01/15 අධ්‍යක්ෂ ජනරාල් ජනරාල් ජනරාල් ජනරාල් ජනරාල්

(b) Panel B

Notes: The figures above show advertisements placed by agency owners on their websites.

Appendix A.10 Data Appendix

We describe the main variables used in our analyses below. All variables are measured at the agency-month level, unless stated otherwise.

Average salary (USD): Mean salary of migrants departing per month, deflated to 2015 US dollars, winsorized at the 5% level.

Contract renewal: Share of migrants in a given month who renew their contract with the same employer or self-migrate to the same country and sector in a subsequent trip. We are able to match migrant trips by the same individual using unique national identifiers available in our data.

Complaint rate: Fraction of migrants departing in a given month who file a complaint at anytime. For example, if agency A sent 10 migrants in March 2013 and 3 of these migrants complain at a later date, the complaint rate for the agency in March 2013 is 0.3.

Solve time for complaints: Average number of months taken to resolve complaints filed in a given month

Job orders: Number of job orders received by an agency in a month, after approval by the regulator. We do not have data for years 2005 and 2014.

Good Orders: Number of job orders from foreign employers whose pre-program (prior to 2010) complaint rate is below the median. This value is only generated for employers who recruit at least 50 migrants during the pre-period (accounting for 93% of all migrants). We do not have data for years 2005 and 2014.

Migrants: Number of migrants recruited per month, winsorized at the 5% level.

Construction of Agency Quality Measures

We use the formula used by the SLBFE to compute a Quantitative Score. This is computed both at the yearly level and at the monthly level as we detail below. First, we detail how the government computes the Quantitative Score:

How the Quantitative Score is Computed:

The Quantitative Score is computed by assigning letter grades to continuous values for 9 assessment variables listed below. For each variable, the value of the variable was converted into a letter grade (A=5, B=4, C=3, D=2, E=1, Z=0). Letter grades were assigned in the following way: the range of the variable is divided into 5 equal bins; then letter grades are assigned based on these bins where the highest-scoring bin receives letter grade "A" and the lowest receives "Z." If a score could not be calculated for the criterion, a letter grade "Z" was assigned. A total score was then calculated by summing the numeric values associated with all letter grades.

We use this same formula for 9 assessment criteria observed in our administrative data to

construct a Quantitative Score. These criteria used are:

- Total migrant recruitment
- Professional recruitment
- Middle level and clerical recruitment
- Skilled recruitment
- Semi-skilled recruitment
- Unskilled recruitment
- Female domestic worker recruitment
- Dispute settlement ratio
- Business sustainability (new markets entered into)

Quantitative Score 2005-08 * Month-Year FE: In order to control for differential agency quality trends, we compute the average yearly Quantitative Score for an agency between 2005 and 2008 and then interact this average with month-year fixed effects.

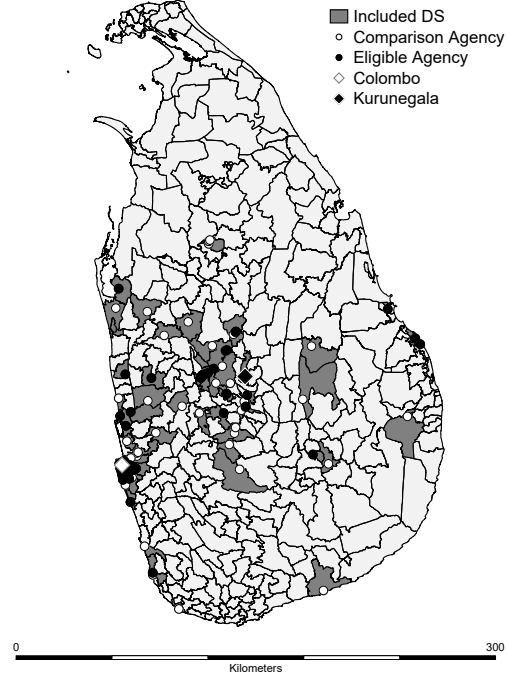
Quantitative Score 2009: In order to look at heterogeneity by pre-program agency quality, we instead use the Quantitative Score computed for 2009. Note, this is also the Quantitative Score used by the government to compute the final rating in 2012.

Predicted Rating: We construct a Predicted Rating using the fact that the government based the actual 2012 ratings on a Quantitative Score that only uses (pre-program) data from 2009 and third-party audit data collected after the program was introduced. We regress an (eligible) agency's actual continuous rating (i.e. their Combined Score) on their 2009 Quantitative Score. The fitted values from this regression are the 'Predicted Rating' for eligible agencies. In addition, we use the estimated coefficient from this regression to compute a 'Predicted Rating' for comparison agencies as well.

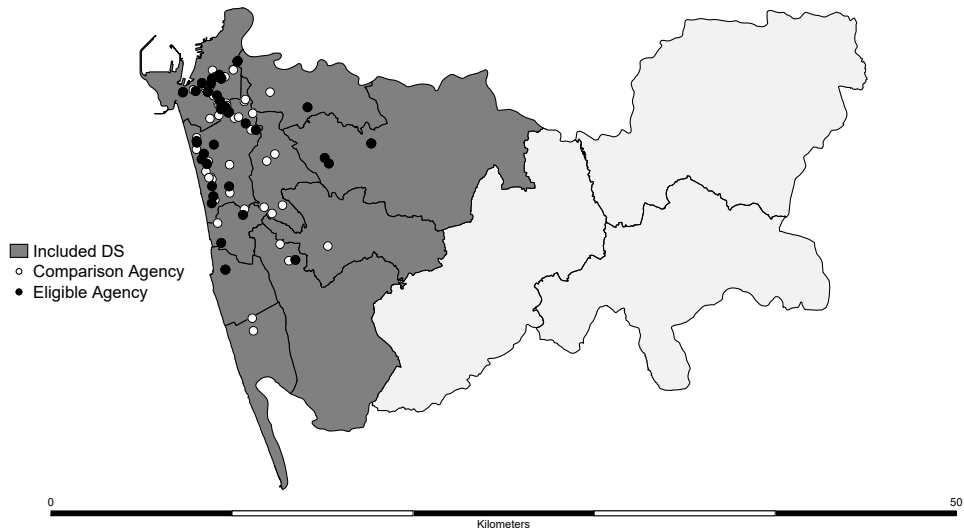
Monthly Quantitative Score: In Table 2, for example, we also use a 'Quant Score' variable to summarize investment in the rating criteria. We do so by computing the *monthly* Quantitative Score for each agency-month. We use the same formula described above, except at a month-level, to compute this score.

Appendix A.11 Agency Locations

Panel A: Agency Locations Across Sri Lanka



Panel B: Agency Locations Across Colombo District



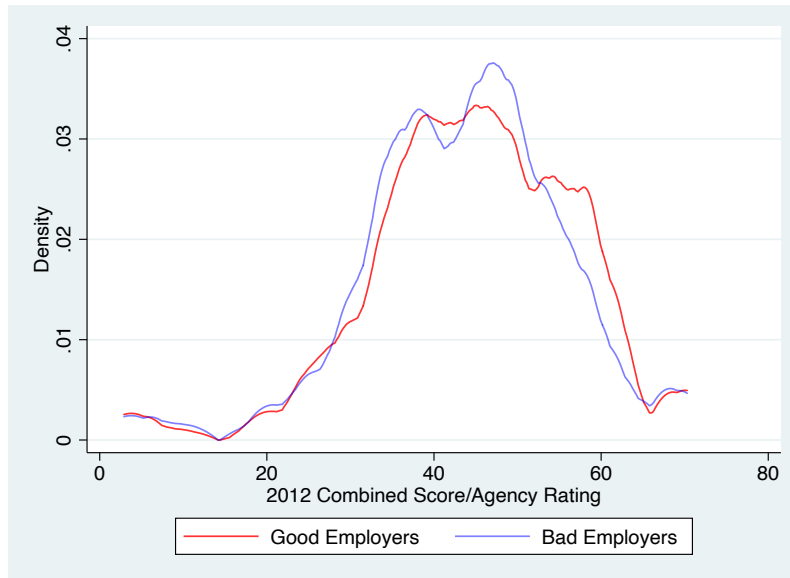
Notes: Panel A shows the location of eligible and comparison agencies included in the study across Sri Lanka. Panel B shows the location of agencies in Colombo District only.

Appendix A.12 Job Types

(1) Job Category	(2) Median Salary	(3) % Full	(4) % Study
1. Accountants, Engineers	720	0.36	0.28
2. Drivers, Machine Operators	223	21.06	18.56
3. Sales / Retail	237	1.07	0.83
4. Draughtsmen	297	2.73	3.82
5. Supervisory Labor	274	4.03	7.79
6. Cleaner / General Laborer	175	19.68	22.28
7. Domestic Worker	164	51.08	46.43

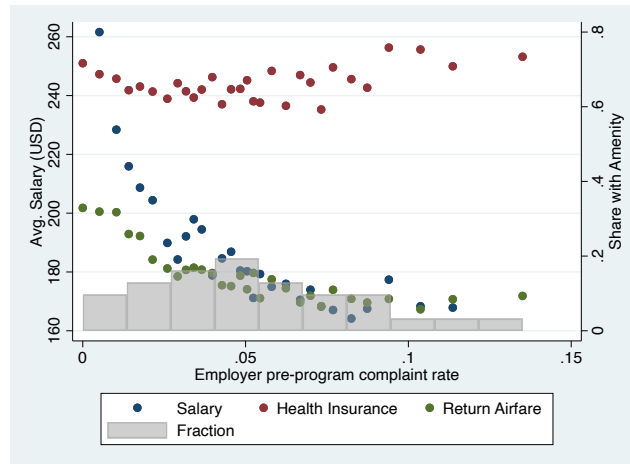
Notes: The table above reports the median salary for seven different skill classifications from the SLBFE. Column 3 reports the percentage of these jobs in the full sample while column 4 reports the composition in the restricted sample used for the study.

Appendix A.13 Pre-Program Employer to Agency Matches

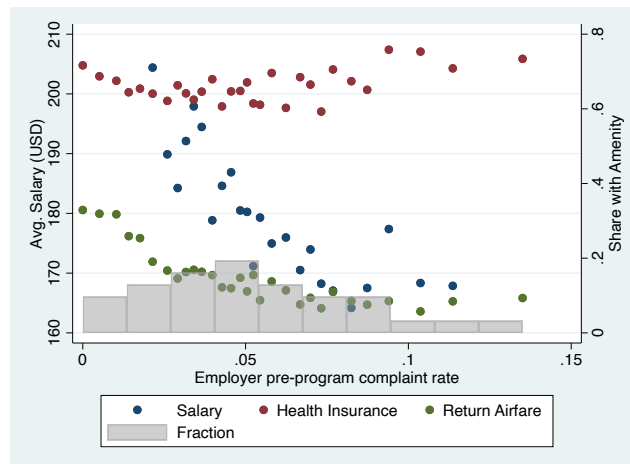


Notes: This figure plots two probability density functions that describe the pre-program matches between recruitment agencies, 'Good Employers' and 'Bad Employers': defined as employers who are below or above the median pre-program (< 2010) complaint rate, respectively. The domain is the continuous Combined Score/2012 agency rating which is only known to the public after 2012.

Appendix A.14 Wages and Amenities by Employer Complaint Rates



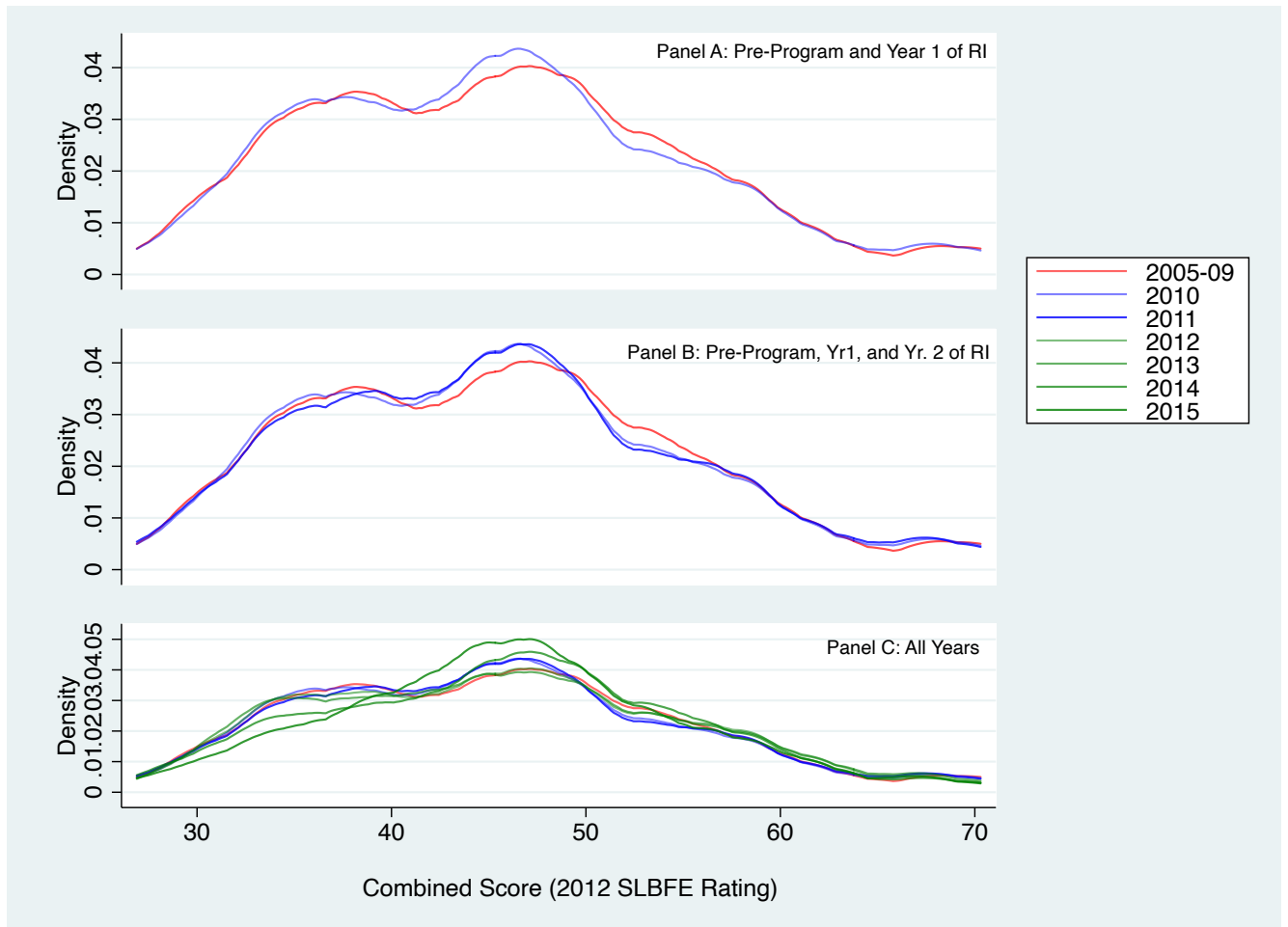
(a) Panel A: All Employers



(b) Panel B: Employers & Agencies Specializing in Domestic Work

Notes: This figure displays average salaries (2015 USD) and the share of amenities by pre-program employer complaint rates. The histogram shows the fraction of employers that fall into each complaint rate bin. Panel A restricts attention to employers who recruited at least 50 migrants prior to the program, which accounts for 94% of all pre-program recruitment. Panel B additionally focuses restricts attention to employers (and/or foreign agencies) for whom at least 75% of recruited migrants were domestic workers.

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



Notes: This figure plots a series of probability density functions that describe which recruitment agencies are matched to 'Bad Employers': defined as employers who have an above median pre-program (< 2010) complaint rate. A match here is defined as whether an agency sends a migrant to a 'Good Employer' in a specific 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 first 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).

Appendix B Additional Results

Appendix B.1 Results from Agency Survey

Panel A: Learning about Employer Reputation

(1) Source	(2) N	(3) Percent
Check if Blacklisted	83	76%
Consult Other Agencies	65	60%
Consult Returned Migrants	73	67%
Contact SL Consulate	48	44%
Contact SLBFE	46	42%

Panel B: Reasons for Rejecting Job Order

(1) Source	(2) N	(3) Percent
Job Wasn't Safe	29	52%
Employer Reputation	17	30%
Foreign Agent Reputation	18	32%
Too Many Amenities	11	20%
Low Commission	7	12%
Hard to find Workers	19	35%

Notes: Panel A asks agencies how they learn about the reputation of an employer. Panel B reports the reasons an agency states for rejected a job order in the past year. 55 out of 109 agencies reported rejecting a job order. Both tables use data collected through the agency survey in 2019 with a sample size of 109 agencies.

Appendix B.2 Characteristics of Job Orders

	(1) # Employers	(2) # New Employers	(3) # Good Orders	(4) Good Orders (%)	(5) # Bad Orders	(6) Bad Orders (%)
<i>Panel A: Eligible Agencies vs. Comparison</i>						
Elig*2010-12	0.5900 (0.500)	-0.0236 (0.422)	0.0526** (0.025)	0.0270* (0.015)	0.0021 (0.022)	-0.0051 (0.015)
Elig*2012-15	0.7266 (0.718)	0.4437 (0.706)	0.0351 (0.025)	0.0185 (0.016)	-0.0037 (0.022)	-0.0116 (0.015)
<i>Panel B: Effects by Agency Quality</i>						
Elig*High*Post-2010	0.1205 (1.087)	-0.2765 (1.182)	0.0764* (0.045)	0.0459* (0.027)	-0.0377 (0.031)	-0.0287 (0.019)
Elig*Low*Post-2010	1.0869* (0.652)	0.2546 (0.527)	0.0498** (0.025)	0.0245 (0.016)	0.0193 (0.025)	0.0016 (0.018)
<i>Panel C: Effects by Agency Quality and Program Phase</i>						
Elig*High*2010-12	-0.0810 (0.846)	-0.4967 (0.828)	0.0614 (0.044)	0.0392 (0.028)	-0.0362 (0.035)	-0.0250 (0.020)
Elig*High*2012-15	0.3159 (1.512)	-0.0660 (1.606)	0.0900* (0.052)	0.0519* (0.030)	-0.0388 (0.031)	-0.0321 (0.021)
Elig*Low*2010-12	1.1489* (0.592)	0.0386 (0.417)	0.0678** (0.029)	0.0329* (0.019)	0.0279 (0.027)	0.0077 (0.020)
Elig*Low*2012-15	1.0393 (0.776)	0.4361 (0.668)	0.0348 (0.026)	0.0175 (0.018)	0.0124 (0.028)	-0.0033 (0.021)
N	18365	18365	18365	18365	18365	18365
Comparison Mean	5.48	2.53	0.17	0.13	0.17	0.13

Notes: This table reports difference-in-difference estimates on the characteristics of job orders. ‘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 both for the period where agencies were solely 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. Panel C reports effects by program phase and agency quality. ‘# Employers’ reports the number of unique employers with whom an agency did business, ‘# New Employers’ refers to an employer with whom the agency had not done business prior to the introduction of the program. ‘Good Orders’ (‘Bad Orders’) are those from an employer with a below median (above median) pre-program complaint rate. All specifications cluster at the agency level and include agency fixed effects. Panels B & C contain month-year fixed effects interacted with average migrant recruitment (2005-08). Panel A contains the same controls as Panel B and month-year fixed effects interacted with the average yearly quantitative score (2005-08).

Appendix B.3 Components of Contract Quality

	(1)	(2)	(3)	(4)	(5)
	Return Ticket	Health Insurance	Accommodation	Food	Avg. Salary (USD)
<i>Panel A: Effects by Program Phase</i>					
Elig*2010-12	0.0009 (0.017)	0.0244 (0.024)	0.0207 (0.024)	0.0233 (0.026)	9.9086 (9.104)
Elig*2012-15	0.0087 (0.021)	0.0226 (0.027)	0.0186 (0.027)	0.0125 (0.028)	9.7274 (12.714)
<i>Panel B: Effects by Pre-Program Agency Quality</i>					
Elig*High*Post-2010	-0.0291 (0.036)	-0.0195 (0.044)	-0.0143 (0.044)	-0.0026 (0.042)	1.4770 (23.784)
Elig*Low*Post-2010	0.0289 (0.019)	0.0560** (0.026)	0.0487* (0.026)	0.0474 (0.030)	17.1335* (9.597)
N	18365	18365	18365	18365	18365
Comparison Mean	0.10	0.31	0.32	0.30	100.70

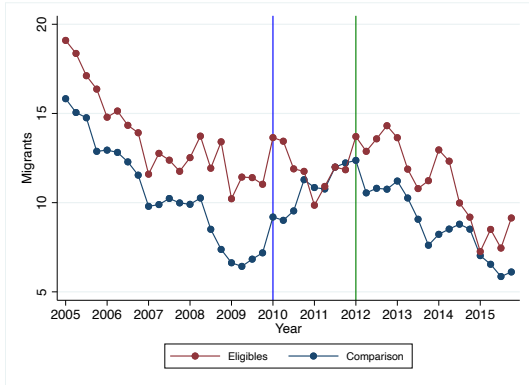
Notes: This table reports difference-in-difference estimates of effect of the program on individual components of the contract quality index. Panel A compares eligible agencies to comparison agencies. Panel B computes heterogeneous effects by pre-program agency quality; 'High' denotes above median pre-program quality, while 'Low' is below median pre-program quality, where quality is based on the 'Predicted Rating' variable. The outcome variables are binary variables that record whether a job order included the provision of return airfare (col. 1), health insurance (col. 2), accommodation (col. 3) and/or food (col. 4). The 'Avg. Salary' variable is either the midpoint of a salary range or the advertised salary (in 2015 USD) depending on what the job order included. All specifications cluster at the agency level and include agency fixed effects, month-year fixed effects interacted with average 2005-2008 migrant recruitment and month-year fixed effects interacted with the average quantitative score between 2005-2008.

Appendix B.4 Composition of Migrants Recruited

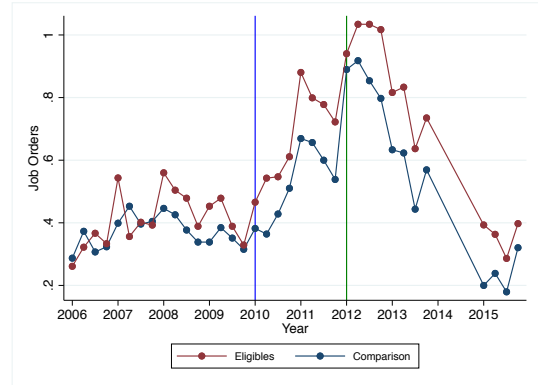
	(1)	(2)	(3)	(4)	(5)
	% Female	% Pre-Complaints	% Under 30	% Over 50	% Saudi
<i>Panel A: Effects by Program Phase</i>					
Elig*2010-12	0.0263 (0.031)	0.0054 (0.004)	0.0010 (0.011)	0.0115 (0.013)	-0.0238 (0.025)
Elig*2012-15	-0.0160 (0.036)	0.0015 (0.004)	0.0050 (0.014)	0.0034 (0.013)	-0.0136 (0.032)
<i>Panel B: Effects by Pre-Program Agency Quality</i>					
Elig*High*Post-2010	-0.0082 (0.061)	-0.0073 (0.006)	-0.0137 (0.019)	0.0103 (0.018)	-0.0403 (0.044)
Elig*Low*Post-2010	0.0079 (0.035)	0.0057 (0.005)	0.0032 (0.014)	0.0065 (0.015)	0.0037 (0.029)
N	21865	21865	21865	21865	21865
Comparison Mean	0.69	0.05	0.06	0.20	0.13

Notes: This table reports difference-in-difference estimates on the composition of migrants recruited by agencies. ‘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 both for the period where agencies were solely 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 A20 for details). All variables condition on whether non-zero migrants were recruited by an agency in a given month. Column 1 reports the share of migrants who are women, column 2 reports the share who had previously (i.e. pre-program) made complaints, column 3 reports the share of migrants under the age of 30, column 4 reports the share over 50, and, finally, column 5 reports the share whose destination was Saudi Arabia. All specifications cluster at the agency level and include agency fixed effects. Panel B contains month-year fixed effects interacted with average migrant recruitment (2005-08). Panel A contains the same controls as Panel B and month-year fixed effects interacted with the average yearly quantitative score (2005-08).

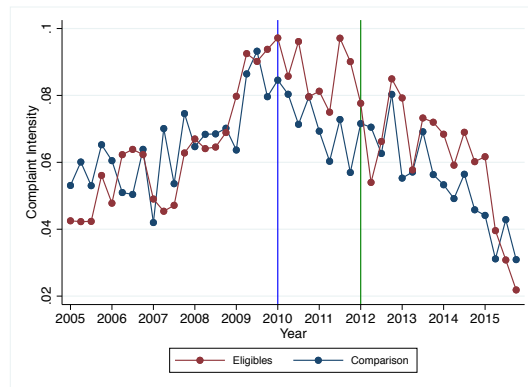
Appendix B.5 Raw Trends



Panel A: Migrant Recruitment



Panel B: Job Orders



Panel C: Complaint Intensity

Notes: These figures show quarterly averages in migrants recruited (Panel A), job orders received (Panel B), and the complaint rate (Panel C). The sample is restricted to agencies who recruit between 45 and 155 migrants in 2009 and who had a license that was valid for at least 1 year in 2009. The vertical blue line is the point prior at which agencies became aware of the program. The vertical green line shows when ratings were released to the public.

Appendix B.6 Components of Quantitative Score

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Professional	Clerical	Middle	Skilled	Semi-Skilled	Operators	Domestic	Solved Rate
<i>Panel A: Effects by Program Phase</i>								
Elig*2010-12	-0.0076 (0.015)	-0.0821 (0.062)	-0.0951 (0.102)	-0.0769 (0.723)	-0.0289 (0.051)	0.6417 (0.757)	0.3538 (0.994)	0.0096 (0.007)
Elig*2012-15	0.0133 (0.015)	0.6275 (0.669)	0.5709 (0.546)	0.8724 (0.666)	0.0496 (0.044)	1.8421 (1.210)	0.2613 (0.901)	-0.0184 (0.019)
<i>Panel B: Effects by Pre-Program Agency Quality</i>								
Elig*High*Post-2010	0.0046 (0.042)	0.9968 (1.044)	1.1373 (1.276)	-0.5032 (1.251)	0.0463 (0.088)	2.5717 (1.792)	-0.5258 (1.766)	-0.0040 (0.017)
Elig*Low*Post-2010	0.0057 (0.005)	-0.0225 (0.032)	-0.0236 (0.040)	1.1760** (0.567)	0.0215 (0.047)	0.7765 (1.158)	0.9329 (1.030)	-0.0099 (0.013)
<i>Panel C: Effects by Agency Quality and Program Phase</i>								
Elig*High*2010-12	-0.0199 (0.048)	-0.2919 (0.334)	-0.1554 (0.169)	-2.1400 (1.586)	-0.0847 (0.111)	-0.0286 (0.627)	-0.6396 (1.963)	0.0055 (0.010)
Elig*High*2012-15	0.0184 (0.055)	1.7253 (1.603)	1.8680 (1.965)	0.4219 (1.486)	0.1204 (0.097)	4.0414 (2.683)	-0.4615 (1.708)	-0.0137 (0.037)
Elig*Low*2010-12	0.0003 (0.008)	-0.0235 (0.033)	-0.0395 (0.044)	1.0342 (0.646)	-0.0050 (0.055)	0.8787 (1.055)	1.2434 (1.125)	0.0086 (0.009)
Elig*Low*2012-15	0.0087* (0.005)	-0.0219 (0.037)	-0.0145 (0.044)	1.2561** (0.567)	0.0365 (0.046)	0.7188 (1.245)	0.7574 (1.053)	-0.0241 (0.023)
N	27324	27324	27324	27324	27324	27324	27324	6759
Comparison Mean	0.03	0.08	0.12	1.62	0.11	1.73	3.94	1.00

Notes: This table reports difference-in-difference estimates on the components of the ‘Quantitative Score’ measure. ‘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 both for the period where agencies were solely 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. Columns 1-7 are count variables for migrants placed in varying skill categories. Column 8 reports estimated effects on the complaint solved rate (or dispute settlement rate) which conditions on the set of agencies who had non-zero complaints in a given month. All specifications cluster at the agency level and include agency fixed effects, month-year fixed effects interacted with average 2005-2008 migrant recruitment and month-year fixed effects interacted with average quantitative score 2005-2008.

Appendix B.7 The Effect of Receiving a Higher Rating

In order to separately identify the effect of receiving a higher rating, we instead use the sample of all eligible agencies and compare agencies that are just above and below the star-specific cutoffs. Appendix A.6 shows the distribution of ratings. The domain reports the Combined Score which is the continuous score that the *Quantitative Score* and *Qualitative Score* on which agencies were rated. In order to estimate the separate effect of receiving a higher rating, we argue that agencies on either side of the star-cutoff are similar aside from the fact that they are arbitrarily assigned to receive either more or less stars.

Empirical Strategy: Regression Discontinuity Design

To operationalize this identification strategy we use a regression discontinuity design. To maximize power, we center the treatment variable at the thresholds and assign a value of 1 if the score is above the threshold and 0 if it is below the threshold.

For agency i in month m and year t :

$$Y_{imt} = \delta_t + \beta_1 \text{Combined Score}_i + \beta_2 \text{Treat}_i + \beta_3 \\ \text{Treat} * \text{Combined Score}_i + \epsilon_{imt}$$

In this regression, Y_{imt} is our outcome variable measured for agency i in month m and year t . Treat_i is a dummy that takes a value of 1 when the score is above the star cutoff. Combined Score_i is the centered forcing variable and $\text{Treat} * \text{Combined Score}_i$ is an interaction between the treatment and forcing variable. The interaction term allows the effect from the rating to vary on either side of the cutoff. We also include year fixed effects, δ_t , and cluster standard errors at the agency level. We use a uniform kernel and the mean-square error optimizing bandwidth.

Appendix B.8 Heterogeneous Effects by Local Market Competition

To explore how the effects of the program vary with local market competition, we construct a proxy based on the density of nearby agencies owing to the geographical clustering of agencies discussed in Section 7.1. We use GPS coordinates to compute the number of agencies in a 100 foot radius for each agency and denote those located near an above median number of agencies as ‘competitive’ and those with a below median number of agencies as ‘not competitive’.

In the table below, we compute effect heterogeneity by local market competition using a specification that includes a local market fixed effect and its interaction with a set of year fixed effects. When comparing eligible agencies to comparison agencies that are located in areas with a lot of competition, we find that eligible agencies have significantly more job orders—both on average and from Good Employers—and are both more likely to recruit migrants and less likely to exit the market. We also note that the point estimates are quantitatively similar for the analogous comparisons for agencies facing less competition. We find similar results when using the Herfindahl-Hirschman index to measure local competition.

	(1) Migrants	(2) Agency Exit	(3) Job Orders	(4) Good Orders
Not Competitive*Elig*Post-2010	3.1065 (6.137)	-0.0664 (0.072)	-0.1567 (0.099)	0.0761* (0.040)
Competitive*Elig*Post-2010	5.2195* (2.890)	-0.0843* (0.044)	0.2357** (0.108)	0.0785** (0.035)
N	25872	25872	21168	17650
Comparison Mean	7.61	0.16	0.34	0.17

Notes: This table reports effect heterogeneity by local market competition. ‘Comparison Mean’ reports the pre-program (2005-2009) mean of the dependent variable for the comparison group. ‘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 specific 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. To proxy for local competition we establish agency density by counting the number of active agencies in 2009 in a 100 foot radius. ‘Competitive’ refers to a binary variable coded as 1 for an above median score of the measure of competition, while ‘Not Competitive’ is the opposite. This specification is clustered at the agency level and includes agency fixed effects, month-year fixed effects interacted with average 2005-2008 migrant recruitment, month-year FE interacted with average Quantitative Score 2005-2008, and local market (i.e. Divisional Secretariat) x month-year fixed effects.

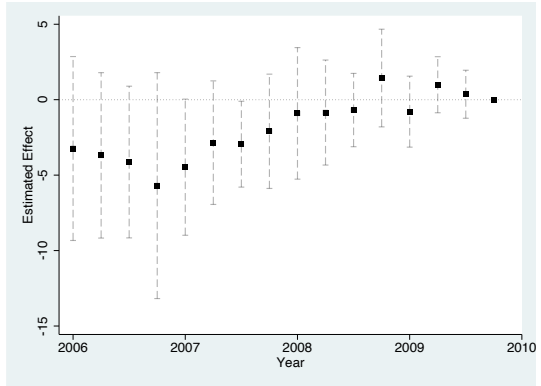
Appendix C Robustness Tests

Appendix C.1 Pre-Trends

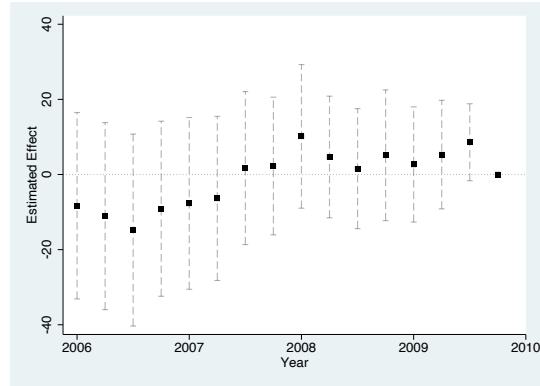
(1) Variable	(2) Sample	(3) F-Stat	(4) p-value
Migrant	All	1.0658	0.3868
	Low	1.2370	0.2862
	High	0.8457	0.5543
Job Orders	All	1.0715	0.3830
	Low	1.4312	0.1970
	High	0.2126	0.9812
Complaint Rate	All	0.9286	0.4853
	Low	1.0334	0.4104
	High	0.9956	0.4435
Quant Score	All	1.0374	0.4060
	Low	1.1191	0.3542
	High	1.7375	0.1177
Avg. Salary	All	0.8198	0.5718
	Low	0.5606	0.7869
	High	1.2228	0.3049
Contract Renewal	All	0.3316	0.9389
	Low	0.6856	0.6841
	High	1.6407	0.1419
Good Orders	All	1.0799	0.3775
	Low	1.9224	0.0700
	High	0.3333	0.9356
Contract Quality	All	0.5033	0.8314
	Low	0.3106	0.9483
	High	0.9315	0.4892

Notes: This table reports joint F-tests of lags for all the main outcomes analyzed in the paper. For each outcome, we separately report the F-test for all agencies, only pre-program low quality agencies, and only pre-program high quality agencies. In each case, we test the joint significance of 15 lags (each a period corresponding to a quarter) for the period 2006-2009 where the reference period is always the quarter preceding Jan 2010. Standard errors are clustered at the agency level.

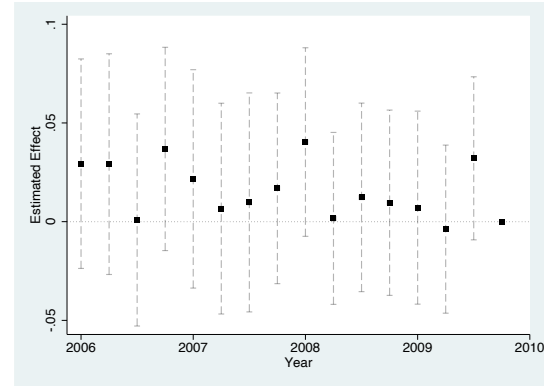
Appendix C.2 Plots of Lagged Effects



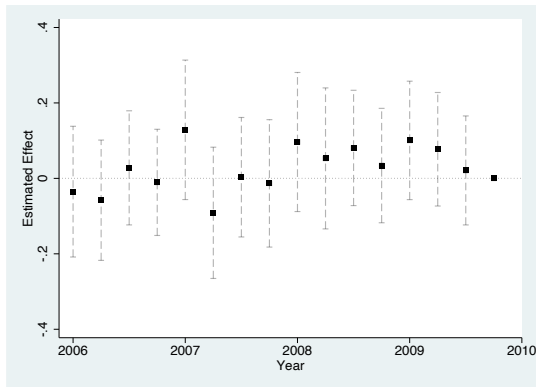
Panel A: Migrants Recruited



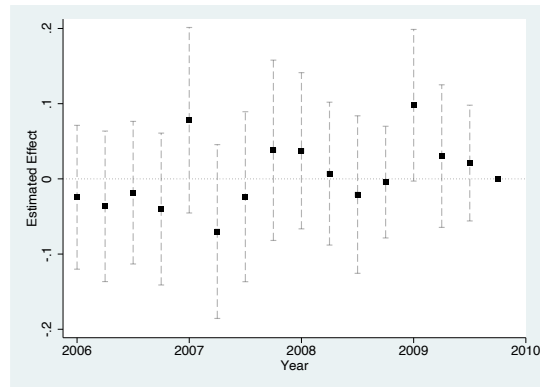
Panel B: Avg. Salary



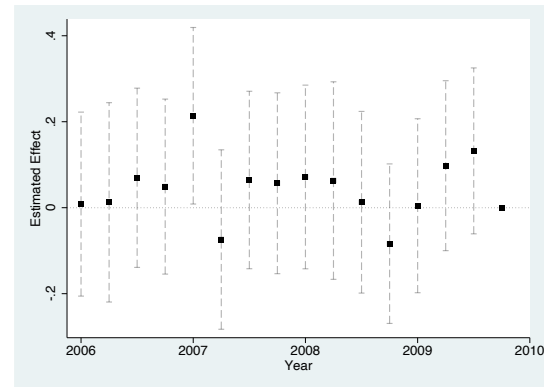
Panel C: Contract Renewal



Panel D: Job Orders



Panel E: Good Orders



Panel F: Contract Quality

Notes: These figures show event study graphs for lagged outcomes. They plot the interaction coefficients—an indicator for an eligible firm interacted with a series of quarterly indicators for 2006-2010. The capped dotted lines are the 95% confidence intervals. In each case the reference is the final quarter of 2009. Standard errors are clustered at the agency level.

Appendix C.3 Agency Exit

	(1)	(2)	(3)
	Agency Exit (6 mo)	Agency Exit (1 yr)	Agency Exit (2 yr)
<i>Panel A: Effects by Program Phase</i>			
Elig*2010-12	-0.0666*** (0.026)	-0.0701*** (0.026)	-0.0666*** (0.026)
Elig*2012-15	-0.0791* (0.044)	-0.0826* (0.044)	-0.0741* (0.043)
<i>Panel B: Effects by Pre-Program Agency Quality</i>			
Elig*High*Post-2010	0.0013 (0.072)	0.0067 (0.074)	0.0133 (0.073)
Elig*Low*Post-2010	-0.1196*** (0.038)	-0.1256*** (0.039)	-0.1173*** (0.038)
N	27324	27324	27324
Comparison Mean	0.17	0.16	0.15

Notes: This table reports difference-in-difference estimates for agency exit. ‘Comparison Mean’ reports the post-program (2010-2015) mean of the dependent variable for the comparison group. Panel A compares eligible agencies to comparison agencies both for the period where agencies were solely 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. ‘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 6 month (column 1), 12 month (column 2) and 24 month (column 3) sequence of zero recruitment after the program was introduced. All specifications cluster at the agency level and include agency fixed effects. Panel B contains month-year fixed effects interacted with average migrant recruitment (2005-08). Panel A contains the same controls as Panel B and month-year fixed effects interacted with the average yearly quantitative score (2005-08).

Appendix C.4 Total Salaries

	(1) Total Salaries	(2) Total Salaries	(3) Total Salaries
Elig*2010-12	0.6958** (0.289)		
Elig*2012-15	0.7053** (0.341)		
Elig*High*Post-2010		0.7056 (0.537)	
Elig*Low*Post-2010		1.0911*** (0.396)	
Any Elig*Post-2010			0.9067 (0.720)
N	25412	25412	14674
Comparison Mean	5.33	5.33	5.33

Notes: This table reports program effects on total salaries, a proxy for match surplus. Col. 3 restricts the sample to comparison agencies. 'Comparison Mean' reports the pre-program (2005-2009) mean of the dependent variable for the comparison group. 'Total Salaries' is computed by summing all salaries of migrants placed by an agency in the relevant month, and this variable is transformed using the inverse hyperbolic sine. The specification in col. 2 contains month-year fixed effects interacted with avg. migrant recruitment (2005-08), clusters at the agency level, and includes agency fixed effects. Col. 1 contains the same controls and month-year fixed effects interacted with the average yearly quantitative score (2005-08). 'Any Elig' is a dummy variable coded as 1 if there is an eligible agency within 100 ft of a comparison agency. The estimate in col. 3 includes the same controls as col. 1 but is clustered at the local market level and includes local market x month-year fixed effects.

Appendix C.5 Complaint Rate

	(1) Complaint Rate	(2) % Breach	(3) % Harassment
<i>Panel A: Effects by Program Phase</i>			
Elig*2010-12	0.006 (0.007)	-0.005 (0.023)	0.012 (0.018)
Elig*2012-15	-0.002 (0.006)	0.051** (0.024)	-0.020 (0.019)
<i>Panel B: Effects by Agency Quality</i>			
Elig*High*Post-2010	-0.003 (0.008)	0.013 (0.040)	-0.005 (0.035)
Elig*Low*Post-2010	0.004 (0.007)	0.025 (0.022)	-0.004 (0.018)
<i>Panel C: Effects by Agency Quality and Program Phase</i>			
Elig*High*2010-12	0.003 (0.009)	-0.034 (0.046)	0.054 (0.036)
Elig*High*2012-15	-0.008 (0.011)	0.060 (0.053)	-0.065 (0.039)
Elig*Low*2010-12	0.008 (0.009)	-0.000 (0.025)	0.002 (0.020)
Elig*Low*2012-15	0.001 (0.007)	0.048* (0.026)	-0.009 (0.022)
N	19659	8470	8470
Comparison Mean	0.08	0.40	0.14

Notes: This table reports difference-in-difference estimates on complaints. Panel A compares eligible agencies to comparison agencies. 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. Panel C reports estimates by agency quality and program phase. 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. % Breach reports the fraction of complaints that are due to breach of contract (including non-payment of wages) while % Harassment refers to the fraction of complaints that are due to physical harassment. Column 1 conditions on non-zero migrants recruited, while columns 2 & 3 additionally condition on non-zero complaints. All specifications cluster at the agency level and include agency fixed effects. The specification reported in Panel B & C contain month-year fixed effects interacted with average migrant recruitment (2005-08). Panel A contains the same controls as Panel B and month-year fixed effects interacted with the average yearly quantitative score (2005-08).

Appendix C.6 Robustness to Agency Exit

	(1)	(2)	(3)
	Quant Score	Migrants	Job Orders
<i>Panel A: Effects by Program Phase</i>			
Elig*2010-12	0.2681 (0.211)	-0.2830 (1.633)	0.0760 (0.061)
Elig*2012-15	0.2854 (0.202)	3.6344 (3.540)	0.1123 (0.081)
<i>Panel B: Effects by Pre-Program Agency Quality</i>			
Elig*High*Post-2010	-0.2284 (0.279)	2.6620 (5.906)	-0.0296 (0.119)
Elig*Low*Post-2010	0.5720*** (0.207)	1.6925 (2.186)	0.1793** (0.082)
N	25412	25412	20869
Comparison Mean	6.25	6.92	0.34

Notes: This table reports difference-in-difference estimates for main outcomes that are adjusted for agency exit. ‘Quant Score’ summarizes an agency’s investment in the rating criteria introduced by the government (See Appendix A.10). ‘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 specific month from employers abroad (i.e. demand). Note, data for 2005 and 2014 is missing for the job order data. All specifications cluster at the agency level and include agency fixed effects. The specification reported in Panel B contains month-year fixed effects interacted with average migrant recruitment (2005-08). Panel A contains the same controls as Panel B and month-year fixed effects interacted with the average yearly quantitative score (2005-08).

Appendix C.7 Robustness to Trend Controls

	(1)	(2)	(3)	(4)
	Quant Score	Avg. Salary (USD)	Job Orders	Agency Exit
<i>Panel A: Quality Trend Only</i>				
Elig*2010-12	0.2357 (0.154)	12.5657 (8.765)	0.1074* (0.061)	-0.0803*** (0.028)
Elig*2012-15	0.0805 (0.127)	16.7593 (10.394)	0.1330* (0.078)	-0.0895** (0.045)
<i>Panel B: Recruitment Trend Only</i>				
Elig*2010-12	0.3426 (0.209)	12.1132 (8.810)	0.1121* (0.063)	-0.0703*** (0.026)
Elig*2012-15	0.2537 (0.195)	15.2904 (10.434)	0.1508* (0.079)	-0.0846* (0.044)
<i>Panel C: No Controls</i>				
Elig*2010-12	0.1254 (0.222)	12.3014 (8.732)	0.1020* (0.061)	-0.0821*** (0.029)
Elig*2012-15	-0.0585 (0.235)	16.6366 (10.253)	0.1247 (0.079)	-0.0889* (0.046)
N	27324	21865	22356	27324
Comparison Mean	5.80	178.18	0.34	0.17

Notes: This table assesses the robustness of the difference-in-difference estimates reported in the paper to varying combinations of trend controls. ‘Comparison Mean’ reports the pre-program (2005-2009) mean of the dependent variable for the comparison group. Panel A reports the estimates including only month-year fixed effects interacted with the average yearly quantitative score (2005-08). Panel B uses month-year fixed effects interacted with average migrant recruitment (2005-08). Panel C only includes month-year fixed effects. ‘Quant Score’ summarizes an agency’s investment in the rating criteria introduced by the government (See Appendix A.10). ‘Job Orders’ are the number of orders that an agency received in a specific month from employers abroad (i.e. demand). Note, data for 2005 and 2014 is missing for the job order data. ‘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 specifications cluster at the agency level and include agency fixed effects.

Appendix C.8 Mean Reversion and Alternative Measures of Agency Quality

One concern with our measure of pre-program agency quality might be that it captures something transient as a result of measurement error and that agency quality may exhibit mean reversion thereafter. Put otherwise, it is not that the program improves outcomes for low quality agencies, but that these were simply high quality agencies that were improperly classified and who then trended back to their mean. We believe this concern is unlikely to explain our estimates for a number of reasons.

First, it is important to note that high and low quality eligible agencies are always being compared to high and low quality comparison agencies (see Appendix A.7).⁴⁷ As such, if there is measurement error in how we specify quality and, for example, a subset of low quality eligible agencies are trending upward because of measurement error, this should also be the case for low quality comparison agencies. The estimated heterogeneous treatment effects are the differential effect for high/low eligible agencies relative to high/low comparison agencies, whose trend is computed by the High*Post term in equation (6).⁴⁸ As such, our estimates 'net out' any potential measurement error. If the measurement error were itself classified with treatment assignment then it would bias our estimates; however this amounts to a test of pre-trends for pre-program agency quality. Using five years worth of (monthly) pre-program data, we find that we cannot reject the null for the joint F-test of the lags for the Quantitative Score variable (p-value = 0.4807).⁴⁹

Nevertheless, we take concerns about measurement error in pre-program agency quality seriously and in Appendix C.8 we assess the robustness of our main results to alternative measures of pre-program agency quality. In Panel A, we report our preferred measure, the Predicted Rating, while in Panel B we report heterogeneous treatment effect estimates by pre-program agency quality defined using the Quantitative Score from 2009. In Panel C we do likewise with the average yearly Quantitative Score between 2007 and 2009. The point estimates are similar in both magnitude and direction across the three measures of pre-program agency quality. We do, however, note that the precision of the estimates is somewhat lower when pre-program agency quality is measured using the average Quantitative Score between 2007-2009 (Panel C).

Our view here is that the Predicted Rating and the 2009 Quantitative Score yield agency quality measures that are more salient from the perspective of an eligible agency. As the program is announced in early 2010 and lays out the criteria for a rating that is forward looking, an agency that was high quality three years ago but is low quality in 2009 can't expect to receive a good rating in the future. In addition, we believe the Quantitative Score only captures a part of the government's ultimate rating which also depends on the 'Qualitative Score' which is in turn a function of third-party inspections. By using the Predicted Rating, we weight the qualitative and quantitative scores by how important they are for the actual rating yields a better measure of actual quality as conceived by the government.

⁴⁷There are 207 agencies in the sample, 78 eligible agencies (30 high quality, 48 low quality) and 129 comparison agencies (30 high quality, 99 low quality).

⁴⁸Note, both High*Post and Low*Post cannot be included because of the inclusion of month-year fixed effects.

⁴⁹Though not a formal test, it is also useful to note that the point estimates in Table 2 & 3 do not suggest a systematic trend in the estimates for high and low agencies; in some cases the estimated effect is positive for low quality agencies and negative for high quality agencies, in others it is the opposite and in still others the effects are in the same direction. It is, therefore, surprising why some outcomes would revert to the mean but not others unless the measurement error were, for some reason, weighted towards specific outcomes. We have no reason to believe that this is the case.

Robustness of Estimates to Varying Measures of Pre-Program Agency Quality

	(1) Quant Score	(2) Migrants	(3) Job Orders	(4) Agency Exit	(5) Good Orders	(6) Contract Quality
<i>Panel A: Predicted Rating</i>						
Elig*High*Post-2010	-0.3422 (0.265)	2.1827 (5.407)	-0.0105 (0.114)	0.0067 (0.074)	0.0764* (0.045)	-0.0629 (0.151)
Elig*Low*Post-2010	0.6574*** (0.200)	1.8596 (2.093)	0.2204*** (0.083)	-0.1256*** (0.039)	0.0498** (0.025)	0.1669** (0.078)
<i>Panel B: Quantitative Score 2009</i>						
Elig*High*Post-2010	0.0603 (0.261)	1.9859 (3.951)	0.1353 (0.099)	-0.0524 (0.056)	0.0686** (0.034)	0.1168 (0.109)
Elig*Low*Post-2010	0.6837*** (0.220)	2.4841 (1.941)	0.1627* (0.087)	-0.1136** (0.052)	0.0470 (0.030)	0.0810 (0.093)
<i>Panel C: Avg. Quantitative Score 2007-2009</i>						
Elig*High*Post-2010	0.2727 (0.206)	-2.0280 (2.111)	0.0929 (0.106)	-0.0618 (0.055)	0.0773** (0.034)	0.1303 (0.105)
Elig*Low*Post-2010	0.6494*** (0.207)	7.5747 (4.638)	0.1887** (0.082)	-0.1052** (0.051)	0.0281 (0.027)	0.0427 (0.091)
N	25956	25956	21684	25956	21684	21684
Comparison Mean	6.56	8.05	0.37	0.16	0.13	0.07

Notes:

This table reports heterogeneous difference-in-difference estimates using alternative definitions of pre-program agency quality. Panel A uses our preferred specification, i.e. the 'Predicted Rating'. Panel B uses the Quantitative Score from 2009, while Panel C uses the average Quantitative Score between 2007 and 2009 (see Appendix A.10 for details). 'Comparison Mean' reports the pre-program (2005-2009) mean of the dependent variable for the comparison group. 'Quant Score' summarizes an agency's investment in the rating criteria introduced by the government (See Appendix A.10). '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 specific month from employers abroad (i.e. demand). Note, data for 2005 and 2014 is missing for the job order data. '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. '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 specifies 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 specifications cluster at the agency level and include agency fixed effects and month-year fixed effects interacted with average 2005-2008 migrant recruitment.

Appendix C.9 Program Effects with Alternative Samples

	(1)	(2)	(3)
	Agency Exit	Job orders	Migrants
<i>Panel A: Recruited 45-155 Migrants in 2009</i>			
Elig*2010-12	-0.0701*** (0.026)	0.1067* (0.062)	0.1991 (1.585)
Elig*2012-15	-0.0826* (0.044)	0.1451* (0.078)	3.3784 (3.266)
N	27324	22356	27324
<i>Panel B: Recruited 25-175 Migrants in 2009</i>			
Elig*2010-12	-0.0521** (0.022)	0.1453*** (0.055)	0.5659 (1.306)
Elig*2012-15	-0.0856** (0.039)	0.1897*** (0.069)	1.7608 (2.787)
N	39204	32076	39204
<i>Panel C: All Agencies</i>			
Elig*2010-12	-0.0657*** (0.019)	0.2951*** (0.046)	2.1962 (1.651)
Elig*2012-15	-0.1354*** (0.030)	0.3889*** (0.061)	-1.0664 (2.242)
N	86064	70416	86064

Notes:

This table reports difference-in-difference estimates of the overall program effect using varying intervals about the 100 migrant program threshold. Panel A reports the 'optimal bandwidth' used in the paper (45-155 migrants recruited in 2009), Panel B restricts the sample to 25-175 migrants recruited in 2009 and Panel C places no restrictions on the sample. '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 specific month from employers abroad (i.e. demand). Note, data for 2005 and 2014 is missing for the job order data. All specifications cluster at the agency level and include agency fixed effects, month-year fixed effects interacted with average migrant recruitment (2005-08) and month-year fixed effects interacted with average yearly quantitative score (2005-08).