Physicians' Occupational Licensing and the Quantity-Quality Trade-off*

Juan Pablo Atal   Tomás Larroucau   Pablo Muñoz   Cristóbal Otero

March 20, 2025

PRELIMINARY
PLEASE DO NOT CIRCULATE

Abstract: Occupational licensing is a quality regulation that increases quality but reduces quantity. We provide a framework to empirically quantify this trade-off and apply it to physicians' licensing, where both access to care and quality of care are a primary concern. Using quasi-exogenous variation driven mostly by a recent and unprecedented migration of physicians to Chile, we show that more physicians improve access, but that their performance in the licensing exam matter for patient outcomes, including mortality. Given this trade-off, we assess the optimality of the current licensing policy in the country by quantifying the effects of locally changing the current licensing threshold. We find that lowering the threshold would improve outcomes on the net.

*This version: March 20, 2025. Latest version is here. Atal: University of Pennsylvania and NBER, ataljp@econ.upenn.edu. Larroucau: University of Arizona, tomas.larroucau@asu.edu. Muñoz: Universidad de Chile, pablomh@uchile.cl. Otero: Columbia Business School, c.otero@columbia.edu. We would like to thank Francesco Agostinelli, Hal Cole, Matt Wiswall, and seminar participants at Fordham University, Johns Hopkins University, University College London, Princeton University, Queen Mary University, Universidad de los Andes-Chile, Pontificia Universidad Católica de Chile, Universidad Chile, and Universidad de Santiago, for valuable comments and suggestions. Pablo Muñoz thanks FONDECYT Iniciación (ANID-FONDECYT-11230049) and the Millenium Nucleus MIGRA (ANID-MILENIO-NCS2022051) for financial support. Nicolas Bozzo, Simon Andrade, and Sofía Pontigo provided superb research assistance. We thank the Health Ministry of Chile for access to data. All remaining errors are our own.

1 Introduction

The shortage of physicians is a long-standing and increasingly urgent concern. Providing adequate healthcare access is estimated to require at least an additional 7 million physicians globally, and a deficit of 125 thousand physicians is projected by 2034 for the United States alone (Haakenstad et al., 2022; Association of American Medical Colleges, 2024).

Licensing requirements in the medical profession have been routinely signaled as a key contributor to physician shortages with unclear benefits from increased quality.[1] However, assessing the quality-quantity trade-off embedded in licensing requirements has proven difficult; it requires credible estimates of its effects on the quantity and quality of physicians, as well as their impacts on access and quality of care. At the same time, a growing migration of physicians worldwide has made licensing an increasingly binding policy in destination countries like the US, potentially changing the nature and magnitude of its associated trade-offs.[2]

In this paper, we provide a simple framework to highlight the key economic fundamentals that govern the quantity-quality trade-off when deciding the stringency of licensing requirements. This framework reveals a set of sufficient statistics needed to evaluate the optimality of the licensing requirements in place, by quantifying the net benefits of locally changing them. We then use rich administrative data from Chile to estimate those sufficient statistics and evaluate the optimality of the licensing policy currently used to appoint public sector physicians in the country.

Chile provides an ideal setting to study physician licensing and how its design may depend on changing labor market fundamentals like increased immigration. The number of physicians taking the licensing exam increased fivefold between 2009 and 2018, largely driven by an unprecedented increase in the number of foreign test-takers. In tandem with this surge, passing rates decreased from 92% to 58% during this period. The migration generates useful variation to credibly estimate the effects of quantity and quality on patient outcomes, as well as to study policies —like changing the licensing stringency– that impact both quantity and quality concurrently.

Despite the large migration influx, the healthcare system is still strained and 25% of the country's annual mortality is explained by individuals who die while on a waiting list.[3] Finally, the Chilean

[1] A highly influential criticism of licensing in the medical profession can be found in Friedman (1962). Svorny (2004) provides a more recent literature review.

[2] In the United States, the number of physicians born and educated abroad has increased by 30% since 2004 and currently represents 20% of the total count (Association of American Medical Colleges, 2023). Similarly, in OECD countries, an increasingly growing share of physicians are foreign-trained physicians and now constitute 30% of the physician workforce (Socha-Dietrich and Dumont, 2021). Still, it is estimated that only one-third of employed foreign physicians are on track to practice medicine and that licensing plays a critical role in explaining this fact (Federal Reserve Bank of Minneapolis, 2022).

[3] As of 2019, Chile had 17.5 physicians per 10,000 inhabitants—nearly half the average of countries with a com-

licensing exam has characteristics and purposes similar to those applied in different countries, including the US.[4]

Our theoretical framework highlights the quantity-quality trade-off embedded in the design of licensing policies. There is a production function for hospital care, where the quantity and quality of physicians are key inputs. The planner faces an exogenous distribution of scores and has a licensing technology that provides a noisy signal of a physician's quality. The planner grants licenses to those with a signal above a threshold. Changing the licensing threshold affects outcomes depending on two sets of sufficient statistics; the output elasticities with respect to the quantity and quality of labor, and the elasticity of those inputs with respect to the licensing threshold. The input elasticities, in turn, depend on the fraction of test-takers at the threshold, their score relative to the score of supra-marginal physicians, and the precision of the licensing score as a signal of quality (i.e., the signal-to-noise ratio).

The empirical part of the paper is devoted to estimating the sufficient statistics needed to evaluate the effects of locally changing the licensing threshold on health outcomes in the Chilean public health care sector. We begin by estimating the latent quality of physicians with a model of licensing scores and retaking behavior. The model enables us to leverage the observed test score histories to infer the precision of the licensing test and to predict test-taking behavior under counterfactual thresholds, which impacts the number of test-takers over time. We find that the test has a relatively high signal-to-noise ratio, especially for foreigners. As a consequence, scores are informative proxies for physicians' latent qualities, and licensing policies can indeed affect the quality distribution across hospitals. To estimate the resulting elasticities of the quantity and quality of labor with respect to the licensing threshold, we estimate a labor matching function between physicians and hospitals. Guided by theory, we approximate the matching function as a flexible hospital-specific function of physicians' latent quality and other observables, as well as of observable proxies for labor-market tightness. As such, the labor-market elasticities vary with counterfactual licensing thresholds as tightness depends on how many physicians pass the exam. Our estimated input elasticities reveal the increase in quality and decrease in quality that would follow from a marginal decrease in the licensing threshold and the quantity-quality tradeoff at the margin. Empirically, we find that in the average hospital, lowering the threshold to increase quantity by 1 percent would reduce average

parable burden of diseases, injuries, and risk factors (33.4 physicians per 10,000). This figure also falls below the minimum threshold of 20.7 physicians per 10,000 needed to achieve an effective Universal Healthcare coverage score of 80 out of 100 at the global level proposed by Haakenstad et al. (2022). The shortage of physicians is also reflected in the current long waiting lists, with approximately 3 million individuals—15% of the population—awaiting medical attention. An estimated 40,000 individuals die annually while on waiting lists, representing 25% of the country's annual mortality.

[4] Similar tests are the USMLE in the US, the NCE in France, and the NMLE in Japan.

quality by 0.32 percent.

We proceed by providing novel estimates of the output elasticities with respect to the quantity and quality of physicians. We focus on access and quality of care as key outcomes, which jointly determine the value added provided by the healthcare sector. We measure access primarily as the number of patients seen as a fraction of potential patients (i.e., the "service rate"), and we measure quality of care by patient mortality. As a key contribution, we are able to address the identification challenges related to the endogeneity of inputs in the estimation of production functions with an instrumental variables (IV) approach. Specifically, we use a shift-share IV design leveraging the increase in the labor supply of physicians brought primarily by the immigration wave as well as by the expansion of medical schools in the country (Altonji and Card, 1989; Autor et al., 2013; Borusyak et al., 2022). Our estimates reveal that access to care significantly increases with the number of physicians with an estimated elasticity of the service rate with respect to the number of physicians of 1. On the other hand, the service rate is not affected by quality. However, both the quantity and quality of physicians matter for quality of care. We document that the elasticity of the in-hospital death rate with respect to the number of physicians is 0.8 and that one point (0.23 SD) increase in the average physicians' quality decreases death rates by 0.04%. We show that these results are not driven by changes in patient composition and are robust to alternative measures of hospital performance.

Equipped with estimates for the sufficient statistics, we quantify the net benefits of marginally lowering the licensing threshold in the Chilean healthcare context. We begin by assessing the short-term effects of lowering the threshold in 2018, the end of our sample period, and five years after the onset of the large migration wave. Our results imply that lowering the threshold would have increased access with minimal detrimental effects on quality. On net, population health would have improved as a consequence. We then investigate how the short-run effects differ across the different phases of the migration wave. Overall, we find a robust positive short-run effect of the policy throughout our sample period, despite the large changes in market fundamentals produced by the migration wave. Still, the policy effects are non-monotonic over time; the largest returns would have occurred in the middle of the migration wave when the number of marginal physicians was large and the stock of physicians already hired was still relatively low. We finish by estimating the long-run effects of the policy, considering that retaking can dampen the effects of licensing thresholds as scores may improve over attempts. In fact, in our context, more than 80% of test-takers who failed their first attempt in 2013 passed by 2018. As a consequence, we find that the effects of lowering the threshold decrease with time. However, there are still net benefits in 2018 of permanently lowering the threshold in 2013.

Our paper contributes to several strands of the literature. We add to a long-standing debate about the desirability of occupational licensing in healthcare (Friedman and Kuznets, 1945; Friedman, 1962; Svorny, 2004; Kleiner, 2014; Kleiner et al., 2016) and other settings, particularly in the public sector (Kleiner and Wang, 2023; Angrist and Guryan, 2008; Kleiner, 2011; Larsen et al., 2020). Relative to this literature, we provide a framework for understanding the quantity-quality trade-off embedded in licensing and empirically evaluate the stringency of licensing policies in relevant outcomes. Close to our work, Kleiner and Soltas (2023) provides a general equilibrium framework to estimate the welfare effect of licensing policies in the US under competitive product and labor markets. In their framework, labor quality is inferred from wages and equilibrium conditions in the labor market. In contrast, we leverage direct measures of relevant outcomes (i.e. access and quality of care) to derive the net effect of licensing on them.

We also contribute by providing quasi-experimental evidence on the impact of physician quantity on health outcomes. Previous work has shown a correlation between location mortality and the number of physicians (Finkelstein et al., 2021) or infant health and the number of primary care physicians (Carrillo and Feres, 2019). Instead, we use quasi-exogenous variation to show direct evidence that physician quantity matters for health outcomes in tertiary care. Moreover, we complement previous research studying whether physician quality matters. Examples in the literature include the impacts of elite medical training (Doyle et al., 2010), medical school exit exams (Guarin et al., 2021), and physicians' value-added (Fletcher et al., 2014; Ginja et al., 2024). However, to our knowledge, there is no empirical evidence showing that licensing scores are associated with health outcomes. Finally, we contribute to a scant literature estimating production functions in healthcare (Grieco and McDevitt, 2017; Arnold, 2025) by exploiting exogenous variation in inputs to estimate the elasticity of service rates and mortality with respect to the quantity and quality of physicians.

The remainder of the paper is organized as follows. Section 2 describes the setting and the data. In Section 3, we provide descriptive evidence to illustrate the main variation leveraged to assess the quantity-quality trade-off. Section 4 presents a stylized model of the licensing problem. In Sections 5 and 6, we present the model and estimates of the input and output elasticities. We present our main counterfactual exercises in Section 7 and conclude in Section 8.

2 Setting and Data

2.1 Institutional Setting

The Chilean Public Healthcare System The healthcare system in Chile is divided into public and private insurers and providers. Our focus is on public providers, which mainly serve patients with public insurance. Public insurance covers approximately 80% of the population and is financed through monthly contributions deducted from labor income, cost-sharing mechanisms, and resources from the general government.[5]

The public health insurance provides coverage within the network of public providers, with varying levels of copayment determined by income and family size. Individuals who cannot afford to pay are granted free access to the public system, ensuring nearly universal health coverage in public hospitals. Beneficiaries of the public insurer may also opt for private providers, although the copayment for such "out-of-network" services is significantly higher than those in the public network.

The network of public hospitals is composed of 181 hospitals that belong to one of 29 different referral regions called Health Services (*Servicios de Salud*).[6] The referral and counterreferral system also operates at this level. Individuals must register with their local primary healthcare provider, and those requiring specialized care are referred to one of the public hospitals within their region. Referrals follow strict and predetermined guidelines based on diagnosis, location, and other patient demographics.[7] Hospitals are categorized into three levels of complexity, low, medium, and high, depending on their size and the range of medical services they offer.

Physician Licensing To work in the healthcare sector, physicians are required to pass the EUNACOM exam (*Examen Único Nacional de Conocimientos de Medicina*).[8] The exam was established in 2009 and follows the characteristics and purposes of medical licensing exams in other

[5] Individuals also have the option to redirect their health contributions toward purchasing private insurance, which covers 15% of the population. The remainder 5% is covered through schemes exclusively for the police and armed forces. In practice, public insurance primarily serves the relatively more disadvantaged population, while wealthier, healthier, and younger individuals tend to opt for private insurance (Pardo, 2019).

[6] The number of 181 hospitals corresponds to hospitals that are present every year throughout our study period 2011-2019.

[7] Patients can also be admitted directly to hospitals through the ER in cases of emergency.

[8] Although the exam is not technically mandatory for physicians in private healthcare organizations, it is effectively binding across all healthcare institutions because passing it is legally required to treat patients covered by public health insurance, regardless of the treatment location. This explains why, *de facto*, most private healthcare institutions only consider applications admissible if candidates include EUNACOM approval. The performance in the EUNACOM is also determinant for accessing training opportunities, as it plays a key role in admissions to residency programs in the country.

countries.[9] It comprises a theoretical section and a practical section. The theoretical section corresponds to a multiple-choice test with 180 questions covering various areas of medical knowledge. The score reflects the candidate's absolute performance and is not standardized relative to other test-takers.[10] To pass, candidates must achieve a score of 51 or higher out of 100 points.

Upon passing the theoretical section, the candidates must complete a pass-or-fail practical section, which involves an examination in a real or simulated clinical environment. However, this requirement is waived for candidates with medical degrees from local universities and for physicians with a medical degree from a select group of countries with which the Ministry of Health has bilateral agreements.[11] The practical portion of the exam is largely not binding (Kunakov et al., 2018).

Importantly, approving the EUNACOM automatically validates the medical degrees of foreign-trained physicians, granting them the same job and training opportunities as locally trained physicians.[12]

Physician's Hiring and Wages in the Public Sector: Physicians can work in both the private and public sectors. In the private sector, employment operates under standard market dynamics, where wages, benefits, and working conditions are negotiated directly between employers and employees. In contrast, in public hospitals, wages are legally regulated and follow a public-sector wage schedule, with annual adjustments based on sector-wide wage revisions.[13] Hiring in public hospitals is decentralized and managed by the regional Health Services within the budgetary constraints established by the Ministry of Health. Health Services determine staffing needs for hospitals and hospital directors oversee the recruitment process. Although passing the EUNACOM exam is sufficient for eligibility, hospitals reportedly consider the candidate's EUNACOM score an important factor in hiring decisions.

[9] Comparable exams include the Medical Licensing Examination (USMLE) in the U.S., the Medical Council of Canada Qualifying Examination (MCCQE) in Canada, the National Competency Exam (NCE) in France, the National Medical Licensing Examination (NMLE) in Japan, and the Korean Medical Licensing Examination (KMLE) in Korea. For an in-depth discussion of the exam and evidence of its validity and reliability, see Mena (2021).

[10] A key objective of the EUNACOM designers is to make the difficulty of the exam comparable across years. Empirically, we observe that although the scores are not standardized, the distribution of scores at local universities is very similar across years.

[11] Argentina, Brazil, Colombia, Ecuador, Spain, the United Kingdom, and Uruguay

[12] Countries that allow all prospective physicians hoping to work in a given jurisdiction to take a licensing exam include Canada, Hong Kong, Japan, Korea, the United Arab Emirates, and the United States (Archer et al., 2017).

[13] This is similar to the pay scales for physicians in the NHS, which are determined nationally through negotiations between government bodies and medical unions.

2.2 Data

Data Sources Our main analysis combines four data sources: a matched employer-employee data for all public hospitals, the National registry of public health care provides, data on licensing scores, and individual-level discharge data. We combine these data with data on waitlists in some additional empirical exercises.

Employer-employee Data for Public Hospitals: We use matched employer-employee administrative records managed by the Ministry of Health. The records cover the universe of physicians employed in all public hospitals in Chile between 2011 and 2019 (SIRH, 2019). These data consolidate information from various public health organizations into a unified registry, providing comprehensive records for payroll processing and workforce management, including detailed wage data, and employment information. We complement these data with hospital-level characteristics, including size, location, and referral area, among others (DEIS, 2020).

National Registry of Public Healthcare Providers: We use a registry of all physicians (and other healthcare professionals) legally authorized to practice in Chile (RNPI, 2024). The registry is managed by Chile's Superintendency of Health (*Superintendencia de Salud*), and provides detailed information on registered healthcare professionals, including their nationality and details of their professional degree including title, date of issuance, and the name and country of the granting institution. For degrees obtained abroad, the registry includes the date of revalidation in Chile.

Licensing Exam Scores: We use confidential score records for all physicians who have taken the national licensing exam, EUNACOM. These data were provided directly by ASOFAMECH (*Asociación de Facultades de Medicina de Chile*), the organization responsible for administering and overseeing the exam (ASOFAMECH, 2019). The dataset includes the date and scores of the theoretical portion of the EUNACOM for every attempt made by the universe of physicians who have taken the exam.

Individual-level Discharge Data: We measure outcomes using administrative records of individual-level inpatient events in all public hospitals in Chile from 2011 to 2019 (DEIS, 2019). The data include diagnoses (ICD-10 codes), discharge or death dates, and patient characteristics such as birth date, gender, residence, and health insurance type. We link these data at the individual level with the universe of death records processed by the Vital Records Office, which includes deaths outside public hospitals.[14]

Waiting Lists: We access individual-level administrative records of waiting lists for surgical proce-

[14]We only have access for these records until 2018.

dures and specialist consultations for non-prioritized conditions (SIGTE, 2019), obtained through a Freedom of Information Act (FOIA) request.[15] In Chile, primary care physicians refer patients to hospitals for specialist consultations or surgical procedures, placing them on a waitlist. We observe each patient's entry and exit dates, along with the assigned hospital.

Data Aggregation and Outcomes We construct a hospital-by-year panel dataset using all the sources of information described above. We assess hospital performance in terms of access and quality. Our primary measure of access is the yearly hospital service rate, defined as the number of admissions in a given year divided by the eligible population in the hospital's referral region in the corresponding year. We also use surgery rates (inpatient surgeries as a proportion of admissions) and exits from the waitlist as additional access measures in some robustness exercises.

Our primary measure of quality is the hospital's yearly average death rate, defined as the ratio between deaths and admissions in a given year. We complement this in-hospital death rate with the death rate resulting from counting deaths within 28 days of a patient's admission—regardless of the place of death—following Gaynor et al. (2013). In robustness exercises, we also replace death rates with complication rates. We calculate complication rates as the number of patients discharged and later readmitted for inpatient care (within 3 months) due to an ICD-10 code related to infections, hemorrhage, or other complications, divided by the total number of admissions. Panel A in Appendix Table A.1 summarizes key patient and hospital characteristics.

Since EUNACOM was introduced in 2009, we impute scores for physicians who did not take the exam.[16] Our imputation procedure assigns scores to a physician from region or origin r working at hospital h by combining the average score of all physicians at hospital h (the grand mean) and the differential score of physicians from region r working at hospital h (the group mean). To estimate this differential score, we model licensing scores as a function of region-of-origin fixed effects at the hospital level, imposing sum-to-zero constraints on the fixed effects. We then adjust these estimates using empirical Bayes, such that group differentials are shrunk towards zero as estimates' precision decreases (Efron and Morris, 1973; Walters, 2024). In Appendix A, we present additional details on our imputation procedure and conduct robustness checks of the main results using LASSO as an alternative imputation method (Murdoch et al., 2019). Panel B in Appendix Table A.1 describes

[15]Waitlists are divided into prioritized and non-prioritized conditions. The "Explicit Healthcare Guarantees" program, established by law, prioritizes specific diseases with evidence-based procedures and timelines for diagnosis and treatment. See Menares and Muñoz (2025) for details.

[16]Before the introduction of EUNACOM, Chile had a voluntary National Medical Examination (EMN) administered in Chilean medical schools from 2003 to 2008. Prior to the EMN, licensing requirements varied: local medical graduates were required to pass the Medical Surgeon Degree Examination, while foreign-trained physicians had to complete a Foreign Medical Qualification Revalidation Examination.

the evolution of test scores across multiple years, and we discuss them in further detail in the next section.

3 Descriptive Evidence

In line with similar trends in other OECD countries, the number of physicians per capita in Chile has increased in the last two decades, driven by a growing number of domestic graduates and greater reliance on foreign-trained physicians (OECD, 2019). Panel A in Figure 1 shows that the number of new physicians enrolled in the National Registry of Healthcare Providers rose from around 1,000 in 2001 to over 4,000 in 2019.

One-third of the increase in newly registered physicians is attributed to the growing number of locally trained physicians, driven by an expansion in the local supply of medical schools. The remaining two-thirds of the increase is explained by an unprecedented inflow of foreign-trained physicians into the system. The share of the stock of physicians Chile who are foreign-trained increased from 15% to 24% between 2014 and 2019, and is currently similar to that of Canada and the US, and higher than the OECD average (OECD, 2015, 2019).[17][18] Interestingly, we do not find evidence of an increase in the share of foreign healthcare workers beyond physicians.[19]

The increasing number of physicians allowed to work in the country was also reflected in the number of physicians taking the licensing exam and their scores. In 2013, relatively few foreign-trained physicians took and passed the exam (i.e., had a score ≥ 51), as shown in Panel B of Figure 1. By 2018, the number of foreign-trained physicians taking the licensing exam had increased nearly tenfold, and they outnumbered local graduates by 2.5 to 1, as illustrated in Panel C.

Importantly, not only did the number of foreign-trained physicians increased, but the whole distribution of scores shifted to the right, and by 2018 the distribution of foreign-trained physicians was centered around the licensing threshold. This indicates that many successfully validated their

[17] While immigrants represented 2% of the population in 2011, this figure rose to 8% by 2022 (INE, 2024), largely driven by the mass migration of Venezuelans to Chile and other countries in the region, which began in late 2015 following the collapse of the Venezuelan economy. The Venezuelan migration to Chile and other countries in the region has been leveraged as a shock to explore the impacts of migration on various outcomes, including crime (Ajzenman et al., 2023), labor markets (Lebow, 2022; Olivieri et al., 2022; Bahar et al., 2024), and discrimination (Groeger et al., 2024), among others.

[18] Differences in remuneration and non-wage amenities are well-documented "push" and "pull" factors in physician migration (OECD, 2019). Various pieces of evidence highlight Chile as an attractive destination for medical professionals in Latin America, driven by significant wage differentials and better working conditions. Physicians in Chile can earn up to eight times more than their counterparts in Argentina, fueling temporary migration among Argentine physicians (Castro, 2024). Similarly, professional insecurity, low salaries, and limited social recognition have prompted physicians to emigrate from Venezuela (Hernández and Ortiz Gómez, 2011).

[19] We discuss this in greater detail in Appendix B.

medical degrees. At the same time, it also indicates that many were unable to do so, and that the licensing threshold was binding constraint for a significant share of the new physicians available to work in Chile.[20]

In Figure 2 we shift our focus to the aggregate effects of the increase in physicians in public hospitals inputs and outcomes. Panel A shows the number of full-time equivalent physicians working in public hospitals. By 2019, the influx of newly registered physicians had led to a 60% increase in full-time physicians in public hospitals compared to 2011.[21]

In terms of licensing scores, the aggregate increase in the number of test takers who are closer to the threshold, suggests that the new physicians hired in public hospitals likely decreased the average scores of the hospitals they were hired in. In Panel B of Figure 2, we plot the distribution of the average hospital scores across all public hospitals in the country. We observe that the whole distribution shifted to the left. On average, the average scores of physicians in public hospitals decreased by 3 points between 2013 and 2018.

To examine whether the increase in the number of physicians with a score close to the threshold working in public hospitals had an effect on hospital quality, in Panel C of Figure 2 we plot the aggregate death rates across public hospitals over time. At face value, we do not observe a visible change in aggregate trends. The fact that aggregate hospital death rates did not change suggests that there could be a quantity-quality trade-off at play. While the increase in the number of physicians could have, in principle, decreased hospital mortality, this effect may have been outweighed by its impact on the average quality of the hospitals they work in. This is exactly the trade-off that physician licensing faces and which we study in the section below.

4 A Simple Licensing Problem

We begin with a simple theoretical framework to study the economic fundamentals that determine the optimal licensing policy in our context. We focus on a particular margin, namely the stringency of the requirement as determined by the passing score of the licensing exam.

The licensing technology is based on a noisy signal s of quality θ. The output of interest (e.g.

[20]Panel B of Appendix Table A.1, shows that the proportion of all test takers achieving a passing score declined from 92% in 2009 to 58% in 2018. Meanwhile, the number of individuals scoring within the [40–51] range increased substantially, from 87 in 2009 to 1,552 in 2018, reflecting the growing mass of physicians who would qualify under a slightly lower passing score.

[21]Despite this significant growth in the workforce, hourly wages remained flat after accounting for sector-wide remuneration adjustments. Appendix Figure A.1 presents the average hourly wages for physicians in public hospitals from 2011 to 2019. This wage rigidity aligns with established wage-setting policies and is an important feature of our setting.

hospital production) is an increasing function of labor $L(\underline{s})$ and a quality index $\mathring{\theta}(\underline{s})$,

$$Y(\underline{s}) = F\left(L(\underline{s}), \mathring{\theta}(\underline{s})\right),$$

where $\underline{s}$ is the licensing threshold (passing score), $F_L > 0$ and $F_{\mathring{\theta}} > 0$. This production function generates a quality-quantity trade-off in the licensing policy, as long as there is positive mass in the support of s and the quality input is positively related to the cutoff; $\mathring{\theta}(\underline{s})' > 0$. Specifically, the elasticity of the outcome with respect to the licensing threshold, $\eta_{\underline{s}}^{Y} \equiv \frac{\partial Y}{\partial \underline{s}} \frac{\underline{s}}{Y}$ equals

$$\eta_{\underline{s}}^{Y} = \underbrace{\eta_L^Y \eta_{\underline{s}}^L}_{\text{Licensing Quantity Effect}} + \underbrace{\eta_{\mathring{\theta}}^Y \eta_{\underline{s}}^{\mathring{\theta}}}_{\text{Licensing Quality Effect}}, \tag{1}$$

where η_L^Y and $\eta_{\mathring{\theta}}^Y$ are the output elasticities with respect to the quantity and quality inputs, respectively, and $\eta_{\underline{s}}^L$ and $\eta_{\underline{s}}^{\mathring{\theta}}$ denote the elasticities of each input with respect to the licensing threshold. The impact of the policy variable $\underline{s}$ on the outcome $Y(\underline{s})$ is fully characterized by these four *sufficient statistics*.[22] On the one hand, increasing the threshold decreases output by reducing the quantity of labor. This *Licensing Quantity Effect* depends on the output elasticity with respect to quantity, and the elasticity of labor with respect to the threshold. On the other hand, increasing the threshold increases output by improving quality. This *Licensing Quality Effect* depends on the output elasticity with respect to quality and the elasticity of quality with respect to the threshold.

Moreover, the ratios of input and output elasticities are sufficient to determine the sign of the effect of changing the licensing threshold on the outcome as

$$\eta_{\underline{s}}^{Y} > 0 \iff \eta_L^Y / \eta_{\mathring{\theta}}^Y < -\eta_{\underline{s}}^{\mathring{\theta}} / \eta_{\underline{s}}^L. \tag{2}$$

Equation (2) states, simply, that it is beneficial to increase the licensing threshold if and only if the ratio of output elasticities between quality and quality is smaller than the associated gains from improving quality.

Parameterizing the model allows us to gain further insights into the microfoundations for the elasticities governing the quality-quantity trade-off in licensing. Let the signal s have a distribution

[22] Chetty (2009) provides a general framework for welfare analysis from a set of sufficient statistics that are derived using envelope conditions under the assumption that agents make optimal decisions. However, Chetty (2009) also notes that the assumption of optimizing behavior is not necessary as long as the researchers can estimate the terms included in the derivative of welfare with respect to the policy variable of interest. That is indeed our case. For the interested reader, in Appendix C we recast the licensing problem as a maximization problem and derive Equation (1) in that framework using envelope conditions.

with density $h(s)$ and total mass m. Physicians who pass the exam can either work at the hospital or go to an outside option. We let $p(s|\underline{s})$ the share of physicians with score s who match with the hospital; which may depend on the licensing score through equilibrium effects in the labor market. The total labor function is therefore $L(\underline{s}) = m \int_{\underline{s}}^{\infty} h(s)p(s|\underline{s})ds$. Let the quality index be equal to the exponential of average quality, $\mathring{\theta} = \exp(\bar{\theta}(\underline{s}))$, with $\bar{\theta}(\underline{s}) = 1/L(\underline{s}) \int_{\underline{s}}^{\infty} \theta(s)h(s)p(s|\underline{s})ds$. Finally, the production function has a Cobb-Douglas form with $F\left(L, \bar{\theta}\right) = L^{\alpha_L} \cdot \exp(\bar{\theta})^{\alpha_{\bar{\theta}}}$.

Under this parametrization, the elasticity of output with respect to the licensing threshold $\underline{s}$ can be written as

$$\eta_{\underline{s}}^{Y} = \alpha_L \cdot \eta_{\underline{s}}^{L} + \alpha_{\bar{\theta}} \cdot \tilde{\eta}_{\underline{s}}^{\bar{\theta}},$$

where $\eta_{\underline{s}}^{L}$ is the elasticity of labor with respect to the licensing threshold and $\tilde{\eta}_{\underline{s}}^{\bar{\theta}}$ is the semi-elasticity of quality with respect to the threshold.

In the absence of equilibrium effects in the labor market, i.e. when $\partial p(s|\underline{s})/\partial \underline{s} = 0$, the elasticities can be expressed as simple functions of the underlying model parameters. The elasticity of labor with respect to the threshold is

$$\eta_{\underline{s}}^{L} = \frac{-m \cdot h(\underline{s}) \cdot p(\underline{s}|\underline{s}) \cdot \underline{s}}{L}, \tag{3}$$

which depends on the mass of workers at the threshold and the fraction of them matching with the hospital. In turn, each marginal worker lowers average *scores* by an amount equal to $\mathbb{E}[s|s > \underline{s}]$. The degree to which average *quality* is affected depends on the precision of the licensing score as a signal of quality. Let $s = \theta + \epsilon$, with $\epsilon \sim N(0, \sigma_{\epsilon}^2)$ and $\theta \sim N(\mu_{\theta}, \sigma_{\theta}^2)$. Defining the signal-to-noise ratio as $\text{SNR} \equiv \frac{\sigma_{\theta}^2}{\sigma_{\theta}^2 + \sigma_{\epsilon}^2}$, the semi-elasticity of quality with respect to the licensing threshold can be written as

$$\tilde{\eta}_{\underline{s}}^{\bar{\theta}} = -\eta_{\underline{s}}^{L} \cdot \text{SNR} \cdot \left(\mathbb{E}[s|s > \underline{s}] - \underline{s}\right). \tag{4}$$

On net, the elasticity of outcome with respect to the threshold is therefore

$$\eta_{\underline{s}}^{Y} = \frac{-m \cdot h(\underline{s}) \cdot p(s|\underline{s}) \cdot \underline{s}}{L} \cdot \left(\alpha_L - \alpha_{\bar{\theta}} \cdot \text{SNR} \cdot \left(\mathbb{E}[s|s > \underline{s}] - \underline{s}\right)\right) \tag{5}$$

The first term measures the number of marginal workers, whereas the second term is the net effect of a marginal worker on production. The quantity effect of an extra worker is α_L. However, the marginal worker decreases average quality by $\text{SNR} \cdot (\mathbb{E}[s|s > \underline{s}] - \underline{s})$. This term, multiplied by the return on quality, gives the quality effect of the marginal worker.

Under this parametrization, Equation (2) can be stated as

$$\frac{\alpha_L}{\alpha_{\bar{\theta}}} > \text{SNR} \cdot \left(\mathbb{E}[s|s > \underline{s}] - \underline{s}\right).$$

Whether decreasing the licensing threshold increases output depends on whether the output elasticity with respect to quantity over the output elasticity with respect to quality is higher than the effect of average quality of a marginal worker. The later increases with the distance between the average score and marginal score, and the precision of the signal.

The optimal licensing threshold $\underline{s}^*$ can be implicitly characterized with a simple expression:

$$\underline{s}^* = \mathbb{E}[s|s > \underline{s}^*] - \frac{\alpha_L}{\alpha_{\bar{\theta}}} \cdot \frac{1}{\text{SNR}}.$$

The optimal threshold increases with the mean score. A higher mean allows the planner to have more quantity for any given level of quality. Also, the optimal threshold decreases with the output elasticity with respect to labor and increases with the output semielasticity with respect to quality. It is also lower when the licensing technology is imprecise (i.e., low signal-to-noise ratio), as in such cases a higher threshold can only weakly differentiate between high- and low-quality physicians.

Objective Function Let patient's health in the absence of treatment (e.g., probability of survival if not treated) be equal to $m(X)$ such that all heterogeneity across patients is captured by their characteristics X; and let $\Delta m(X, \bar{\theta}(\underline{s}), L(\underline{s}))$ be the value added of health care treatment, e.g., the reduction in mortality risk if treated under a healthcare system characterized by inputs $\bar{\theta}(\underline{s})$ and $L(\underline{s})$. Denoting the unconditional distribution of X by dG_X, the set of patients who receive treatment under the licensing cutoff $\underline{s}$ by $\mathcal{I}(\underline{s})$, and the conditional distribution of their characteristics by $dG'_{X|\underline{s}}$, we can write population health as:

$$\begin{aligned} H(\underline{s}) &= \underbrace{\int_X m(X) dG_X}_{H^0} + \underbrace{|\mathcal{I}(\underline{s})|}_{\text{service rate}} \times \underbrace{\int_X \Delta m(X, \bar{\theta}(\underline{s}), L(\underline{s})) dG'_{X|\underline{s}}}_{\text{per-patient treatment value added}} \\ &= H^0 + Y(\underline{s}), \end{aligned}$$

where H^0 represents baseline health and $Y(\underline{s})$ captures the impact of the healthcare system in improving health outcomes, which depends on the licensing threshold through its impact on the number of patients seen (service rate), and on its impact on the per-patient value added, as well as the distribution of characteristics of treated patients. In our empirical analysis, we will focus on lives saved as the main health outcome. Thus, we will proxy per-patient treatment value added with the *decrease* in patients' mortality as a function of physicians' quantity and quality while also controlling for patients' case mix to account for heterogeneous treatment effects due to variations in patients' characteristics. Δm thus represents risk-adjusted decreases in mortality risk, which is a common measure of quality used in the literature evaluating hospital quality (Chandra et al., 2016; Arnold, 2025).

Given this objective function, we can approximate the elasticity of healthcare's impact with respect to the licensing threshold as the difference between the elasticity of the service rate and the elasticity of per-patient mortality risk, as follows:[23]

$$\eta_{\underline{s}}^{Y} \simeq \eta_{\underline{s}}^{\text{service rate}} - \eta_{\underline{s}}^{\text{mortality}}. \tag{6}$$

Thus, the licensing threshold may affect lives saved in two ways. First, it may affect access through the service rate. Second, it may affect treatment value added by changing patients' mortality.

5 Input Elasticities

To estimate hospital-specific input elasticities, we augment the model of the previous section by including many hospitals and the corresponding endogenous sorting of physicians across them. We consider a social planner who, in period $t_0 = 0$, determines the licensing threshold $\underline{s}$ considering its impact on outcome Y in a period $R \geq 0$, Y_R:

$$Y_R \equiv \prod_{j \in \mathcal{J}} y_{jR}^{p_j},$$

where p_j are hospital Pareto weights. Restating Equation (1) in the context of the functional form assumptions for the production function of Section 4, the elasticity of Y_R with respect to the licensing threshold is given by:

$$\eta_{\underline{s}}^{Y_R} = \underbrace{\alpha_L \cdot \bar{\eta}_{\underline{s}}^{L,R}}_{\text{Licensing Quantity Effect}} + \underbrace{\alpha_\theta \cdot \bar{\eta}_{\underline{s}}^{\bar{\theta},R}}_{\text{Licensing Quality Effect}},$$

where $\bar{\eta}_{\underline{s}}^{L,R} = \sum_{j \in \mathcal{J}} p_j \cdot \eta_{\underline{s}}^{L_j,R}$, is the average elasticity of labor with respect to $\underline{s}$ in period R, and

[23] Formally, $\eta_{\underline{s}}^{Y} = \eta_{\underline{s}}^{\text{service rate}} - \eta_{\underline{s}}^{\text{mortality}} - \eta_{\underline{s}}^{\text{service rate}} \eta_{\underline{s}}^{\text{mortality}}$. Our approximation ignores the cross-term. Empirically, we find that $\eta_{\underline{s}}^{\text{mortality}} \simeq 0$.

$\bar{\eta}_{\underline{s}}^{\bar{\theta},R} = \sum_{j \in \mathcal{J}} p_j \cdot \eta_{\underline{s}}^{\bar{\theta}_j,R}$, is the average semi-elasticity of quality with respect to $\underline{s}$ in period R.

5.1 Scores and Labor Matching Model

As shown in section 5, the elasticity of inputs depends on the mass of test-takers at the threshold, the mapping between scores and quality, and the matching of physicians and hospitals. We recover these objects with a dynamic model of scores and a labor-matching model.

Dynamic Model of Scores Modeling the dynamics of scores serves two purposes. First, it allows us to infer physicians' latent quality by leveraging the entire history of their scores across attempts. Second, it will allow us to predict the dynamic effects on the mass of test-takers when evaluating the long-run effects of changing the threshold.

Each physician i is characterized by her latent quality θ_i. The vector of scores is determined by the realization of scores in each attempt as well as retaking decisions after failing. The score at the n^{th} attempt is a noisy measure of latent quality and test-taking ability, Γ_{ni}, which can change over attempts.[24] We let the mean and variance of quality, as well as the variance of the noise to depend on physician's origin, $\tau_i \in \{\text{National,Foreigner}\}$, to allow for the possibility that the precision of the test differ across these groups:[25]

$$
\begin{aligned}
s_{in} = &\quad \theta_i + \Gamma_{in} + \epsilon_{in} \\
\theta_i \sim &\quad N(\mu_{\tau_i}, \sigma^2_{\theta,\tau_i}) \\
\epsilon_i \sim &\quad N(0, \sigma^2_{\epsilon,\tau_i})
\end{aligned}
$$

In the data, average score gains over attempts are positive, decreasing, and convex (see Figure A.2). We assume, therefore, that test-taking ability improves with exponential decay, such that for any attempt number $n \geq 1$, test-taking ability is

$$\Gamma_{in} = \sum_{k=0}^{n-1} \gamma \cdot \exp(-\rho \cdot k), \tag{7}$$

[24] We assume that quality is constant over time which implies that physicians cannot experience quality gains over attempts. To support this assumption, in Appendix D, we use regression discontinuity designs to show that physicians who retake the exam around the passing threshold do not experience gains on quality proxies.

[25] There are several reasons to believe that the test's precision may be depend on the physician's region of training. Nationals have been trained in a system that often prepares them to take the test; which could reduce the precision of the signal. On the other hand, foreigners normally take the test several years after graduating. Also, although most foreigners speak Spanish as their mother tongue, not all of them do, and even for Spanish-speaking foreigners there could be difficulties due variations of the language in the region.

where γ and ρ are parameters to be estimated. The parameter γ governs the average improvement in the first retake, while ρ governs the average rate at which improvements decrease over subsequent attempts.[26]

Physicians who fail the exam in attempt n, retake it with a probability that is a function of the distance between their average past score and the passing threshold $\bar{s}_{in} - \underline{s}$, the number of attempts, n, physician's type, τ_i, and the licensing threshold, $\underline{s}$:

$$P(\text{retake}|\bar{s}_{in}, n_i, \tau_i) = \frac{e^{\beta_{0,\tau_i} + \beta_{n,\tau_i} n_i + \beta_{s,\tau_i}(\bar{s}_{in} - \underline{s})}}{1 + e^{\beta_{0,\tau_i} + \beta_{n,\tau_i} n_i + \beta_{s,\tau_i}(\bar{s}_{in} - \underline{s})}}. \tag{8}$$

By allowing the retaking probability to depend on both the distance between an individual's average score and the licensing exam threshold, and the number of attempts (a sufficient statistic for predicting future scores and posterior quality), this specification captures the net benefits of retaking. These net benefits arise from two factors: (i) the probability of passing the threshold in subsequent attempts (determined by the gap between expected future scores and the threshold), and (ii) the expected value of the match conditional on passing the exam (which depends on physicians' type and posterior quality). In addition, explicitly incorporating the licensing threshold in the specification allows retaking behavior to change under counterfactual scenarios.

The posterior of quality for each physician given a sequence of scores over attempts $\{s_{i1}, s_{i2}, ..., s_{in}\}$ is equal to:

$$\mathbb{E}[\theta_i \mid s_{i0}, s_{i1}, \ldots, s_{in}] = \mu_{\theta,\tau_i} + \frac{\sigma^2_{\theta,\tau_i}}{\sigma^2_{\varepsilon,\tau_i} + (n+1)\sigma^2_{\theta,\tau_i}} \left(\sum_{t=0}^{n} (s_{it} \underbrace{-\Gamma_{t,\tau_i}}_{\text{de-trending}} - \mu_{\theta,\tau_i}) \right), \tag{9}$$

Equations (8) and (9) are the basis for inferring physician's quality from their vector of scores.

Labor Market Matching Physicians who pass the exam match with hospitals in a decentralized labor market. There are two main challenges in modeling this matching process. First, for a counterfactual that lowers the licensing threshold, we want to predict matches outside the support of the score distribution for which we observe matches in the data. Second, the model should capture potential changes in the matching process in scenarios with different licensing thresholds. In particular, relaxing the licensing threshold increases the number of (lower-quality) physicians

[26] In principle, γ and ρ could be type-specific. However, as most nationals pass the test in the first attempt, we do not have enough identifying power for estimating parameters for nationals and foreigners separately. This limitation is largely inconsequential for our counterfactual analysis as even fewer nationals would retake when we lower the threshold.

seeking jobs, which may impact the equilibrium matching probabilities due to competition in the labor market.

We address these challenges by positing a flexible reduced-form matching function that includes the key variables suggested by economic theory (Chetty, 2009). We favor this approach over estimating a structural model of the labor market for two main reasons: First, as described by Equation (2), the key sufficient statistic for signing the derivative of the outcome with respect to the threshold is the ratio between the quantity and quality elasticity; which under our parameterization is independent of matching probabilities and only depends on the difference between the quality of supramarginal physicians and the quality of marginal individuals. Second, our counterfactual exercise is a *local* reform (i.e. a small change in the licensing threshold), and is therefore less sensitive to the underlying approximations of our statistical approach.

Denote by $M_t(\underline{s})$ the mass of physicians who pass the licensing exam in period t when the threshold is $\underline{s}$, and let $\kappa_t = \{\kappa_{1t}, ..., \kappa_{|\mathcal{J}|t}\}$ be the vector of posted vacancies in period t. Consider a physician of posterior quality $\hat{\theta}_i$ within a range $r(\hat{\theta}_i)$ seeking a job in period t. We specify their matching probability with hospital j to depend on the number of other job seekers within their quality range or above, $M_{it}(\underline{s})$, relative to the posted vacancies κ_t.[27] We also specify matching probabilities to depend flexibly on observables to capture physicians' and hospitals' preferences. Specifically, conditional on passing the licensing exam in period t, each physician i matches with a hospital j or opts for their outside option, $\emptyset$, based on a conditional matching probability function denoted by $CMP(\cdot)$ that we model as:

$$CMP_{ijt} \equiv CMP\Big([x_{i1t}, \ldots, x_{i|\mathcal{J}|t}],\ s_i,\ \tau_i,\ M_{it},\ \kappa_t,\ \underline{s},\ j\Big) = \frac{e^{v\left(x_{ijt}, \hat{\theta}(s_i,\tau_i)\right) + g(M_{it}(\underline{s}), \kappa_t | \tau_i, j)}}{1 + \sum_{j'} e^{v\left(x_{ij't}, \hat{\theta}(s_i,\tau_i)\right) + g(M_{it}(\underline{s}), \kappa_t | \tau_i, j')}}, \tag{10}$$

The function $v(\cdot)$ captures the determinants of matching probabilities based on physicians' and hospitals' preferences. In turn, the function $g(\cdot)$ captures the potential equilibrium effects driven by competition in the labor market; the key channel through which matching probabilities depend on the licensing threshold. Although our approach to defining this matching function is reduced form in nature, in Appendix H we propose a theoretical model that can rationalize it. In the model therein, hospitals with fixed capacities have vertical preferences for physicians, and wages are rigid. The equilibrium allocation is pairwise stable and is characterized by a set of hospital-specific quality cutoffs (Azevedo and Leshno, 2016). In equilibrium, no pair physician-hospital would like

[27] Our approach is flexible enough to account for displacement effects, which are a natural concern in this setting (Crépon et al., 2013). A change in the mass of physicians in a specific quality range can have a direct effect on the CMPs within that range and can also have indirect effects on the matching probabilities of physicians of lower quality.

to deviate from their current assignment to match together. In the model, equilibrium quality cutoffs are jointly determined by the mass of physicians, their joint distribution of preferences and posterior qualities, and hospitals' vacancies. However, we show that under vertical preferences, physicians' choice sets and their matching probabilities can only be affected by changes in the mass of physicians ranked above in their posterior qualities and not below them. Two key empirical facts motivate this model. First, wages in our setting follow the wage schedule of public workers, and have remained stagnant in spite of the large migration wave (see Figure 1). Second, sorting patterns of physicians across hospitals are in line with the predictions of a model with vertical preferences, i.e., during our period of study, the variance of physician's quality *between* hospitals accounts for 70% of the total variance in physicians' quality.

Hospital-specific Elasticities: The labor-matching probabilities and the licensing policy determine the hospital-specific inflows of physicians and their respective quality. For a physician $i \in \mathcal{I}$ with type $\tau \in \mathcal{T}$, exam score s, and a matrix of physician–hospital characteristics $X \in \mathbb{R}^{K\times|\mathcal{J}|}$ (whose jth column is the vector $\boldsymbol{x}_{ijt} \in \mathbb{R}^K$ for hospital $j \in \mathcal{J}$), in a market characterized by (M_t, κ_t) in period t, we define the conditional probability that such a physician matches with hospital $j \in \mathcal{J}$ in period t as:

$$CMP_j\big(X, s \,\big|\, \tau;\, M_t;\, \kappa_t;\, \underline{s}\big) \;\equiv\; CMP\Big([\,x_{i1t}, \ldots, x_{i|\mathcal{J}|t}\,] = X,\; s_i = s,\; \tau_i = \tau,\; M_{it} = M_t,\; \kappa_t,\; \underline{s},\; j\Big).$$

Letting $h_t^\tau(s, X)$ be the type- and time-specific density of the joint distribution of observables and scores, the labor inflow of type τ in hospital j in period t is given by:

$$\Delta L_{jt}^\tau(\underline{s}) \;=\; m_{\tau,t} \int_X \int_{s \geq \underline{s}} CMP_j\big(X, s \mid \tau;\, M_t;\, \kappa_t;\, \underline{s}\big)\, h_t^\tau(s, X)\, ds\, dX.$$

In addition, the average quality of physicians inflow of type τ in hospital j at time t is given by:

$$\Delta\bar{\theta}_{jt}^\tau(\underline{s}) \;=\; \frac{\displaystyle\int_X \int_{s \geq \underline{s}} CMP_j\big(X, s \mid \tau;\, M_t;\, \kappa_t;\, \underline{s}\big)\, \hat{\theta}(s, \tau)\, h_t^\tau(s, X)\, ds\, dX}{\displaystyle\int_X \int_{s \geq \underline{s}} CMP_j\big(X, s \mid \tau;\, M_t;\, \kappa_t;\, \underline{s}\big)\, h_t^\tau(s, X)\, ds\, dX}$$

The labor in hospital j at time t and its corresponding average quality are given by

$$L_{jt}(\underline{s}) = L_{j,t-1} + \sum_{\tau \in \mathcal{T}} \Delta L_{jt}^{\tau}(\underline{s}) \tag{11}$$

$$\bar{\theta}_{jt}(\underline{s}) = \frac{1}{L_{jt}} \left(\bar{\theta}_{j,t-1} L_{j,t-1} + \sum_{\tau \in \mathcal{T}} \Delta L_{jt}^{\tau}(\underline{s}) \cdot \Delta \bar{\theta}_{jt}^{\tau}(\underline{s}) \right) \tag{12}$$

The elasticity of quantity and semi-elasticity of quality with respect to the licensing threshold follow directly from the expressions above.

5.2 Estimation and Results

Latent quality, matching probabilities, and the mass of test-takers around the licensing threshold are the key ingredients to estimate the input elasticities with respect to the licensing threshold.

Latent Quality Table 1, Panel A, presents the maximum likelihood estimation results for the retaking model specified in Equation (8). As expected, the probability of retaking decreases with the number of attempts and with the distance of the score to the cutoff. Using the estimated retaking probabilities, we estimate the model of score gains via Simulated Method of Moments (SMM).[28] We match the following moments by physician type (national or foreign): the mean over attempts, the mean of gains over attempts, the covariance between attempts, and the variance of the first attempt. Panel B of Table 1 shows the estimated coefficients. We observe that foreigners have larger variances for the estimated quality distribution and exam noise than nationals. The estimated first gain $\hat{\gamma}$ is 9.6, but score gains decay at an exponential rate of $\hat{\rho} = 0.285$, i.e. decrease by approximately 25% between two consecutive gains.

With the estimated quality distribution and score gain parameters, and given the history of score realizations for each physician, we use Equation (9) to estimate the individual posterior quality means.[29]

Matching Probabilities We estimate the conditional matching probabilities by maximum likelihood following Equation (10). We specify $v(x_{ijt})$ as a function of the distance between the university of training and the centroid of location j (Distance$_{ij}$), and the percentage of physicians from that

[28] This estimation strategy allows us to model the selection behavior in the retaking process.

[29] Alternatively, we could use only data of initial scores, and leverage our instruments to correct for the associated measurement error, without the need to specify the model of scores (Agostinelli and Wiswall, 2016). However, specifying the model of scores allows us to perform the long-run counterfactuals, and thus we also use it for the estimation of the returns to quality for the sake of parsimony.

region who worked at hospital j in $t-1$, Share_{ijt-1}.[30] We also include alternative-specific time fixed effects, alternative-specific shifters of the matching probability based on physician's posterior quality ($\mathbb{E}(\theta_i)$), whether the physician is foreign (Foreign_i) and whether the physician is a specialist (Specialist_i):

$$\begin{aligned} v_{ijt} = \alpha^d \text{Distance}_{ij} + \alpha^h \text{Share}_{ijt-1} + \alpha_{jt} + \alpha_j^f \text{Foreign}_i \\ + \alpha_j^q \mathbb{E}(\theta_i) + \alpha_j^{fq} \mathbb{E}(\theta_i) \times \text{Foreign}_i + \alpha_j^s \text{Specialist}_i. \end{aligned} \tag{13}$$

The coefficients α_{jt} represent alternative-specific year fixed effects, and α_j^f, α_j^q and α_j^{fq} and α_j^s allow those mean effects to vary with physicians' characteristics. We normalize $v_{ijt} = 0$ for the outside option.

As specified in Equation (10), the matching probabilities depend on the mass of physicians who approve the licensing exam across different quality ranges and the vacancies in each location. For our empirical application, we specify the general equilibrium effects as a simple linear function of the ratio between job seekers and vacancies (i.e. the inverse of the "labor market tightness"):

$$g(M_t(\underline{s}), \kappa_t | \tau, r_{\hat{\theta}}(i), j) = \boldsymbol{\beta}_j \frac{M_{it}(\underline{s})}{\kappa_{jt}}, \tag{14}$$

where $\boldsymbol{\beta}_j$ is a vector of hospital-specific coefficients and $M_{it}(\underline{s}) = [M_{it}^0(\underline{s}), M_{it}^+]$ is a vector composed by the mass of physicians in i's quality range; $M_{it}^0(\underline{s})$, and the mass of physicians above i's quality range; M_{it}^+.[31] We construct quality ranges to construct the vector of masses M_{it} using four quality quantiles. We proxy the number of vacancies κ_{jt} with the ratio of beds and the stock of physicians in the previous period, that is, $\tilde{\kappa}_{jt} = \frac{\text{Beds}_{jt}}{\text{Stock of Physicians}_{j,t-1}}$.

The results are presented in Table 2, where columns (1)-(3) vary how we incorporate the labor market tightness on the matching probabilities. Column (1) shows a specification where we specify the function $g()$ as depending only on other physicians with the same quality rage (M_{it}^0). Column (2) allows the function $g()$ to also depend on the mass of physicians with higher quality range

[30] The inclusion of this variable is predicated upon the idea that past immigrant settlement/job patterns predict where new immigrants will locate/work (Altonji and Card, 1989). We also use this variable as part of the shift-share instrument used to identify output elasticities in the next section.

[31] Let m_t be the number of test takers in period t and $h_t(s, X)$ be the joint distribution of scores and characteristics in period t. The mass of physicians who pass the exam in year t and have expected quality within a range $r \in \mathcal{R}$ is:

$$M_{r,t}(\underline{s}) \equiv m_t \int_X \int_{s > \underline{s}:\ \hat{\theta}(s) \in r} h_t(s, X) ds dX. \tag{15}$$

The mass of physicians in i's quality range is $M_{it}^0 = M_{r(\hat{\theta}_i),t}$ and the mass of physicians above i's quality range, $M_{it}^+ = \sum_{r' \succ r(\hat{\theta}_i)} M_{r',t}$. Notice that only M_{it}^0 depends on the licensing score as the licensing score only affects the mass at the bottom of the distribution.

(M^{+}_{it}). Column (3) allows the effects of these masses to vary across three different hospital-quality tiers; which we construct based on the average quality of their physicians. As a placebo check, in column (4) we show that the mass of physicians in a lower quality range (which we denote M^{-}_{it}) does not affect the matching probabilities.

Correlates of the Input Elasticities Our estimated matching probabilities imply that the average (across years and hospital) quantity and quality elasticities are -0.072 and 1.4, respectively. To investigate patterns of heterogeneity in the elasticities, Table 3 present the results from estimating a series of OLS regressions where we project the labor elasticity (Columns 1-4) and the quality semi-elasticity (Columns 5-9) on the following hospital observables: i) the variable "High phys/pat", which indicates hospitals with above-the-median ratio of physicians per patients in the previous year, ii) the variable "High average score" indicating hospitals with above-the-median score, and North$_j$ indicating hospitals at a latitude below the median.[32] In all specifications we include year fixed effects. We find that the labor elasticity is higher (in absolute terms) in hospitals with a low ratio of physicians per patient and in hospitals with a lower average score. As expected, the quality semi-elasticity displays an opposite pattern.

6 Output Elasticities

In this section, we introduce the hospital production function for healthcare, present our instrumental variables approach to estimate the output elasticities, and discuss our estimation results.

6.1 Production Function Model

In period t, each hospital j produces an outcome k that we denote by y^k_{jt}, by combining a quantity of physicians L_{jt} with quality index $Q(F_{\theta_{jt}})$, and other determinants of hospital output such as capital, patient case-mix, and productivity, collectively represented by A_{jt}. We assume the production function is a Cobb-Douglas, such that:

$$y_{jt} = A_{jt} L_{jt}^{\alpha_L} exp(Q\left(F_{\theta_{jt}}\right))^{\alpha_\theta}, \tag{16}$$

where $Q(F_{\theta_{jt}}) = \int q(\theta)\, f_{\theta_{jt}}(\theta)\, d\theta$ and $q(\cdot)$ is a *nondecreasing* function of θ. Intuitively, $Q(F_{\theta_{jt}})$ captures how the distribution of quality at hospital j in period t. For our empirical application, we

[32]Latitudes in Chile are negative. Due to the particular geography of the country, latitude captures most of the location.

adopt $q(\theta) = \theta$, which implies $Q(\theta_{jt})$ is simply the (unconditional) mean of the quality distribution in hospital j in period t.[33]

The parameters of interest are α_L and α_θ which correspond to the output y-specific elasticities with respect to the quantity and quality of physicians, respectively.

6.2 Estimation and Results

Using the posterior quality $\theta_i = \mathbb{E}(\theta_i|s_i) + \nu_i$, we can express the empirical analog of Equation (16) as:

$$\ln(y_{jt}^k) = \alpha_L^k \ln(L_{jt}) + \alpha_\theta^k \frac{1}{L_{jt}} \sum_{i \in J_t} E(\theta_i|\mathbf{s}_i) + \gamma_{f(j)t}^k + \rho_j^k + \beta^k X_{jt} + \underbrace{\omega_{jt}^k + \alpha_\theta^k \frac{1}{L_{jt}} \sum_{i \in J_t} \nu_{it} + \varepsilon_{jt}^k}_{\mu_{it}}. \quad (17)$$

Equation (17) is our estimating equation to recover α_L^k and α_θ^k, the output-specific elasticities with respect to quantity and quality, respectively. This specification implicitly models logged productivity as being influenced by time-invariant unobservable characteristics, which we capture using hospital fixed effects, ρ_j^k. These fixed effects also capture other time-invariant unobservables, such as the capital stock, and we further account for time-varying changes in capital by including the number of beds per patient in each HRR in the vector X_{jt}. Additionally, following Propper and Van Reenen (2010) and Gaynor et al. (2013), X_{jt} includes demographic controls to account for changes in patient composition that might influence hospital outcomes. Specifically, these case-mix controls include the shares of female inpatients, foreign inpatients, and inpatients across eight age bands (0–29, followed by 10-year increments up to 90+), as well as the shares of inpatients with different insurance types categorized by co-payment levels.

To account for potential differential time trends by hospital complexity, the specification also includes a vector of year-fixed effects that vary with hospital complexity, $\gamma_{f(j)t}^k$. Finally, the error term μ_{it} consists of three components: an unobserved productivity shock potentially known before input choices, ω_{jt}^k (Olley and Pakes, 1996; Levinsohn and Petrin, 2003; Ackerberg et al., 2015); measurement error in quality, $\alpha_\theta^k \frac{1}{L_{jt}} \sum_{i \in J_t} \nu_{it}$; and an unobserved productivity shock that occurs after input decisions, ε_{jt}^k.

To address the problem of measurement error and the identification challenge that physicians'

[33] In Appendix E, we show that an alternative specification—defining the quality index as the share of physicians in hospital j during period t with quality below the median of the overall distribution—yields very similar empirical results. We also discuss the impact of physician quantity and quality on health outcomes using a Translog production function that includes a linear interaction between the two.

quantity and quality could be correlated with unobserved factors affecting hospital outcomes, we implement a two-stage least squares (2SLS) estimation strategy that uses two shift-share (or Bartik) instruments (Altonji and Card, 1989; Autor et al., 2013) to recover the *causal* impact of physician quantity and quality on hospital-level outcomes. Our instruments leverage the increase in the labor supply of physicians brought about by immigration flows from other countries and, to a lesser extent, by the expansion of medical schools in Chile. The instrument for physician quantity, Z^L_{jt}, is constructed by summing the percentage change in the number of physicians clearing the cutoff of the licensing exam from each region of training c (i.e., the shift component ΔS_c), weighted by the percentage of physicians from that region who worked at hospital j the year before (i.e., the share component $\text{Share}_{i(c)jt-1} \equiv \frac{L_{cjt-1}}{\sum_c L_{jct-1}}$). Similarly, the instrument for physician quality, Z^θ_{jt}, sums the change in the average quality of eligible test-takers from each training region c, weighted by the percentage of physicians from that region who worked at hospital j the year before.

Table 4 presents the results from our two-stage least squares (2SLS) estimation, which examines the causal impact of physicians' quantity and quality on different hospital-level outcomes of interest. We present point estimates alongside exposure-robust standard errors clustered at the region of origin (i.e., the shock) level in brackets (Adao et al., 2019; Borusyak et al., 2022),[34] and we complement our analysis by showing the Anderson and Rubin (1949) p-values (Lee et al., 2022).

Panel A focuses on access. Column (1) reports the results for the hospital service rate, which is our main measure of access. We find an elasticity of service rate with respect to the number of physicians of 1. Column (2)-(4) show the impacts on additional measures of access. Column (2) focuses on the number of inpatient surgeries, for the subset of hospitals performing surgeries. We find a surgery elasticity of five. We also find that physicians' quality does not affect the service rate or surgeries. Columns (3) and (4) analyze the impact on the completion of wait-listed surgical procedures and specialist consultations, respectively. Column (3) shows that increasing the number of physicians leads to more exits from the surgical waiting list, with no effect from changes in physician quality. Similarly, Column (4) indicates that more physicians increase the number of patients exiting the specialist waiting list, while physician quality remains inconsequential.

Overall, our findings suggest that the quantity of physicians significantly influences access, consistent with the notion that physician scarcity is the primary binding constraint in expanding hospital capacity and addressing patient demand (e.g., Carrillo and Feres, 2019). Our results also show that the quality of physicians does not impact a hospital's ability to admit more patients, suggesting that the availability of physicians, rather than their expertise, is the primary determinant of hospital

[34] To obtain these standard errors—which are shown to be asymptotically valid (Adao et al., 2019)—we estimate the transformed regression proposed in Borusyak et al. (2022).

utilization rates.

Panel B of Table 4 focuses on quality, where our main measure is the in-hospital death rate. Column (1) shows the results. We find that: i) increasing the number of physicians by 1% decreases the in-hospital death rate by 0.8%, and ii) a one-point increase in the average quality of physicians—which corresponds to a 0.23 standard deviations increase—decreases death rates by 0.04%.

To address the concern that a potential correlation between increases in the number of physicians and the admission of healthier patients might be driving our results (i.e., marginal, previously non-admitted patients might be healthier), Column (2) shows the impact on a hospital death rate that is predicted based upon patients' characteristics.[35] Neither physician quantity nor quality significantly affects predicted mortality rates, suggesting that changes in patient composition are not driving our findings.

Columns (3) and (4) present the results for additional quality metrics. In Column (3), we use the 28-day mortality rate as the outcome and find effects consistent with those reported in Column (1). This suggests that the impact of physicians' quantity and quality remains robust even when accounting for out-of-hospital deaths among discharged patients. Finally, Column (4) focuses on the in-hospital complications rate. We define complications rate as the number of patients discharged from the hospital but later readmitted for inpatient care (at any hospital within 3-months) due to an ICD-10 code related to infections, hemorrhage, or other complications, divided by the total number of admissions. Consistent with the impacts on mortality, we find that physicians' quantity and quality help to reduce the complications rate.[36]

Robustness Checks Identification of the returns to quantity and quality using shift-share instruments is predicated upon the assumption that either the "shifts" or the "shares" components are as good as random and not correlated with factors that would affect the outcomes of interest (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022). Identification "from the shifts" can be understood as leveraging a shift-level natural experiment, while identification "from the shares" can be viewed as pooling together multiple difference-in-differences designs leveraging heterogeneous shock

[35] We calculate the expected death rate by fitting a logit model for death outcomes at the inpatient level, controlling for patient demographics and diagnoses group, as per the enhanced Elixhauser comorbidity index (Elixhauser et al., 1998; Quan et al., 2005). The total number of *predicted deaths* at each hospital and year is then divided by the number of admissions.

[36] Several mechanisms could explain why quantity and quality affect health outcomes. On the one hand, more physicians can facilitate better patient monitoring and quicker responses to health complications, both of which can improve patient outcomes and reduce mortality. More physicians also mean that inpatients are less likely to experience delays in receiving necessary treatments, which is particularly critical in life-threatening situations. On the other hand, physicians of higher quality may possess greater skill in accurately diagnosing and treating complex cases, making timely decisions, and effectively applying evidence-based practices. This expertise can decrease complications and enhance patient survival by reducing misdiagnosis, treatment errors, and adverse health outcomes.

exposure (Borusyak et al., 2024). In Appendix F, we present different robustness checks assessing the exogeneity of shifts and shares components. We find that the quantity and quality shocks do not predict predetermined variables related to hospitals' workforce and patients' demographics, as it should be the case if shocks are as-good-as-randomly assigned. We also find that differential exposures to common shocks (i.e., the "shares") are not correlated with changes in our outcomes of interest. This result speaks to concerns regarding confounders that could simultaneously affect both the composition of physicians from different origins within hospitals and hospital outcomes.

We also evaluate the role of other hospital inputs in our results. A concern is that the quantity instrument is also capturing an increase in other inputs in the hospital, as would be the case if the migration wave also shifted the number of foreign healthcare workers. In Appendix G, we first examine the complementarity between physicians and other healthcare workers. We do not find a change in the mix of physicians and healthcare workers over time, and this ratio remains constant across hospitals. Furthermore, we show that if the instrument is picking up variation in other healthcare workers, this covariance is negligible conditional on the number of physicians. Consistent with the visual descriptive evidence presented in Appendix Figure A.4, we do not find evidence that the shock directly affected other inputs.

Finally, in Appendix E we examine the impact of physicians' quantity and quality on various hospital-level outcomes using variants of our preferred model. First, we obtain qualitatively similar results when the quality index is defined as the share of physicians in hospital j whose quality falls below the median of the overall quality distribution—instead of the average quality of the physicians working at the hospital. Also, the effects of physician quantity and quality on health outcomes also remain qualitatively similar when considering a translog production function—instead of a Cobb-Douglas—albeit the first stage is weaker in this case.[37]

6.3 Discussion

As described in Section 3, the migration wave increased the number of physicians but also decreased average scores. Our identification strategy exploits this variation to identify the effects of quantity and quality to study the effects of the licensing threshold. In this subsection, we show how the effects of the migration wave can be directly linked to our estimates of the output elasticities, and discuss the extent to which our counterfactual exercise relates to this identifying variation.

We can estimate the effect of the migration wave on inputs and mortality by estimating reduced

[37] In this case, we include the interaction between physicians' quantity and quality as an additional endogenous variable and the interaction of our shift-share instruments as an additional instrumental variable.

form equations of the form:

$$Y_{jt} = \rho_j + \gamma_{f(t)} + \beta Z_{tj} + \epsilon_{jt}$$

where ρ_j and $\gamma_{f(t)}$ are defined as in Equation (17), and Z_{jt} is the quantity shift-share IV.

We consider the following outcome variables Y_{jt}: logged labor $ln(L_{jt})$, average quality $\bar{\theta}_{jt}$, logged death rate, and logged service rate. Reduced-form estimate are presented in Table 5. Column (1) shows that the migration wave increased the number of physicians; while Column (2) shows that it decreased average quality. Finally, Columns (3) and (4) suggest that the migration wave had no significant impacts on hospital death rates but it increased access. The migration wave can therefore be understood as an exogenous shock that simultaneously increased quantity and reduced quality, with increases in service rate but no net effects on mortality.

The relative magnitudes in the changes in quantity and quality that resulted in no mortality effects are informative. Note that a unit increase in the instrument is associated with a 0.504 decrease in average quality and a 0.028 increased in logged labor; or, equivalently, the migration changed the quality per-unit of (logged) labor by $-\frac{d\bar{\theta}}{dZ}/\frac{dln(L)}{dZ} = \frac{0.504}{0.028} = 18.$[38]

By the properties of the IV estimator, and the structural Equation (17), it follows that:

$$\frac{dln(Y)}{dZ} = \hat{\alpha}_L \frac{dln(L)}{dZ} + \hat{\alpha}_{\bar{\theta}} \frac{d\bar{\theta}}{dZ} \simeq 0 \tag{18}$$

Consistent with the fact that $\frac{dln(Y)}{dZ} \simeq 0$, our IV estimates of the return to quantity and quality imply that:

$$\hat{\alpha}_L / \hat{\alpha}_{\bar{\theta}} = 19 \simeq -\frac{d\bar{\theta}}{dZ} / \frac{dln(L)}{dZ}$$

This result illustrates how the ratio of output elasticities is identified using the instrumental variable approach. In other words, the migration wave identifies the slope of the "isomortality curve" in the space of logged labor and average quality.[39] It is also worth noting that the migration wave generated a change in the market that is qualitatively the same, and quantitatively similar to the counterfactual policy of lower the threshold. The migration wave decreased average quality and increased logged quality at a ratio of 19. Our policy counterfactual is an equivalent policy but that decreases quality per unit of logged quality at a ratio of 23. The counterfactual therefore is close to the identifying variation in the data, and makes us more confident that it is robust to

[38] At an average quality of 70.2; this means per each percent unit increase in labor, the migration wave decreased average quality by 0.26 percent.

[39] Note that Equation (18) is also satisfied with $\hat{\alpha}_L = 0$ and $\hat{\alpha}_{\bar{\theta}} = 0$. The additional instrument Z^{θ}_{jt} allows us to discard that solution.

misspecification of the production function.

7 The Impact of Lowering the Licensing Threshold

Taken together, the input and output elasticities we estimated in the previous sections enable us to measure the elasticities of hospital-specific health outcomes to changes in the licensing threshold.

The possibility of exam retaking generates a distinction between the immediate and future impacts of a change in the licensing threshold, as test-takers in any period t' include retakers who failed the exam in a previous period $t < t'$. As such, lowering the threshold in a period t_0 not only impacts who passes in period t_0 but also the set of test takers in any period $t' > t_0$.

We define two time windows for evaluating the effects of changing the licensing threshold:

(i) *Short-Run* ($R = 0$): Elasticities reflect the immediate effects of licensing policy changes.

(ii) *Long-Run* ($R > 0$): Elasticities incorporate the dynamic effects of the policy considering exam retaking.

Leveraging the regime-specific formulation of Y_R and $\eta_{\underline{s}}^{Y_R}$, presented in Section 5, we proceed to evaluate counterfactual scenarios and compare the immediate and dynamic impacts of licensing policy changes.

7.1 Short-Run Impacts of Lowering the Licensing Threshold

The first set of statistics needed for evaluating the effects of changing the licensing threshold are the elasticity of the quantity and quality of physicians in each hospital with respect to the licensing threshold. We compute these elasticities numerically, by simulating the hospital-physician matches under the baseline licensing threshold $\underline{s}$ and under a counterfactual threshold $\underline{s}' = \underline{s} - \Delta\underline{s}$, and compute the resulting qualities and quantities using the expressions in Equation (11).[40]

Figure 3 summarizes our main results for the short-run elasticies in 2018. For ease of interpretation, we plot the *negative of* the elasticities to quantify the impact of a counterfactual *decrease* in the licensing threshold. The size of the marker is proportional to the number of patients in each hospital.

There is generally a quantity-quality trade-off of lowering the licensing threshold as $-\eta_{\underline{s}}^{L_j} > 0$

[40] We provide more details on how we perform this calculation in Appendix I.

and $-\eta_{\underline{s}}^{\bar{\theta}_j} < 0$, although there is ample heterogeneity across hospitals in the magnitudes of the elasticities.

The black dashed line in Figure 3 corresponds to the "iso-mortality curve", depicting the increase in logged quantity (-η_L) that is need to keep mortality constant given the change in average quality (-η_θ) when the threshold is reduced. For most hospitals we find that the increase in quantity is slightly outweighed by the decrease in quality when the threshold is reduced. The solid diamond shows the weighted averages of the quantity and quality elasticities; which lies slightly below the iso-mortality locus.

The green dashed line in Figure 3 corresponds to the "iso-service rate curve", where we replace hospital mortality for service rate as the outcome. Since the quality of physicians do not affect the hospital service rate, the iso-service rate curve is a horizontal line at zero. For service rate, the weighted average elasticity depicted by the solid diamond is above the iso-service locus.

To evaluate the overall effects of lowering the licensing threshold, we contrast these elasticities with the quantity and quality output elasticities noting that, combining Equation (1) with Equation (6), the elasticity of hospital j's value added (lives saved) with respect to the licensing threshold is:

$$\eta_{\underline{s}}^{Y_j} = (\alpha_L^{\text{service}} - \alpha_L^{\text{mortality}}) \cdot \eta_{\underline{s}}^{L_j} + (\alpha_{\bar{\theta}}^{\text{service}} - \alpha_{\bar{\theta}}^{\text{mortality}}) \cdot \tilde{\eta}_{\underline{s}}^{\bar{\theta}_j}.$$

The iso-value added curve is depicted by the red dashed line. On net, we find that healthcare's value added would increase in the short run with a policy that marginally decreases the licensing threshold.

Time Series Evolution of Short-Run Impacts Given the large changes in market fundamentals, mostly resulting from the migration wage, we turn to analyze whether the policy effects found for 2018—5 years after the onset of the migration wave—differ from those in earlier years.

Figure 4-A shows the evolution of the short-run input elasticities over time. For the quantity elasticity (in red), we find an inverse U pattern. The quantity elasticity is the smallest in 2013 when there are few physicians at the margin of passing (see Table A.1). Over time, migration increases the number of marginal physicians, which increases the quantity elasticity. However, after 2016, the elasticity decreases. Even if the number of marginal physicians increases throughout the sample period, more physicians were already hired when the later cohorts arrived. This decreases the elasticity mechanically as the baseline is higher, but also through the lower capacity in the public sector which increases the share of physicians matching with the outside option as captured

by the labor matching functions. Conversely and by the same arguments, the quality semi-elasticity has a U shape.

Figure 4-B shows the resulting elasticities of mortality, service rate, and total healthcare value added, that result in all years in our sample. The elasticity of mortality stays mostly flat over time, although it becomes slightly negative in 2016. Overall, per-patient mortality increases due to the quality reductions outweighing the quantity gains of lowering the threshold. However, the increase in the service rate more than compensates for this effect, and overall value added increases when the threshold is reduced.

7.2 Long-Run Impacts of Lowering the Licensing Threshold

We now turn to analyze the dynamic effects of permanently changing the licensing threshold. The key mechanism we stress in this section is that retaking mitigates the relevance of the licensing threshold over time. For example, in our sample, 83% of test-takers who fail their first attempt in 2013 pass by 2018.

We quantify the dynamic effects of changing the threshold by simulating individual histories for each cohort of test-takers —defined by the year they first take the exam— using the estimated model of scores and retaking from section 5. We use this simulated model to compute yearly elasticities with respect to the threshold that we set permanently lower in 2013.

To gain intuition for the results, Table 6 shows the simulated passing rates for cohort 2013, under the status-quo threshold ($\underline{s} = 51$), as well under a counterfactual policy with a threshold set permanently lower at $\underline{s}' = 41$. In the status quo, 86% of test takers pass in their first attempt (in 2013), compared to 94% under the counterfactual policy. The 8% gap in passing rates determines the short-run elasticities. However, the magnitude of the gap in passing rates across both scenarios shrinks over time. By 2018, almost 94.7% of physicians from the 2013 cohort pass under the status quo, compared to 98.5% under the counterfactual.

The ratio between the simulated short-run and long-run elasticities is shown in Figure 5. By definition, the short and long-run elasticities coincide for 2013 when the policy is implemented. For the following years, the long-run elasticities are smaller, as some of the marginal physicians who contribute to the short-run elasticities in any year $t > 2013$ are retakers who pass in a year $t' < t$ when the threshold is set permanently lower in 2013. By 2018, five years after the policy is implemented, the long-run elasticities are approximately 20% of the short-run elasticities. However, lowering the threshold in 2013 still generates positive outcomes by 2018.

8 Conclusion

Occupational licensing is a widely used regulation in the labor market to ensure a minimum quality standard. However, it comes at the cost of reducing labor supply. This quantity-quality trade-off is particularly relevant in healthcare, where licensing is ubiquitous and could have first-order welfare implications. In this paper, we first analytically characterize this trade-off in an optimal licensing problem. We then propose and estimate an empirical model of physician licensing that allows us to quantify the quantity-quality trade-off in the context of physician licensing in Chile.

Our estimation is built around estimating two sets of sufficient statistics to evaluate the quantity-quality trade-off: The elasticities of inputs—the quantity and quality of labor—with respect to the threshold, and the elasticities of hospital outputs—service rate and mortality rate—with respect to those inputs. To estimate the elasticities of inputs, we infer the latent quality of physicians using individual histories of exam scores and introduce a microfounded labor matching function to estimate the sorting of physicians with different quality levels into hospitals. We then estimate production functions to provide novel estimates of the elasticity of healthcare outcomes with respect to the quantity and quality of physicians. Key to our analysis, we leverage exogenous shocks to these inputs coming from a large migration wave of physicians.

We find that both the quantity and quality of physicians affect health outcomes. As a consequence, lowering the licensing threshold generally entails a quantity-quality trade-off. Still, on net, population health would improve in the short run. Notably, this result is robust to the large changes in the labor market fundamentals coming from the migration wave. We also investigate the long-run policy effects, taking into account exam retaking. While retaking dampens the net effects of lowering the threshold over time, it would not have fully offset them five years later, had the policy been implemented at the start of our sample period.

Other policies can also be effective tools to address physician shortages. Expanding nurses' scope of practice, which involves similar trade-offs between increasing access and potentially affecting quality, may be preferable to lowering licensing thresholds for physicians. Additionally, improving access to medical education can help increase the supply of high-quality physicians. Comparing these alternative policies with the reduction in licensing thresholds presents an important avenue for future research.

References

Abdulkadiroğlu, A. and T. Sönmez (1998). Random serial dictatorship and the core from random endowments in house allocation problems. *Econometrica 66*(3), 689–701.

Ackerberg, D. A., K. Caves, and G. Frazer (2015). Identification properties of recent production function estimators. *Econometrica 83*(6), 2411–2451.

Adao, R., M. Kolesár, and E. Morales (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics 134*(4), 1949–2010.

Agostinelli, F. and M. Wiswall (2016). Estimating the technology of children's skill formation. Technical report, National Bureau of Economic Research.

Ajzenman, N., P. Dominguez, and R. Undurraga (2023). Immigration, crime, and crime (mis) perceptions. *American Economic Journal: Applied Economics 15*(4), 142–176.

Altonji, J. G. and D. Card (1989, September). The Effects of Immigration on the Labor Market Outcomes of Natives. Working Paper 3123, National Bureau of Economic Research.

Anderson, T. W. and H. Rubin (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *The Annals of mathematical statistics 20*(1), 46–63.

Angrist, J. D. and J. Guryan (2008). Does Teacher Testing Raise Teacher Quality? Evidence from State Certification Requirements. *Economics of Education Review 27*(5), 483–503.

Archer, J., N. Lynn, L. Coombes, M. Roberts, T. Gale, and S. Regan de Bere (2017). The medical licensing examination debate. *Regulation & Governance 11*(3), 315–322.

Arnold, J. (2025). *Merger Efficiencies in the US Hospital Industry.* Ph. D. thesis, University of Pennsylvania. Unpublished PhD thesis.

ASOFAMECH (2009–2019). EUNACOM. https://www.eunacom.cl. Asociación de Facultades de Medicina de Chile (ASOFAMECH).

Association of American Medical Colleges (2023). 1 in 5 U.S. Physicians Was Born and Educated Abroad: Who Are They and What Do They Contribute? Accessed: 2024-07-30.

Association of American Medical Colleges (2024). The Complexities of Physician Supply and Demand: Projections From 2021 to 2036. Technical report, AAMC, Washington, DC.

Autor, D. H., D. Dorn, and G. H. Hanson (2013). The China Syndrome: Local Labor Market Effects of Import Competition in the United States. *American economic review 103*(6), 2121–2168.

Azevedo, E. M. and J. D. Leshno (2016). A supply and demand framework for two-sided matching markets. *Journal of Political Economy 124*(5), 1235–1268.

Bahar, D., I. Di Tella, and A. Gulek (2024). Formal Effects of Informal Labor and Work Permits: Evidence from Venezuelan Refugees in Colombia. Working paper.

Borusyak, K., P. Hull, and X. Jaravel (2022). Quasi-Experimental Shift-Share Research Designs. *The Review of economic studies 89*(1), 181–213.

Borusyak, K., P. Hull, and X. Jaravel (2024). A practical guide to shift-share instruments. Technical report, National Bureau of Economic Research.

Borusyak, K. and X. Jaravel (2017). Revisiting event study designs. *Available at SSRN 2826228*.

Carrillo, B. and J. Feres (2019). Provider Supply, Utilization, and Infant Health: Evidence from a Physician Distribution Policy. *American Economic Journal: Economic Policy 11*(3), 156–96.

Castro, M. (2024). Salarios hasta ocho veces más altos: el nuevo fenómeno de los "médicos gaviota" argentinos en Chile. *El País*.

Chandra, A., A. Finkelstein, A. Sacarny, and C. Syverson (2016). Health care exceptionalism? performance and allocation in the us health care sector. *American Economic Review 106*(8), 2110–2144.

Chetty, R. (2009). Sufficient statistics for welfare analysis: A bridge between structural and reduced-form methods. *Annu. Rev. Econ. 1*(1), 451–488.

Crépon, B., E. Duflo, M. Gurgand, R. Rathelot, and P. Zamora (2013). Do labor market policies have displacement effects? evidence from a clustered randomized experiment. *The quarterly journal of economics 128*(2), 531–580.

DEIS (2019). Egresos Hospitalarios. Available at http://www.deis.cl/estadisticas-egresoshospitalarios/. Accessed: 2020-02-09.

DEIS (2020). Listado de Establecimientos de Salud. Available at https://deis.minsal.cl/#datosabiertos. Accessed: 2020-02-09.

Doyle, J. J., S. M. Ewer, and T. H. Wagner (2010). Returns to Physician Human Capital: Evidence from Patients Randomized to Physician Teams. *Journal of Health Economics 29*(6), 866–882.

Efron, B. and C. Morris (1973). Stein's estimation rule and its competitors—an empirical bayes approach. *Journal of the American Statistical Association 68*(341), 117–130.

Elixhauser, A., C. Steiner, D. R. Harris, and R. M. Coffey (1998). Comorbidity measures for use with administrative data. *Medical care*, 8–27.

Fack, G., J. Grenet, and Y. He (2019). Beyond truth-telling: Preference estimation with centralized school choice and college admissions. *American Economic Review 109*(4), 1486–1529.

Federal Reserve Bank of Minneapolis (2022). Occupational licensing requirements can limit employment options for immigrants. Accessed: 2024-08-02.

Finkelstein, A., M. Gentzkow, and H. Williams (2021). Place-Based Drivers of Mortality: Evidence from Migration. *American Economic Review 111*(8), 2697–2735.

Fletcher, J. M., L. I. Horwitz, and E. Bradley (2014). Estimating the Value Added of Attending Physicians on Patient Outcomes. Working Paper 20534, National Bureau of Economic Research.

Friedman, M. (1962). *Capitalism and Freedom.* University of Chicago Press.

Friedman, M. and S. Kuznets (1945). *Income from Independent Professional Practice.* NBER.

Gaynor, M., R. Moreno-Serra, and C. Propper (2013, November). Death by market power: Reform, competition, and patient outcomes in the national health service. *American Economic Journal: Economic Policy 5*(4), 134–66.

Gilraine, M. and J. Penney (2023, 02). Focused Interventions and Test Score Fade-Out. *The Review of Economics and Statistics*, 1–27.

Ginja, R., J. Riise, B. Willage, and A. Willén (2024). Does Your Doctor Matter? Doctor Quality and Patient Outcomes. *Journal of Political Economy Microeconomics*.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2020). Bartik instruments: What, when, why, and how. *American Economic Review 110*(8), 2586–2624.

Grieco, P. L. and R. C. McDevitt (2017). Productivity and Quality in Health Care: Evidence from the Dialysis Industry. *The Review of Economic Studies 84*(3), 1071–1105.

Groeger, A., G. León-Ciliotta, and S. Stillman (2024). Immigration, Labor markets and Discrimination: Evidence from the Venezuelan Exodus in Perú. *World Development 174*, 106437.

Guarin, A., C. Posso, E. Saravia, and J. Tamayo (2021). The Luck of the Draw: The Causal Effect of Physicians on Birth Outcomes. Technical report.

Haakenstad, A., C. M. S. Irvine, M. Knight, C. Bintz, A. Y. Aravkin, P. Zheng, V. Gupta, M. R. Abrigo, A. I. Abushouk, O. M. Adebayo, et al. (2022). Measuring the availability of human resources for health and its relationship to universal health coverage for 204 countries and territories from 1990 to 2019: a systematic analysis for the global burden of disease study 2019. *The Lancet 399*(10341), 2129–2154.

Haakenstad, A., C. M. S. Irvine, M. Knight, C. Bintz, A. Y. Aravkin, P. Zheng, V. Gupta, M. R. M. Abrigo, A. I. Abushouk, O. M. Adebayo, G. . Agarwal, and R. Lozano (2022). Measuring the Availability of Human Resources for Health and Its Relationship to Universal Health Coverage for 204 Countries and Territories from 1990 to 2019: A Systematic Analysis for the Global Burden of Disease Study 2019. *The Lancet 399*(10341), 2129–2154.

Hernández, T. and Y. Ortiz Gómez (2011). La migración de médicos en venezuela. *Revista Panamericana de Salud Pública 30*(2), 177–181.

INE (2024). Estimación de la población extranjera residente en Chile 2022: Resultados Instituto Nacional de Estadísticas. Accessed: December 11, 2024.

Kleiner, M. M. (2011). Enhancing Quality or Restricting Competition: The Case of Licensing Public School Teachers. *Journal of Law and Public Policy 5*(2), 1–15.

Kleiner, M. M. (2014). Occupational Licensing in Health Care. In A. Culyer (Ed.), *Encyclopedia of Health Economics*. Elsevier.

Kleiner, M. M., A. Marier, K. W. Park, and C. Wing (2016, 5). Relaxing Occupational Licensing Requirements: Analyzing Wages and Prices for a Medical Service. *The Journal of Law and Economics 59*(2), 261–291.

Kleiner, M. M. and E. J. Soltas (2023, 02). A Welfare Analysis of Occupational Licensing in U.S. States. *The Review of Economic Studies 90*(5), 2481–2516.

Kleiner, M. M. and W. Wang (2023). The Labor Market Effects of Occupational Licensing in the Public Sector. Working Paper 31213, National Bureau of Economic Research.

Kunakov, N., L. Moraga, and L. Ortiz (2018). Revalidación de títulos médicos extranjeros: eficacia y eficiencia de un examen colaborativo y estandarizado. *Revista médica de Chile 146*(2), 232–240.

Larsen, B., Z. Ju, A. Kapor, and C. Yu (2020). The Effect of Occupational Licensing Stringency on the Teacher Quality Distribution. Working Paper 28158, National Bureau of Economic Research.

Lebow, J. (2022). The labor market effects of venezuelan migration to colombia: Reconciling conflicting results. *IZA Journal of Development and Migration 13*(1), 1–49.

Lee, D. S., J. McCrary, M. J. Moreira, and J. Porter (2022). Valid t-ratio inference for iv. *American Economic Review 112*(10), 3260–3290.

Levinsohn, J. and A. Petrin (2003). Estimating production functions using inputs to control for unobservables. *The review of economic studies 70*(2), 317–341.

Mena, B. (2021). Validity of the single national examination of medical knowledge (eunacom). *Validity of educational assessments in Chile and Latin America*, 353–369.

Menares, F. and P. Muñoz (2025). The impact of standardized disease-specific healthcare coverage. *Journal of Public Economics*, 105312.

Murdoch, W. J., C. Singh, K. Kumbier, R. Abbasi-Asl, and B. Yu (2019). Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences 116*(44), 22071–22080.

OECD (2015). OECD Health Statistics 2015. Accessed: 2024-06-17.

OECD (2019). *Recent Trends in International Migration of Doctors, Nurses and Medical Students*. Paris: OECD Publishing.

OECD (2019). Recent trends in international migration of doctors, nurses and medical students. Accessed: 2024-06-17.

Olivieri, S., F. Ortega, A. Rivadeneira, and E. Carranza (2022). The labour market effects of venezuelan migration in ecuador. *The Journal of Development Studies 58*(4), 713–729.

Olley, G. S. and A. Pakes (1996). The dynamics of productivity in the telecommunications equipment industry. *Econometrica 64*(6), 1263–1297.

Pardo, C. (2019). Health care reform, adverse selection and health insurance choice. *Journal of health economics 67*, 102221.

Propper, C. and J. Van Reenen (2010). Can pay regulation kill? panel data evidence on the effect of labor markets on hospital performance. *Journal of Political Economy 118*(2), 222–273.

Quan, H., V. Sundararajan, P. Halfon, A. Fong, B. Burnand, J.-C. Luthi, L. D. Saunders, C. A. Beck, T. E. Feasby, and W. A. Ghali (2005). Coding algorithms for defining comorbidities in icd-9-cm and icd-10 administrative data. *Medical care*, 1130–1139.

RNPI (2024). Registro Nacional de Prestadores Individuales. https://rnpi.superdesalud.gob.cl/. Superintendencia de Salud de Chile.

SIGTE (2015–2019). Sistema de Gestión de Tiempos de Espera. https://sigte.minsal.cl/. Ministerio de Salud de Chile.

SIRH (2011–2019). Sistema de Información de Recursos Humanos. http://sirh.minsal.cl. Ministerio de Salud de Chile.

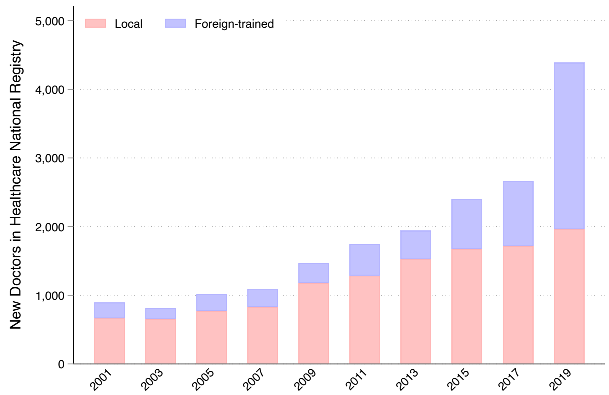
Socha-Dietrich, K. and J. Dumont (2021). International Migration and Movement of Doctors to and Within OECD Countries - 2000 to 2018: Developments in Countries of Destination and Impact on Countries of Origin. Technical Report 126, OECD Publishing, Paris.

Svorny, S. (2004). Licensing doctors: do economists agree? *Econ Journal Watch 1*(2).

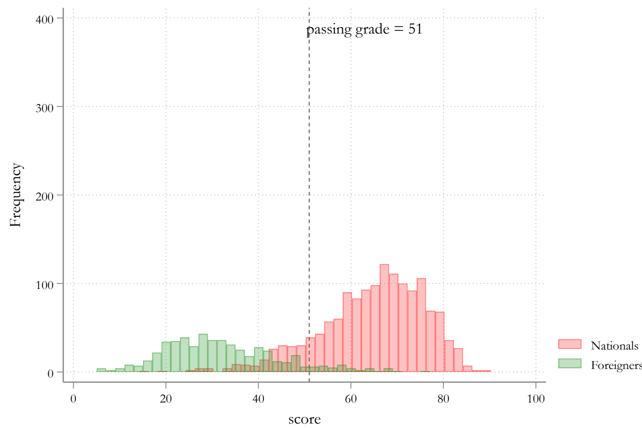
Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology 58*(1), 267–288.

Walters, C. (2024). Empirical bayes methods in labor economics. In *Handbook of Labor Economics*, Volume 5, pp. 183–260. Elsevier.

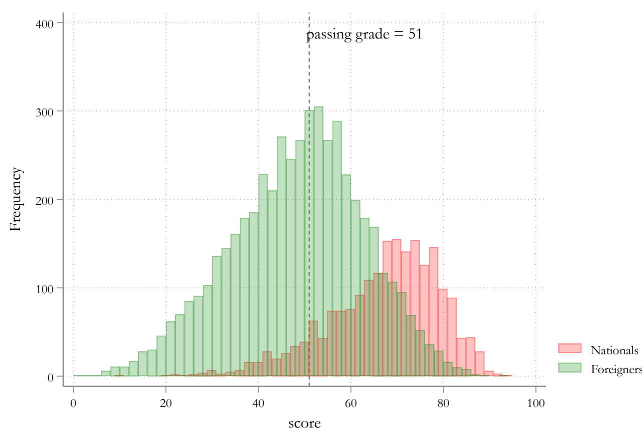
Figure 1: Physician's Labor Market



A. Registered Physicians



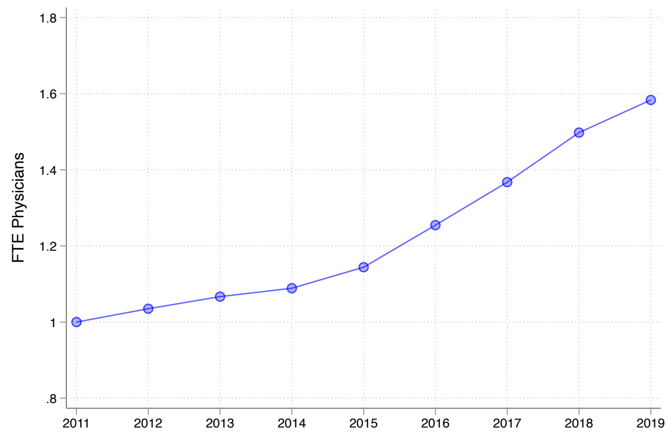
B. Scores in 2013



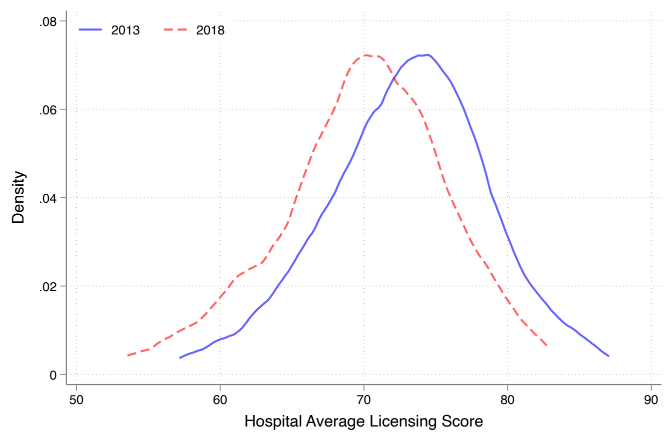
C. Scores in 2018

Notes: Panel A shows the number of newly registered physicians in the National Registry of Healthcare Providers between 2001 and 2019. This registry includes all healthcare workers legally allowed to practice in Chile. The bars are divided into locally-trained physicians (light red) and foreign-trained physicians (light blue). Panels B and C display the distribution of test scores for locally-trained (red) and foreign-trained (green) physicians in 2013 and 2018, respectively. The dashed line in both graphs represents the minimum passing score.

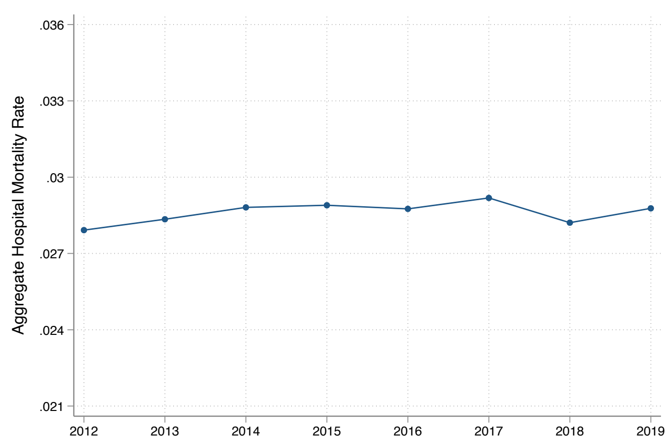
Figure 2: Physicians in Public Hospitals



A. Quantity



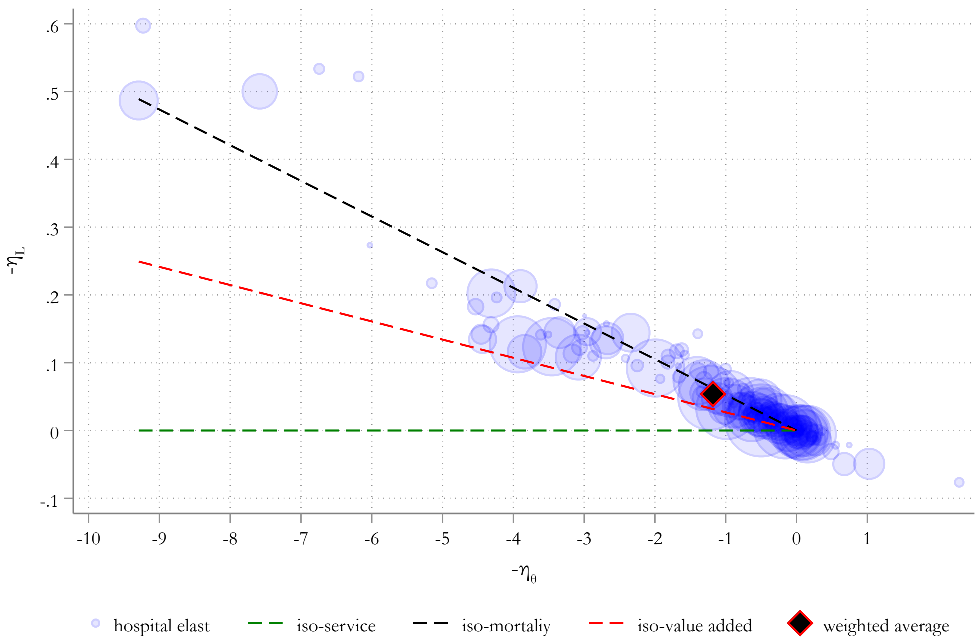
B. Scores



C. Hospital Mortality

Notes: Panel A plots the growth of full-time equivalent physicians in public hospitals, indexed to 2011, which we set equal to 1. Panel B presents the distribution of public hospitals' average licensing scores of physicians in 2013 and 2018. Panel C depicts the average aggregate mortality rate in public hospitals throughout our analysis period.

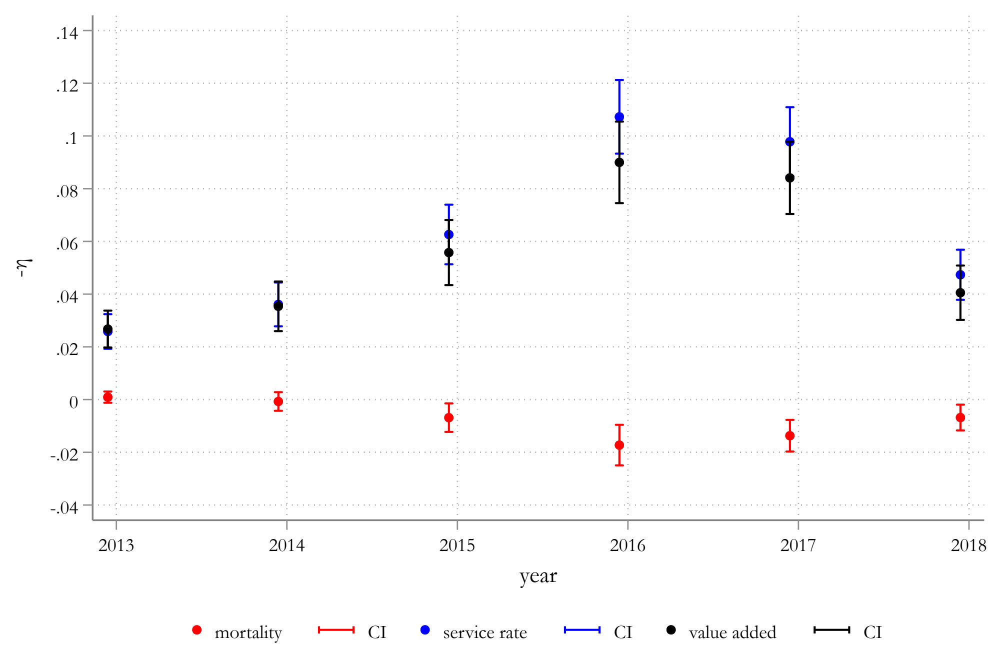
Figure 3: Short-run Elasticity of Quantity and Semi-elasticity of Quality by Hospital in 2018



Notes: This figure shows a scatter plot with the estimates for the (negative of) the elasticity of quantity and the semi-elasticity of quality with respect to the licensing threshold. Each dot represents a hospital. The size of the marker is proportional to the number of patients. Combinations of elasticities above the dashed lines are such that patient outcomes for the outcomes depicted by that line improve when the threshold is reduced.

Figure 4: Time Series of Short-Run Elasticities

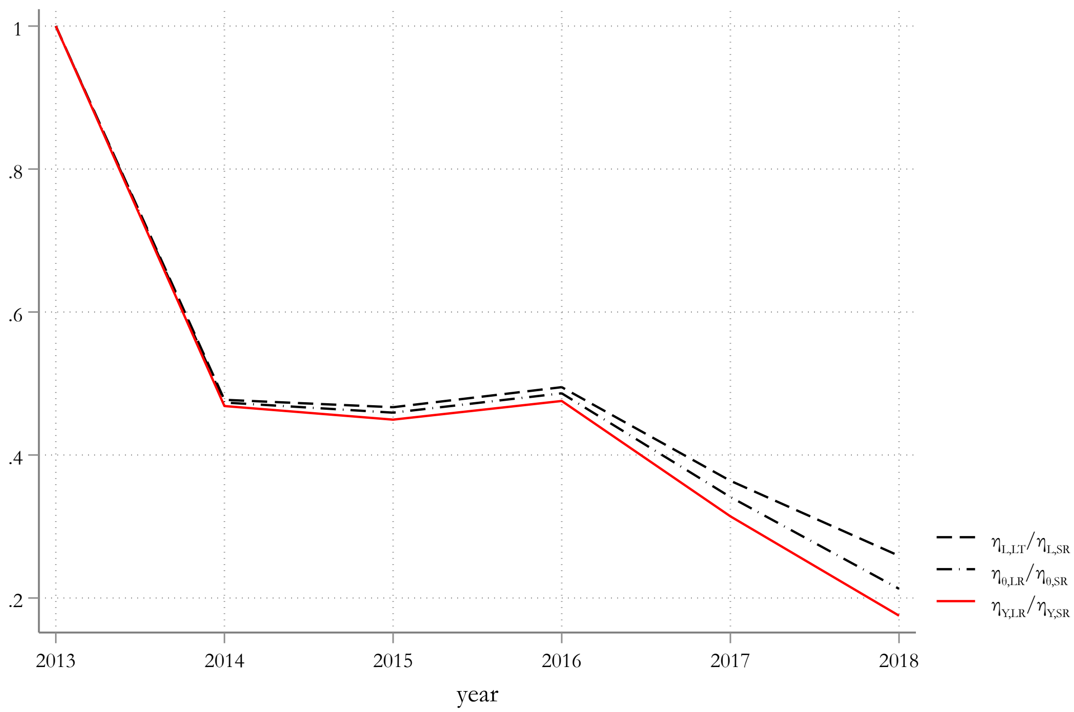
A. Labor Quantity and Quality



B. Mortality, Access and Value Added

Notes: Panel A shows the evolution of the average short-run elasticities of quantity and quality of physicians with respect to the licensing threshold. Panel B shows the resulting average effects on patient mortality, access, and on the value added of the healthcare sector.

Figure 5: Ratio of Simulated Short- and Long-Run Elasticities



Notes: The figure shows the ratio between the short- and long-run elasticities of mortality, service rate, and value added.

Table 1: Model Estimates

Panel A: Retaking Model Estimates		
	Foreigners	Nationals
n_{it}	-0.231	-0.163
	(0.020)	(0.059)
$\bar{s} - s_{it}$	-0.036	-0.060
	(0.003)	(0.009)
Constant	2.592	1.595
	(0.077)	(0.139)
Observations	8,221	1,340

Panel B: Scores Model Estimates		
	Foreigners	Natives
$\hat{\mu}_\theta$	46.002	65.474
	(0.864)	(0.107)
$\hat{\sigma}_\theta$	14.723	8.679
	(0.174)	(0.113)
$\hat{\sigma}_\epsilon$	9.227	8.677
	(0.190)	(0.072)
SNR	0.718	0.500
	(0.012)	(0.009)
	Common	
$\hat{\gamma}$	9.599	
	(0.541)	
$\hat{\rho}$	0.285	
	(0.050)	

Notes: The table shows estimates associated with the dynamic model of scores. Panel A presents maximum likelihood estimates for the coefficients of the logit model for retaking, as specified in Equation (8), with standard errors in parentheses. Panel B presents the coefficients estimated via Simulated Method of Moments (SMM) with bootstrapped standard errors, computed using the empirical standard deviation of the estimates across 10 simulations, in parentheses.

Table 2: Conditional Matching Probabilities Estimates

	Alternative Models			Placebo
	(1)	(2)	(3)	(4)
Distance_{ij}	-0.228	-0.228	-0.228	-0.228
	(0.014)	(0.014)	(0.014)	(0.014)
Share_{ijt-1}	0.651	0.651	0.653	0.652
	(0.146)	(0.146)	(0.146)	(0.146)
$(M_{it}^{0})/\kappa_{jt}$	-0.637	-0.676	-0.726	
	(0.152)	(0.159)	(0.186)	
$(M_{it}^{+})/\kappa_{jt}$		0.017	0.050	
		(0.021)	(0.038)	
$(M_{it}^{0})/\kappa_{jt} \times \mathbb{1}[r_j = 2]$			0.004	
			(0.097)	
$(M_{it}^{0})/\kappa_{jt} \times \mathbb{1}[r_j = 3]$			0.115	
			(0.122)	
$(M_{it}^{+})/\kappa_{jt} \times \mathbb{1}[r_j = 2]$			-0.037	
			(0.045)	
$(M_{it}^{+})/\kappa_{jt} \times \mathbb{1}[r_j = 3]$			-0.067	
			(0.056)	
$(M_{it}^{-})/\kappa_{jt}$				0.022
				(0.019)
Log likelihood	-15276.84	-15276.53	-15275.33	-15285.23

Notes: The table shows results from a multinomial logit model for the matching probabilities between physicians and hospital referral regions. The functions $\mathbb{1}[r_j = k]$ indicate the hospital's quality tier. As shown in Equation (13), all specifications include alternative-specific coefficients for the following variables: Year, Foreign indicator, physician's posterior quality, an interaction between Foreign indicator and posterior quality, and a Specialist indicator. We display only the estimates for the coefficients that do not depend on the alternative.

Table 3: Correlates of Elasticities

	Labor elasticity $\eta_{\underline{s}}^{L_{jt}}$				Quality Semi-elasticity $\eta_{\underline{s}}^{\bar{\theta}_{jt}}$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
High (phys/pat)$_{j,t-1}$	0.019			0.023	-0.294			-0.416
	(0.007)			(0.007)	(0.133)			(0.138)
High average score$_{j,t-1}$		0.033		0.029		-0.236		-0.163
		(0.006)		(0.006)		(0.135)		(0.126)
North$_j$			-0.025	-0.028			0.557	0.630
			(0.007)	(0.007)			(0.139)	(0.143)
Mean Dep. Var.	-0.072	-0.072	-0.072	-0.072	1.401	1.401	1.401	1.401
Observations	1086	1086	1086	1086	1086	1086	1086	1086

Notes: The table shows OLS regression coefficients relating the estimated elasticities quantity and quality elasticities to hospital observables. All specifications include year fixed effects.

Table 4: Impact of Physicians' Quantity and Quality on Access and Quality of Care

Panel A: Access

	Ln service rate	Ln inpatient surgeries	Ln exits from waiting list: Surgical	Ln exits from waiting list: Medical
	(1)	(2)	(3)	(4)
Ln Physicians ($\hat{\alpha}_L^{\text{service}}$)	1.01	4.97	3.69	3.00
	[0.25]	[1.96]	[0.69]	[1.02]
Avg. Physicians' Quality ($\hat{\alpha}_\theta^{\text{service}}$)	0.01	0.11	-0.00	0.02
	[0.01]	[0.10]	[0.04]	[0.06]
Observations	1,402	744	738	942
Mean Dep. Var.	0.015	3,803	1,534	8,403
F-stat (First-stage)	22	12.3	9.9	15.9
Anderson-Rubin (χ^2)	0.00	0.00	0.00	0.00

Panel B: Quality

	Mortality: In-Hospital: Ln death rate	Mortality: In-Hospital: Pred. death rate	Mortality: 28-days: Ln death rate	In-hospital Complications: Ln complications rate
	(1)	(2)	(3)	(4)
Ln Physicians ($\hat{\alpha}_L^{\text{mortality}}$)	-0.83	0.13	-0.74	-0.58
	[0.19]	[0.11]	[0.20]	[0.23]
Avg. Physicians' Quality ($\hat{\alpha}_\theta^{\text{mortality}}$)	-0.04	-0.00	-0.04	-0.04
	[0.01]	[0.01]	[0.01]	[0.02]
Observations	1,402	1,402	1,402	1,402
Mean Dep. var.	3.284	3.494	5.075	3.272
F-stat (First-stage)	22	34.90	22	22
Anderson-Rubin (χ^2) p-value	0.00	0.02	0.00	0.01

Notes: This table presents the impact of the quantity and quality of physicians on public hospital performance. Panel A focuses on utilization, which we proxy through the service rate, inpatient surgeries, and exits from the waiting list. Panel B focuses on patients' mortality and complications. Estimates come from the two-stage least squares estimation of Equation (17). We present the exposure-robust standard errors clustered at the region of origin (i.e., the shock) level in brackets (Adao et al., 2019; Borusyak et al., 2022).

Table 5: The Effects of the Migration Wave on Inputs and Mortality

	Ln # Physicians	Average Quality	Ln Death Rate	Service Rate
	(1)	(2)	(3)	(4)
Z_{jt}^{L}	0.028	-0.504	-0.002	0.026
	(0.003)	(0.033)	(0.006)	0.005
Observations	1,402	1,402	1,402	1,402

Notes: The table shows the reduced-form effects of the quantity-shift-share IV on number of physicians, average quality and the log of the death date.

Table 6: Simulated Passing Year for 2013 Cohort

Year	$\underline{s} = 51$		$\underline{s} = 41$	
	passing rate	cumulative passing rate	passing rate	cumulative passing rate
2013	86.0	86.0	94.0	94.0
2014	6.8	92.8	3.5	97.5
2015	1.4	94.2	0.7	98.2
2016	0.3	94.6	0.2	98.4
2017	0.1	94.6	0.0	98.4
2018	0.1	94.7	0.1	98.5

Notes: The table shows the simulated passing rates over time for the cohort of physicians that take the exam for the first time in 2013, under a (baseline) scenario where the threshold is kept at 51 points, and under a counterfactual scenario where the threshold is reduced to 41 points.

ONLINE APPENDIX

Physicians' Occupational Licensing and the Quantity-Quality Trade-off

Juan Pablo Atal, Tomás Larroucau, Pablo Muñoz, and Cristóbal Otero

List of Figures

List of Tables

A Score Imputation

The EUNACOM exam was introduced in 2009. Before that, physicians were not required to take a standardized test to practice medicine. To account for the quality of physicians licensed before EUNACOM, we impute their hypothetical scores. This appendix outlines the imputation procedure.

Formally, let y_{ih} denote the licensing score of physician i in hospital h. Our baseline model for scores is:

$$y_{ih} = \mu_h + \alpha^h_{r(i)} + \varepsilon_{ih},$$

where μ_h represents the hospital-level mean, $\alpha^h_{r(i)}$ captures the region-of-origin effect (or "differential score") for physicians from region r working in hospital h, and ε_{ih} is an idiosyncratic error term with mean of zero.

To improve out-of-sample predictions, which is particularly important when hospitals have very few physicians from a particular region, we apply an empirical Bayes shrinkage procedure (Efron and Morris, 1973; Walters, 2024). Specifically, we regress scores on region-of-origin fixed effects for each hospital h using OLS with sum-to-zero constraints (i.e., $\sum_r \alpha^h_r = 0; \forall h$). We then treat each α^h_r and its standard error s_h as coming from a prior distribution centered at 0—reflecting the sum-to-zero constraint—with limiting variance s^2_h. Given the OLS estimate $\widehat{\alpha}^h_r$ and its sampling variance $\widehat{\text{Var}}(\widehat{\alpha}^h_r)$, the empirical Bayes estimate (posterior mean) is a weighted average of the OLS estimate and the prior mean (zero). Specifically:

$$\widetilde{\alpha}^h_r \;=\; \underbrace{\frac{\widehat{\text{Var}}(\widehat{\alpha}^h_r)}{\widehat{\text{Var}}(\widehat{\alpha}^h_r) + s^2_h}}_{\text{shrinkage factor}} \;\widehat{\alpha}^h_r.$$

Since the shrinkage factor ranges between 0 and 1, each estimated effect is pulled ("shrunk") toward zero, with the degree of shrinkage increasing as the precision of $\widehat{\alpha}^h_r$ decreases.

Imputation of Missing Scores: For a physician i at hospital h from region $r(i)$ who does *not* have an observed licensing score, we impute her score using the empirical Bayes-adjusted model:

$$\widehat{y}_{ih} \;=\; \widehat{\mu}_h \;+\; \widetilde{\alpha}^h_{r(i)},$$

where $\widehat{\mu}_h$ is the average score of physicians working at hospital h. Thus, the imputed score is a combination of the grand mean for hospital h ($\widehat{\mu}_h$), and the "shrunk" region-of-origin differential at hospital h ($\widetilde{\alpha}^h_{r(i)}$).

In very few cases (1.6%), the data does not include a region of origin for the physician. In these cases, we predict the individual score using a LASSO model. This model incorporates hospital indicators as well as controls for physicians' age and gender.

A.1 Alternative Imputation Method:

As an alternative to our Empirical Bayes approach to imputation, we also employ a Least Absolute Shrinkage and Selection Operator (LASSO) model for prediction (Tibshirani, 1996; Murdoch et al., 2019). Specifically, we estimate a LASSO regression at the physician level to predict EUNACOM scores based on physicians' gender, age, age squared, and a full set of hospital by region-of-origin fixed effects.

Panel A of Figure A.3 plots the mean cross-validation (CV) error against the regularization parameter λ. Each point represents the average error from k-fold CV for a specific value of λ. We highlight the λ that minimizes the CV error, denoted λ_{CV}, and use it to shrink coefficients toward zero, effectively performing variable selection. Panel B of Figure A.3 presents the density of prediction errors from each model. For this comparison, we focus on physicians whose EUNACOM scores we observe and contrast their actual scores with those predicted by both LASSO and our Empirical Bayes approach.

Finally, Table A.2 replicates the analysis presented in Table 4 in the main body of the paper but using the LASSO imputation method. We find quantitative and qualitative similar results, albeit with a weaker first stage.

B Other Labor Inputs

To work in a public hospital, foreign-trained healthcare professionals must validate their degree in Chile. This process is overseen by specific recognized Chilean universities and involves submitting academic credentials, with additional coursework or exams required in some cases. For physicians, passing the EUNACOM is also necessary, as it qualifies them to practice in public healthcare settings.[41] After validation, they must register with the National Registry of Individual Health Providers by providing their validated degree and other required personal documentation.

We observe the impact of the migration wave only in the introduction of foreign-trained physicians. As described in the main text, there is a significant increase in the share of foreign-trained physicians, both among those newly enrolled in the National Registry of Healthcare Providers and those working in public hospitals. However, we do not find similar effects for other healthcare workers. As shown in Panel A of Figure A.4, the share of newly registered healthcare workers in the National Registry of Individual Health Providers remained fairly stable between the 2000s and 2019, except for physicians. Panel B highlights the share of foreign-trained healthcare workers employed in public hospitals, and shows a significant increase for physicians—from 7% in 2011 to 18% in 2019. In contrast, there is no noticeable change in the share of foreign-trained workers in other healthcare categories.

C A mapping with Chetty (2009)

Chetty (2009) provides a framework to compute the welfare effects of policies from a set of sufficient statics rather than from the entire set of model primitives. Chetty's main framework provides a rubric of steps to derive the sufficient statics in the context of a static single-agent model; where the agent takes actions (e.g. decides consumption and leisure) to maximize utility subject to constraints that are affected by government policies (e.g. budget constraints affected by taxes and transfers). However Chetty (2009) also notes that assuming optimizing behavior is not needed to derive sufficient statistics, as long as one can directly estimate the objects determining the derivative

[41]Passing the EUNACOM simultaneously qualifies physicians to practice and validates their degree, which reduces the incentive to pursue a separate degree validation process before taking the EUNACOM.

of welfare with respect to the policy variable. We follow that approach in the paper.

For completeness, we discuss below how one could apply the optimization framework in our context. We recast our licensing problem by considering the problem of a representative hospital optimally choosing labor and quality subject to constraints that are affected by the licensing threshold. We follow the same rubric of steps as in Chetty (2009), and show how to use it to derive the sufficient statics formula in the text (Equation 1).

Using the notation of Section 4 the structure of the model would be given by:

$$\begin{aligned} \max_{L,\bar{\theta}} \quad & F(L,\bar{\theta}) \\ \text{subject to} \quad & G_1(L,\underline{s}) \equiv L - m\int_{\underline{s}}^{\infty} h(s)\,p(s|\underline{s})\,ds = 0, \\ & G_2(\bar{\theta},L,\underline{s}) \equiv L\,\bar{\theta} - \int_{\underline{s}}^{\infty} \theta(s)\,h(s)\,p(s|\underline{s})\,ds = 0, \end{aligned}$$

where $F(L,\bar{\theta})$ is the hospital's output as a function of quantity and quality of labor. $G_1(L,\underline{s})$ and $G_2(L,\underline{s})$ are the constraints linking quantity and quality with the licensing threshold. Note that these constraints include the matching functions. The constraints therefore result, in part, from the matching process.

In the paper, we focus on the elasticity of hospital production with respect to the threshold, which is the planner's welfare criterion in Chetty's context. Welfare as a function of the licensing threshold is given by

$$Y(s) \equiv \max_{L,\bar{\theta}} F(L(s),\bar{\theta}(s)) + \lambda G_1(L,s) + \mu G_2(\bar{\theta},L,s).$$

As in Chetty (2009), we can use the envelope conditions and differentiate Y to get

$$\frac{dY}{ds} = \lambda\frac{\partial G_1}{\partial s} + \mu\frac{\partial G_2}{\partial s}. \tag{A.1}$$

In Chetty's framework, Lagrange multipliers are recovered from marginal utilities. In our framework, Lagrange multipliers are recovered from the output elasticities in the production function. Note that optimization implies that the marginal products of inputs are equated to linear combinations of the Lagrange multipliers:

$$\frac{\partial F}{\partial L} = \lambda\frac{\partial G_1}{\partial L} + \mu\frac{\partial G_2}{\partial L}, \tag{A.2}$$

and

$$\frac{\partial F}{\partial \bar{\theta}} = \mu\frac{\partial G_2}{\partial \bar{\theta}}. \tag{A.3}$$

Also, note that differentiating $G_1(L,s)$ with respect to s yields:

$$\frac{dL}{ds} = \frac{\partial G_1}{\partial s}\cdot\frac{1}{\frac{\partial G_1}{\partial L}}. \tag{A.4}$$

Similarly, differentiating $G_2(L, s, \bar{\theta})$ with respect to s yields:

$$\frac{d\bar{\theta}}{ds} = \frac{\partial G_2}{\partial s} \cdot \frac{1}{\frac{\partial G_2}{\partial \theta}} - \frac{\partial G_2}{\partial L} \cdot \frac{1}{\frac{\partial G_2}{\partial \theta}} \cdot \frac{\partial G_1}{\partial s} \cdot \frac{1}{\frac{\partial G_1}{\partial L}}. \tag{A.5}$$

Combining equations A.1-A.5 yields the sufficient statistic formula (Equation (1) in the text).

We also note that Chetty (2009) proposes to use the first order conditions to derive the marginal utilities from observed choices. In our context, it would entail linking the marginal product of inputs to hospital's input decisions (e.g. using factor shares). Instead, we rely on exogenous variation in inputs to estimate those parameters without imposing optimality.

D Quality Gains

In general, our model could allow for scores to improve over attempts due to increased test-taking ability as well as due to improvements in quality. With data on scores alone, it is not possible to disentangle both mechanisms. In this section, we show that auxiliary data is consistent with no quality improvements over attempts. This justifies the assumption of no quality gains over attempts.

We leverage the discontinuity in retaking probability around the passing threshold to show that retaking does not improve labor-market outcomes that proxy for quality. Figure A.5-A shows the probability of ever retaking the test as a function of the distance to the score in the first attempt to the threshold. It shows that the retaking probability drops substantially around the cutoff. Figure A.5-B shows that the maximum achieved scored changes discontinuously around the threshold. Note that score gains in panel B are a combination of gains in test-taking ability, gains in quality, and selection around cutoff (Gilraine and Penney, 2023).

We investigate how two proxies of quality change around the threshold. Figure A.5-C shows that the probability of being appointed chief general practitioner increases substantially with the score in the first attempt. However, there are no discernible increases around the threshold. Similarly Figure A.5-D shows a (more modest) increase in the hours of contract with the score in the first attempt, with no discernible increases around the threshold. Together, we take this as evidence against the idea that there are quality improvements from retaking.

A key feature of the above framework is to assume that agents maximize utility although, as pointed out by Chetty (2009); it is not a necessary feature of the sufficient statistic framework as long as one can identify the relevant elasticities. However, a common challenge in identifying the elasticities without imposing maximization is that marginal utilities are unobserved. Imposing optimizing behavior allows researchers to recover marginal utilities from observed choices. One feature of our application is that we focus on an *observable* outcome (hospital production) and therefore can measure marginal utilities (the marginal products of labor and quality) directly from the data.

E Production Function

E.1 Translog Production Function:

Given its parsimony and tractability, we opt for the Cobb-Douglas specification in our main analysis. Table A.5 below assesses the impacts of physician quantity and quality on health outcomes when using a Translog production function, which allows for a linear interaction between physicians' quantity and quality. The table shows the results from the main model in columns (1) and (3) and the results from the alternative model in columns (2) and (4). At the bottom of the table, we present the overall effect of physicians' quantity and quality on access and mortality. The results are similar, but we obtain a weak first stage for the Translog model, an additional reason in favor of our preferred specification.

E.2 Alternative Quality Index:

Table A.6 presents estimates of the health production function using an alternative specification. In this specification, the quality index is defined as the share of physicians in the hospital j during period t whose quality falls below the median of the overall distribution. Reassuringly, the results remain qualitatively similar: quality does not affect access, and worse quality is associated with poorer health outcomes.

F Shift-share Instruments

To assess the robustness of our instrumental variables approach, we build on a recent econometric literature which suggests two distinct paths to identification. One path, developed by Borusyak and Jaravel (2017) and Adao et al. (2019), leverages many exogenous shifts while making no assumption on the exogeneity of the shares. The second path, proposed by Goldsmith-Pinkham et al. (2020), instead focuses on share exogeneity. As pointed out in Borusyak et al. (2024), identification "from the shifts" can be understood as leveraging a shift-level natural experiment, while identification "from the shares" can be viewed as pooling together multiple difference-in-differences designs leveraging heterogeneous shock exposure. In this appendix, we present different robustness checks assessing the exogeneity of our shifts and shares.

F.1 Identification "from the shifts"

Shift-based identification stems from the observation that a share-weighted average of random shifts is itself as-good-as-random (Borusyak et al., 2024). This is true even if the shares are econometrically endogenous, in the sense that units with different shares may have systematically different unobservables. Indeed, Borusyak et al. (2022) shows that classical shift-share IV regression coefficients are numerically equivalent to those obtained from a regression where the outcome and treatment variables are first averaged, using exposure shares as weights, and the shocks are directly used as instrument for the aggregated treatment.

The fact that shift-share estimates can be equivalently obtained by a shock-level IV procedure suggests ways to establishing their consistency. Following Borusyak et al. (2022), we begin by showing statistics of the shocks. Table A.3 shows the distribution of the quantity and quality shocks. As shown by columns (1) and (3), the distribution of the quantity shocks has an average of 3.0, a standard deviation of 7.6, and an interquartile range of 1.8; while the distribution of the quality shocks has an average of -0.5, a standard deviation of 2.2, and an interquartile range of 3.9. Columns (2) and (4) show that there is residual shock variation even conditional on period fixed effects. The inverse HHI of the exposure shares is 28.8 across region by period cells. The largest shock weights are 6% across region-by-periods.

To corroborate the plausibility of the conditional quasi-random shock assignment we perform a shock balance test. If the migration shocks are as-good-as-randomly assigned, we expect them to not predict predetermined variables related to hospitals' workforce and patients' demographics. Figure A.6 reports the results of our balance tests on potential confounders. Panel A focuses on the quantity shocks and Panel B on the quality shocks. Reassuringly, we find that there is no statistically significant relationship between most variables and the shocks. There is some evidence of unbalance between the quantity shock and a few patient and hospital characteristics. Nonetheless, these are variables we control for in our main analysis.

F.2 Identification "from the shares"

Share-based identification stems from the work by Goldsmith-Pinkham et al. (2020) showing that the Bartik 2SLS estimator can be decomposed into a weighted sum of the just-identified instrumental variable estimators that use each entity-specific share as a separate instrument.

The fact that the exogeneity of the shift-share instruments might also rely on the exogeneity of the shares implies that for our empirical strategy to be valid, we require the differential exposures to common immigration shocks (the "shares") to be independent of differential changes in our outcome of interest during the pre-period. In our case, this would not be the case if there are endogenous mechanisms affecting both the composition of immigrants within hospitals and patients' outcomes.

To assess the plausibility of the identification assumption we follow Goldsmith-Pinkham et al. (2020) and proceed in two steps. First, we compute the Rotemberg weights for each country-specific instrument. Rotemberg weights indicate which country-specific exposure gets a larger weight in the overall Bartik-2SLS estimate, and thus which nationality-share effects are more important for testing. In our data, Ecuador has the highest weight (0.585), followed by Colombia (0.262), mid- and low-tier universities in the central region of Chile (0.22), and Cuba (0.182). Second, we test for "parallel trends". Table A.4 presents our results on the relationship between differential exposures to common immigration shocks (the "shares") and differential changes in our outcome of interest during the pre-period. Specifically, we assess "pre-trends" by regressing the outcome of interest against the region of origin-shares during the (pre) period 2012-2015, controlling by hospital and year fixed effects as well our main battery of control variables. Generally, we find that the differential exposure to immigrants from different regions do not statistically or economically predict differential utilization or health outcomes. An exception is the share of European physicians, which seem to be statistically associated with surgeries and infection rates. However, the Rotemberg weight of Europe is among the smallest (0.007). Figure A.7 complements our analysis by visually showing the point estimate and confidence interval associated to each share, for all our outcomes of interest.

G Do other inputs affect the results?

In this Appendix, we first document the existence of complementarities between physicians and other healthcare workers at the hospital level and then discuss the implications for identifying physicians' returns on hospital outcomes.

Panel A in Appendix Figure A.8 presents a scatter plot of the number of physicians and other FTE healthcare workers at the hospital. An observation is a hospital-by-year. The figure shows a strikingly linear relationship between physicians and other inputs. We further examine whether the mix of other inputs relative to physicians changed with the increase in the number of physicians by differentiating between the pre- and post-migration wave and do not find differences between them. Since both sets of observations overlap, this provides evidence that as hospitals hired more physicians, they also hired additional inputs at a similar rate.

A related question is whether the mix of physicians and other healthcare workers varies according to physician quality. In Panel B, we plot the ratio of other healthcare workers to physicians against average physician quality. We do not find that hospitals with higher-quality physicians substitute away from other inputs.

Given this evidence, it is natural to ask about the implications of incorporating other inputs into the production function as a function of the number of physicians. Since the ratio of other inputs to physicians does not change with physician quality, we assume away quality for ease of exposition. The production function takes the following form: $Y_i = AL_i^{\alpha_L} O_i^{\alpha_O}$, with $O_i = e^c L_i^{\gamma}$. Thus,

$$\ln Y_i = \phi + \underbrace{(\alpha_L + \gamma \alpha_O)}_{\tilde{\alpha}_L} \ln L_i.$$

Given the complementarities between physicians and other inputs, the impact of physicians on the outcome of interest, $\tilde{\alpha}_L$, is a bundled effect: a direct effect of an extra doctor, α_L, and an indirect effect from the increase in other inputs, $\gamma \alpha_O$.

We next discuss the implications for identification in our setting where the quantity of physicians and the number of other inputs are endogenous. Assume that we specify the following estimating equation:

$$y_i = \phi + \alpha_L l_i + \alpha_O o_i + \epsilon_i,$$

where y_i is a logged outcome (e.g., death rate, service rate, etc.), l_i is the logged quantity of doctors, and o_i is the logged number of other healthcare workers. Assume that l_i and o_i are endogenous (i.e., $\text{Cov}(l_i, \epsilon_i) \neq 0$, and $\text{Cov}(o_i, \epsilon_i) \neq 0$). Further, assume a plausibly exogenous instrument z_i for l_i, such that $\text{Cov}(l_i, z_i) \neq 0$, but that might be correlated with o_i ($\text{Cov}(z_i, o_i) \neq 0$), which could violate the exclusion restriction. The IV estimator of α_L is given by:

$$\begin{aligned}\hat{\alpha}_L^{IV} &= \frac{\text{Cov}(y_i, z_i)}{\text{Cov}(l_i, z_i)} \\ &= \alpha_L + \alpha_O \frac{\text{Cov}(o_i, z_i)}{\text{Cov}(l_i, z_i)},\end{aligned}$$

where the second step follows from the assumption that $\text{Cov}(\epsilon_i, z_i) = 0$. Incorporating the complementarity between physicians and other inputs as $o_i = c + \gamma l_i + \nu_i$,

$$\hat{\alpha}_L^{IV} = \alpha_L + \alpha_O \frac{\text{Cov}(c + \gamma l_i + \nu_i, z_i)}{\text{Cov}(l_i, z_i)}$$
$$= \underbrace{\alpha_L + \gamma\alpha_O}_{\tilde{\alpha_L}} + \underbrace{\alpha_O \frac{\text{Cov}(\nu_i, z_i)}{\text{Cov}(l_i, z_i)}}_{bias}.$$

Thus, $\hat{\alpha_L}^{IV}$ is the bundled effect of the direct and indirect effects of more physicians on the outcome, plus *bias* if $\text{Cov}(\nu_i, z_i) \neq 0$ (i.e., if innovations shifting o_i correlate with z_i). We can empirically gauge the magnitude of the bias by estimating the linear projection of the instrument on other inputs:

$$o_i = \alpha + \beta z_i + e_i.$$

Using the fact that $\beta = \frac{Cov(o_i,z_i)}{Var(z_i)}$, it is straightforward to show that:

$$\hat{\text{Cov}}(\nu_i, z_i) = \hat{\beta}\hat{\text{Var}}(z_i) - \gamma\hat{\text{Cov}}(l_i, z_i).$$

Empirically, we obtain $\hat{\beta} = 0.0207$, $\hat{\text{Var}}(z_i) = 1.042$, $\hat{\gamma} = 0.710$, and $\hat{\text{Cov}}(l_i, z_i) = 0.030$. Adding these numbers, we find that $\hat{\text{Cov}}(\nu_i, z_i) \approx 0$, which implies that the bias is also close to zero. Complementary evidence for these results comes from the trends in labor supply of L_i and O_i over time, discussed in Appendix B. The evidence therein shows that migration shocks, which are the main quasi-exogenous variation leveraged by our instrument, mostly shifted the supply of physicians, with negligible effects on other inputs.

H Microfoundation of CMPs

We now provide a micro-foundation for the reduced-form CMPs presented in the main body of the paper. We consider a static version of the model and fix t. Consider a continuum of physicians indexed by $i \in \mathcal{I}$. Each physician i belongs to a type $\tau_i \in \{N, F\}$, with a total mass m_τ for each type τ. Let $\mathcal{J}$ be a finite set of hospitals, each indexed by j and endowed with $\kappa_j \in \mathcal{K}$ vacancies, and let by ϕ physicians' outside option.

Physician preferences. The physician's indirect utility from matching with hospital $j \in \mathcal{J}$ is given by

$$u_{ij} \;=\; \tilde{v}\left(x_{ij}, \hat{\theta}(s_i, \tau_i)\right) \;+\; \varepsilon_{ij}, \tag{A.6}$$

where $\tilde{v}(\cdot)$ is a known function that captures the systematic part of physicians utilities; x_{ij} represents observable characteristics at the physician-hospital level; $\hat{\theta}(s_i, \tau_i)$ is physician i's posterior mean quality, which is a function of their licensing score s_i and type τ_i; and ε_{ij} are i.i.d. Type-I Extreme Value shocks. Physicians matched with the outside option, ϕ, receive a systematic utility normalized to zero.

Hospital preferences. Hospitals maximize an objective function that is strictly increasing in the expected quality $\hat{\theta}$ of the physicians they match with and find all to be admissible. Strict monotonicity in $\hat{\theta}$ implies that if a hospital j can replace a physician of lower expected quality with one of higher expected quality, it will strictly prefer to do so.

Equilibrium. An equilibrium allocation satisfies:

(i) *Individual rationality*: no matched physician prefers their outside option to their allocation.

(ii) *Pairwise stability*: no hospital-physician pair wants to deviate from their current match and match with each other (no blocking pairs).

Following Azevedo and Leshno (2016), the equilibrium can be characterized by a system of demand and supply equations determining a set of–expected quality–cutoffs. In particular, each hospital j admits any physician i with $\hat{\theta}_i$ above their cutoff $\bar{\theta}_j$, while each physician i chooses the hospital j that maximizes her indirect utility among those that would admit her. Mathematically, i matches with j, i.e., $\mu(i) = j$, if and only if

$$\mu(i) = \arg\max_{j \in \mathcal{J}_i} u_{ij},$$

where $\mathcal{J}_i = \{j \in \mathcal{J} : \hat{\theta}_i \geq \bar{\theta}_j\} \cup \{\phi\}$.

Given our distributional assumptions on physicians' unobserved preferences, the conditional matching probability for a physician i matching with hospital j can be written as:

$$CMP^m\big(\mu(i) = j\big) \;=\; \frac{e^{\tilde{v}\left(x_{ij}, \hat{\theta}(s_i, \tau_i)\right)} \, \mathbf{1}\{\hat{\theta}(s_i, \tau_i) \geq \underline{\hat{\theta}}_j\}}{1 \;+\; \sum_{j'} e^{\tilde{v}\left(x_{ij'}, \hat{\theta}(s_i, \tau_i)\right)} \, \mathbf{1}\{\hat{\theta}(s_i, \tau_i) \geq \underline{\hat{\theta}}_{j'}\}}, \tag{A.7}$$

where the indicator $\mathbf{1}\{\hat{\theta}(s_i, \tau_i) \geq \underline{\hat{\theta}}_j\}$ enforces that we only consider hospitals that are ex-post feasible for physician i (Fack et al., 2019), and the equilibrium cutoffs are determined by the following system of demand and supply equations:

$$\kappa_j \geq M_j\left(\underline{\hat{\theta}}\right) \forall j \in \mathcal{J}, \tag{A.8}$$

where

$$M_j\left(\underline{\hat{\theta}}\right) = \sum_{\tau} m_\tau \int_{\underline{\hat{\theta}}_j}^{\bar{\hat{\theta}}} \int_{\mathcal{X}} \frac{e^{\tilde{v}(X, y|j)} \, \mathbf{1}\{y \geq \underline{\hat{\theta}}_j\}}{1 \;+\; \sum_{j'} e^{\tilde{v}(X, y|j')} \, \mathbf{1}\{y \geq \underline{\hat{\theta}}_{j'}\}} h\left(X, y|\tau\right) dX dy \tag{A.9}$$

with $\underline{\hat{\theta}}$ denoting the vector of posterior quality cutoffs, $h\left(X, y|\tau\right)$ denoting the joint density over observable characteristics X and posterior quality y given type τ, and $\bar{\hat{\theta}}$ denotes the upper bound in the domain of posterior mean qualities.

The previous equations make explicit the dependency of equilibrium cutoffs on hospitals' vacancies, and the mass and joint distribution of physicians' preferences and posterior qualities.

Approximation of CMPs. We now provide conditions under which our reduced form CMPs can be seen as approximations of the micro-founded CMPs. The reduced-form CMPs for this simplified model, specified in Equation (10), would be given by:

$$CMP^r\left(\mu(i)=j\right) \;=\; \frac{e^{v\left(x_{ij},\hat{\theta}(s_i,\tau_i)\right)+g(M_i,\kappa|\tau_i,j)}}{1+\sum_{j'} e^{v\left(x_{ij'},\hat{\theta}(s_i,\tau_i)\right)+g(M_i,\kappa|\tau_i,j')}}. \tag{A.10}$$

Therefore, the reduced-form CMPs given in Equation (A.10) can approximate the micro-founded CMPs from Equation (A.7) if:

$$e^{\Delta_v\left(x_{ij},\hat{\theta}(s_i,\tau_i)\right)+g(M_i,\kappa|\tau_i,j)} \;\approx\; \mathbf{1}\{\hat{\theta}(s_i,\tau_i)\geq \underline{\hat{\theta}}_j\} \quad \forall j\in\mathcal{J},$$

where $\Delta_v\left(x_{ij},\hat{\theta}(s_i,\tau_i)\right)\equiv v\left(x_{ij},\hat{\theta}(s_i,\tau_i)\right)-\tilde{v}\left(x_{ij},\hat{\theta}(s_i,\tau_i)\right)$.[42]

To establish under which conditions Equation (H) might hold, notice that the indicator function can be smoothed out by using the Gompertz function,

$$e^{-\,e^{-\lambda\left(\hat{\theta}(s_i,\tau_i)-\underline{\hat{\theta}}_j\right)}} \approx \mathbf{1}\{\hat{\theta}(s_i,\tau_i)\geq \underline{\hat{\theta}}_j\} \quad \forall j\in\mathcal{J},$$

which converges to an indicator as $\lambda\to\infty$. Thus, our reduced-form CMPs can approximate their micro-founded counterparts if

$$\Delta_v\left(x_{ij},\hat{\theta}(s_i,\tau_i)\right)+g(M_i,\kappa|\tau_i,j)\approx e^{-\lambda\left(\hat{\theta}(s_i,\tau_i)-\underline{\hat{\theta}}_j\right)}. \tag{A.11}$$

Equation (A.11) suggests that for our reduced-form CMPs to be a reasonable approximation to their micro-founded counterparts, the function $g(\cdot)$ must capture the determinants of the equilibrium cutoffs $\underline{\hat{\theta}}$. As shown in Equations (A.8) and (A.9), these equilibrium cutoffs are the solutions to a system of demand–supply equations that depend on the distribution of preferences, the masses of physicians, and hospitals' capacities. In principle, changes in the mass of physicians with different preferences and qualities can affect these cutoffs in counterfactual analyses. However, under the assumption of vertical sorting, it is straightforward to see that the (ex-post feasible) choice set—and, consequently, the matching probabilities—of a physician with posterior quality $\hat{\theta}_i$ cannot be affected by a change in the mass of physicians with qualities below $\hat{\theta}_i$; otherwise, a blocking pair would emerge.

To see this more concretely, consider a large matching market with a finite set of agents. When the stable equilibrium is unique, the equilibrium allocation can be obtained by performing the Serial Dictatorship (SD) mechanism (see, e.g., Abdulkadiroğlu and Sönmez (1998)). Under SD, physicians are ordered in descending order by their posterior quality $\hat{\theta}_i$, and in that order, each physician chooses her most preferred hospital from those with remaining vacancies. Because hospitals have vertical preferences, a physician's choice is determined solely by her own quality relative to the hospital's cutoff. Crucially, when physician i makes her choice, the set of hospitals that have not yet been filled is determined only by the competition among physicians with quality higher than $\hat{\theta}_i$. Therefore, any change in the mass of physicians with lower quality than $\hat{\theta}_i$ does not impact the set of hospitals available to i; her allocation and conditional matching probability remain invariant.

[42]Notice that because $v\left(x_{ij},\hat{\theta}(s_i,\tau_i)\right) = \Delta_v\left(x_{ij},\hat{\theta}(s_i,\tau_i)\right)+\tilde{v}\left(x_{ij},\hat{\theta}(s_i,\tau_i)\right)$, in our reduced-form CMPs we will not be able to separately identify how much of the determinants of matching probabilities (v) are affected by physicians' preferences ($\tilde{v}$), versus the effect of their observable characteristics on their choice-sets (Δ_v).

Using this result from our micro-founded model, we specify the function $g(\cdot)$ to depend on programs' vacancies κ and on the mass of physicians with quality above i's posterior quality, denoted by M_i.

I Computation of numerical elasticities

To compute the elasticity of labor and quality, we simulate $R = 100$ draws for the EV1 shocks in the estimated CMPs, and compute the resulting matches for physicians who take the test in year t and have a score above the corresponding threshold. If $\Delta^r L_{jt}(\underline{\mathrm{s}})$ physicians of average quality $\Delta^r \theta_{jt}(\underline{\mathrm{s}})$ match with hospital j in simulation r; we compute the resulting quantity and quality of each hospital by adding the marginal physicians that match with hospital j in year t to the stock of physicians of the hospital j in year $t-1$. That is, the resulting simulated physicians is $L^r_{jt}(\underline{\mathrm{s}}) = L_{jt-1} + \Delta^r L_{jt}(\underline{\mathrm{s}})$ and the resulting simulated average quality is $\bar{\theta}^r_{jt}(\underline{\mathrm{s}}) = \frac{(\bar{\theta}_{jt-1} L_{jt-1} + \Delta^r L_{jt} \Delta^r \theta_{jt}(\underline{\mathrm{s}}))}{L^r_{jt}(\underline{\mathrm{s}})}$.

We compute the quantity elasticity numerically as

$$\eta^r_{L_{jt},\underline{s}} \approx -\frac{\left(L^r_{jt}(\underline{s} + \Delta\underline{\mathrm{s}}) - L^r_{jt}(\underline{\mathrm{s}})\right)}{\Delta\underline{\mathrm{s}}} \cdot \frac{\underline{s}}{L^r_{jt}(\underline{\mathrm{s}})}$$

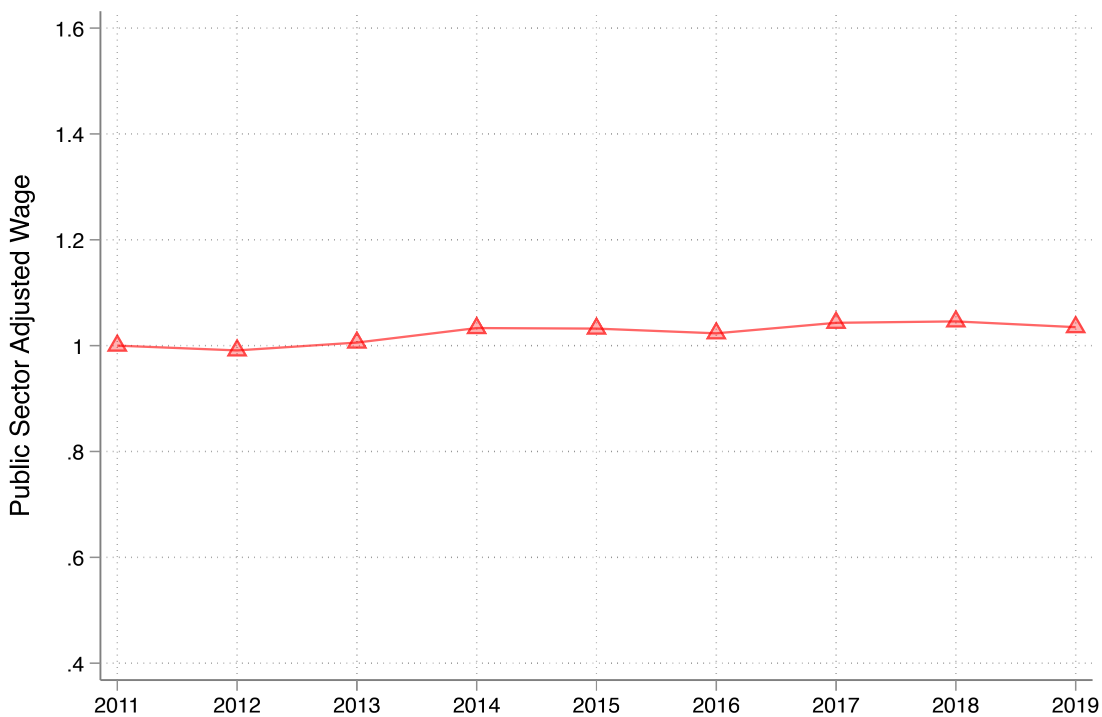
and the quality semi-elasticity as

$$\tilde{\eta}^r_{\bar{\theta}_{jt},\underline{s}} \approx -\frac{\left(\bar{\theta}^r_{jt}(\underline{s} + \Delta\underline{\mathrm{s}}) - \bar{\theta}^r_{jt}(\underline{\mathrm{s}})\right)}{\Delta\underline{\mathrm{s}}} \cdot \underline{s}$$

.

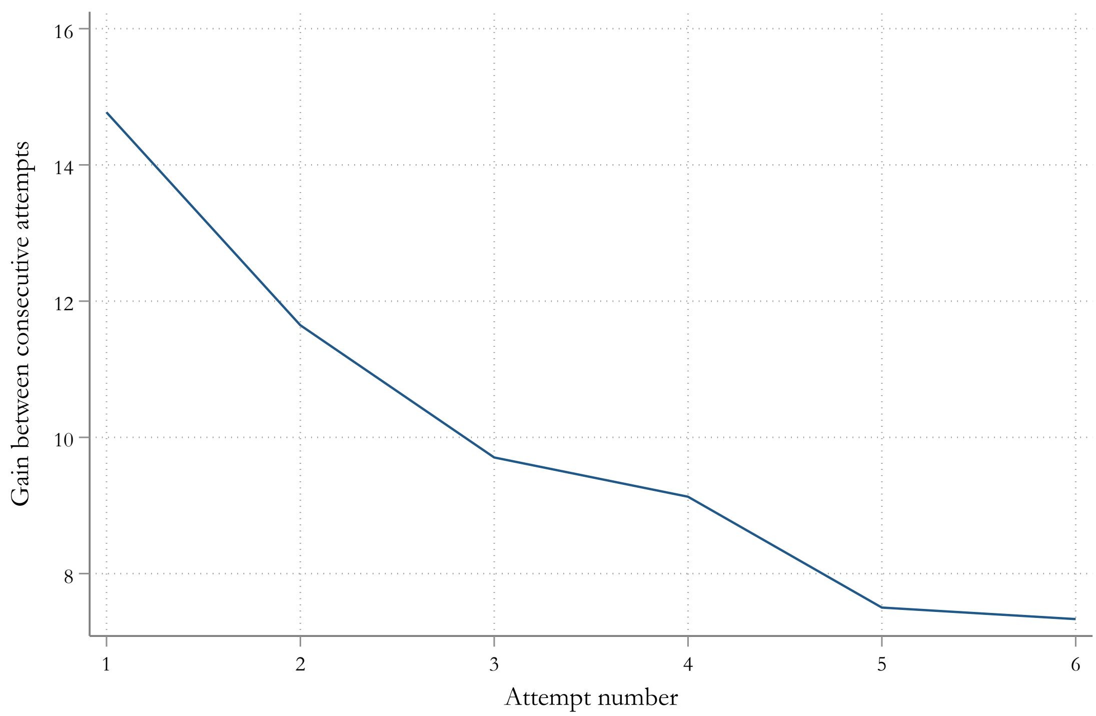
We then compute the estimated elasticities by averaging out across simulations.

Figure A.1: Physician's Wages in Public Hospitals



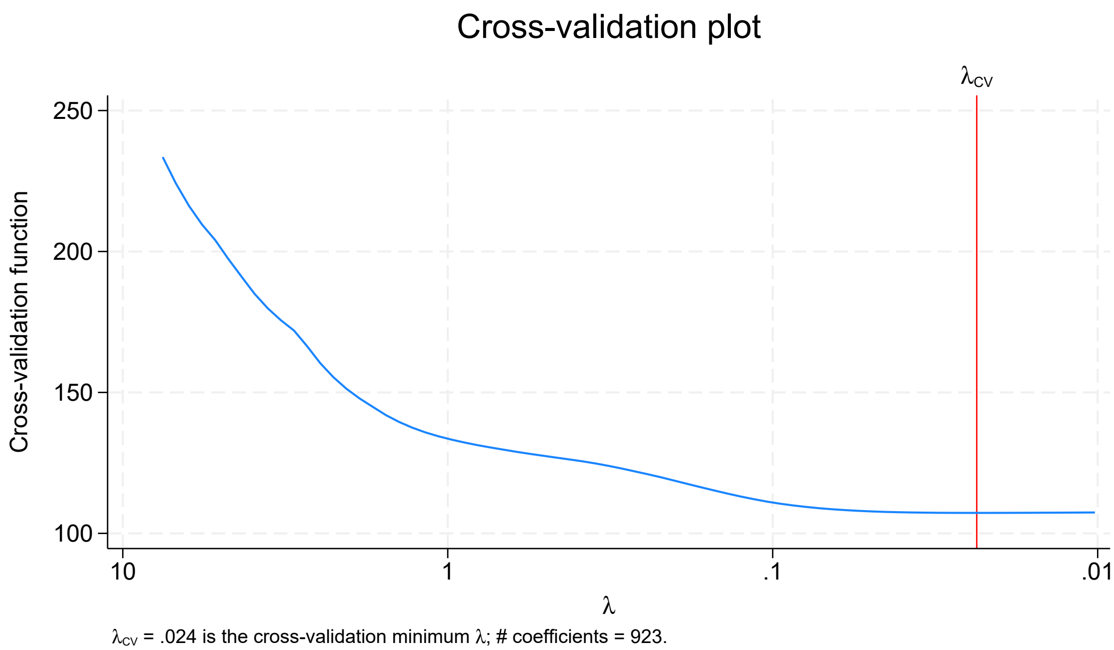
Notes: This figure illustrates the growth of physician wages in public hospitals. Wages are adjusted using the public sector-wide adjustment and indexed to 2011, which is set to 1. The data are sourced from SIRH (2019).

Figure A.2: Score Gains Over Attempts

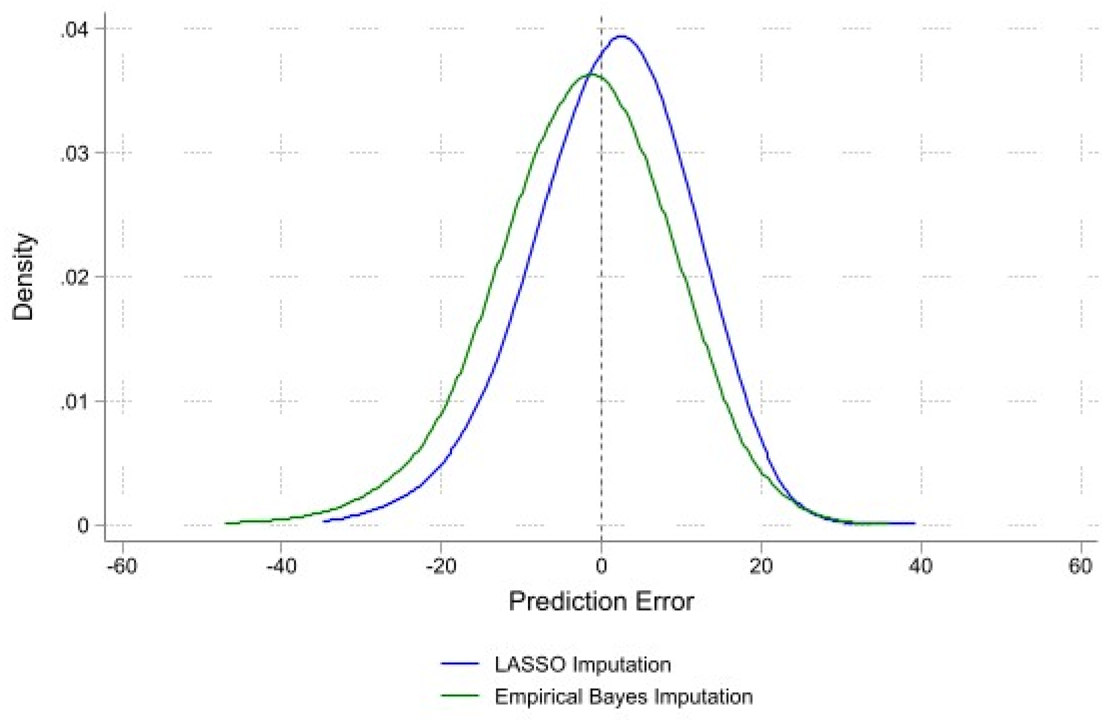


Notes: This figure presents the score gains after consecutive test-taking attempts by the same individual. It shows that on average score gains diminish with each additional attempt. Data come from ASOFAMECH (2019) and the figure includes the universe of test-takers between 2009 and 2018.

Figure A.3: Alternative Imputation Methods



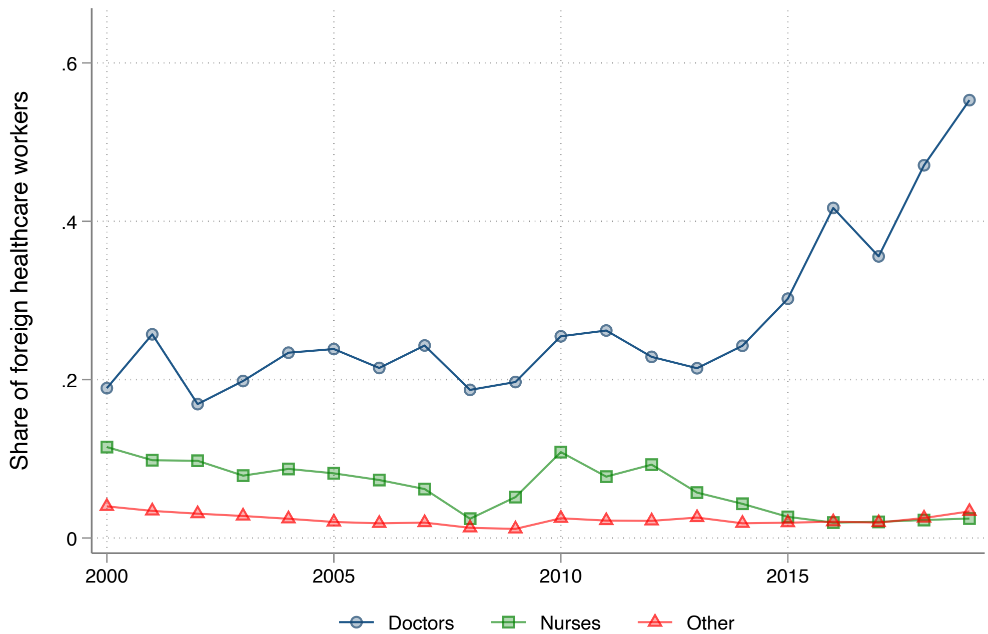
A. Cross-validation



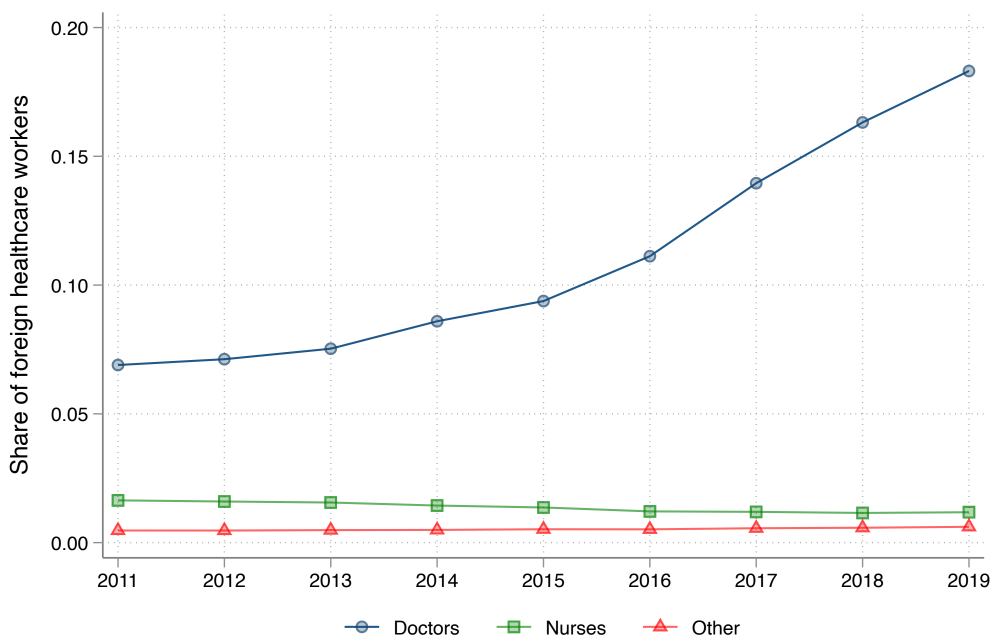
B. Prediction Error

Notes: This figure shows statistics related to the imputation methods. Panel A shows the mean cross-validation (CV) error against the regularization parameter λ. Panel B shows the density of prediction errors from the LASSO and the Empirical Bayes imputation methods.

Figure A.4: Migration Wave: Other Labor Inputs



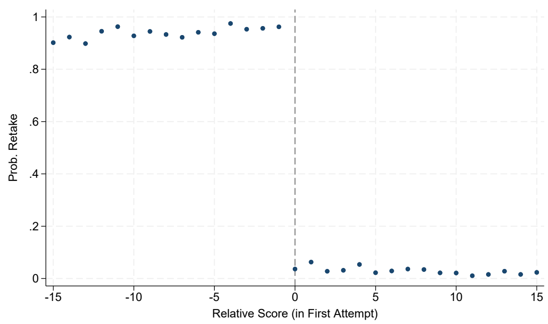
A. Newly Registered Providers



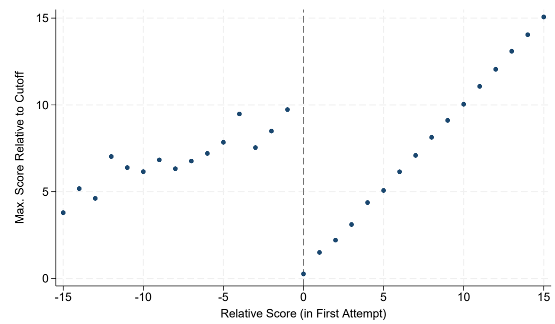
B. Stock in Public Hospitals

Notes: This figure shows the evolution of the share of foreign-trained healthcare providers in Chile. Panel A examines trends from 2000 to 2019 using data from the National Registry of Healthcare Providers (RNPI, 2024). Panel B focuses on healthcare providers working in public hospitals from 2011 to 2019, based on data from SIRH (2019).

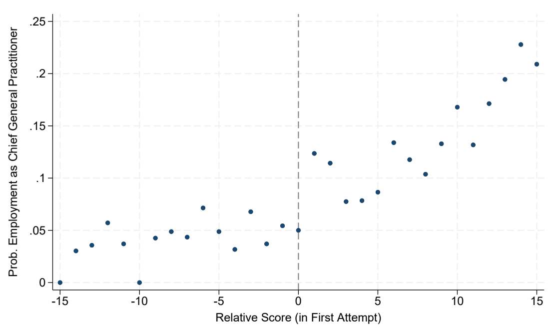
Figure A.5: Testing for Quality Gains



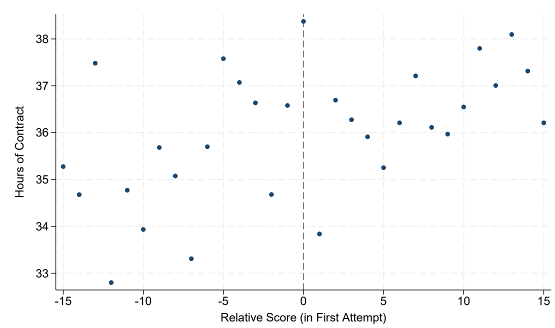
A. Retaking Probability



B. Maximum achieved score



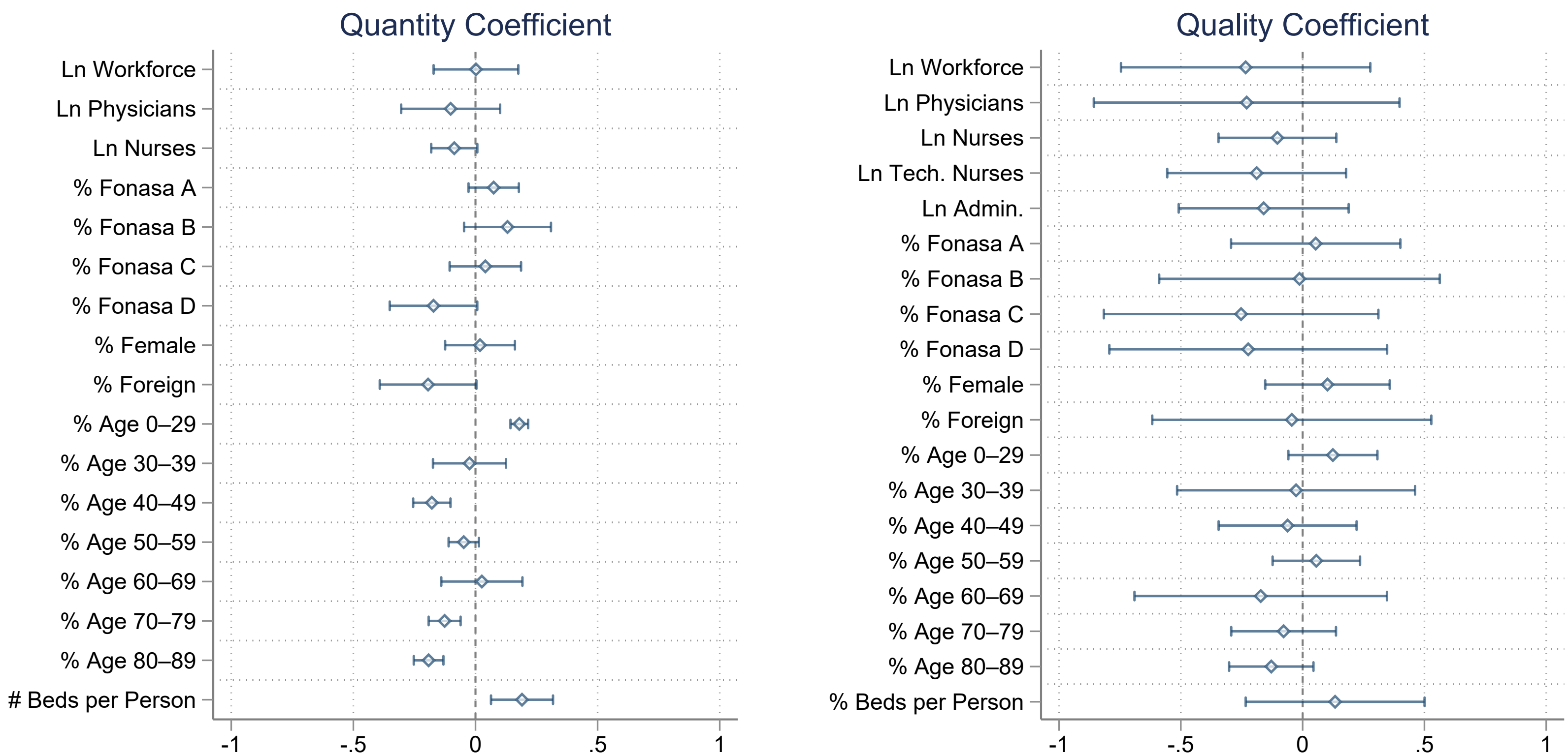
C. Appointed Chief General Practitioner



D. Hours of Contract

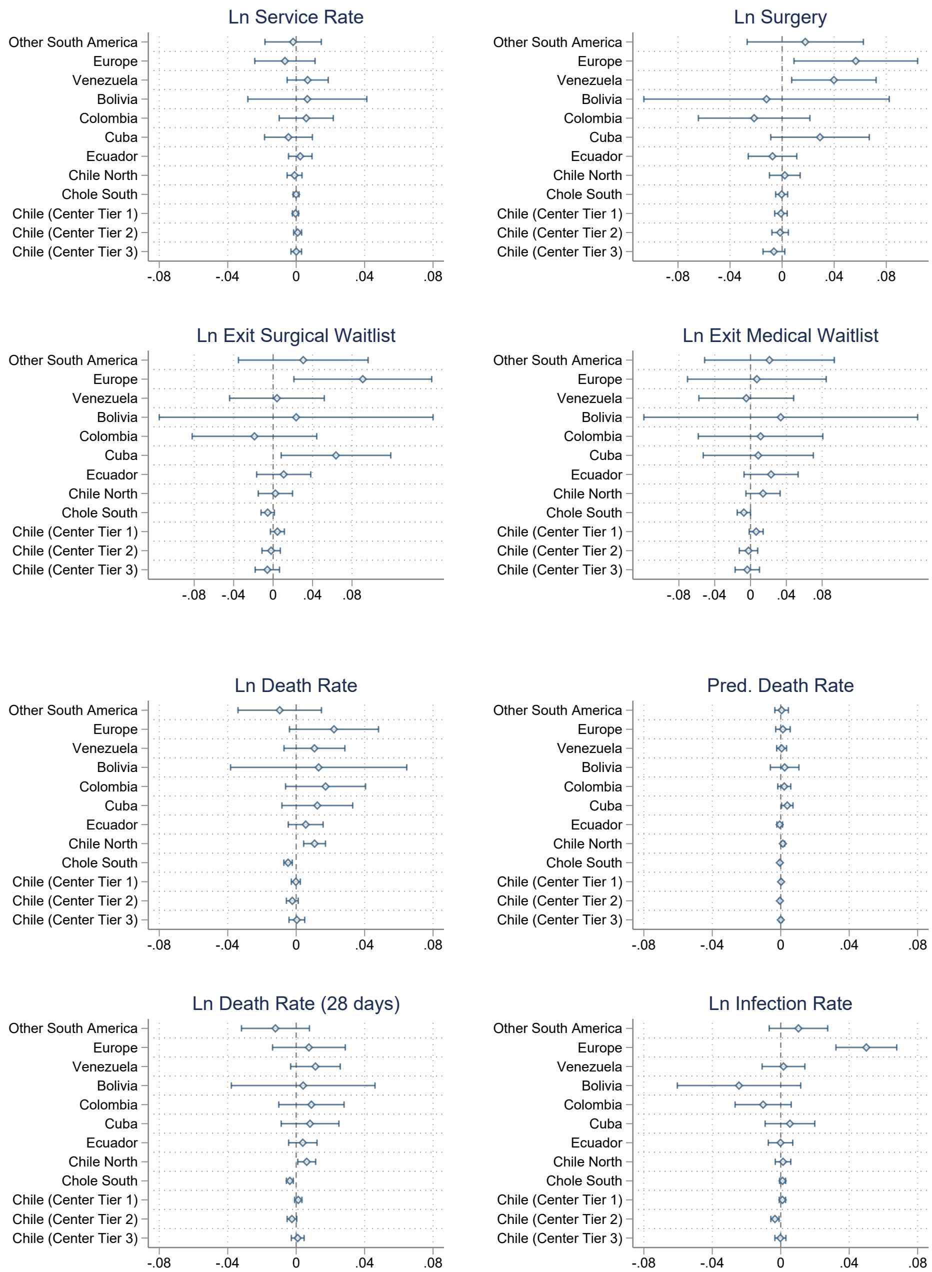
Notes: This figure illustrates the effect of exam retaking on physician-quality improvements. Panel A depicts the probability of individuals retaking the exam based on whether they scored above the passing threshold on their first attempt. Panel B presents the maximum additional points individuals gained across all their retake attempts, relative to the passing threshold. The line on the right closely follows a 45-degree trajectory, consistent with the fact that individuals who pass the exam do not retake it. In contrast, the line on the left indicates that individuals just below the threshold improve their scores by approximately 10 points on average, with smaller improvements observed for those further below the threshold. Panel C examines the probability of becoming Chief General Practitioner as a proxy for physician quality. There is no visible discontinuity at the threshold on their first test-taking attempt, which is consistent with the absence of quality gains from repeated test-taking. Panel D use weekly working hours as an additional quality measure and similarly there is no visible effect at the threshold. Data on scores come from ASOFAMECH (2019). Panels A and B use the full sample of test-takers, while Panels C and D incorporate data from SIRH (2019), restricting the sample to physicians observed working in public hospitals.

Figure A.6: Shocks Balance Test



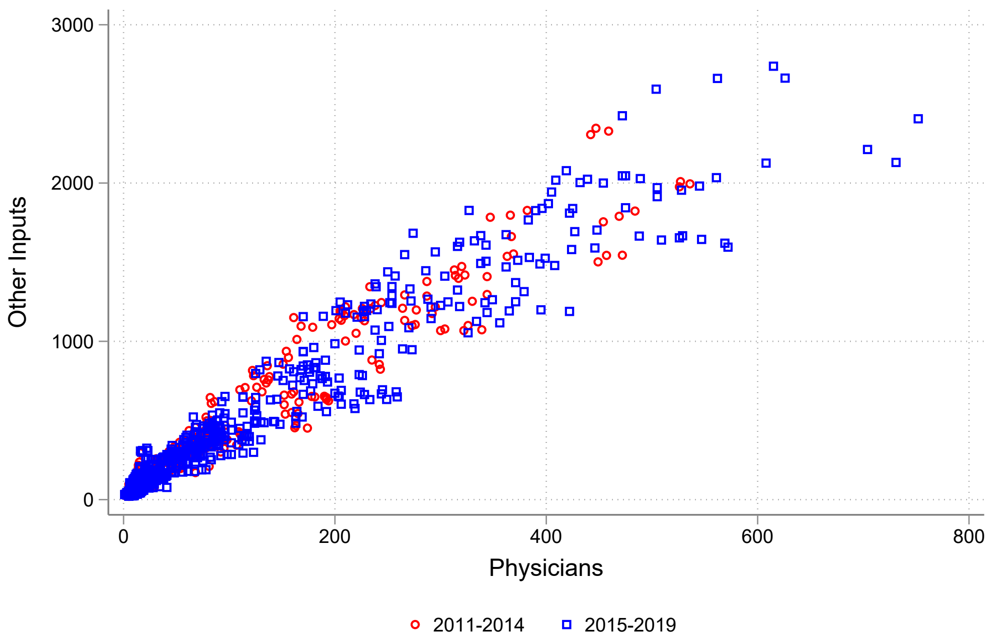
Notes: These figures assess the plausibility of the conditional quasi-random assignment of shocks. Panels A and B show the point estimate and confidence interval obtained from separate regressions of predetermined variables (as of 2012) on the quantity and quality shocks, respectively. All variables are standardized to have a mean of zero and a standard deviation of one.

Figure A.7: Share Balance Test

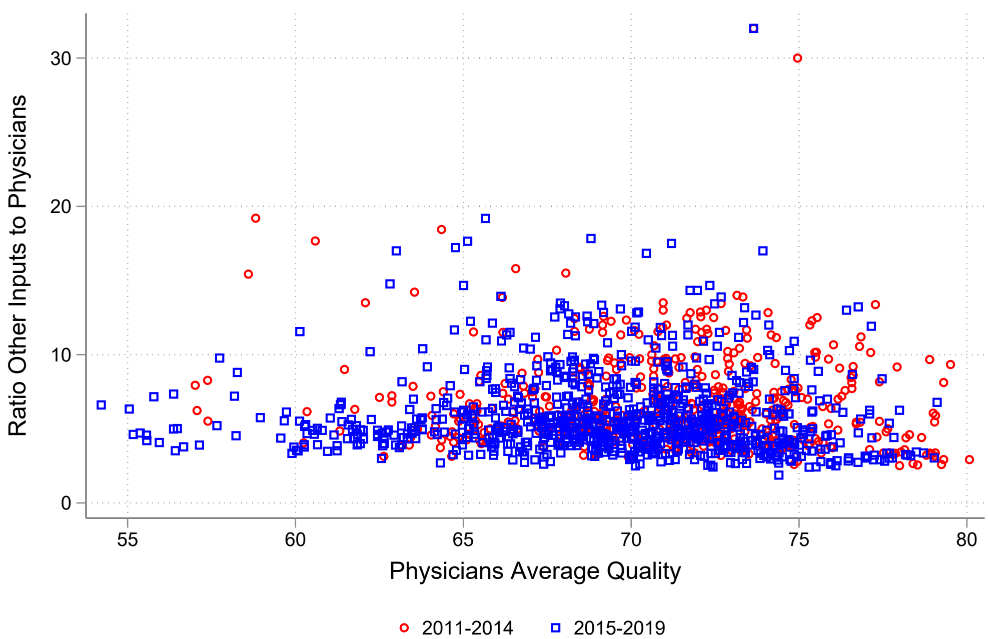


Notes: These figures show the results from regressing the outcome of interest against the region of origin-shares during the (pre) period 2012-2015, controlling by hospital and year fixed effects as well our main battery of control variables.

Figure A.8: Other Hospital Healthcare Workers



A. Ratio of other inputs by physicians



B. Mix by physician quality

Notes: This figures XXXX. Panel A... Panel B...

Table A.1: Descriptive Statistics

Panel A: Health Outcomes

	Mean	Std. Dev.	Median (p50)	# of Obs.
	(1)	(2)	(3)	(4)
Patient Characteristics:				
% Female	0.57	0.08	0.58	1402
% Age $<$ 29	0.30	0.15	0.31	1402
% Age $\in$ (30,29)	0.10	0.04	0.10	1402
% Age $\in$ (40,49)	0.09	0.03	0.09	1402
% Age $\in$ (50,59)	0.11	0.03	0.11	1402
% Age $\in$ (60,69)	0.12	0.04	0.12	1402
% Age $\in$ (70,79)	0.14	0.06	0.13	1402
% Age $\in$ (80,89)	0.11	0.06	0.10	1402
% Age $>$ 89	0.03	0.02	0.02	1402
% Public Insurance	0.97	0.04	0.98	1402
Hospital Characteristics:				
In-hospital Death Rate	3.28	1.82	2.92	1,402
28-day Death Rate	5.07	2.71	4.51	1,402
Patients (# Admissions)	5,656	7,686	1,964	1,402
Service Rate (# Admissions/Beneficiaries)	0.02	0.02	0.01	1,402
Average Length of Stay	4.03	5.66	3.00	1,402
Complication Rate	11.41	4.25	11.05	1,402
Number of Surgeries	2,018	3,332	6.00	1,402
Physicians	77.64	119.64	20.00	1,402

Panel B: Test Scores

Year	Number Tests Takers	Average score	% Approved (score $\geq$ 51)	Average score if score $\geq$ 51	# Tests $\in [40-51)$
2009	1,389	71.8	92	74.3	87
2010	1,535	65.1	80	72.1	142
2011	1,748	66.6	81	73.3	160
2013	2,003	56.1	66	67.5	231
2014	2,557	55.8	65	67.5	335
2015	3,641	54.7	60	66.5	651
2016	4,999	53.0	54	66.9	1,012
2017	6,014	52.1	55	64.9	1,233
2018	7,121	53.9	58	65.0	1,552

Notes: This table presents descriptive statistics for the data used in the primary analysis. Panel A summarizes patient and hospital characteristics for all public hospitals included in the analysis. These data are derived from individual-level inpatient records reported by DEIS (2019) and restricted administrative records on public hospital employees from SIRH (2019). Panel B provides statistics on EUNACOM scores from 2009 to 2018. For cases where the exam was taken twice in the same year, all data are pooled. The dataset includes all records of test takers and comes from ASOFAMECH (2019).

Table A.2: Impact of Physicians' Quantity and Quality: LASSO Imputation

Panel A: Access

	Ln service rate	Ln inpatient surgeries	Ln exits from waiting list	
			Surgical	Medical
	(1)	(2)	(3)	(4)
Ln Physicians ($\hat{\alpha}_L$)	1.06	6.90	3.62	3.24
	(0.40)	(3.54)	(3.05)	(2.20)
Avg. Physicians' Quality ($\hat{\alpha}_\theta$)	0.01	0.16	-0.01	0.02
	(0.02)	(0.15)	(0.12)	(0.09)
Observations	1,402	744	738	942
Mean Dep. Var.	0.015	3,803	1,534	8,403
F-stat (First-stage)	11.18	3.544	2.979	5.978
Anderson-Rubin (χ^2)	0.00	0.00	0.00	0.00

Panel B: Quality

	Mortality			In-hospital
	In-Hospital		28-days	Complications
	Ln death rate	Pred. death rate	Ln death rate	Ln complications rate
	(1)	(2)	(3)	(4)
Ln Physicians ($\hat{\alpha}_L$)	-1.21	0.10	-1.11	-1.44
	(0.47)	(0.11)	(0.42)	(0.43)
Avg. Physicians' Quality ($\hat{\alpha}_\theta$)	-0.05	-0.00	-0.04	-0.05
	(0.02)	(0.01)	(0.02)	(0.02)
Observations	1,402	1,402	1,402	1,373
Mean Dep. var.	3.28	3.51	5.08	11.65
F-stat (First-stage)	11.18	19.22	11.18	10.60
Anderson-Rubin (χ^2) p-value	0.00	0.03	0.00	0.00

Notes: This table presents the impact of the quantity and quality of physicians on public hospital performance. Panel A focuses on utilization, which we proxy through the service rate, inpatient surgeries, and exist from the waiting list. Panel B focuses on patients' mortality and complications. Estimates come from the two-stage least squares estimation of Equation (17), and we use a quality measure based on the scores imputation using LASSO.

Table A.3: Shock Summary Statistics

	Quantity Shock		Quality Shock	
	(1)	(2)	(3)	(4)
Mean	3.0	0.0	-0.5	0.0
Standard deviation	7.6	7.5	2.2	1.3
Interquartile range	1.8	1.2	3.9	1.3
Residualizing on period FE	No	Yes	No	Yes
Effective sample size ($1/HHI$ of s_{nt} weights)	28.83	28.83	28.83	28.83
Largest s_{nt} weight	0.06	0.06	0.06	0.06
No. of regions-period shocks	126	126	126	126
No. of regions	16	16	16	16

Notes: This table summarizes the distribution of migration shocks across regions of origin and periods. Quantity shocks are measured as the percentage change in the number of physicians from each region of origin clearing the cutoff of the licensing exam. Similarly, quality shocks are measured as the change in the average quality of eligible test-takers from different regions of origin. All statistics are weighted by the region of origin exposure shares. Columns (1) and (2) consider the quantity shocks and columns (3) and (4) consider the quality shocks. Columns (2) and (4) residualize the migration shocks on period indicators. As in Borusyak et al. (2022), we also report the effective sample size (the inverse re-normalized Herfindahl index of the weights) as well as the largest shares.

Table A.4: Rotemberg Weights and Pre-trends

	Region of Origin										
Other S. America	Europe	Venezuela	Bolivia	Colombia	Cuba	Ecuador	Chile (north)	Chile (south)	Chile (tier 1)	Chile (tier 2)	Chile (tier 3)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Ln Service Rate											
-0.002	-0.007	0.007	0.007	0.006	-0.004	0.002	-0.001	-0.000	-0.000	0.001	0.000
(0.008)	(0.009)	(0.006)	(0.018)	(0.008)	(0.007)	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)
Ln Surgery											
0.018	0.057	0.040	-0.012	-0.021	0.029	-0.007	0.002	-0.000	-0.001	-0.002	-0.006
(0.023)	(0.024)	(0.017)	(0.048)	(0.022)	(0.019)	(0.009)	(0.006)	(0.002)	(0.002)	(0.003)	(0.004)
Ln Exit Surgical Waitist											
0.031	0.091	0.004	0.023	-0.019	0.064	0.011	0.002	-0.005	0.004	-0.002	-0.006
(0.033)	(0.036)	(0.024)	(0.071)	(0.032)	(0.028)	(0.014)	(0.009)	(0.003)	(0.004)	(0.005)	(0.006)
Ln Exit Medical Waitist											
0.021	0.007	-0.005	0.034	0.011	0.009	0.023	0.014	-0.007	0.006	-0.002	-0.004
(0.037)	(0.039)	(0.027)	(0.078)	(0.035)	(0.031)	(0.015)	(0.010)	(0.004)	(0.004)	(0.005)	(0.007)
Ln Death Rate											
-0.010	0.022	0.011	0.013	0.017	0.012	0.006	0.011	-0.005	-0.000	-0.002	0.000
(0.012)	(0.013)	(0.009)	(0.026)	(0.012)	(0.011)	(0.005)	(0.003)	(0.001)	(0.001)	(0.002)	(0.002)
Ln Death Rate (28-days)											
-0.012	0.007	0.011	0.004	0.009	0.008	0.004	0.006	-0.004	0.001	-0.002	0.001
(0.010)	(0.011)	(0.007)	(0.021)	(0.010)	(0.009)	(0.004)	(0.003)	(0.001)	(0.001)	(0.001)	(0.002)
Ln Infection Rate											
0.010	0.050	0.002	-0.024	-0.010	0.005	-0.000	0.001	0.001	0.001	-0.003	-0.000
(0.009)	(0.009)	(0.006)	(0.018)	(0.008)	(0.007)	(0.004)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)

Notes: This table presents robustness checks in the spirit of Goldsmith-Pinkham et al. (2020). Panel A shows the Rotemberg weights associated to each region of origin. Panel B assesses pre-trends by regressing each of the outcomes of interest against the region of origin shares in the (pre) period 2012-2015. Point estimates reflect the differential effect of region-specific shares on the dependent variable.

Table A.5: Impact of Physicians' Quantity and Quality: Translog Production Function

	Ln Death Rate		**Ln Service Rate**	
	(1)	(2)	(3)	(4)
Ln Physicians ($\hat{\alpha}_L$)	-0.83	0.43	1.01	1.08
	(0.31)	(1.64)	(0.29)	(1.34)
Avg. Physicians' Quality ($\hat{\alpha}_\theta$)	-0.04	0.07	0.01	0.01
	(0.02)	(0.15)	(0.02)	(0.12)
Interaction ($\hat{\alpha}_{L\theta}$)		-0.02		-0.00
		(0.03)		(0.03)
Observations	1,402	1,402	1,402	1,402
Model	2SLS	2SLS	2SLS	2SLS
Year FE	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes
Mean dep var	3.28	3.28	0.015	0.015
First-stage F-stat	22	2.332	22	2.332
Translog Quantity and Quality Impacts:				
Quantity Impact		-1.280		0.987
		(0.460)		(0.309)
Quality Impact		-0.010		0.008
		(0.002)		(0.001)

Notes: This table presents the impact of the quantity and quality of physicians on public hospital performance. Columns (1) and (2) focus on utilization, which we proxy through the service. Columns (3) and (4) focus on health outcomes, which we proxy with in-hospital deaths. Estimates in columns (1) and (3) come from the two-stage least squares estimation of Equation (17). Estimates in columns (2) and (4) come from alternative models that include the interaction between physicians' quantity and quality as an additional endogenous variable and the interaction of our shift-share instruments as an additional instrumental variable.

Table A.6: Impact of Physicians' Quantity and Quality: Alternative Quality Index

Panel A: Access

	Ln service rate	Ln inpatient surgeries	Ln exits from waiting list	
			Surgical	Medical
	(1)	(2)	(3)	(4)
Ln Physicians ($\hat{\alpha}_L$)	0.98	4.36	3.71	2.94
	(0.25)	(1.29)	(1.32)	(1.19)
% Low Quality Physicians ($\hat{\alpha}_\theta$)	-0.05	-1.00	0.05	-0.17
	(0.18)	(0.80)	(0.78)	(0.82)
Observations	1,376	740	736	934
Mean Dep. Var.	0.0155	3,819	1,537	8,467
F-stat (First-stage)	16.25	10.33	9.29	9.96
Anderson-Rubin (χ^2)	0.00	0.00	0.00	0.00

Panel B: Quality

	Mortality			In-hospital
	In-Hospital		28-days	Complications
	Ln death rate	Pred. death rate	Ln death rate	Ln complications rate
	(1)	(2)	(3)	(4)
Ln Physicians ($\hat{\alpha}_L$)	-0.68	0.13	-0.61	-0.45
	(0.28)	(0.07)	(0.25)	(0.27)
% Low Quality Physicians ($\hat{\alpha}_\theta$)	0.49	0.03	0.48	0.48
	(0.20)	(0.06)	(0.18)	(0.19)
Observations	1,376	1,376	1,376	1,376
Mean Dep. var.	3.30	3.50	5.09	11.65
F-stat (First-stage)	16.25	21.57	16.25	15.46
Anderson-Rubin (χ^2) p-value	0.00	0.03	0.00	0.00

Notes: This table presents the impact of the quantity and quality of physicians on public hospital performance. Panel A focuses on utilization, which we proxy through the service rate, inpatient surgeries, and exist from the waiting list. Panel B focuses on patients' mortality and complications. Estimates come from the two-stage least squares estimation of Equation (17), where we use the share of physicians in each hospital and time with quality below the median of the entire distribution as quality index. For this exercise, we only consider physicians with EUNACOM scores.