Expertise Overlap and Team Productivity: Evidence from the Hospital Industry*

Danyelle Branco Bladimir Carrillo Di Fang Wilman Iglesias

April 5, 2023

Abstract

Many organizations use teamwork in the workplace. A longstanding question in economics is whether teams should consist of heterogeneous specialists or members with similar expertise. Diversity in specialized skills may generate productive complementarities, but it also may increase the costs of coordination and complicate teamwork. We study this question in the context of a common heart procedure with high mortality and significant spending, where doctors are assigned to teams in a quasi-random fashion. Using a clean research design and novel data from Brazil, we find that teams of doctors with shared expertise achieve lower patient mortality rates while using fewer medical inputs. These effects are larger when team members have less accumulated experience working together and perform more complex tasks, consistent with an improved coordination channel. These findings suggest that, at least in a high-pressure environment, the coordination costs of diversity in specialized skills are large enough to outweigh any beneficial complementarities. Counterfactual policy simulations suggest considerable efficiency gains from reallocating certain physicians across regular workdays.

JEL codes: I10, I11, J24, M12, M54

Keywords: healthcare production, teams, doctors, expertise

*Contact information: Branco: Department of Economics, Universidade Federal de Pernambuco, Caruaru, PE, 55014-900 (e-mail: danyelle.branco@ufpe.br). Carrillo: Department of Economics, Universidade Federal de Pernambuco, Recife, PE, 50670-420 (e-mail: bladimir.carrillo@ufpe.br). Fang: University of Florida, Gainesville, FL 32611, United States (e-mail: difang@ufl.edu). Iglesias: University of Nebraska-Lincoln, Lincoln, NE 68588, United States (e-mail: wilman.iglesias@huskers.unl.edu). Financial support for this research from the Brazilian National Council for Scientific and Technological Development (CNPq —# 04/2021) is gratefully acknowledged. An earlier version of this paper was circulated under the title "Expertise Overlap, Team Productivity, and Patient Mortality: Evidence from Doctors in Brazil." We are especially grateful to Xiaoxi Zhao, the discussant at the American Economic Association Meeting for her insightful comments and suggestions. We also thank Ana Almeida, Daniel Araujo, Sonia Bhalotra, John Cawley, Yiqun Chen, Carlos Charris, Joseph Doyle, Fernanda Estevan, Bruno Ferman, Marcelo Fernandes, Emanuel Ornelas, Eduardo Pontual, Mark Rush, Breno Sampaio, Perihan Saygin, and participants at various conferences and seminars for helpful comments and suggestions. We thank Roberto Cardenas, Billie Maciunas, and Alexis Skinner for excellent research assistance. We are indebted to the staff of the Department of Regulation, Evaluation, and Control of Systems at the Brazilian Ministry of Health for assistance with the data. We are solely responsible for this paper's contents. The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

1 Introduction

Teamwork is a pervasive aspect of modern organizations. Physicians, judges, teachers, scientists, and many other types of workers perform tasks that often require collaboration with multiple peers. According to calculations by Bandiera et al. (2013), more than 50 percent of firms in the United States employ some type of teamwork, and Lazear and Shaw (2007) document that this phenomenon is becoming increasingly common across both public and private spheres. Understanding the production process of teams is therefore important for understanding both how to make firms more productive and how to improve state effectiveness in areas as essential as education, justice, and health care.

This paper empirically studies the implications of variation in expertise across team members for performance. The question of whether teams should be composed of heterogenous specialists or members with similar expertise dates back to at least Adam Smith (1776). In his most celebrated Book, Wealth of Nations, Adam Smith (1776) stresses that diversity in specialized knowledge creates complementarities between workers that are beneficial to the production process, noting that "*the most dissimilar geniuses are of use to one another*." Closer to our times, Arrow (1974) argues that a diverse set of perspectives and expertise is essential for problem solving and effective decision-making. Under this line of thinking, more heterogeneity in productive skills is better than less. Notwithstanding this powerful intuition, subsequent literature has emphasized the costs of teamwork among heterogeneous workers. It is often argued that coordination in teams with diverse specialists is complicated because different experts may diverge when interpreting the same problems, own different communication styles, and have distinct approaches to perform similar tasks. Becker and Murphy (1992)'s landmark paper on specialization and division of labor articulates this point formally and asserts that:

> *"[T]he chances of a breakdown in production due to poor coordination of the tasks. . . , or to communication of misleading information among members, also tends to expand as the number of separate specialists grows"* (p. 1138).

This tradeoff between the productive benefits and costs of variety in expertise is succinctly summarized in the illuminating paper of Lazear (1999) who argues that, to offset the coordination costs of diversity, *"there must be complementarities between the workers that are sufficiently important."*

Despite their long history in economics, little careful research has investigated these ideas empirically. The few existing studies are at the level of broad correlations and thus subject to the standard challenges of selection and omitted variable bias (e.g., Bunderson and Sutcliffe, 2002; Hoisl et al., 2017). A series of studies using randomized experiments provide compelling and clean evidence that differences between team members in basic demographic characteristics, including ethnicity and gender, matter for team performance (Hjort, 2014; Lyons, 2017; Marx et al., 2021). But because these studies do not rely on measures of expertise or skills *per se*, they do not examine whether team members with different areas of knowledge perform better or worse than more homogenous teams. As emphasized by Hjort (2014), differences in performance between teams that are more and less diverse in dimensions like ethnicity may be driven by a "taste for discrimination" mechanism and not necessarily by differences in specialized skills or expertise affecting the coordination of interdependent tasks.

The core contribution of this paper is to provide the first rigorous analysis of how overlap in team members' expertise affects team performance. We study this question in the context of the hospital industry, paying close attention to the structure, nature, and functioning of heart teams. We use a clean research design and exceptionally high-quality data to show that teams of doctors with similar expertise achieve better patient outcomes and this appears to occur primarily through improved team coordination.

We study teamwork during percutaneous coronary interventions (PCI), an expensive treatment performed in patients with acute heart diseases and whose post-procedure mortality rate is among the highest.[1] This treatment is performed by a well-defined team consisting of one proceduralist who executes the procedure and one or several physicians who provide pre- and post-procedure care during the patient's hospital stay. These doctors operate in a fast-paced and high-pressure environment where quick decision-making and effective collaboration between them are critical for patient outcomes (Mazzocco et al., 2009). We explore the outcomes of patients of assigned to teams where proceduralists and physicians have higher and lower degree of overlapping medical specialties, or hereafter, expertise overlap. Because doctors typically complete more than one specialty during the course of their careers, there is considerable variation in our measure of expertise overlap. This setting provides a natural laboratory to study this question for several reasons. Remarkably, since most patients undergoing PCI are those requiring urgent care and thus assigned to the doctors sequentially available, the scope for patients sorting into doctor teams is largely limited. Moreover, team performance can be accurately measured in terms of patient survival. From a policy perspective, this is a highly relevant context: the healthcare sector relies heavily on teamwork, and thus the performance of doctor teams is central to questions of efficiency.

Our analysis is organized into three parts.

In the first part, we present our benchmark estimates of the effect of team members' expertise overlap on patient mortality. To conduct this analysis, we introduce a new and uniquely rich compilation of administrative microdata on healthcare from Brazil. These data allow us to track doctors' careers with information on all specialties completed, observe each medical procedure that doctors provide to different patients, and examine the mortality outcomes of patients undergoing PCI. Data with this rich level of detail have been available only for the United States and a few European countries. The rising costs of healthcare are a major fiscal challenge in Brazil, similar to the United States and other affluent governments, but aggravated by the political, economic and capacity difficulties that characterize developing countries (Banerjee et al., 2021).[2] An important contribution of this paper is therefore to document new facts about the production of health care in arguably one of the most relevant yet understudied settings.

Our identification strategy exploits within-proceduralist variation by comparing emergency patients treated by the same proceduralist but different physicians.[3] Identification requires that the assignment of physicians to patients be determined in a quasi-random manner, even if the assignment of proceduralists is not. This is reasonable in our setting because of the unpredictability of emergency patient admissions and the predetermined

[1] In fact, acute heart disease is the leading cause of death globally. According to World Health Organization, approximately 17.9 million people die each year due to a cardiovascular disease (see , last accessed on December 21, 2022). In the United States, the expected costs of a patient undergoing PCI is $16000 (in 2014 dollars) with an aggregate cost of $10 billion each year, the highest aggregated cost of any cardiovascular procedure (Amin et al., 2017). In Brazil, the medical costs per patients are much lower at R$10000 (or US$1941 at an exchange rate of 5.15R$/US$) but still sizeable, accounting for approximately 50 percent of all public spending on cardiovascular diseases (Oliveira et al., 2022).

[2] In Brazil, healthcare spending amounted to R$710 billion in 2019 (or US$137 billion at an exchange rate of 5.15R$/US$) or 8 percent of GDP (Brazilian Ministry of Health, 2022). For comparison, this figure is 16 percent in the United States and 12 percent in OECD countries (see https://data.worldbank.org/indicator/SH.XPD.CHEX.GD.ZS?locations=BR, last accessed on December 21, 2022).

[3] Chen (2021) uses a similar within-proceduralist identification strategy to explore the effects of team-specific experience on patient mortality in the United States. We build on and refine this strategy using information on the specialty that motivated the assignment of a given physician to a patient and incorporating a full set of physician's case-related specialty fixed effects, as described in Section 4.

character of doctor work schedules. Some physicians could be assigned to a patient based on the type of care and specialty required. We take advantage of the fact that our data provide information on the specialty that motivated the assignment of a given physician to a patient and incorporate a full set of physician's case-related specialty fixed effects to tighten identification.[4] We document extensively that, conditional on physician's case-related specialty and proceduralist effects, our measure of team members' expertise overlap is strongly balanced along a wide array of patient and physician pre-determined background characteristics.

By implementing this within-proceduralist identification strategy, we find that greater overlap in team members' expertise leads to a statistically significant and economically meaningful decline in patient 30-day mortality.[5] This finding is robust to a variety of specification checks, subsample analyses, and alternative constructions of the expertise overlap. We show that these estimates are unlikely to be driven by the number of cardiovascular-related specialists in the team, selection of patients into the procedure, selective team formation, or other factors that vary across teams such as team-specific experience. We find similar results when we implement an alternative identification strategy that simulates the composition of a team based on the physicians available on the date that patients arrive in the hospital in an intent-to-treat approach. Since doctor work schedules are set in advance and the exact date that emergency patients arrive at the hospital is unpredictable, conditional only on time and proceduralist fixed effects, these results provide strong evidence that our estimates do not simply reflect some form of unobserved sorting or an artifact of the baseline design. Quantitatively, our estimates imply that moving from the 25th to 75th percentiles of the expertise overlap distribution would reduce 30-day mortality by 1 percentage point, or 8 percent of the gap between high- and low-mortality risk patients.[6]

In the second part of the paper, we investigate the likely mechanisms behind our results. Our reading of the evidence is most consistent with an improved coordination mechanism: proceduralists and physicians with similar backgrounds have more knowledge about how to better collaborate with each other, and this improves patient outcomes. We provide diverse pieces of evidence consistent with this interpretation. First, we find that team members' expertise overlap reduces treatment intensity, as measured by the number of tests exams and the length of hospital stay, with a significant decline in overall hospital spending. That is, proceduralist-physician teams improve patient outcomes using fewer medical inputs, a gain in productivity that most likely results from better collaboration. Second, the mortality decline is significantly larger when the complexity of care is greater, exactly the cases where the largest returns to improved coordination would be expected.

Third, we examine the interplay between past collaboration experience and expertise overlap. The coordination hypothesis establishes that doctor teams with shared expertise collaborate more efficiently because having overlapping expertise allows them to better understand each other's practice style. If this mechanism is at work, the returns to shared expertise should be lower for teams that have extensive past collaboration experience and have already learned each other's practice style. We find evidence in line with this hypothesis: the effects of

[4]For example, a patient experiencing a heart attack with a health history of cancer could be assigned a physician with an oncology specialty. Note that the case-related specialty varies across physicians between and within patients. A same patient could be assigned to different physicians with different specialties motivating the assignment, and two different patients could be assigned the same physician with different motivating specialties. Because most physicians and proceduralists typically complete more than one specialty, controlling for the specialty that motivated the assignment of a given physician to a case leaves us still with substantial variation in our measure of team members' expertise overlap across cases. In Section 4, we document and discuss this variation.

[5]Our focus is on 30-day mortality. This metric has been used widely in the medical literature as a measure of care quality (Menees et al., 2013; Stehli et al., 2019). In our sample, approximately, 5.6 percent of PCI patients died within 30 days after the procedure.

[6]For example, the mortality gap between patients above and below age 85 is about 11.5 percentage points. This implies $1/11.5 \approx 8$ percent.

expertise overlap on mortality tend to be smaller for patients treated by proceduralists and physicians with more accumulated experience working with each other.

Overall, we interpret our results as showing robust evidence that doctor teams with shared expertise achieve better patient outcomes via improved coordination. In the final part of the paper, we use a simple welfare framework to conduct counterfactual policy simulations and explore the efficiency gains from reorganizing teams. We consider two alternative counterfactual policies that alter physicians' work schedules in a way that increases team members' expertise overlap: i) increasing the number of hours worked of certain physicians —*intensity-based allocation*; and ii) reallocating physicians across days of the week —*nonintensity-based allocation*. Our simulations suggest that the nonintensity-based allocation intervention is much more cost-effective. The key reason for this is that the costs of implementing a policy that significantly increases physicians' labor supply are extremely high under reasonable assumptions regarding the labor supply elasticity (Chetty, 2012). In contrast, the costs of reallocating physicians across days of the week are relatively low, such that the gains in hospital spending reductions alone would cover a sizeable portion of the costs of policy implementation. As a result, this intervention would be justifiable on the grounds of economic efficiency even under very conservative assumptions on the value of a statistical life.

This paper is related to several contributions in the literature. We build on early work in the healthcare literature documenting substantial heterogeneity in treatment intensity across markets and across providers within markets (Fisher et al., 2003a,b; Skinner, 2011). Several prominent studies have emphasized that differences in skills and comparative advantage across providers explain in part why many providers use fewer medical inputs without necessarily leading to worse health outcomes (Chandra and Staiger, 2007, 2020; Chan et al., 2022).[7] Our findings suggest that providers' productivity is contingent on the form how teams are organized, even holding skills and comparative advantage constant. This is consistent with a growing body of research examining the importance of teams for the production of health care. Notable examples include Chan (2016, 2021), Chen (2021), Silver (2021), and Doyle and Staiger (2022).[8] Our key contribution to this research agenda is to document the implications of variation in provider expertise within teams, an important source of variation in care quality and spending that has been largely overlooked in the literature. Our findings suggest that healthcare spending could be significantly reduced, without harm to patients, if teams are restructured to increase the overlap of their members' expertise.

More broadly, our results contribute to a large literature in economics on the optimal organization of teams. A body of work has documented the importance of team incentives (Bandiera et al., 2005, 2013; Friebel et al., 2017; Delfgaauw et al., 2022), transmission of information (Sandvik et al., 2020; Battiston et al., 2021), team-specific human capital (Jaravel et al., 2018; Chen, 2021), moral hazard (Holmstrom, 1982; Chan, 2016), and

[7] In explaining the sources of variation in treatment intensity, a large body of work has also emphasized the role of provider incentives. Clemens and Gottlieb (2014) exploit quasi-experimental variation in reimbursement rates and document that providers increase treatment intensity when they are reimbursed at higher rates. Einav et al. (2018) reach a similar conclusion in the context of long-term care hospitals. Finkelstein et al. (2016) examine more generally the importance of both place-specific demand-versus-supply factors. Using a mover-style design, they document that approximately 40-50 percent of the differences in healthcare utilization can be explained by demand-side factors.

[8] Chan (2016) investigates the incentives to engage in moral hazard in organizational systems where physician teams have more and less autonomy to manage care. Chan (2021) explores the impacts of teams on decision-making in the context of medical residency, documenting that teams lead to concentrated influence among senior trainees. Doyle and Staiger (2022) and Silver (2021) provide clean evidence that a physician's treatment intensity depends on the peers with whom they collaborate. Chen (2021) documents that doctor teams whose members have more experience working together develop team-specific human capital and produce better health outcomes.

spillover effects (Mas and Moretti, 2009; Cornelissen et al., 2017; Silver, 2021; Jarosch et al., 2021). Our study is most closely related to research on the effects of diversity among team members on performance (Lazear, 1999; Hjort, 2014; Lyons, 2017; Marx et al., 2021), which has focused almost exclusively on diversity in demographic characteristics. Though the costs and benefits of expertise diversity have been extensively discussed from a theoretical perspective, little research has tackled these questions empirically. Our paper provides the first rigorous, empirical evidence showing that, at least in a time-pressured environment, the coordination costs of diversity in specialized skills are large enough to override any beneficial complementarities.

The rest of the paper is organized as follows. Section 2 provides brief background information on our setting. Section 3 provides details on the sample and key variable definitions. Section 4 outlines our baseline research design, while Sections 5, 6 and 7 present the key results of the paper. Section 8 explores the quantitative implications of the results by simulating the consequences of allocation policies. Finally, Section 9 concludes.

2 Background

This section describes the main features of Brazil's health system and the Percutaneous Coronary intervention.

2.1 Brazilian Health System —SUS

Under the 1988 Constitution, Brazil established universal access to health care as a right and created the Unified Health System (SUS, for its acronym in Portuguese). The SUS provides access to preventive and curative care through public institutions directly managed by the government as well as through a network of SUS-affiliated private facilities. Using a standardized fee system managed by the Ministry of Health, these affiliated facilities provide care to any individual and the government reimburses them according to the type and number of procedures they perform. Individuals must present only an identification card to access these services free of charge in any SUS health facility (no co-payments, coinsurance, deductibles, or other fees).[9]

People who opt for private care outside the SUS can receive ambulatory and hospital services either through out-of-pocket payments or private health insurance plans. The demand for health insurance is mostly from companies offering insurance coverage to their employees. In practice, most of the population receives care through the SUS, with only 26 percent of the population having a private health plan (Paim et al., 2011). According to our administrative data, more than 90 percent of all percutaneous coronary interventions were performed in SUS hospitals in 2019.

2.2 Percutaneous Coronary Intervention

A Percutaneous Coronary Intervention (PCI) is a non-surgical procedure to open an occluded vessel and restore blood flow through the blocked arteries. During the intervention, a thin tube called catheter with a balloon and stent attached is inserted through a blood vessel in the groin or arm and guided to the obstructed coronary artery. When the catheter is at the optimal point, the balloon is inflated to widen the narrowed artery and expand the stent. The balloon is taken out, and the stent is permanently left in the artery to hold it open. PCIs are mostly

[9]The creation of the SUS represented an unprecedented change in health policy. Prior to the SUS, only formal workers received care through the Ministry of Health, while the other segments of the population depended largely on philanthropic institutions and out-of-pocket expenses.

used to treat heart attacks and angina pectoris.[10,11] In Brazil, PCI is the single most common invasive procedure to treat cardiovascular diseases, accounting for 75-80 percent of all invasive cardiovascular procedures and 50 percent of SUS spending on cardiovascular diseases (Oliveira et al., 2022).

Teamwork. PCIS are performed by teams in specialized hospitals with appropriate catheterization laboratories. A team consists of the proceduralist or operator who executes the procedure and physicians who provide pre/post-procedure care during the patient's hospital stay. While a patient sees only one proceduralist, more than one physician may provide care to that patient during the hospital stay. In Brazil, there are, on average, 1.5 physicians per team, with 70 percent of teams having only one physician (see Appendix Table A.1).

The proceduralist's and physician's tasks require inputs from each other. The proceduralist must decide on where to insert the catheter, the number of artery segments in which to intervene, and how many stents to insert. If these decisions are not made optimally, the risk of another adverse event such as a heart attack and mortality may increase. Since the physician has better information on a patient's clinical status, the inputs that the proceduralist receives from the physician are critical for these decisions. At the same time, the physician requires inputs from the proceduralist on events during the procedure to define subsequent treatment and monitoring plans.

Team members' areas of expertise may influence team performance. A team with heterogeneous expertise may benefit from productive complementarities during the production process. For example, a proceduralist may be more likely to learn something relevant to the intervention strategy when the physician is specialized in an area outside of the proceduralist's competence. However, heterogeneity in expertise could increase the costs of coordination, and this may be particularly important in the healthcare industry (Risser et al., 1999; Lingard et al., 2004; Haig et al., 2006). Proceduralists may know better how to interpret the physician's messages if they have common expertise, potentially reducing the risk of medical errors and complications. Moreover, a proceduralist's style may depend on the type of complementary specialties acquired. Physicians who share areas of competence with the proceduralist may have a better understanding of the proceduralist's practice style and adjust their post-procedure strategy. Whether and how team members' overlap in areas of expertise affects team performance is ultimately an empirical question.

2.3 Medical Specialties

In Brazil, there are more than 60 different areas of specialty. Obtaining a specialist degree in a given area requires completing a residency program of between 2 and 6 years, depending on the specialty. Proceduralists and physicians typically have more than one specialty.[12] This is not surprising because some specialties are pre-requisite for others, as in the United States and the United Kingdom. For example, to become a cardiologist

[10] In our sample, patients with these health conditions account for more than 90 percent of cases, with further 10 percent corresponding mostly to other acute and chronic cardiovascular diseases. See Appendix Table A.1 for details.

[11] An alternative and more invasive treatment for coronary heart diseases is Coronary Artery Bypass Grafting (CABG). CABG surgery restores blood flow by creating a new path for blood to flow around the affected artery. It is particularly recommended when multiple arteries are blocked. In Brazil, CABG is not common. According to our administrative data, CABG is performed only in 2 percent of all patients experiencing a heart attack. This low prevalence of CABG in Brazil is the reason why we do not consider this procedure for our analysis.

[12] Appendix Figure A.4 shows the 20 most common specialties in our sample. Cardiology, clinical medicine, and cardiovascular surgery are the three most common specialties among proceduralists, while cardiology, clinical medicine, and intensive care medicine are the most prevalent specialties among physicians.

or oncologist, a doctor must first complete a clinical medicine residency program. Some residency programs do not have pre-requirements of this sort. Consider cardiovascular surgery. Someone interested in this medical area could obtain this degree directly in a training institution affiliated with the Brazilian Society of Cardiovascular Surgery, but for students in institutions not affiliated with this society, it would be necessary to become a general surgeon before applying for a residency program in cardiovascular surgery.

These institutional characteristics explain the considerable variation in areas of expertise across proceduralists and physicians and thus in the degree of expertise overlap between them. Moreover, it is not obvious that other characteristics of doctors beyond specialties, such as experience, can predict the combination of physician-proceduralist specialties —and thus the degree of expertise overlap —in a quasi-randomly formed team. This is a key aspect of our identification strategy, which we formally discuss in Section 4.

3 New Data on Health Care

This section introduces a new patient-doctor-procedure dataset from the entire set of SUS hospitals, which we assemble by combining information from several administrative registries managed by the Ministry of Health.[13] We provide a brief description of these data here, but Appendix F provides further details.

3.1 Sources, Samples, and Variable Definitions

3.1.A Data Sources

Matched patient-doctor-procedure data. We have access to administrative data on the universe of patients admitted to hospitals matched to doctors and procedures since 2008 from the *Sistema de Informações Hospitalares* registers. The matching rate between the patient and doctor-procedure files is 100 percent. These matched data include information on the date of birth, date of admission, date of discharge, diagnosis codes, race, sex, and place of residence (ZIP code) for all patients.[14]Importantly for our research design, the data include information on whether the patient was admitted to the hospital due to emergency health conditions.[15] As discussed later, our analysis focuses on emergency patients who are unable to select into treating teams. The data also contain information on the amount paid per procedure to the providers under the national fee schedule managed by the Ministry of Health. We use this information to create a measure of total hospital spending on each patient, converted into R$2022 using the consumer price index.

A distinctive feature of these data is that we can observe all procedures provided to a given patient and the unique identifier code of the health professional who performed a procedure. The longitudinal nature of the data allows us to track doctors across patients and procedures over time. There exists one main procedure, and

[13]While we do have access to data covering patients from hospitals that are not affiliated with the SUS, these data provide limited information for our analysis. For example, these data provide no information on the health professionals or teams who provide care to a given patient.

[14]The diagnosis codes are divided into a primary and up to nine secondary diagnoses. The former is the health condition responsible for the admission and the latter are diseases that coexist at the time of the admission or develop during the hospital stay. We use secondary diagnoses to create controls for comorbidities that commonly affect patients.

[15]Classification of a patient's emergency status depends on how the health professional who first treats the patient interprets her signs and symptoms. Given the inherent subjectivity in this classification process, there is naturally scope for measurement error. In supplementary analyses, we show that the results are essentially the same if we focus on patients with myocardial infarctions, a health condition whose emergency classification is less ambiguous.

there could be several secondary procedures per patient. The main procedure corresponds to the basic reason for the treatment, whereas the secondary procedures complement the main one.[16] We focus on patients whose main procedure is a PCI treatment.

Doctor characteristics. We use data on providers from the *Cadastro Nacional de Estabelecimentos de Saúde*. It is a rich source of data collected monthly since 2005 that covers all health professionals providing services in any private or public health facility in Brazil. A major strength of these data is their universal nature and high-frequency observations, with minimal underreporting. They contain key information on doctor characteristics, including specialty, weekly hours of work, and the identification codes of the establishments for which the doctor provides services, among others. Using the unique doctor identifiers, we matched these data to the hospital records, with a matching rate of 98 percent.[17]

Mortality data. Our main outcome is an indicator for mortality within 30 days of undergoing PCI, a widely-used metric of care quality in the medical literature (Menees et al., 2013; Stehli et al., 2019). Data on mortality come from the *Sistema de Informações de Mortalidade*. These data contain comprehensive information on all deceased individuals in Brazil, including date of death and birth, race, sex, and place of residence. Unfortunately, there are no identifiers that allow us to link mortality and hospital records directly. Our strategy is to match these data based on the exact date of birth, location of residence, gender, and race (see Appendix F.5 for details and discussion). We generate a mortality indicator variable for whether a patient is matched to the mortality data within 30 days after the PCI treatment. Of course, this matching procedure is imperfect and introduces measurement error. However, since few patients undergoing PCI were born on the same date, reside in the same location, and are of the same gender and race, we believe this measurement error should be negligible.[18] Consistent with this idea, the 30-day mortality rate in our matched data —approximately 5.6 percent —is extremely similar to that estimated by a number of longitudinal studies of PCI patients in Brazil (D'Avila et al., 2015; Machado et al., 2018).

While mortality status will be measured with error for some patients, there is no reason to expect that it is correlated with doctors' expertise overlap. In this case, such measurement error would make it less likely to detect precisely estimated effects, but without causing bias. Consistent with the idea that the measurement error in a patient's mortality status is unlikely to be severe, the data strongly suggest that the mortality indicator has useful empirical content: it is highly correlated with age, gender, and diagnosis severity, with patterns consistent with those documented in the medical literature.[19] Moreover, as shown below, we obtain very precise estimates

[16]For example, consider a patient with a heart attack who was treated with PCI. In this case, the PCI treatment would be the main procedure and any tests, exams, or physician visits during the hospital stay would be classified as secondary procedures.

[17]The matching rate is not 100 percent because there is some underreporting in the database on doctor characteristics.

[18]In our sample, we observe unique location-of-residence$\times$date-of-birth$\times$gender$\times$race cells in 76 percent of cases. In 20 percent of cases, we observe clusters of exactly two people. The remaining 6 percent include clusters of more than three people. Among patients matched to a death record in the mortality database within 30 days of the hospital admission, 99 percent correspond to unique matches.

[19]Most notably, the data exhibit a remarkably sharp age gradient: mortality rates are significantly higher among older patients, with patients over age 95 have mortality rates 6 times as high as those under age 40. Similarly, patients admitted for myocardial infarction, the most severe condition, have much higher mortality rates than those admitted for angina pectoris, the least severe condition, with differences on the order of 67 to 120 percent. Among patients experiencing a heart attack, those with more severe diagnoses experience higher mortality rates than those with less severe diagnoses: patients with ST elevation myocardial infarction exhibit a 30-day mortality rate of almost 8 percent, which is as much as 40 percent higher than that of patients with non-ST elevation myocardial infarction. The data also reveal that the 30-day mortality rate is systematically higher for women than for men, consistent with the medical evidence

of the relationship between expertise overlap and 30-day mortality.

3.1.B Measuring Expertise Overlap

Our independent variable of interest is the degree of expertise overlap between the proceduralist and physicians. We define expertise overlap as the extent to which the proceduralist j and physician k have the same specialties:

$$z_{jk} = \frac{\sum_{s\in\mathcal{S}} \mathbb{1}_j(\tau = s) \times \mathbb{1}_k(\tau = s)}{\sum\limits_{n\in j,k}\sum\limits_{s\in\mathcal{S}} \mathbb{1}_n(\tau = s) - \sum\limits_{s\in\mathcal{S}} \mathbb{1}_j(\tau = s) \times \mathbb{1}_k(\tau = s)} \tag{1}$$

where $\mathcal{S}$ is the full set of specialties, and $\mathbb{1}(\cdot)$s are indicators for whether the proceduralist and physician have specialty $s \in \mathcal{S}$. Note that the numerator is a count of specialties that overlap between the proceduralist and physician, while the denominator captures the total number of distinct specialties in a given proceduralist-physician pair. To avoid double counting when generating the number of distinct specialties, we subtract the number of specialties that overlap (right-hand side of the denominator) from the sum of all the proceduralist's and physician's specialties (left-hand side of the denominator). By definition, z_{jk} ranges from 0 to 1, with 1 denoting perfect overlap and 0 denoting no overlap.[20]

Since a patient is treated by only one proceduralist but often multiple physicians, our benchmark measure of expertise overlap among the doctors treating patient i admitted in month-year t is an average-weighted expression:

$$\text{Expertise overlap}_{ijt} \equiv Z_{ijt} = \sum_{k\in\mathcal{K}(i)} \underbrace{\left(\frac{q_k}{\sum\limits_{k\in\mathcal{K}(i)} q_k}\right)}_{\substack{\text{share of visits}\\ \text{by physician } k}} \times \underbrace{z_{jkt}}_{\substack{\text{proceduralist-physician } (j,k)\\ \text{expertise overlap}}} \tag{2}$$

where $\mathcal{K}(i)$ is the set of all physicians treating patient i during the hospital stay, and q_k is the total number of visits provided by physician k. The share of hospital visits accounts for the fact that each physician has a differential contribution of care. In the robustness checks described below, we consider alternative constructions of the expertise overlap.

3.1.C Baseline Sample

There are 337,065 patients undergoing PCI treated by a proceduralist-physician team between 2008 and 2022. We apply several restrictions on this starting sample. We first drop cases in which the proceduralist performed both the PCI procedure and provide care during the hospital stay ($\approx$ 19 percent), as it is not possible to identify interactions between the proceduralist and physicians in teams formed of a single member. Second, we limit the sample to patients admitted before November 2020 since our main outcome focuses on 30-day mortality, and we do not observe those patients who died in 2021 or later.[21]This restriction leaves us with 243,132 patients. Third,

that women have a higher incidence of adverse outcomes following PCI than men (Cowley et al., 1985; Maynard et al., 1997; Kelsey et al., 1993). This is striking given that the mortality rate in the general population is higher for men, suggesting that the gender-specific pattern in our data is unlikely to be an artifact of our merging strategy.

[20]As an example, consider a proceduralist with specialties in cardiology and general medical clinic, and a physician with specialties in general medical clinic and surgery. In this case, the expertise overlap between them is 0.5 ($= \frac{1}{3-1} = \frac{1}{2}$).

[21]The Department of Regulation, Evaluation, and Control of Systems at the Brazilian Ministry of Health takes about 2 years to collect and prepare vital statistics before making these data available. Therefore, we did not have any information for 2021 or later at the time of preparation of this manuscript.

we limit the sample to observations with no missing values on any of the covariates. This implies excluding patients admitted in 2008 (19,418) because one covariate is physician experience in the past 12 months, and we do not observe this covariate for physicians treating patients in 2008. Finally, we restrict the analysis to patients admitted to the hospital for emergency health conditions (≈ 79 percent) to reduce the risk of non-random sorting of patients into teams.

In the end, the final sample contains 176,108 patients from 201 hospitals. Appendix Table A.1 shows key sample characteristics. Patients undergoing PCI have, on average, 62 years of age, 65 percent are males, and about 5.6 percent died within 30 days of the procedure. There are 966 proceduralists and 9,628 physicians, with a mean degree of overlap in specialties of 0.42. Approximately 80 percent of the proceduralists perform PCI treatments in more than one hospital within a year. For ease of readability, we use the terms hospital-proceduralist and proceduralist fixed effects indistinctly when describing the empirical strategy and results unless stated otherwise.

4 Research Design

This section presents our research design, identification conditions, and empirical specification.

4.1 Overview

Our research design relies on the comparison of patients assigned to the same proceduralist but different physicians. This within-proceduralist approach is similar in spirit to that used by Chen (2021). It requires that the assignment of patients to physicians working with the same proceduralist is as good as random, even if the assignment of cases to proceduralists is not. This assumption is plausible as patients are unlikely to select physicians and as physicians are unlikely to deliberately sort themselves into cases.

A basic step to reduce the risk of non-random sorting of patients to physicians is to focus on patients admitted to the hospital for emergency health conditions. These patients require immediate care and thus have no choice but to see the physician sequentially available. The focus on emergency patients should also limit the scope for physicians to select patients, given idiosyncratic variation in the arrival of emergency patients and the availability of physicians working at the patient's visit. Physicians typically follow a predetermined work schedule and are therefore unable to control the types of patients arriving in emergency situations. Once a patient is assigned, physicians cannot refuse to treat her under any circumstances by law. While some hospitals could try to implement the "heart team" concept, under which the patient is treated by a predetermined team of doctors with specific specialties, this is difficult to be operationalized in practice due to logistical constraints. As vividly described by CEO and Chief Cardiac Surgeon Yadava (2018, p. 3):

> *[t]here are a large number of practical issues in implementing the Heart Team concept. The logistics of the availability of all the constituents of the Heart Team... all at an anointed time, and obvious funding requirements for implementing this concept, are important bottlenecks.*

Our baseline specification includes hospital-proceduralist-year (hereafter, proceduralist-year) fixed effects, so we exploit comparisons between patients treated by the same proceduralist within the same year. When several physicians are available, those with a specialty directly related to a case are naturally more likely to be assigned to that case. For example, an available physician with an oncology specialty may be assigned to PCI patients with cancer. We take advantage of the fact that the inpatient data provide information on the specialty related to the case for each physician providing care during the hospital stay. We therefore refine identification controlling for a full set of physician's case-related specialty indicators. Since patients often see more than one physician, we construct weighted averages of the physicians' case-related specialty indicators for a given patient following equation (2).[22] By controlling for physician's case-related specialty and proceduralist-year fixed effects, the parameter of interest is identified from a relatively narrow source of variation: any omitted factor that generates bias would have to operate within proceduralist-year and across physicians with similar case-related specialties in proportion to the expertise overlap. In principle, it is difficult to think of plausible scenarios in which this could occur.

Identifying variation. While our research design exploits very fine-grained comparisons, Appendix Figure A.1 documents that there is still substantial variation in the degree of expertise overlap across cases within proceduralists. The median standard deviation in expertise overlap within proceduralist-year cells is 0.20, comparable to the overall standard deviation of 0.29. Proceduralist-year and physicians' case-related fixed effects account for approximately 67 percent of the overall variation in expertise overlap, leaving a fair amount of variation for identification.

4.2 Identification Conditions and Balance

To make our design-based approach as transparent as possible, we next formalize the identification conditions and provide direct evidence of their plausibility.

Condition 1.1 (Independence). *Conditional on proceduralist-year ξ_{jt}, and physician's case-related specialty $\mathcal{S}(\mathcal{K}(i))$, potential outcomes $\tilde{y}_{ijt}$ of patient i are independent of the expertise overlap $Z_{ijt}(j, \mathcal{K}(i))$:*

$$\tilde{y}_{ijt} \perp\!\!\!\perp Z_{ijt}(j, \mathcal{K}(i)) \mid \xi_{jt}, \mathcal{S}(\mathcal{K}(i))$$

This identification condition implies that the degree of expertise overlap varies quasi-randomly across patients, conditional on the minimal set of controls. This is not a strong assumption in view of the limited scope for emergency patients to choose physicians and for physicians with predetermined work schedules to choose patients, as discussed above. To assess the plausibility of this assumption, Panel A of Figure 1 examines whether predetermined patient demographic and health characteristics are correlated with the expertise overlap. It reports the coefficients from separate regressions where patient characteristics are used as dependent variables and the expertise overlap is the independent variable of interest, conditional on proceduralist-year and physician's cases-related specialty effects. Each coefficient is standardized for ease of comparison and readability. The expertise overlap appears to be balanced with respect to patient demographics and comorbidities, despite the fact that these observable characteristics are strong predictors of 30-day mortality.[23] Out of 22 coefficients,

[22] For simplicity, we use the terms physicians' case-related specialty indicators or physicians' case-related specialty fixed effects interchangeably.

[23] We regress 30-day mortality on patient characteristics and find a F-test of the joint significance of these observable characteristics of 48.9 (p-value<0.01). Even after including our demanding set of physician's case-related specialty and proceduralist-year fixed effects,

only 1 is statistically significant at the 5 percent level, most likely due to sampling error. Note that this lack of significance is not driven by large standard errors. In all cases, the coefficients are very close to zero and precisely estimated.[24]

To further check for balance, we combine all pre-determined characteristics into a single measure based on how they affect mortality. This provides a higher-powered balance test because the combined measure uses the most relevant information from the set of patient characteristics. To generate this measure, we use the coefficients obtained from estimating a logit model of 30-day mortality on all the characteristics from Panel A of Figure 1. Panel A of Figure 2 presents a binned scatter plot of the relationship between predicted mortality and expertise overlap after conditioning on the basic set of controls. The estimated coefficient is not only statistically insignificant but small in magnitude, and the binned scatter plot is virtually flat without a clear tendency toward improving or worsening mortality. Comparing patients treated by teams above and below the median (residualized) expertise overlap, we find a difference in predicted mortality of less than 0.003 percentage points (Panel B of Figure 2).

Condition 1.2 (Excludable Characteristics). *Conditional on proceduralist-year, and physician's case-related specialty, unobserved doctor's characteristics $\mu_{ijt}(j, \mathcal{K}(i))$ are independent of the expertise overlap:*

$$\mu_{ijt}(j, \mathcal{K}(i)) \perp\!\!\!\perp Z_{ijt}(j, \mathcal{K}(i)) \mid \xi_{jt}, \mathcal{S}(\mathcal{K}(i))$$

The intuition behind condition 1.2 is that even if the expertise overlap varies randomly across cases, it is not simply proxying for other characteristics of doctors that affect patient outcomes. This is clearly a stronger assumption than the independence condition 1.1. However, it is not implausible, both because there is a relatively large number of different specialties and because the number and types of specialties vary not only among physicians but also among proceduralists. In this case, small changes in physicians' and proceduralists' specialty bundles can lead to very different degrees of overlap. This makes it difficult to predict the expertise overlap in a given quasi-random proceduralist-physician draw.

To evaluate the plausibility of condition 1.2, Panel B of Figure 1 checks for balance in predetermined physicians' characteristics. These characteristics are measured as a weighted average of physicians' attributes in a given case, following the same logic as in equation (2). We consider gender, experience, workload, hospital affiliations, and type of employment relationship, among others, which are jointly highly significant predictors of 30-day mortality.[25] We measure individual experience as the number of patients for which she or he provided care during the hospital stay in the past 12 months.[26] There is no consistent relationship between these characteristics and the expertise overlap, with estimated coefficients that are very small in magnitude and

these covariates continue to be strong predictors of 30-day mortality, with the associated F-test of the joint significance of 24.30 (p-value<0.01).

[24]For example, the largest coefficient (in absolute value), the 95 percent confidence interval allows us to rule out effects as small as 0.002 standard deviations.

[25]The F-test for joint significance of physician characteristics on 30-day mortality is 42.28 (p-value <0.001) without additional controls, 34.11 controlling for patient characteristics, and 11.69 (p-value <0.001) controlling additionally for physician's case-related specialty and proceduralist-year fixed effects.

[26]The choice of this timing is motivated by the evidence in previous studies suggesting that experience in the very distant past may not be so relevant for the current case (Benkard, 2000). The conclusions are essentially the same if we consider a window of 24 or 36 months. We focus on the 12-month cutoff because we will include this variable in our baseline specification, and using a higher cutoff implies a larger reductions in the number of observations available.

far from significant. One could think of stories in which patients with worse health conditions see a larger number of physicians, and the number of physicians may be correlated with the degree of expertise overlap. Nevertheless, the relatively large variation in the number and types of specialties makes this story difficult to construct. The last row of Figure 1, Panel B indeed shows that the (conditional) correlation between the expertise overlap and number of physicians providing care to the patient during the hospital stay is essentially zero.

We also estimate a logit model of mortality on the physicians' characteristics displayed in Panel B of Figure 1 to check for balance in predicted mortality. Consistent with the evidence above, the expertise overlap is not significantly associated with predicted mortality. Panel B of Figure 2 illustrates this visually by plotting a binned scatter plot of both variables after accounting for the basic set of controls. Moving from the 10th to 90th percentiles of the expertise overlap distribution would lead to a statistically insignificant increase in the predicted mortality rate of less than 0.07 percentage points.

Overall, the evidence presented in this section suggests that the expertise overlap is balanced across cases and is not a proxy for other features of doctors.

4.3 Estimating Equation

We estimate the effect of expertise overlap on case outcomes using the following specification:

$$Y_{ijt} = \alpha \ + \ \beta \text{ Expertise overlap}_{ijt} \ + \ \mathbf{X}'_{ijt}\Psi \ + \ \mathcal{S}'_{ijt}\Omega \ + \xi_{jt} + \eta_{ijt} \tag{3}$$

where Y_{ijt} is the outcome of interest for patient i admitted in year t and treated by proceduralist j. The term *expertise overlap* is our independent variable of interest as defined in equation (2). Our baseline specification includes the detailed set of proceduralist-year fixed effects ξ_{jt} and the vector $\mathcal{S}'_{ijt}$ of physicians' case-related specialty indicators. Moreover, we add controls for patient and physician characteristics, included in the vector $\mathbf{X}'_{ijt}$. This vector also includes day-of-the-week and month fixed effects interacted with hospital indicators. These hospital-time fixed effects allow for time-variant organizational features of hospitals across days and months (e.g., fewer staff members on the weekend and in typical vacation months). The residual term η_{ijt} is clustered at the hospital level to allow for correlation across cases treated in the same hospital.

Under the identification conditions 1 and 2, our estimates of β can be given a causal interpretation. While the evidence presented in the previous subsection supports the validity of these identifying assumptions, we discuss possible threats to identification and empirically evaluate their relevance after presenting the basic findings.

5 Results: Expertise Overlap and Mortality

5.1 Main Findings

We begin by presenting a non-parametric binned scatter plot of the relationship between expertise overlap and 30-day mortality. We first residualize both variables with respect to proceduralist-year and physician's case-related specialty fixed effects. We then divide the residualized expertise overlap into 20 equal-sized groups (vintiles) and plot the mean value of the residualized 30-day mortality in each bin. As can be seen from Panel A of Figure 3, patients treated by teams with greater overlap in specialties are less likely to die 30 days after the procedure. Except for a few outliers in the right tail of the expertise overlap distribution, the data points lie very close to the fitted line. Panel B of the figure shows that patients above the median residual expertise overlap have a 30-day mortality rate that is 1.1 percentage points lower than those below the median residual expertise overlap.

Table 1 presents these results in regression format. In column 1, we present estimates from the most parsimonious version of equation (3), which controls only for proceduralist-year and physician's case-related specialty effects. The coefficient of interest, β, is estimated at -0.042 and significant at less than 1 percent. It implies that a 10-percentage point increase in expertise overlap is associated with a decline of 0.42 percentage points in 30-day mortality, an effect that represents a 7-percent decline relative to the sample mean.

The remaining columns add more covariates. Column 2 incorporates day-of-the-week and month indicators interacted with hospital fixed effects. Columns 3 and 4 control for the full set of patient and physician characteristics respectively. Column 5 includes all additional controls simultaneously. If the identification conditions discussed in the previous section hold, the inclusion of these controls should leave the coefficient estimate largely unchanged. This is indeed the case. Adding either hospital-time fixed effects, patient characteristics, physician characteristics, or all of these variables simultaneously has no material impact on our results. The point estimate goes from -0.042 in the benchmark model to -0.041 in the most extensive specification and remains highly significant.

To further evaluate the robustness of the results to control choice, we use a machine learning technique, namely least absolute shrinkage and selection operator (LASSO), to select covariates from among the set of patient and physician characteristics. This procedure chooses the subset of covariates that better predicts mortality and has a relatively low penalty in terms of the model size (Belloni et al., 2014). In addition to the baseline set of covariates, we consider a full set of interactions between gender, race, age, and comorbidities as well as between all physician covariates. Appendix Table B.3 shows that the results from the specifications identified by LASSO are extremely similar to those reported in Table 1, with the estimated coefficient ranging between -0.0412 and -0.0405.[27]

We can evaluate the potential for selection on unobservables more formally using the test developed by Oster (2019) based on Altonji et al. (2005). This test compares the relative changes in the coefficient β with movements in the R^2.[28] The more stable the coefficient of interest relative to increases in the model's fits, the less likely it is that it is spuriously driven by omitted variables.[29] Comparing the model with the minimal set of controls to the most extensive specification based on this test suggests that selection on unobservables would have to be approximately 50 times larger than selection on observables to explain our results, a figure well above the rule-of-thumb cutoff of 1. Comparing the benchmark model with respect to the most saturated regression selected by LASSO (Table B.3, column 6), this figure becomes somewhat smaller at 26 but still remains far from 1. It seems unlikely that our estimates are fully driven by omitted determinants of mortality.

[27]The results from the specifications selected by LASSO tend to be more precisely estimated, which is unsurprising since LASSO penalizes models with poor predictors of mortality and thus focuses on those that significantly reduce sampling variation.

[28]This statistic is calculated as $\frac{\beta_{full}}{\beta_{base}-\beta_{full}} \times \frac{R^2_{full}-R^2_{base}}{\lambda R^2_{full}}$, where β_{full} is the coefficient from the specification with the full set of controls, β_{base} is the baseline coefficient estimate, and R^2_{full} and R^2_{base} are the R^2 from the corresponding regressions. The parameter λ measures the relative increase in the R^2 from the regression with the full set of observed controls to one hypothetical regression that includes all observed and unobserved controls. Oster (2019) derives the value of λ for which 90 percent of randomized studies would survive the test and finds it to be 0.3. We follow her recommendation and set λ to 0.3.

[29]The intuition is that the proportion of the variance that is left to be explained by omitted variables decreases as the fit of the model increases.

While Figure 3 suggests that our results are not driven by a few outliers in the last vintile of the residualized expertise overlap, we repeat the baseline specification while excluding these observations. If anything, the point estimate becomes larger in magnitude (Appendix Table B.2, column 2). To further ensure that our results are not the product of outliers, we estimate the baseline specification using subsamples that exclude cumulatively the first hospitals with the largest share of patients.[30] These estimates are presented in Appendix Figure B.1, with the coefficients ordered from the smallest to largest group of hospitals excluded. Even when we exclude hospitals that account for approximately 45 percent of observations, our estimates remain very robust. As a further test, Appendix Figure B.2 repeats the baseline specification when each hospital is excluded one by one, with virtually no impact on our results.

Heart-attack patients. Our analysis focuses on patients in emergency. Although we find similar results when including non-emergency patients in the sample (Appendix Table B.2, column 3), a possible concern is misclassification of emergency status since it depends on how the health professional who first treats the patient interprets her signs and symptoms. The extent to which a patient's emergency status is subject to error, it could increase the risk that our estimates are simply reflecting non-random sorting of patients to doctor teams. The fact that the expertise overlap is balanced across patient characteristics suggests that any bias generated by misclassification of emergency patients is not severe.

As a robustness test, we limit the sample to patients experiencing a heart attack. This condition requires immediate treatment due to its acute nature and, once diagnosed, it is considered an emergency case by clinical protocols, which plausibly minimizes the risk of misclassification and any degree of selection into teams. The results obtained from using this subsample are displayed in column 4 of Appendix Table B.2. The point estimate is -0.0508 (standard error=0.0156) and thus somewhat larger in magnitude relative to the baseline.[31] We conclude that any bias due to misclassification of a patient's emergency status in the benchmark sample is unlikely to be a major issue.

Additional checks. Before turning to major identification concerns and providing further insights on the relationship between expertise overlap and mortality, we perform several additional sensitivity tests to alternative empirical choices. To save space, we present and discuss the results from these additional exercises in Appendix B. We demonstrate the robustness of our basic results to:

(i) alternative constructions of the expertise overlap, including an unweighted version of (2), the median proceduralist-physician expertise overlap, and the expertise overlap between the proceduralist and physician with the highest number of hospital visits (Appendix Table B.1).

(ii) excluding either patients treated during the COVID-19 pandemic crisis or patients treated by proceduralists with many patients in our sample (Appendix Table B.2).

[30]PCI treatments are concentrated in some large hospitals, so a few hospitals have a disproportionate share of observations (Appendix Figure A.2, Panel A). The median number of patients per hospital is 442, but hospitals at the 99th percentile treated as many as 6785 patients. Hospitals above the 90th percentile of the distribution account for almost 50 percent of patients in the sample. These differences across hospitals are not a threat to the internal validity of our findings since our baseline specification includes a rich set of hospital-time fixed effects, but they could still affect the interpretation of the estimated relationship in terms of external validity.

[31]This larger result is consistent with the hypothesis that the returns to improved coordination are higher when the tasks are more complex, as heart attack patients have a higher underlying mortality risk and thus are more challenging cases. In Section 7, we return to this discussion and provide further evidence consistent with this hypothesis.

(iii) removing all controls and fixed effects, using proceduralist fixed effects instead of proceduralist-by-year fixed effects, adding hospital-by-day-of-week-by-month fixed effects, and including proceduralist-by-day-of-week and proceduralist-by-month fixed effects (Appendix Table B.4).

(iv) using standard errors clustered either at the proceduralist, municipality, or health region as well as two-way clustered at the hospital and year-month level (Appendix Table B.5).[32] We also perform permutation tests that randomly assign either expertise overlap or 30-day mortality status to patients (Appendix Figure B.3).

Overall, we find that the expertise overlap is robustly associated with patient mortality. Quantitatively, our estimates imply that moving from the 25 to 75th percentiles of the expertise overlap distribution would lead to a decline in 30-day mortality of 1 percentage point or 8 percent of the gap between high- and low-mortality risk patients.[33]

5.2 Validity of Baseline Design

In this subsection, we investigate several major threats to the validity of our research design, including selection into the sample, case severity, persistence in the composition of teams, and role of nursing inputs.

5.2.A Expertise Overlap or More (Cardiovascular-Related) Specialties?

A possible concern with our results is that our measure of expertise of overlap could capture variation in the total number of specialties since it is constructed using the number of specialties on the denominator in equation 1. Doctors with more specialties have more training, so it is possible that our results are capturing the effects of a team's "quality" rather than proceduralist-physician expertise overlap. It is worth noting, that while significant, there is no strong correlation between expertise overlap and number of specialties conditional on our basic set of controls. Indeed, the conditional correlation coefficient between both variables is only -0.12. In Appendix Table C.1, we directly explore the sensitivity of our estimates to controlling for the number of specialties. Controlling for this variable ensures that our estimates are not driven by the comparison of patients treated by teams with fewer and more specialties. Given the weak correlation between expertise overlap and number of specialties, the inclusion of this control has unsurprisingly little impact on our coefficient of interest, suggesting that our results are not simply capturing the effect of having more specialties.

More cardiovascular-related specialties? A related concern is whether our results reflect in part the effect of having more cardiovascular-related specialties, which could be associated with a differentially higher quality of care. We first note that there is no statistically significant correlation between expertise overlap and number of cardiovascular-related specialties conditional on the baseline controls.[34] In column 3 of Appendix Table C.1, we repeat our baseline specification but control for the number of cardiovascular-related specialties. Consistent with

[32] A health region is a group of neighboring municipalities defined strategically by the Minister of Health for the implementation and consolidation of healthcare programs. On average, there are 11 municipalities per health region and 477 health regions.

[33] For example, the mortality gap between patients above and below age 85 is about 11.5 percentage points. This implies $1/11.5 \approx 8$ percent.

[34] Conditional on physician's case-related and proceduralist-year fixed effects, we find that a 10-percentage point increase in expertise overlap is associated with a statistically insignificant decline in the number of cardiovascular-related specialties of 0.032 or less 1 percent of the sample mean.

the lack of correlation between this variable and expertise overlap, our results remain similar to the baseline. Column 4 simultaneously controls for the total number of specialties and the number of cardiovascular-related specialties, with no material effect on the estimated effect of expertise overlap.

5.2.B Selection into the Procedure

Patients undergoing PCI are not chosen at random. Among patients with the most common diagnoses for PCI, females and those with comorbidities such as diabetes and kidney disease are significantly less likely to undergo PCI (Appendix Table C.2).[35] This non-random sorting of patients into PCI raises concerns about the validity of our identification strategy. But note that identification would be threatened only if selection into PCI were based on the identity of team members. For example, if patients with lower mortality risk are more likely to undergo PCI when doctors with similar expertise are available, then it would bias the results toward finding $\beta < 0$ in our estimation of equation (3).

We investigate this concern from a number of different angles. Our first approach is to construct a measure of simulated expertise overlap based on the doctors available when patients arrive in the hospital.[36] We generate this measure as the average expertise overlap over all possible proceduralist-physician teams available on the day patients enter the hospital (see Appendix C.2 for details) and estimate a regression of PCI treatment on this simulated expertise overlap.[37] Reassuringly, the estimated association is statistically and economically insignificant irrespective of whether we adjust for covariates (Appendix Table C.3). The results obtained from the specification with the full set of patient covariates and hospital-time fixed effects imply that moving from the percentile 25th to the 75th percentile would increase PCI rates by only 0.50 percentage points (relative to a mean PCI rate of 45 percent).

Our second approach compares the *actual* expertise overlap between patients with lower and higher predicted probabilities of undergoing PCI. We predict PCI as a function of patient characteristics using the estimates from the probit model reported in Appendix Table C.2. If doctors are more likely to recommend PCI treatment when the available doctors have more areas of expertise in common, then those patients with a high predicted probability of PCI treatment might on average be treated by teams whose members have higher overlap in specialties. We do not find evidence to support this story. The association between predicted PCI propensity and actual expertise overlap is essentially zero from a statistical and practical perspective (Appendix Table C.4).[38]

In Appendix Table C.5, we go one step further and estimate the effect of expertise overlap on mortality controlling flexibly for PCI propensity as suggested by Heckman and Hotz (1989). Column 1 repeats the baseline for ease of comparison. Columns 2 through 5 control linearly and non-linearly for PCI propensity. The coefficient of interest and corresponding standard error are virtually unaffected by the inclusion of these controls. Column 6 rather controls for the inverse Mill's ratio, again leaving the coefficient almost unchanged. As a final

[35] Patient demographics are jointly significant predictors of PCI treatment. The F statistics for joint significance is 49.5 with a p-value=0.00, conditional on state and year indicators. A noteworthy predictor is age. The relationship with this covariate is concave. PCI probability increases smoothly with age until 80 years of age, at which point it declines rapidly.

[36] The reason why we use a measure of potential rather than actual expertise overlap is that the latter is observed only for patients who underwent PCI treatment.

[37] The strong association between simulated and actual expertise overlap motivates an instrumental variable strategy where the former is used as an instrumental variable for the latter. In Section 6, we describe this instrumental variable strategy.

[38] For instance, comparing patients below and above the expertise overlap, after conditioning on proceduralist-year and physician's case-related effects, we find a difference in PCI propensity of less than 0.1 percentage points.

check, column 7 limits the sample to patients who are in the top 20 percent of the PCI propensity distribution and thus have an *ex-ante* high probability of undergoing PCI. The precision falls considerably since the resulting subsample is massively smaller, but the coefficient of interest remains extremely similar to the baseline, going from -0.041 to -0.0407.

We conclude that selection into PCI is very unlikely to be responsible for the association between expertise overlap and mortality we observe.

5.2.C Case Severity

One might think of stories where patients are assigned to different types of physicians depending on their underlying mortality risk. It is possible that physicians with rare specialties see specific patients with higher mortality risk because that rare specialty is the most relevant for these cases. This would generate negative values of β even in the absence of a causal relationship. However, note that the focus on emergency patients and inclusion of the rich set of physician's case-related fixed effects should largely eliminate this kind of bias since we are comparing cases in which physicians see patients based on the same most relevant specialty for the case. Moreover, note that the evidence in Figures 1-2 showing that the expertise overlap is strongly balanced across a broad spectrum of patient characteristics and mortality risk already provides direct evidence against this identification threat.

In this subsection, we provide an additional piece of evidence against this potential source of bias. Specifically, we rerun our baseline specification but include a detailed set of patient primary diagnosis fixed effects. This means that the parameter of interest is identified from the comparisons between patients with similar detailed disease types and thus case severities.[39] As shown in Appendix Table C.6, we find a qualitatively and quantitatively similar effect using this more demanding specification.

5.2.D Persistence, and Team-Specific Experience

Proceduralists tend to work with the same physicians over time. This persistence in team composition raises the question of whether part of the effect we estimate reflects the effect of team members' past collaboration experience. For this story to make sense, proceduralists and physicians with greater overlap in specialties would need to be more likely to form the same team over time, as shared work experience has been shown to be beneficial for patient outcomes.[40] We do not find evidence that this is the case. Column 1 of Appendix Table C.7 shows that the expertise overlap does not significantly predict team members' collaboration experience in the past two years.[41] Not surprisingly, given this weak correlation, the inclusion of this variable in the mortality

[39] The severity of the condition that motivated the hospital admission varies significantly across patients. In our sample, for example, the differences in 30-day mortality rates between patients suffering from a myocardial infarction with ST elevation and non-ST elevation is approximately 60 percent. More generally, 30-day mortality rate goes from 3.3 percent for the primary diagnosis in the 25th percentile of the distribution to 7.7 percent for that in the 75th percentile, a difference of more than 100 percent.

[40] Chen (2021) provides rigorous evidence in the US context that patients assigned to teams whose members have collaborated in the past have better outcomes, including increased survival. We corroborate the same findings in the Brazilian context. These results are not shown in this paper to save on space but are available upon request.

[41] Our measure of team members' collaboration experience is the same as the one used in Chen (2021): the number of times that proceduralists and physicians have worked together in the past two years. The two-year cutoff is motivated by the evidence showing that the effect of experience decay with time (Benkard, 2000; Kellogg, 2011; Dinerstein et al., 2022), a pattern also corroborated in the healthcare setting (Chen, 2021).

equation (3) yields a coefficient of β that is precisely estimated and close to the baseline (Appendix Table C.7, column 3).[42]

This evidence is reassuring, but serial correlation in the proceduralist-physician expertise overlap may still be a concern. While such serial correlation is not very high,[43] one may remain worried that our results are simply capturing some persistent unobserved factor affecting mortality. To examine this possibility formally, we estimate a modified version of model (3):

$$Y_{ijm} = \alpha \ + \ \theta_v \text{ Expertise overlap}_{ijm+v} \ + \ \mathbf{X}'_{ijm}\Psi \ + \ \mathcal{S}'_{ijm}\Omega \ + \xi_{jt} + \eta_{ijm} \tag{4}$$

where the subscript m denotes time in month-year pairs (e.g., January 2010). The second term on the right-hand side of 4 is the average expertise overlap observed for proceduralist j in time $d+v$, where $v \neq 0$ represents a lead or lag. We estimate this model for $v \in \{-4, -3, -2, -1, 1, 2, 3, 4\}$. If our results are capturing some persistent omitted factor affecting mortality, we should observe large and significant estimates of θ_v.

As one can infer from Figure 4, neither past nor future expertise overlap affects current patient mortality. The estimated placebo coefficients are far from significant and very small in magnitude. Only current expertise overlap between proceduralists and physicians matters for current patient outcomes.

In sum, there is no empirical indication that our results are driven by team members' past collaboration or any persistent unobserved factor spuriously correlated with the effects we estimate.

5.2.E The Role of Nurses

Nurses assist doctors on the PCI treatment and thus play an important role during the patient's hospital stay. It is possible that the availability of different types of doctors varies with the number of nurses available across dates. In the presence of a negative shock in the supply of nurses (e.g., strikes), hospitals may respond by reorganizing doctor teams and endogenously altering the proceduralist-physician expertise overlap. Given the importance of nursing inputs for patient outcomes (Gruber and Kleiner, 2012), one may argue that it is not that proceduralist-physician expertise overlap improves patient survival but rather concurrent changes in the supply of nurses.

To evaluate the role of nurses in driving our results, we compute the number of nurses available per hospital beds on the day the patients are admitted to the hospital. Column 1 of Appendix Table C.8 documents that the association between the availability of nurses and proceduralist-physician expertise overlap is essentially zero. Column 3 of Table C.8 shows that controlling for nursing supply has unsurprisingly little impact on the estimated coefficient of interest.

[42]While the model including this control helps rule out biases due to shared work experience, it is not our preferred specification because the sample size decreases significantly. To construct the shared work experience in the past two years, we have to exclude patients treated in the first two years of our sample since it is not possible to compute the measure of shared work experience for these observations.

[43]To test for serial correlation, we first collapse the expertise overlap at the proceduralist-by-year-by-month. We then fit an AR(1) model to the expertise overlap. The autocorrelation coefficient from this model is 0.39.

6 Physician Availability Design

While the results presented so far are very reassuring, one might remain concerned about non-random sorting of physicians into teams. To dispel any residual concern about selection into teams, this section proposes an alternative identification strategy that exploits a plausibly exogenous source of variation coming from changes in the availability of physicians within hospitals over time.[44]

6.1 Empirical Strategy

Overview. The basic idea behind this identification strategy is to simulate what would have been the composition of a team based on the physicians available at the time patients arrive in the hospital. This is an intent-to-treat approach. For a patient assigned to a given proceduralist, we assume that *all* physicians present on the date of admission provided care to that patient during the hospital stay. We then use the expertise overlap between the proceduralist and physicians in the simulated team as an instrumental variable for the actual proceduralist-physician expertise overlap. Since the exact day that an emergency patient arrives in the hospital is unpredictable, and doctor work schedules are set in advance, the degree of overlap in specialties between the proceduralist and physicians treating a patient in this context is arguably independent of patients' health conditions and physicians' background characteristics, conditional on time and proceduralist fixed effects.[45]

Simulated overlap. We identify physicians available on a given date and hospital based on whether they provided any care on that date in that hospital.[46,47] Because several physicians are often available on a given date, several potential physicians could have treated a same patient. But not all physicians are equally likely to see a patient. Those with greater number of scheduled working hours per day are more likely to be present at the exact moment when an emergency patient enters the hospital. To account for these differences, we compute the simulated expertise overlap for a patient as the weighted-average expertise overlap between the proceduralist and physicians in the simulated team using physicians' hours of work as weights:

$$\text{simulated expertise overlap}_{jd} \equiv \hat{Z}_{jd} = \sum_{k \in \mathcal{K}(d)} w_{kd} \times z_{jkd} \quad \text{with} \quad w_{kd} = \frac{h_{kd}}{\sum_{k \in \mathcal{K}(d)} h_{kd}} \tag{5}$$

where z_{jkd} is the degree of overlap in specialties between proceduralist j and physician k available on date d. The weight w_{kd} is computed as a function of the number of hours worked h_{kd}. Note that we calculate the simulated expertise overlap keeping fixed the proceduralist who $actually$ treated patient i. This implies that although the set of available physicians is invariant within dates, it is possible to observe differences in the simulated expertise

[44] This strategy is similar to that employed in Doyle (2020) who estimates the effects of cardiologists on patient outcomes in the United States. We adapt his approach to our setting, generating sources of variation in the proceduralist-physician expertise overlap.

[45] A possible endogeneity concern could arise if physicians are called to the hospital to see specific patients. While this could be the case for elective procedures, it is unlikely to be the case for patients with emergency health conditions requiring immediate care. Remember that our estimation sample is limited to these patients.

[46] Ideally, we would use doctor work scheduling data to identify the set of potential physicians, but these data are not readily available beyond the usual hours worked per week.

[47] Of course, this approach is imperfect and could be subject to measurement error. However, because the hospital files provide information on all procedures performed by physicians in detail, we believe that the potential for measurement error is minimal. It is possible to imagine instances where a physician might not see any patient on a given day even if she or he was present that day, but this is unlikely to be the case given that our sample is composed of high-volume hospitals where the demand for physicians is always high.

overlap across patients admitted on the same date if they were treated by different proceduralists.

Instrumental variable specification. The first-stage relationship between actual and simulated expertise overlap is given by the following model:

$$\begin{aligned}\text{expertise overlap}_{ijd} &= \pi_0 \ + \ \pi_1 \text{ simulated expertise overlap}_{jd} \\ &+ \ \mathbf{X}'_{ijd}\Psi \ + \lambda_d \ + \theta_j + \upsilon_{ijd}\end{aligned} \tag{6}$$

Because the underlying source of variation in the simulated expertise overlap occurs at the date-of-admission-by-proceduralist, we have to control for fixed effects of date-of-admission (λ_d) and proceduralist (θ_j). We also control for the set of patient and physician characteristics contained in vector $\mathbf{X}'_{ijd}$. Physician characteristics are average-weighted variables analog to the simulated expertise overlap. The vector $\mathbf{X}'_{ijd}$ also includes day-of-the-week and month indicators interacted with hospital fixed effects to account for hospital-specific seasonal effects. Standard errors are clustered at the hospital level.

Appendix Figure D.2 provides a visualization of the first-stage relationship between actual and simulated expertise overlap by plotting results from estimating a local linear regression version of equation (6). There is a strong first-stage relationship between these variables that is monotonic and approximately linear. Table 2 presents formal first-stage results from estimating model (6). Simulated expertise overlap is a powerful predictor of actual expertise overlap, with the Kleibergen and Paap (2006) F-statistics well above the conventional weak instrument threshold of 10. The results from the specification with the full set of controls imply that a 10-percentage point increase in the simulated expertise overlap is associated with a 5.4-percentage point increase in the actual expertise overlap.

We then use the predicted values from the first-stage regression (6) to estimate the following second-stage model:

$$\begin{aligned}y_{ijd} &= \alpha \ + \ \beta \ \widehat{\text{expertise overlap}}_{ijd} \\ &+ \ \mathbf{X}'_{ijd}\tilde{\Psi} \ + \tilde{\lambda}_d \ + \tilde{\theta}_j + \eta_{ijd}\end{aligned} \tag{7}$$

The coefficient of interest is β. In addition to the instrument relevance condition, which is supported by the evidence discussed above, this two-stage least square (2SLS) regression recovers a local average treatment effect (LATE) under the three additional conditions formalized below.

Condition 2.1 (LATE Validity). Let $\tilde{y}_{ijd}$ denote the potential outcome for patient i admitted on date d treated by proceduralist j. Conditional on proceduralist and time fixed effects, the following assumptions hold:

(i) Exclusion: $\tilde{y}_{ijd}(\hat{Z}_{jd}, Z_{ijd}) = \tilde{y}_{ijd}(\hat{Z}'_{jd}, Z_{ijd})$ for all $\hat{Z}_{jd}, \hat{Z}'_{jd}$ and for all Z_{ijd}

ii) Independence: $(\tilde{y}_{ijd}, \hat{Z}_{jd})$ is independent of the assigned physicians.

iii) Monotonicity: for all $\hat{Z}_{jd}, \hat{Z}'_{jd}$ either $Z_{ijd}(\hat{Z}_{jd}) \geq Z_{ijd}(\hat{Z}'_{jd})$ for all i, or $Z_{ijd}(\hat{Z}_{jd}) \leq Z_{ijd}(\hat{Z}'_{jd})$ for all i.

Appendix Figure D.1 provides evidence consistent with (i) and (ii). It verifies that the simulated expertise overlap is as good as random conditional on the basic set of controls. Specifically, we estimate regressions of predetermined patient and physician covariates on the simulated expertise overlap controlling for the basic set of proceduralist and time fixed effects. Out of 35 hypotheses tested, none is statistically significant at conventional levels of significance. Moreover, the estimated coefficients are small in magnitude. For example, a one standard deviation increase in the simulated expertise overlap is associated with an insignificant increase in patient age of 0.05 years.

The monotonicity condition requires that, for each patient, the actual expertise overlap would be at least as high if there were a potential team with low expertise overlap as if there were a potential team with high expertise overlap.[48] This identification assumption is untestable by definition, but we can look at the sign of the first-stage relationship across different subsamples. If the monotonicity condition holds, the first-stage coefficient estimates should be non-negative in all subsamples. In line with the monotonicity condition, the first-stage estimates are consistently same-signed and meaningful across all subsamples (see Appendix Figure D.3).

6.2 Results

Table 2 presents the results from the physician availability design. Column 2 shows the reduced form estimates from directly regressing 30-day mortality on the simulated expertise overlap. Column 3 exhibits the 2SLS results in which we instrument actual expertise overlap with the simulated expertise overlap. The 2SLS coefficient β scales the reduced-form coefficient by the first stage estimate.

As one can infer from the table, greater overlap in specialties between the proceduralist and physicians is significantly associated with lower mortality. The 2SLS results are similar to those obtained from our first identification strategy in Section 5. Indeed, the 2SLS coefficient is -0.044, which is slightly larger (in absolute value) than that in column 4 of Table 1, -0.041. The implied magnitudes are comparable. A 10-percentage point increase in the expertise overlap is associated with a 0.44 percentage point decrease in the 30-day mortality, a 7.3 percent change relative to the sample mean. A noticeable difference between the results obtained from both designs is that the 2SLS estimates are somewhat less precise. The standard error of the 2SLS coefficient is 0.02, which is nearly double the standard error reported in column 4 of Table 1, of 0.0106. This is most likely because the specific source of variation we exploit in the physician availability design is much narrower, as there are typically several patients within the same hospital exposed to the same set of available physicians. Nevertheless, it is reassuring that the 2SLS results are still statistically significant at the 5 percent level.

One potential concern with our physician availability design is the presence of time-varying hospital level shocks. For example, some hospitals could be on a specific trend toward expanding or shrinking staffing, which could alter the types of available physicians. However, note that identification would be threatened only if these changes are correlated with the simulated expertise overlap conditional on proceduralist and time fixed effects. To account for possible slow- or fast-moving conditions and practices in each hospital, Appendix Table D.1 repeats the first-stage, reduced-form, and second-stage models but includes a full set of hospital-specific linear time trends.[49] The results from estimating this very demanding specification are similar to the baseline. If anything, the results become stronger.

[48] If this condition were not satisfied, our IV coefficient estimates would represent a weighted average of marginal treatment effects with weights that could be negative (Angrist et al., 1996; Heckman and Vytlacil, 2005). As Angrist et al. (1996) show, the bias due to violations of monotonicity depends, among other things, on the strength of the instrument: the stronger the first-stage relationship, the less sensitive it is to violations of the monotonicity condition. Although we find no evidence of systematic violations of the monotonicity assumption, it is reassuring that the instrument we use has enormous predictive power in the first-stage equation.

[49] We do not estimate models with hospital$\times$date-of-admission fixed effects because they would "over-control" and absorb almost all of the variation we use to help identify the parameter of interest. This problem is less severe —but still important—when we use hospital-specific linear time trends instead.

Additional checks. In Appendix D, we subject the 2SLS estimates to a number of additional robustness checks. Appendix Table D.2 documents that the results are quantitatively and qualitatively similar when we use the unweighted version of the instrument. However, the first-stage relationship is somewhat weaker using this alternative version of the instrument, confirming the strengths of our main approach. Appendix Table D.3 shows the robustness of the results when limiting the sample to heart-attack patients only. Although the precision with which the parameter is estimated falls substantially due to reduced sample size, the coefficient of interest remains almost identical to the baseline. The table also documents that the results are robust to dropping either patients admitted during the COVID-19 crisis or treated by proceduralists who have treated many patients in our sample. Similarly, including nonemergency patients leaves the results unchanged. Appendix Table D.4 verifies that our results are not appreciably affected when we implement alternative inference procedures: i) clustering by proceduralist; ii) clustering by municipality; iii) clustering by health region; iv) two-way clustering by hospital and year-month; v) standard errors corrected using the tF procedure for 2SLS estimates introduced by Lee et al. (2022).

In all cases, we find similar first-stage F-statistics, and the 2SLS coefficient remains negative, significant, and similar to the baseline. The results of this section increase our confidence in the general picture presented so far.

7 How Do Doctors Improve Survival?

There are two major classes of mechanisms that could generate the link between team members' overlap in specialties and patient outcomes. The first channel is effort. Physicians and proceduralists may be less likely to engage in moral hazard and more willing to exert effort when they have shared expertise because they may have better information about how to monitor each other's actions or because they may simply be more motivated when working with "similar" co-workers. The second channel is coordination. Communication and practice styles tend to have a specialty bias, and this naturally influences how doctors interpret the specific messages they receive from each other. When proceduralists and physicians have areas of expertise in common, the noise in the information they exchange may be smaller, and thus medical errors may be less likely to occur.

In this section, we investigate the plausibility of these mechanisms by exploring the different predictions they generate. To preview, we find limited evidence in favor of an increased effort story. Rather, the evidence is consistent the most with a role for the improved coordination hypothesis.

7.1 Treatment Intensity

We start by exploring changes in the use of medical inputs. Previous studies have documented that physicians alter treatment intensity depending on the incentives to engage in moral hazard and the types of peers they are exposed to (Chan, 2016; Silver, 2021). If physicians increase effort because the costs of engaging in moral hazard are higher or because they are more motivated when working with a "similar" co-worker, they could more carefully monitor the patient by requesting more formal test exams and prolonging the length of hospital stay.[50] The effort hypothesis implies that we should observe an increase in the use of medical inputs in teams whose

[50]Consistent with this hypothesis, Chan (2016) convincingly shows that physicians reduce treatment intensity and spend less time with patients when they practice in systems where the incentives to engage in moral hazard are higher.

members have greater overlap in specialties. In contrast, the coordination channel implies the exact opposite effect. Better communication between proceduralists and physicians may avoid unnecessary tests and exams that may otherwise arise when they do not completely understand or misinterpret the specific messages they exchange. Thus, proceduralist-physician teams may be more productive and achieve better performance using medical inputs with the same or even lower intensity when they have more areas of expertise in common.

To examine these hypotheses, we estimate our baseline research design using medical inputs as dependent variables —including the number and types of tests exams as well as the length of hospital stay. As we are considering numerous indicators within a same domain, multiple hypothesis testing could be a concern. To address this issue, we standardize each of these measures and create a composite index of medical inputs by taking the unweighted average of the standardized outcomes. We find that team's member expertise overlap is associated with a significant decline in the use of medical inputs (Table 3). The effects are significant, meaningful, and always negative. The results imply that a 10-percentage point increase in proceduralist-physician expertise overlap would reduce the length of stay by 0.6 days (or 10 percent of the sample mean) and the composite index of medical inputs by 1.3 standard deviations.

These effects translate into lower hospital spending. A 10-percentage point increase in the degree of proceduralist-physician expertise overlap is associated with a statistically significant decline in hospital spending of approximately R$65 or 0.6 percent relative to the sample mean (Table 3, column 8).

Overall, we find no evidence of increased inputs, which is at odds with the prediction of the effort channel. In fact, we observe a meaningful decline in treatment intensity and hospital spending, a pattern that is more in line with the coordination or improved productivity hypothesis.

7.2 Case Complexity and Team Coordination

We next explore heterogeneity with respect to case severity. Care for patients with higher underlying mortality risk is more complex, and step-by-step guidelines are not comprehensive enough for treating these types of patients. A large body of medical research suggests that optimal communication between doctors is particularly important in these cases (Salas et al., 2008). For example, a proceduralist could avoid femoral vein cannulation in patients with high ischemic and bleeding risks to minimize the likelihood of post-procedure complications (Levine et al., 2011). But it is necessary to have an accurate understanding of the patient's clinical status to make this decision, which depends critically on the inputs the proceduralist receives from the physician. We are therefore interested in verifying whether the "returns" to team's member overlap in specialties are higher in teams treating patients with higher underlying mortality risk.

To investigate this hypothesis, we estimate our baseline specification interacting our measure of team expertise overlap with the predicted 30-day mortality. We first estimate a logit model of patient actual mortality on patient characteristics, including indicators for common comorbidities, and then use the estimated coefficients to generate mortality predictions. To increase the predictive power of the model, we include a full set of interactions between all the patient characteristics. With the mortality predictions at hand, we estimate the baseline model but include linear and quadratic interactions between the expertise overlap and mortality predictions to account for possible non-linearities:

$$Y_{ijt} = \alpha \; + \; \sum_{\ell=1}^{2} \kappa_\ell \, \hat{m}_{ijt}^{\ell} \; + \; \sum_{\ell=0}^{2} \beta_\ell \, Z_{ijt} \times \hat{m}_{ijt}^{\ell} \; + \; \mathbf{X}'_{ijt}\Psi \; + \; \mathcal{S}'_{ijt}\Omega \; + \xi_{jt} + \eta_{ijt} \tag{8}$$

where the term $\hat{m}$ is the mortality prediction and Z is expertise overlap, as previously described.[51] Panel A of Figure 5 plots the estimated effect of expertise overlap on 30-day mortality for the 10th, 50th, and 90th percentiles of the predicted mortality distribution. Consistent with an improved coordination channel, we find that patients with higher underlying risk are significantly more likely to benefit from teams with greater overlap in areas of expertise. The estimates imply effects that are approximately 3 times as large for patients in the 90th percentile of the predicted mortality distribution as those in the 10th percentile.

7.3 Team-Specific Experience as a Substitute for Expertise Overlap

The improved coordination hypothesis establishes that teams whose members have similar expertise are more effective because they can better understand each other's practice style and collaborate more efficiently. However, proceduralists and physicians who have worked extensively together in the past have already learned how to collaborate best, irrespective of their shared expertise. In this case, the returns to expertise overlap should be smaller for patients treated by proceduralists and physicians with more accumulated experience working with each other. Verifying this prediction represents a simple test of the coordination channel.

Motivated by this reasoning, we ask whether the effect of expertise overlap decline with team members' past collaboration experience. In testing this hypothesis, we estimate the mortality equation including interactions between our measure of proceduralist-physician expertise overlap and shared work experience in the past two years using the analogous version of the interacted model (8). These results are exhibited in Panel B of Figure 5. The effect of expertise overlap on mortality decreases with the accumulated experience between the proceduralist and physicians. The differences are meaningful. Moving from the 10 to 90th percentile of the shared work experience distribution would reduce the mortality effect of proceduralist-physician expertise overlap by approximately 76 percent, consistent with the improved coordination channel. When physicians and proceduralists have limited experience working together, and thus have little information about each other's ability and style, similarity in areas of expertise serves as a substitute for this lack of team-specific human capital.

7.4 Summary

The general picture presented so far is consistent with the improved coordination channel: team members with shared expertise achieve better outcomes using fewer inputs, particularly when performing more complex tasks and doctors have less accumulated experience working together. Note that this is not to say that there is no space for other mechanisms. Although we find no evidence supporting the effort effect channel, the data do not allow us to completely rule out the possibility that it plays a role. It is possible that doctors put forth effort in dimensions beyond treatment intensity that are not directly observable by the researcher. It is also possible that there exists an interplay between the coordination channel and other mechanisms, such as increased motivation, such that the baseline effects are magnified or dampened. Identifying these effects is beyond the scope of this paper and a possible direction for future work.

[51] To avoid any mechanical correlation between actual mortality and this interaction term due to the correlation between actual and predicted mortality, we include predicted mortality in the regression as an additional covariate.

8 Allocation Policies and Welfare

In this section, we are interested in understanding what types of policies are more cost-effective in improving welfare. We compare the consequences of two different policies that alter physicians' work schedules and thus average proceduralist-physician expertise overlap. We next describe the framework used for this analysis and discuss the normative implications of our findings.

8.1 Framework

We follow recommendations from Finkelstein and Hendren (2020) and use a simple framework that compares the benefits of a policy to its net costs to the government. The benefits include the number of lives saved by the intervention. We assign a monetary value to each life saved using estimates of the value of a statistical life (VSL). Using Brazilian data, Corbi et al. (2006) provide estimates of the value of a life year that range between R\$18599.7 and R\$143074.7 (or US\$3611.5-15696.5 at an exchange rate of 5.15R\$/US\$) in 2022 R\$. We use the mid-point of these estimates as a benchmark. Based on the medical literature, we assume that every patient who lives to 30 days after PCI treatment lives to the age corresponding to the country's life expectancy at birth (Cutlip and Fischman, 2018). Given that the mean age in our sample is 62 and life expectancy at birth during our study period is 75, this represents a gain of 13 years in life expectancy. Therefore, the value of a statistical life for the typical surviving patient is approximately R\$1 million ($\approx$ 80837.2$\times$13).

On the cost side, we consider the "mechanical" costs of policy implementation and any fiscal externalities from changes in teams' performance on the government budget. The costs of policy implementation depend on the type of policy at hand. For example, if the government implements an intervention designed to attract additional physicians by offering higher salaries, the total payments to the additional physicians would represent the costs of policy implementation. For simplicity, we ignore any administrative cost involved in a given intervention. Regarding the fiscal externality, a major source in our setting is hospital spending. As shown in Section 7, team members with similar expertise achieve better outcomes using fewer medical resources. This decline in resource use (and thus hospital spending) represents a positive externality on the government budget.

We compare the current allocation of physicians to counterfactual allocations generated by a given policy. Let denote survival benefit of patient i treated by proceduralist j and a set of physicians $\mathcal{K}$ by $v_i(\cdot)$, hospital spending on a patient by $C_i(\cdot)$, mechanical policy costs by $B(\cdot)$, and current allocation of physicians by Ω. The effect of a policy that generates an alternative allocation of physicians on social welfare $W(\cdot)$ is given by:

$$\begin{aligned} W(\Omega') - W(\Omega) = & \underbrace{\sum_{i\in\mathcal{I}}\Big[v_i(j,\mathcal{K}(\Omega')) - v_i(j,\mathcal{K}(\Omega))\Big]}_{\text{change in patient welfare}} \\ & - \underbrace{\sum_{i\in\mathcal{I}}\Big[C_i(j,\mathcal{K}(\Omega')) - C_i(j,\mathcal{K}(\Omega))\Big]}_{\text{fiscal externality}} \\ & - \underbrace{B(\Delta\Omega)}_{\text{mechanical policy cost}} \end{aligned} \tag{9}$$

Any alternative allocation that leads to greater proceduralist-physician expertise overlap would have positive effects on social welfare by increasing the number of lives saved and reducing hospital spending. If these effects are large enough to outweigh the mechanical costs of policy implementation, then it could improve overall social welfare.

For ease of interpretation, we normalize expression (9) by calculating the marginal value of public funds (MVPF) —in the terminology of Hendren (2016). It is defined as the ratio of marginal benefits to marginal costs of a policy:

$$MVPF = \frac{\sum_{i \in \mathcal{I}} \Big[v_i(j, \mathcal{K}(\Omega')) - v_i(j, \mathcal{K}(\Omega)) \Big]}{\sum_{i \in \mathcal{I}} \Big[C_i(j, \mathcal{K}(\Omega')) - C_i(j, \mathcal{K}(\Omega)) \Big] + B(\Delta\Omega)} \tag{10}$$

It measures the gains in lives saved for every R$1 spent on the intervention of interest. An MVPF above 1 would suggest that the intervention is cost-effective using this standard.

Counterfactual allocations. We consider two different allocation policies: i) increasing the number of hours worked up to 100 percent for a subset of physicians who have above-average overlap in specialties with proceduralists; and ii) reallocating physicians with below-average overlap in specialties with proceduralists across the regular days of work.[52] These alternative allocations are implemented within (and not between) hospitals. Appendix E provides details on how we perform these allocations and calculate the simulated counterfactual mortality and hospital spending reductions. In the most practical terms, the first policy is an *intensity-based allocation* that increases work intensity, whereas the second is a *nonintensity-based allocation* that alters the regular workdays with no consequence on work intensity.

Condition 3.1 (No Spillovers). *Potential 30-day mortality outcomes of non-PCI patients are mean independent of the intensity-based and non-intensity-based allocation policies.*

In simple terms, condition 3.1 implies that, on average, the outcomes of non-PCI patients are unaffected by the counterfactual allocation policies. This assumption is made only for simplicity, and we recognize that it could not hold exactly in some cases. By increasing the availability of doctors with similar expertise, non-PCI patients are also more exposed to such teams. To the extent that non-PCI patients exposed to teams with greater overlap also experience better outcomes just like patients undergoing PCI, then our analysis would *underestimate* the true aggregate gains in patient welfare and so our results should be taken to be lower bounds.

To compute the mechanical costs of policy implementation, we use information from the Census conducted in 2010 to obtain information on physicians' hourly wages. To implement the intensity-based policy, the government would need to increase the hourly wage such that physicians decide to increase their labor supply. This requires that we make assumptions on the labor supply elasticity. Chetty (2012) reviews existing estimates in the literature and finds an approximate intensive margin elasticity of 0.33, which is fairly close to that found in previous studies focusing on physicians' labor supply decisions.[53]We use this estimate as an approximate

[52] As described in Appendix E, for each physician reallocated from day d to d', we randomly choose one corresponding "control" physician to be reallocated from day d' to d. This ensures that the number of available physicians on each day of the week is not affected.

[53] For example, Showalter and Thurston (1997) exploit geographic variation in tax rates in the United States and find a labor supply elasticity for physicians of 0.30.

benchmark. The mechanical policy costs for the nonintensity-based intervention are different because this intervention does not imply an increase in the overall workload. The hourly wages are the same on the regular days of the week but are generally higher on holidays and weekends. To be conservative, we assume that physicians reallocated to holidays and weekends receive a 100-percent increase in their wages on those days.

8.2 Normative Findings and Discussion

The results are presented in Table 4. The intensity-based allocation would save a total of 114 lives (with a societal value of approximately R$119 million) with a positive fiscal externality of R$4.2 million from a decline in hospital spending. However, the costs of implementation are excessively high. Implementing this intervention would cost as much as R$6 billion, yielding an MVPF well below 1. For every R$1 spent on this intervention, the marginal benefit generated is only R$0.02.[54] Even if we used the highest value of a statistical life available, the MVPF would be far from 1 at 0.03. The numbers for the nonintensity-based intervention are more favorable. The gains in terms of VSL and hospital spending reductions are not only larger (R$611 and R$21 million respectively), but the costs of policy implementation are substantially smaller (R$53 million). The money saved on hospital spending alone would be enough to cover almost 40 percent of the program's costs. These calculations imply an MVPF of 19.23, and MVPF would continue to be well above 1 even if we used a VSL in the lower range of available estimates in the literature.

These estimates are far from perfect. They may not completely capture other possible externalities, both negative and positive. For example, patients surviving PCI treatment may generate positive externalities on the government budget if they participate in the labor force and thus contribute to the tax system. Moreover, the welfare gains in terms of VSL depend on the exact estimates and specific assumptions on the gains in life expectancy. Nevertheless, at a broad level, our analysis suggests that policies that alter the specific days of work are much more cost-effective than interventions seeking to increase workload intensity. A relatively small value of a statistical life would be enough to justify the former from a cost-effectiveness perspective, but the latter would require an implausibly large VSL.

9 Conclusion

Teams are a widespread feature of modern organizations. An important question is whether teams should be composed of heterogeneous specialists or members with similar expertise. A powerful and intuitively appealing view asserts that diversity in specialized skills and knowledge generates complementarities that should enhance creativity, task completion, decision-making, and ultimately team performance. But heterogeneity in expertise also has the potential to significantly raise the costs of coordination and reduce the quality of collaboration between team members. Answers to this question have remained elusive due to a lack of detailed microdata and endogeneity challenges. We study this important question in the context of the healthcare industry using newly constructed administrative dataset from Brazil and a within-proceduralist identification strategy. Our results indicate that overlap in specialties between the doctor who performs a heart procedure and those who provide pre- and post-procedure care leads to a lower patient mortality rate. Consistent with the improved coordination

[54] If we assume that physicians are altruists and are willing to increase the number of hours worked without altering their wage rate, the costs of policy implementation would be approximately R$1 billion. This yields an MVPF of 0.12, still well below 1.

hypothesis, we find that teams whose members have areas of expertise in common use medical resources with lower intensity and achieve better patient outcomes, particularly when team members have less accumulated experience working together, and the tasks are more complex.

We use our estimates to evaluate the welfare consequences of two counterfactual policies that alter the allocation of physicians to increase teams' expertise overlap. We consider an intensity-based policy that increases the number of hours worked for a subset of physicians and another non-intensity-based intervention that simply reallocates physicians across days within the same week. These types of policies have been the subject of intense debate in policy circles, not only in the context of the healthcare sector but also in the context of interventions intended to improve managerial practices in general. Our calculations suggest that the non-intensity-based intervention yields better outcomes at a substantially lower cost. These findings contribute to our understanding of team dynamics and have important implications for the organization of teams to achieve higher productivity.

Bibliography

Adam Smith, *An inquiry into the nature and causes of the wealth of nation*, London: W. Strahan and T. Cadell; reprinted, New York: Modern Library, 1937, 1776.

Altonji, Joseph G, Todd E Elder, and Christopher R Taber, "Selection on observed and unobserved variables: Assessing the effectiveness of Catholic schools," *Journal of political economy*, 2005, *113* (1), 151–184.

Amin, Amit P, Mark Patterson, John A House, Helmut Giersiefen, John A Spertus, Dmitri V Baklanov, Adnan K Chhatriwalla, David M Safley, David J Cohen, Sunil V Rao et al., "Costs associated with access site and same-day discharge among medicare beneficiaries undergoing percutaneous coronary intervention: an evaluation of the current percutaneous coronary intervention care pathways in the United States," *JACC: Cardiovascular Interventions*, 2017, *10* (4), 342–351.

Angrist, Joshua D, Guido W Imbens, and Donald B Rubin, "Identification of causal effects using instrumental variables," *Journal of the American statistical Association*, 1996, *91* (434), 444–455.

Arrow, Kenneth J, *The limits of organization*, WW Norton & Company, 1974.

Bandiera, Oriana, Iwan Barankay, and Imran Rasul, "Social preferences and the response to incentives: Evidence from personnel data," *The Quarterly Journal of Economics*, 2005, *120* (3), 917–962.

_ , _ , and _ , "Team incentives: Evidence from a firm level experiment," *Journal of the European Economic Association*, 2013, *11* (5), 1079–1114.

Banerjee, Abhijit, Amy Finkelstein, Rema Hanna, Benjamin A Olken, Arianna Ornaghi, and Sudarno Sumarto, "The challenges of universal health insurance in developing countries: Experimental evidence from Indonesia's national health insurance," *American Economic Review*, 2021, *111* (9), 3035–63.

Battiston, Diego, Jordi Blanes i Vidal, and Tom Kirchmaier, "Face-to-face communication in organizations," *The Review of Economic Studies*, 2021, *88* (2), 574–609.

Becker, Gary S and Kevin M Murphy, "The division of labor, coordination costs, and knowledge," *The Quarterly journal of economics*, 1992, *107* (4), 1137–1160.

Belloni, Alexandre, Victor Chernozhukov, and Christian Hansen, "High-dimensional methods and inference on structural and treatment effects," *Journal of Economic Perspectives*, 2014, *28* (2), 29–50.

Benkard, C Lanier, "Learning and forgetting: The dynamics of aircraft production," *American Economic Review*, 2000, *90* (4), 1034–1054.

Brazilian Ministry of Health, "Contas de saúde na perspectiva da contabilidade internacional: conta SHA para o Brasil, 2015 a 2019," 2022.

Bunderson, J Stuart and Kathleen M Sutcliffe, "Comparing alternative conceptualizations of functional diversity in management teams: Process and performance effects," *Academy of management journal*, 2002, *45* (5), 875–893.

Chan, David C, "Teamwork and moral hazard: evidence from the emergency department," *Journal of Political Economy*, 2016, *124* (3), 734–770.

_ , "Influence and information in team decisions: evidence from medical residency," *American Economic Journal: Economic Policy*, 2021, *13* (1), 106–37.

_ , Matthew Gentzkow, and Chuan Yu, "Selection with variation in diagnostic skill: Evidence from radiologists," *The Quarterly Journal of Economics*, 2022, *137* (2), 729–783.

Chandra, Amitabh and Douglas O Staiger, "Productivity spillovers in health care: evidence from the treatment of heart attacks," *Journal of political Economy*, 2007, *115* (1), 103–140.

_ and _ , "Identifying sources of inefficiency in healthcare," *The quarterly journal of economics*, 2020, *135* (2), 785–843.

Chen, Yiqun, "Team-specific human capital and team performance: Evidence from doctors," *American Economic Review*, 2021, *111* (12), 3923–62.

Chetty, Raj, "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 2012, *80* (3), 969–1018.

Clemens, Jeffrey and Joshua D Gottlieb, "Do physicians' financial incentives affect medical treatment and patient health?," *American Economic Review*, 2014, *104* (4), 1320–49.

Corbi, Raphael, Naércio Menezes-Filho, Rodrigo R Soares, and Sérgio Ribeiro da Costa Werlang, "Avaliação econômica de ganhos sociais na área da saúde–estimativas do valor de uma vida estatística para o Brasil," *Unpublished Manuscript, Pontifícia Universidade Católica do Rio de Janeiro*, 2006.

Cornelissen, Thomas, Christian Dustmann, and Uta Schönberg, "Peer effects in the workplace," *American Economic Review*, 2017, *107* (2), 425–56.

Cutlip, Donald E and David L Fischman, "Mortality after percutaneous coronary intervention: narrowing the knowledge gap," 2018.

de Oliveira, Gláucia Maria Moraes, Luisa Campos Caldeira Brant, Carisi Anne Polanczyk, Deborah Carvalho Malta, Andreia Biolo, Bruno Ramos Nascimento, Maria de Fatima Marinho de Souza, Andrea Rocha De Lorenzo, Antonio Aurélio de Paiva Fagundes Júnior, Beatriz D Schaan et al., "Estatística Cardiovascular–Brasil 2021," *Arquivos Brasileiros de Cardiologia*, 2022, *118*, 115–373.

Delfgaauw, Josse, Robert Dur, Oke Onemu, and Joeri Sol, "Team incentives, social cohesion, and performance: A natural field experiment," *Management Science*, 2022, *68* (1), 230–256.

Dinerstein, Michael, Rigissa Megalokonomou, and Constantine Yannelis, "Human Capital Depreciation and Returns to Experience," *American Economic Review*, 2022, *112* (11), 3725–3762.

Doyle, Joseph J, "Physician characteristics and patient survival: Evidence from physician availability," Technical Report, National Bureau of Economic Research 2020.

_ and Becky Staiger, "Physician Group Influences on Treatment Intensity and Health: Evidence from Physician Switchers," Technical Report, National Bureau of Economic Research 2022.

D'Avila, Alan Castro, Renato Roese Filho, Marcia Moura Schmidt, Karina Melleu, Cristiano de Oliveira Cardoso, Carlos Antonio Mascia Gottschall, Alexandre Schann de Quadros, and Alexandre Damiani Azmus, "Angioplastia coronariana primária em pacientes com mais de 80 anos," *Revista Brasileira de Cardiologia Invasiva*, 2015, *23* (4), 261–265.

Einav, Liran, Amy Finkelstein, and Neale Mahoney, "Provider incentives and healthcare costs: Evidence from long-term care hospitals," *Econometrica*, 2018, *86* (6), 2161–2219.

Finkelstein, Amy and Nathaniel Hendren, "Welfare analysis meets causal inference," *Journal of Economic Perspectives*, 2020, *34* (4), 146–67.

_ , Matthew Gentzkow, and Heidi Williams, "Sources of geographic variation in health care: Evidence from patient migration," *The quarterly journal of economics*, 2016, *131* (4), 1681–1726.

Fisher, Elliott S, David E Wennberg, Thrse A Stukel, Daniel J Gottlieb, F Lee Lucas, and Etoile L Pinder, "The implications of regional variations in Medicare spending. Part 1: the content, quality, and accessibility of care," *Annals of internal medicine*, 2003, *138* (4), 273–287.

_ , _ , _ , _ , _ , and _ , "The implications of regional variations in Medicare spending. Part 2: health outcomes and satisfaction with care," *Annals of internal medicine*, 2003, *138* (4), 288–298.

Friebel, Guido, Matthias Heinz, Miriam Krueger, and Nikolay Zubanov, "Team incentives and performance: Evidence from a retail chain," *American Economic Review*, 2017, *107* (8), 2168–2203.

Gruber, Jonathan and Samuel A Kleiner, "Do strikes kill? Evidence from New York state," *American Economic Journal: Economic Policy*, 2012, *4* (1), 127–57.

Haig, Kathleen M, Staci Sutton, and John Whittington, "SBAR: a shared mental model for improving communication between clinicians," *The joint commission journal on quality and patient safety*, 2006, *32* (3), 167–175.

Heckman, James J and Edward Vytlacil, "Structural equations, treatment effects, and econometric policy evaluation 1," *Econometrica*, 2005, *73* (3), 669–738.

_ and V Joseph Hotz, "Choosing among alternative nonexperimental methods for estimating the impact of social programs: The case of manpower training," *Journal of the American statistical Association*, 1989, *84* (408), 862–874.

Hendren, Nathaniel, "The policy elasticity," *Tax Policy and the Economy*, 2016, *30* (1), 51–89.

Hjort, Jonas, "Ethnic divisions and production in firms," *The Quarterly Journal of Economics*, 2014, *129* (4), 1899–1946.

Hoisl, Karin, Marc Gruber, and Annamaria Conti, "R&D team diversity and performance in hypercompetitive environments," *Strategic management journal*, 2017, *38* (7), 1455–1477.

Holmstrom, Bengt, "Moral hazard in teams," *The Bell journal of economics*, 1982, pp. 324–340.

Jaravel, Xavier, Neviana Petkova, and Alex Bell, "Team-specific capital and innovation," *American Economic Review*, 2018, *108* (4-5), 1034–73.

Jarosch, Gregor, Ezra Oberfield, and Esteban Rossi-Hansberg, "Learning from coworkers," *Econometrica*, 2021, *89* (2), 647–676.

Kellogg, Ryan, "Learning by drilling: Interfirm learning and relationship persistence in the Texas oilpatch," *The Quarterly Journal of Economics*, 2011, *126* (4), 1961–2004.

Kleibergen, Frank and Richard Paap, "Generalized reduced rank tests using the singular value decomposition," *Journal of econometrics*, 2006, *133* (1), 97–126.

Lazear, Edward P, "Globalisation and the market for team-mates," *The Economic Journal*, 1999, *109* (454), 15–40.

_ and Kathryn L Shaw, "Personnel economics: The economist's view of human resources," *Journal of economic perspectives*, 2007, *21* (4), 91–114.

Lee, David S, Justin McCrary, Marcelo J Moreira, and Jack Porter, "Valid t-ratio Inference for IV," *American Economic Review*, 2022, *112* (10), 3260–90.

Levine, Glenn N, Eric R Bates, James C Blankenship, Steven R Bailey, John A Bittl, Bojan Cercek, Charles E Chambers, Stephen G Ellis, Robert A Guyton, Steven M Hollenberg et al., "2011 ACCF/AHA/SCAI guideline for percutaneous coronary intervention: a report of the American College of Cardiology Foundation/American Heart Association Task Force on Practice Guidelines and the Society for Cardiovascular Angiography and Interventions," *Journal of the American College of Cardiology*, 2011, *58* (24), e44–e122.

Lingard, Lorelei, Sherry Espin, Sarah Whyte, Glenn Regehr, G Ross Baker, Richard Reznick, John Bohnen, Beverly Orser, Diane Doran, and Ellen Grober, "Communication failures in the operating room: an observational classification of recurrent types and effects," *BMJ Quality & Safety*, 2004, *13* (5), 330–334.

Lyons, Elizabeth, "Team production in international labor markets: Experimental evidence from the field," *American Economic Journal: Applied Economics*, 2017, *9* (3), 70–104.

Machado, Guilherme Pinheiro, Gustavo Neves de Araújo, Stéfani Mariani, Elvis Pellin Cassol, Felipe Homem Valle, Ana Maria Rocha Krepsky, Luiz Carlos Corsetti Bergoli, Sandro Cadaval Gonçalves, Rodrigo Vugman Wainstein, and Marco Vugman Wainstein, "On-vs. off-hours admission of patients with ST-elevation acute myocardial infarction undergoing percutaneous coronary interventions: data from a tertiary university Brazilian hospital," *Clinical and biomedical research. Porto Alegre. Vol. 38, n. 1 (2018), p. 30-34*, 2018.

Marx, Benjamin, Vincent Pons, and Tavneet Suri, "Diversity and team performance in a Kenyan organization," *Journal of Public Economics*, 2021, *197*, 104332.

Mas, Alexandre and Enrico Moretti, "Peers at work," *American Economic Review*, 2009, *99* (1), 112–45.

Mazzocco, Karen, Diana B Petitti, Kenneth T Fong, Doug Bonacum, John Brookey, Suzanne Graham, Robert E Lasky, J Bryan Sexton, and Eric J Thomas, "Surgical team behaviors and patient outcomes," *The American Journal of Surgery*, 2009, *197* (5), 678–685.

Menees, Daniel S, Eric D Peterson, Yongfei Wang, Jeptha P Curtis, John C Messenger, John S Rumsfeld, and Hitinder S Gurm, "Door-to-balloon time and mortality among patients undergoing primary PCI," *N Engl J Med*, 2013, *369*, 901–909.

Oster, Emily, "Unobservable selection and coefficient stability: Theory and evidence," *Journal of Business & Economic Statistics*, 2019, *37* (2), 187–204.

Paim, Jairnilson, Claudia Travassos, Celia Almeida, Ligia Bahia, and James Macinko, "The Brazilian health system: history, advances, and challenges," *The Lancet*, 2011, *377* (9779), 1778–1797.

Risser, Daniel T, Matthew M Rice, Mary L Salisbury, Robert Simon, Gregory D Jay, Scott D Berns, MedTeams Research Consortium et al., "The potential for improved teamwork to reduce medical errors in the emergency department," *Annals of emergency medicine*, 1999, *34* (3), 373–383.

Salas, Eduardo, Katherine A Wilson, Carrie E Murphy, Heidi King, and Mary Salisbury, "Communicating, coordinating, and cooperating when lives depend on it: tips for teamwork," *The Joint Commission Journal on Quality and Patient Safety*, 2008, *34* (6), 333–341.

Sandvik, Jason J, Richard E Saouma, Nathan T Seegert, and Christopher T Stanton, "Workplace knowledge flows," *The Quarterly Journal of Economics*, 2020, *135* (3), 1635–1680.

Showalter, Mark H and Norman K Thurston, "Taxes and labor supply of high-income physicians," *Journal of Public Economics*, 1997, *66* (1), 73–97.

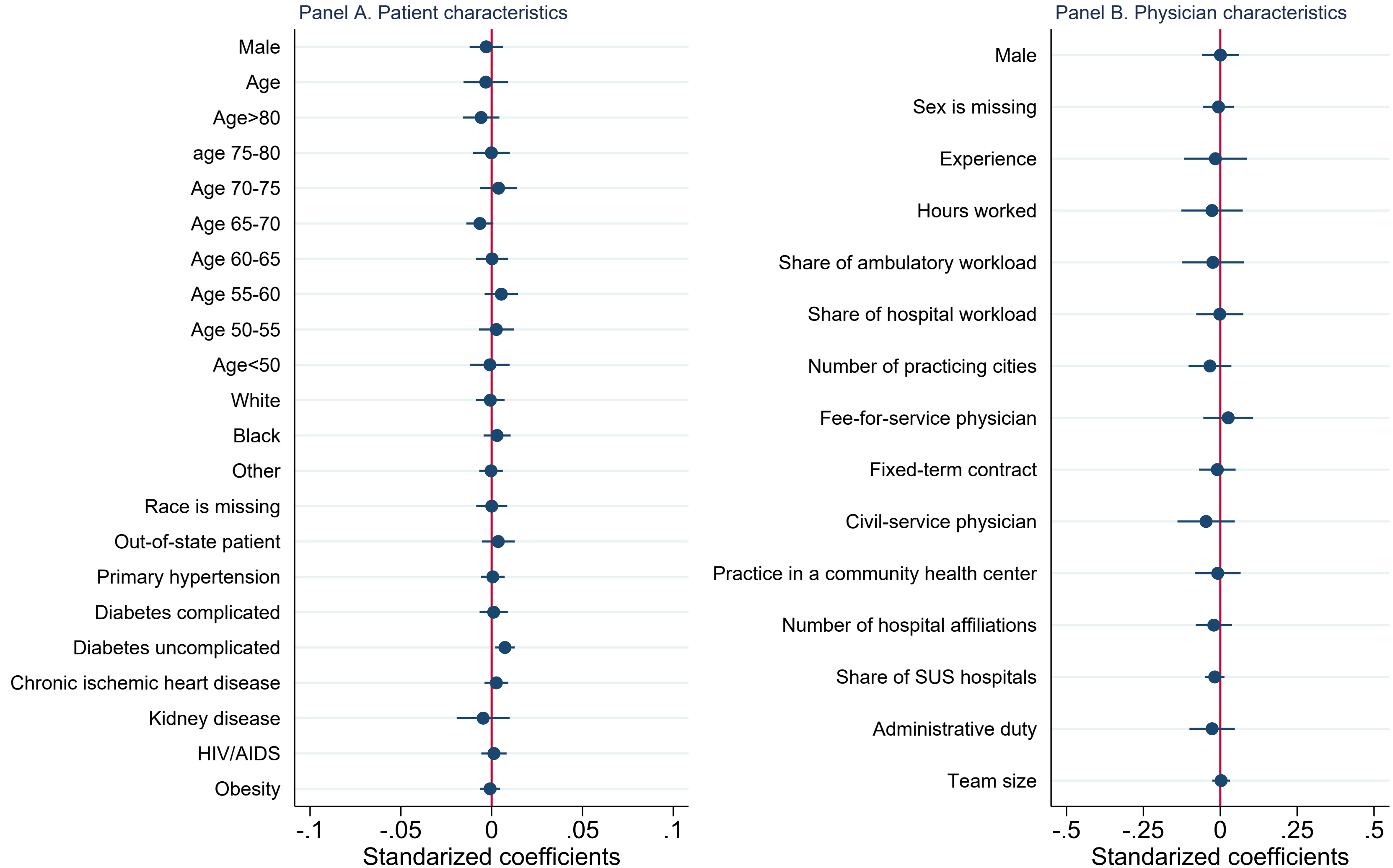
Silver, David, "Haste or Waste? Peer pressure and productivity in the emergency department," *The Review of Economic Studies*, 2021, *88* (3), 1385–1417.

Skinner, Jonathan, "Causes and consequences of regional variations in health care," in "Handbook of health economics," Vol. 2, Elsevier, 2011, pp. 45–93.

Stehli, Julia, Catherine Martin, Angela Brennan, Diem T Dinh, Jeffrey Lefkovits, and Sarah Zaman, "Sex differences persist in time to presentation, revascularization, and mortality in myocardial infarction treated with percutaneous coronary intervention," *Journal of the American Heart Association*, 2019, *8* (10), e012161.

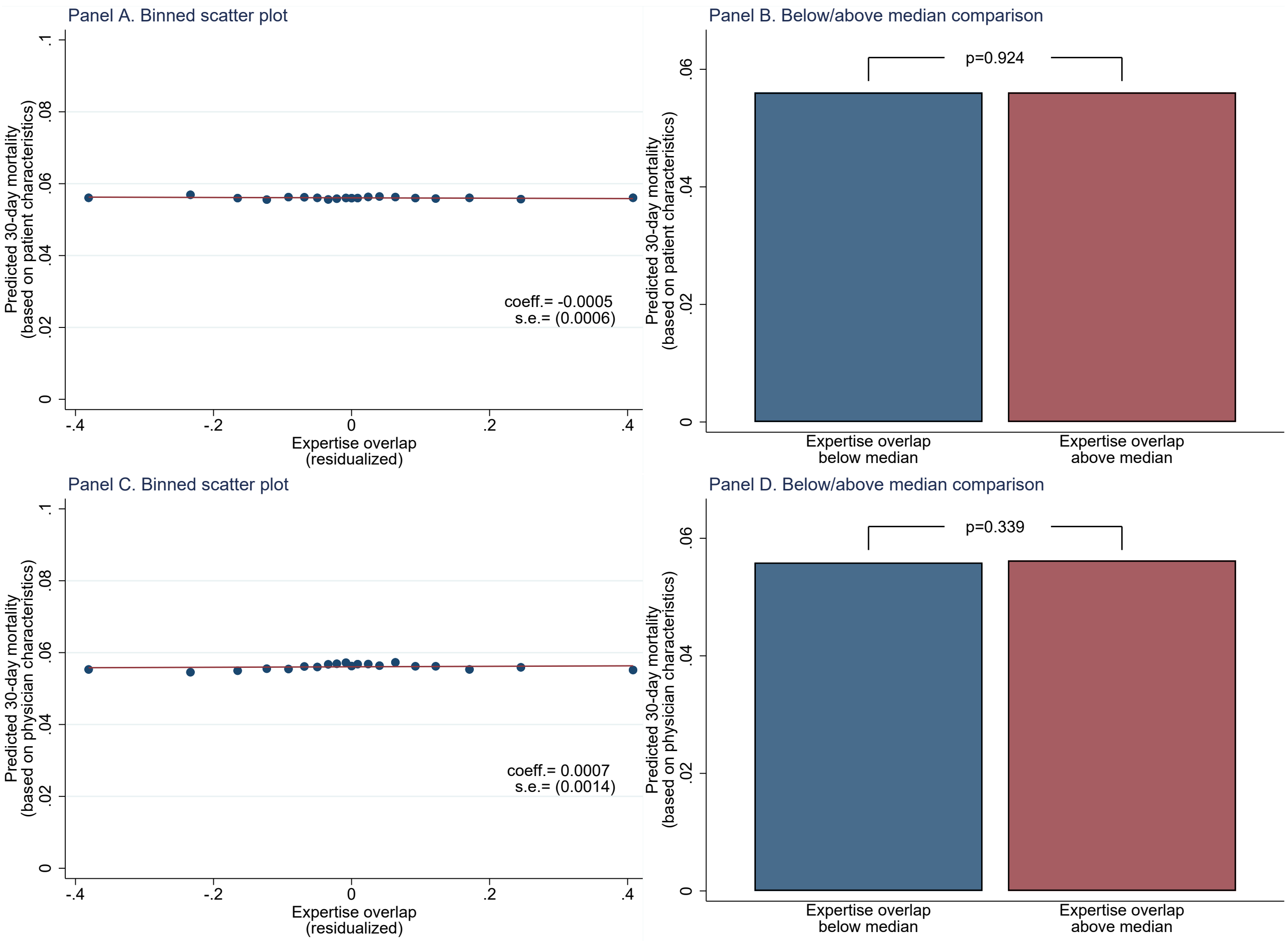
Yadava, OP, "Heart Team-The Reality?," *Brazilian Journal of Cardiovascular Surgery*, 2018, *33*, III–IV.

Figure 1: Covariate Balance



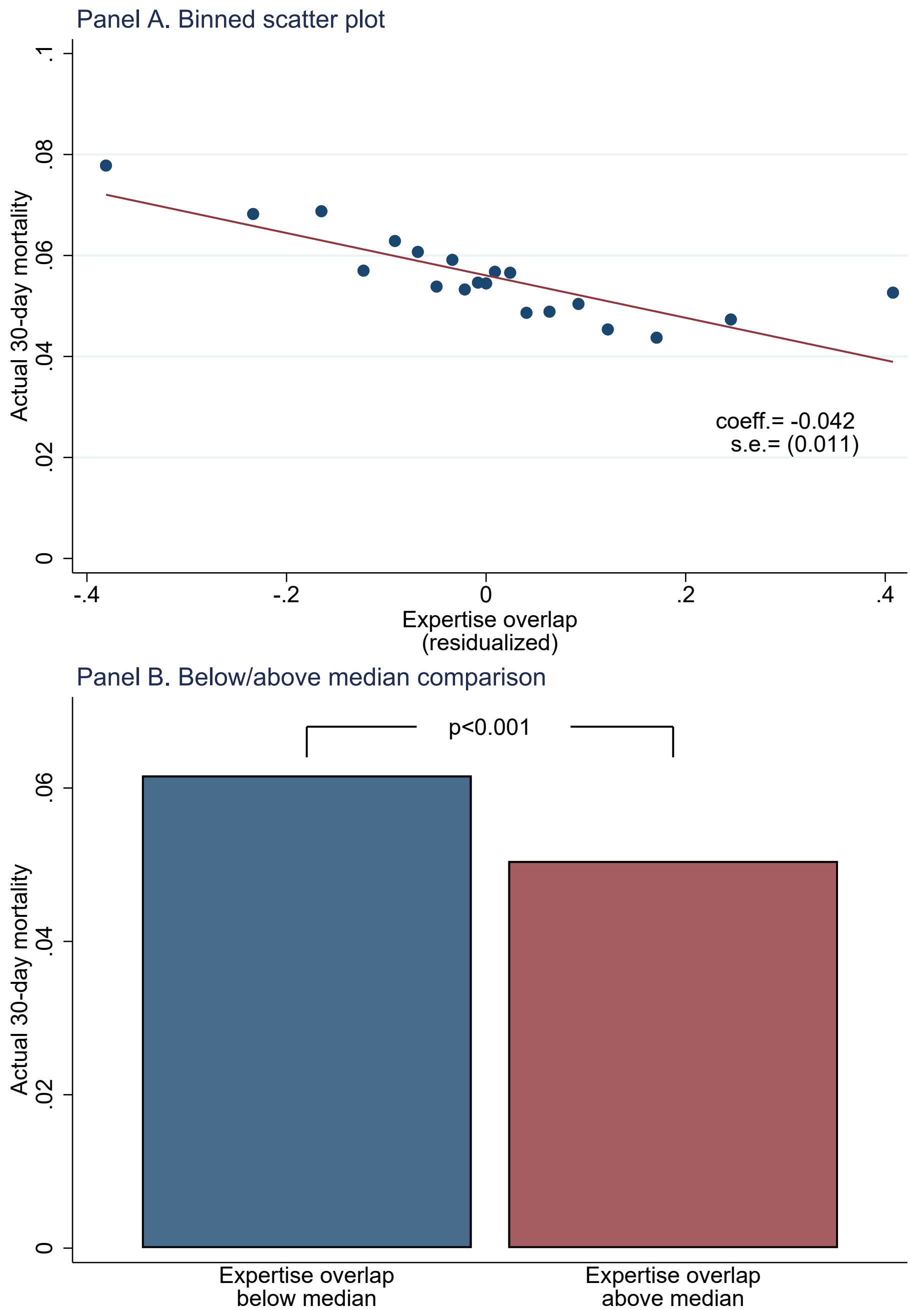
Notes: Each coefficient is from a different regression where expertise overlap is the independent variable of interest and the predetermined characteristics are the dependent variables. The regressions include proceduralist-year and physician's case-related specialty fixed effects. All the coefficients are standardized for ease of readability. Standard errors are clustered at the hospital level.

Figure 2: Balance in Predicted 30-Day Mortality



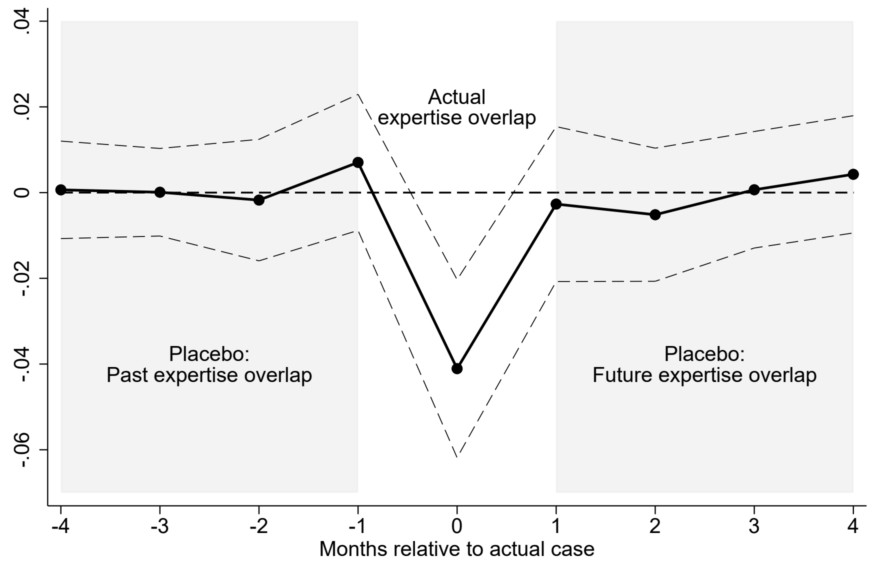
Notes: Panels A and C are binned scatter plots of the relationship between expertise overlap and predicted 30-day mortality. Panels C and D plot average predicted 30-day mortality of patients treated by teams with expertise overlap below versus above the median. In all panels, the predicted 30-day mortality is residualized with respect to proceduralist-year and physician's case-related specialty fixed effects. The unconditional mean of the predicted 30-day mortality is added back for ease of comparison. In panels A and B, the predicted mortality is obtained from a probit model of 30-day mortality on all the patient characteristics displayed in Figure 1. In panels C and D, the predicted mortality is obtained from a probit model of 30-day mortality on all the physician characteristics displayed in Figure 1. Standard errors are clustered at the hospital level.

Figure 3: Relationship Between Expertise Overlap and 30-Day Mortality



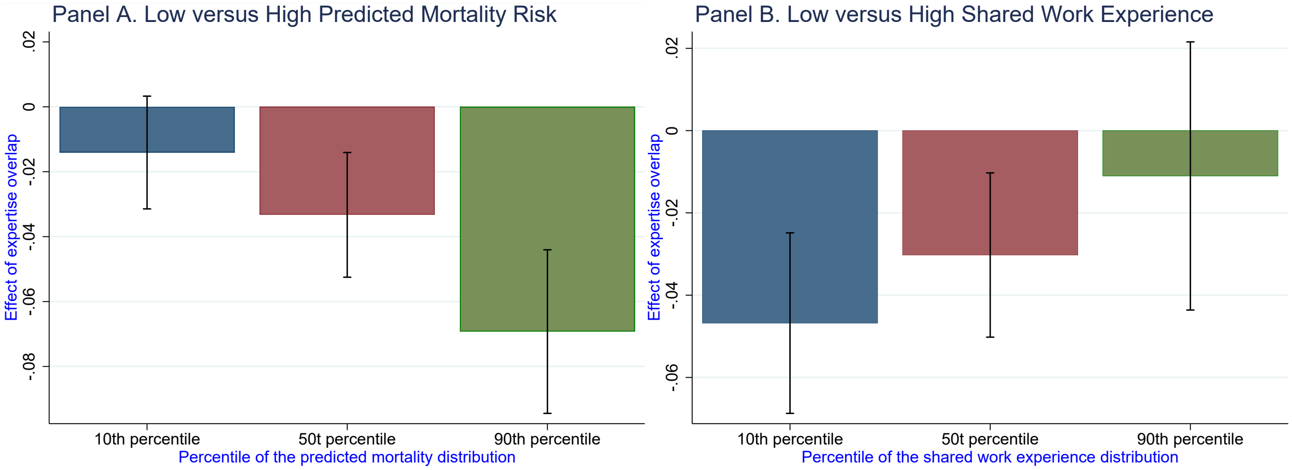
Notes. Panel A is a binned scatter plot of the relationship between 30-day mortality and expertise overlap. Panel B plots average 30-day mortality of patients treated by teams with expertise overlap below versus above the median. The 30-day mortality is residualized with respect to proceduralist-year and physician's case-specialty fixed effects. The unconditional mean of the 30-day mortality is added back for ease of interpretation. Standard error clustered at the hospital level.

Figure 4: Effects of Past, Actual and Future Expertise Overlap



Notes: This figure presents estimates of the effects of past, actual, and present expertise overlap. Each coefficient is from a different regression. All models include the full set of controls (see notes to Table 1 for details). 95 percent confidence intervals are based on standard errors clustered at the hospital level.

Figure 5: Heterogeneity



Notes. This figure reports estimates of the effects of expertise overlap on mortality considering interactions with predicted mortality (Panel A) and shared work experience (Panel B). Estimates are based on specification (8). The predicted mortality is obtained from estimating a logit model of patient actual mortality on patient characteristics, including indicators for common comorbidities. To increase the predictive power of the model, we include a full set of interactions between all the patient characteristics. Shared work experience is defined as in Chen (2021): the number of times that proceduralists and physicians have worked together in the past two years. Since patients often see more than one physician, the shared work experience for a patient is constructed as the weighted average of the shared work experience between the proceduralist and each physicians treating the patient, where the weights are the share of hospital visits provided by each physician to that patient. 95 confidence intervals are based on standard errors clustered at the hospital level.

Table 1: Expertise Overlap and 30-Day Mortality

	Dependent variable is 30-day mortality				
	(1)	(2)	(3)	(4)	(5)
Expertise overlap	-0.0419	-0.041	-0.0415	-0.0425	-0.041
	[0.0107]***	[0.0107]***	[0.0107]***	[0.0107]***	[0.0106]***
Mean of dep. variable	0.056	0.056	0.056	0.056	0.056
Observations	176108	176108	176108	176108	176108
Proceduralist × year FE	✓	✓	✓	✓	✓
Physician's case-related specialty FE	✓	✓	✓	✓	✓
Hospital × time FE		✓			✓
Patient characteristics			✓		✓
Physician characteristics				✓	✓

Notes. This table presents results from estimating different versions of model (3). The expertise overlap is defined in equations (1) and (2). In all columns we control for proceduralist-year and physician's case-related specialty fixed effects. Since patients often see more than one physician, the pysician's case-related specialty indicators are constructed as weighted averages of the physician's case-related specialty indicators, where the weights are the share of hospital visits provided by each physician to that patient. Column (2) includes hospital fixed effects interacted with day-of-week and month indicators. Column (3) adds controls for the patient characteristics: indicators for age groups, race, gender, out-state visitor, and specific comorbidities. These comorbidities include primary hypertension, diabetes complicated, diabetes uncomplicated, chronic ischemic heart disease, kidney disease, HIV/AIDS, and obesity. Column (4) controls for physician characteristics: gender, experience, hours worked, fee-for-service contract, fixed-term contract, civil-service employment, practice in a community health center, administrative duty, share of ambulatory workload, share of hospital workload, number of practicing cities, number of hospital affiliations, share of SUS hospitals, and team size fixed effects. Column (5) control for all variables simultaneously. Standard errors are clustered at the hospital level.

* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

Table 2: Expertise Overlap and 30-Day Mortality:
First Stage, Reduced Form, and 2SLS Estimates

	Dependent variable is:		
	Expertise overlap	30-day mortality	
	(First Stage)	(Reduced-Form)	(2SLS)
	(1)	(2)	(3)
Expertise overlap			-0.0447
			[0.0199]**
Simulated expertise overlap	0.5399	-0.0241	
	[0.0588]***	[0.0113]**	
Kleibergen and Paap (2006) F statistics			84.1518
Mean of dep. variable	0.4288	0.056	0.056
Observations	175349	175349	175349
Patient characteristics	✓	✓	✓
Physician characteristics	✓	✓	✓
Hospital × month FE	✓	✓	✓
Hospital × day-of-week FE	✓	✓	✓
Date-of-admission FE	✓	✓	✓
Proceduralist FE	✓	✓	✓

Notes. This table reports first-stage, reduced form, and 2SLS coefficient estimates. The instrument for expertise overlap is the simulated measure of expertise overlap based on the availability of physicians at the time patients are admitted to the hospital. See Section 6 for details on the construction of the instrument. Patient and physician characteristics include all those displayed in Figure D.1. Standard errors are clustered at the hospital level.

* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

Table 3: Expertise Overlap and Medical Inputs

	Dependent variable is							
		Number of exam tests						
	length of stay	biochemical tests exams	hematology tests exams	laboratory tests exams	radiology exams	electro-cardiogramas	Medical input index	Hospital spending (in R$)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Expertise overlap	-0.5847	-5.4217	-1.2111	-0.5671	-0.5422	-0.3499	-0.1341	-651.8722
	[0.2551]**	[1.3773]***	[0.3720]***	[0.1732]***	[0.1729]***	[0.1594]**	[0.0367]***	[258.9633]**
Mean of dep. variable	5.7428	18.8877	5.0055	1.5795	1.1702	2.5858	1.40e-09	10619.0879
Observations	176108	176108	176108	176108	176108	176108	176108	176108
Proceduralist $\times$ year FE	✓	✓	✓	✓	✓	✓	✓	✓
Physician's case-related specialty FE	✓	✓	✓	✓	✓	✓	✓	✓
Hospital times time FE	✓	✓	✓	✓	✓	✓	✓	✓
Patient characteristics	✓	✓	✓	✓	✓	✓	✓	✓
Physician characteristics	✓	✓	✓	✓	✓	✓	✓	✓

Notes. This table presents estimates of the effects of expertise overlap on the use of medical inputs using model (3). Column 1 considers the length of hospital stay, measured in days. Columns 2 through 6 consider number of specific tests and exams. Column 7 estimates the effect of expertise overlap on the medical input index. This index is constructed by taking the unweighted average of all the standardized inputs across columns 1-6. See notes to Table 1 for a full description of the control variables used. Standard errors are clustered at the hospital level.

* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

Table 4: Welfare Analysis of Allocation Policies

Type of policy	Lives saved	Lives in VSL (in R$)	Fiscal externality: hospital spending reductions (in R$)	Policy mechanical costs (in R$)	Marginal value of public funds (MVPF)
	(1)	(2)	(3)	(4)	(5)
Increased hours of work	114	119800805	4276719	6586961501	0.02
Reallocation across days	582	611614634	21836540	53636085	19.23

Notes. This table reports results from the welfare analysis described in Section 8. Column 1 shows the number of lives saved from each type of policy. Column 2 multiplies the number of lives saved by the value of 13 life years (R$80837.25). Column 3 presents total hospital spending saved from each intervention. Column 4 documents the estimated costs of implementing each intervention. Column 5 reports the marginal value of public funds, calculated as the ratio of column (2) to (4)-(3). Values in 2022R$ (exchange rate to the US$ of approximately 5.15 R$/US$ in 2022).

Online Appendix to "Expertise Overlap and Team Productivity: Evidence from the Hospital Industry"

Danyelle Branco   Bladimir Carrillo   Di Fang   Wilman Iglesias

April 5, 2023

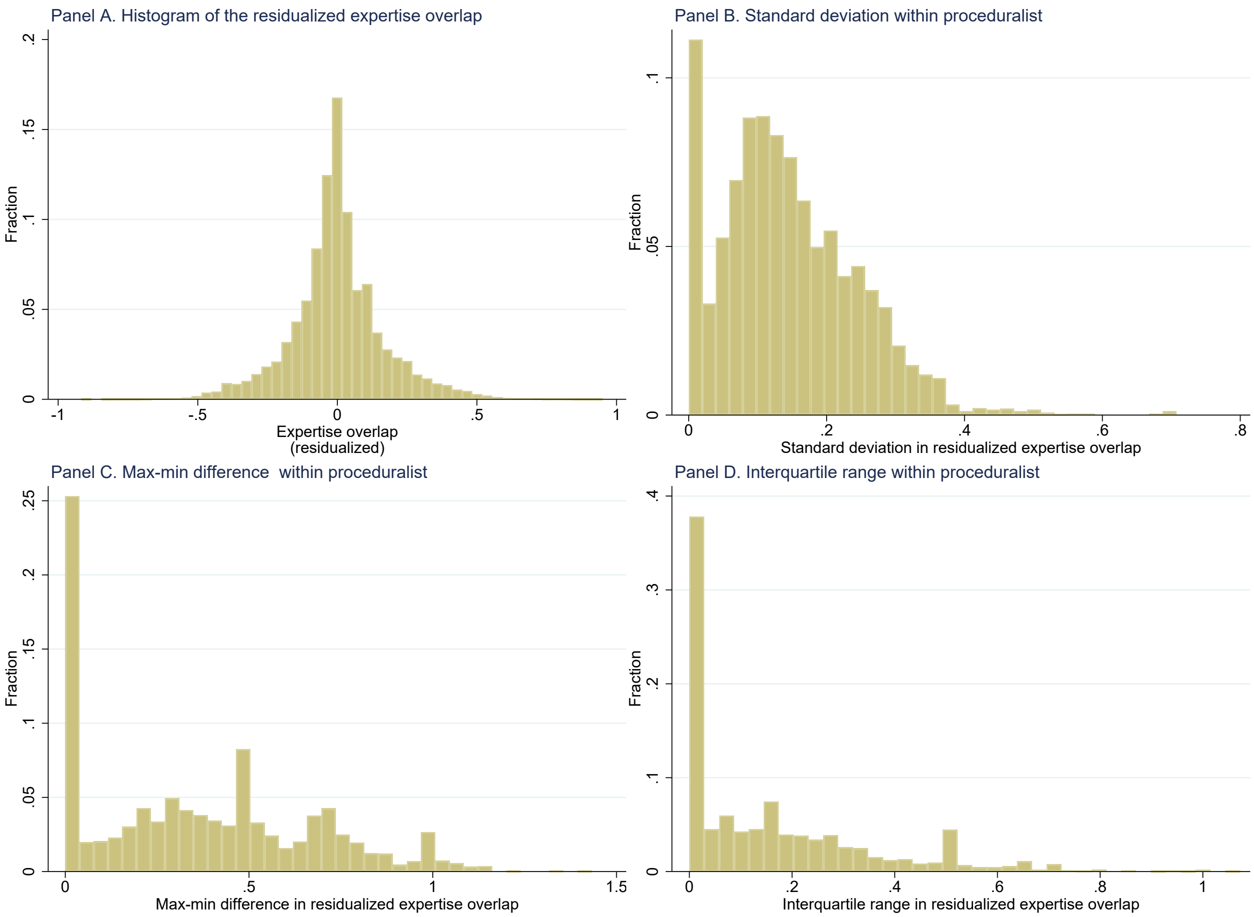
Contact information: Branco: Department of Economics, Universidade Federal de Pernambuco, Av. Marielle Franco, Caruaru-PE, 55014-900 (e-mail: danyelle.branco@ufpe.br). Carrillo: Department of Economics, Universidade Federal de Pernambuco, AV. Prof. Moraes Rego, 1235 - Cidade Universitaria, Recife - PE, 50670-420 (e-mail: bladimir.carrillo@ufpe.br). Fang (e-mail: difang@ufl.edu). Iglesias (e-mail: wi001@uark.edu). The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

List of Figures

List of Tables

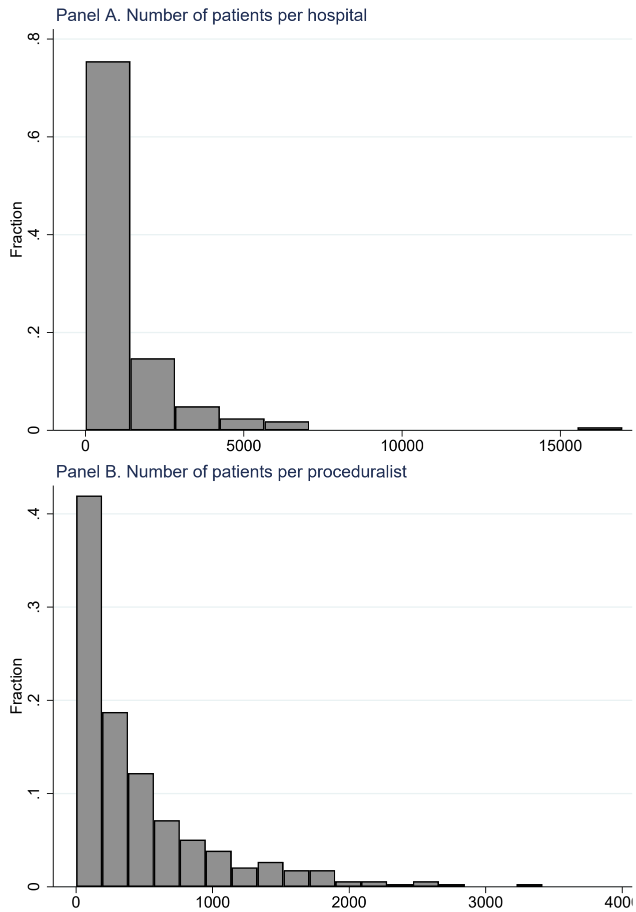
A Descriptive Statistics

Figure A.1: Variation in Expertise Overlap



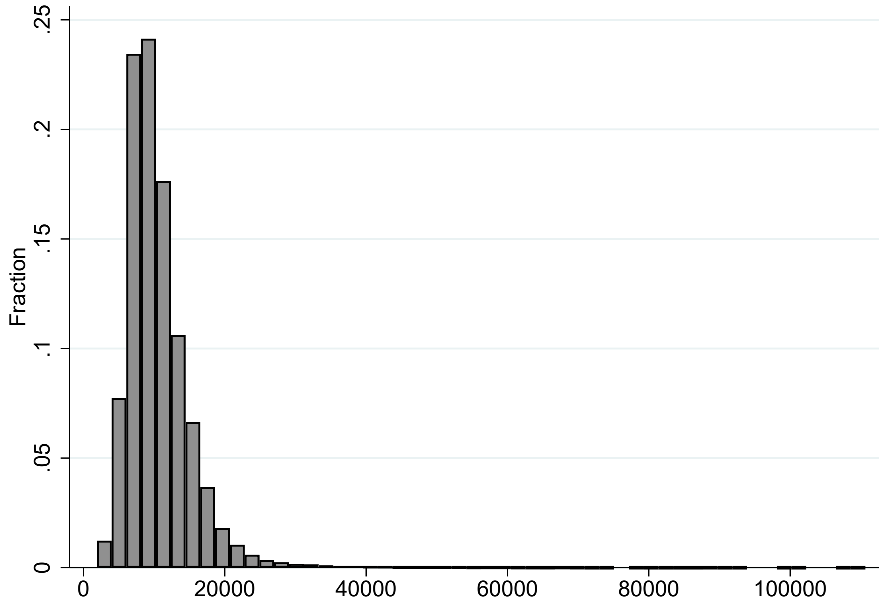
Notes. This figure shows the variation in expertise overlap within proceduralists after accounting for physician's case-related specialty indicators.

Figure A.2: Distribution of Patients across Hospitals and Proceduralists



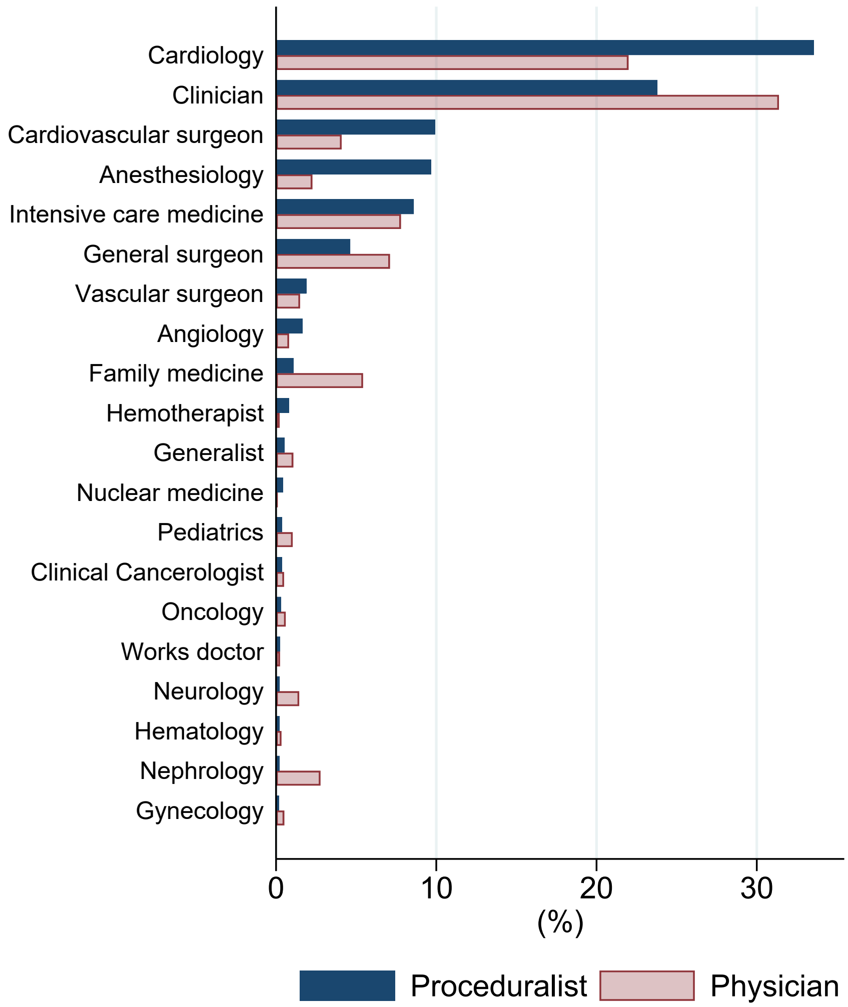
Notes. This figure plots the distribution of patients across hospitals and proceduralists in our sample.

Figure A.3: Hospital Spending on PCI Patients



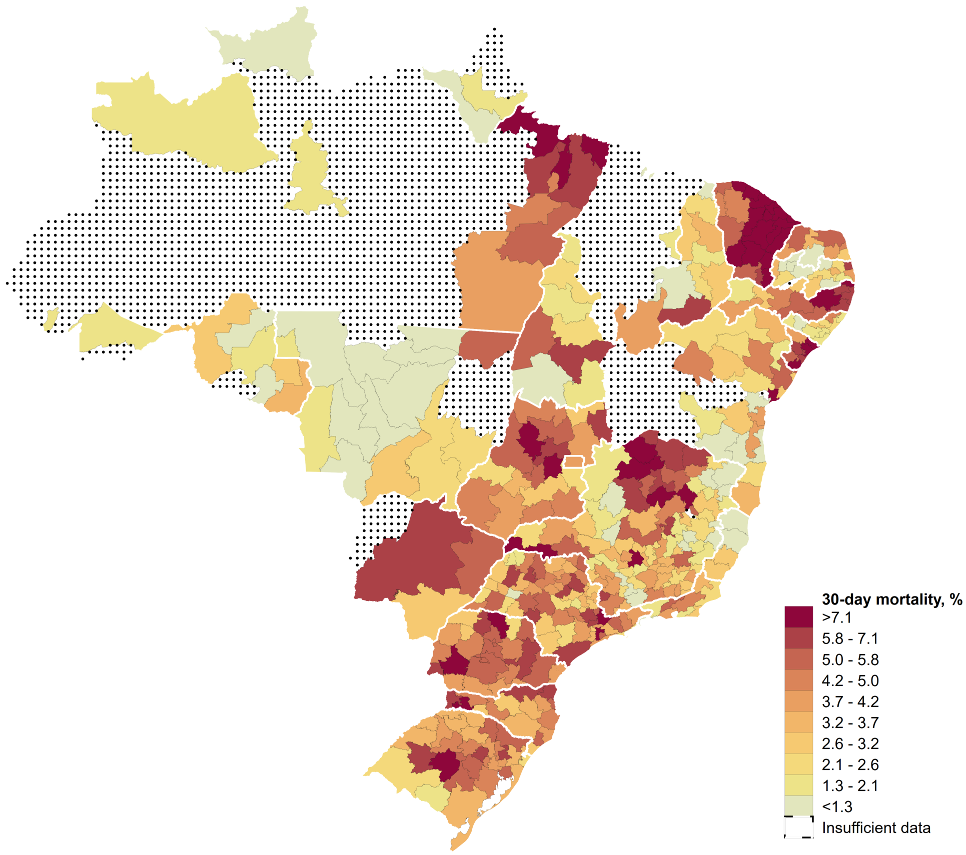
Notes. This figure plots the distribution of hospital spending on PCI patients, expressed in 2022 R$ (exchange rate to the US$ of 5.15 R$/US$).

Figure A.4: Top 20 Specialties



Notes. This figure shows the distribution of the top 20 specialties across physicians and proceduralists in the sample.

Figure A.5: 30-Day Mortality Rate across Health Regions



Notes. This figure shows the distribution of the 30-day mortality rate across health regions. Estimates for regions with fewer than 50 patients are omitted and are shaded with the cross-hatch pattern

Table A.1: Basic Sample Characteristics

	Mean (1)	Standard Deviation (2)
Proceduralist-physician expertise overlap	0.429	0.296
30-day mortality	0.056	0.230
Age	62.215	11.542
Age > 80	0.068	0.251
Age 65-79	0.352	0.478
Age < 65	0.580	0.493
Male	0.657	0.475
Race		
White	0.638	0.481
Black	0.024	0.154
Other	0.197	0.397
Primary diagnosis		
Acute myocardial infarction	0.484	0.500
Angina pectoris	0.431	0.495
Other acute ischemic heart disease	0.031	0.174
Chronic ischemic heart disease	0.047	0.211
Other diagnosis	0.007	0.081
Length of stay	5.743	5.772
Hospital spending (in 2022 R$)	10619.088	4837.194
Counts:		
Number of patients		176108
Number of hospitals		201
Number of proceduralists		966
Number of physicians		9628
Number of physicians per team		1.5

Notes. This table reports key characteristics of the estimation sample.

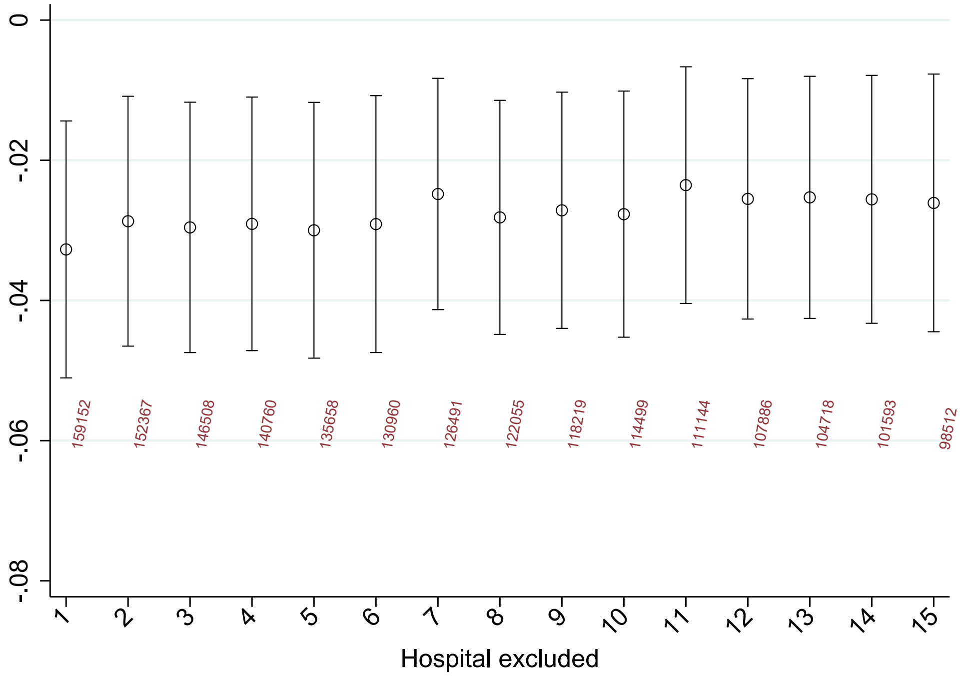
B Additional Robustness Checks

In this section, we present and discuss the results from several exercises designed to evaluate the validity of our basic results.

B.1 Outliers

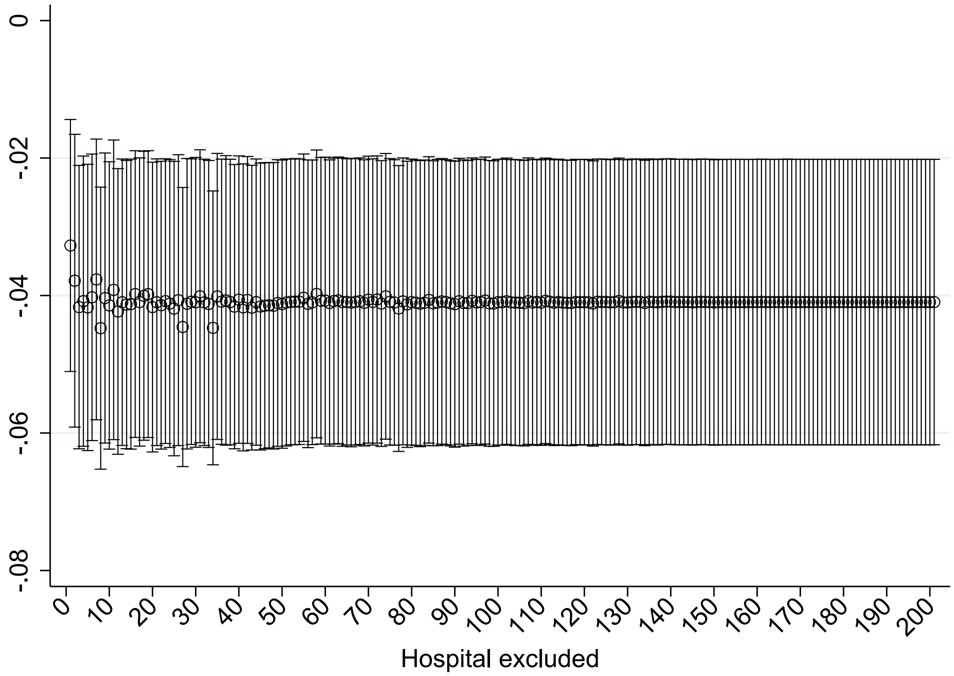
While Figure 3 suggests that our results are not driven by a few outliers in the last vintile of the residualized expertise overlap, we repeat the baseline specification excluding these observations. If anything, the point estimate becomes larger in magnitude (Appendix Table B.2, column 2). To further ensure that our results are not the product of outliers, we estimate the baseline specification using subsamples that exclude cumulatively the first hospitals with the largest share of patients. These estimates are presented in Appendix Figure B.1, with the coefficients ordered from the smallest to largest group of hospitals excluded. Even when we exclude hospitals that account for approximately 45 percent of the observations, our estimates remain very robust. As a further test, Appendix Figure B.2 repeats the baseline specification when each hospital is excluded one by one, with virtually no impact on our results.

Figure B.1: Exclusion of Hospitals One to Many



Notes. This figure presents estimates of the effect of expertise overlap on mortality from samples that exclude cumulatively the first hospitals with the largest share of patients, with the coefficients ordered from the smallest to largest group of hospitals excluded. Each estimated coefficient and confidence interval emanates from a single estimation. The number below each coefficient is the number of observations in the underlying estimation sample. All regressions use the baseline set of covariates reported in Table 1, column 5. See notes to Table 1 for details. Confidence intervals at 95 percent are clustered at the hospital level.

Figure B.2: Exclusion of Hospitals One by One



Notes. This figure presents estimates of the effect of expertise overlap on mortality when each hospital is excluded one by one. The coefficients are ordered from the smallest to largest hospitals excluded. Each estimated coefficient and confidence interval emanate from a single estimation. All regressions use the baseline set of covariates reported in Table 1, column 5. See notes to Table 1 for details. Confidence intervals at 95 percent are clustered at the hospital level.

B.2 Alternative Constructions of Expertise Overlap

Our measure of expertise overlap is an average weighted by the number of visits physicians provide during the hospital stay to account for differences in the contribution of care across physicians. We consider several alternative approaches in Appendix Table B.1. A first natural alternative strategy is to assume that all physicians have an equal contribution of care and use simple averages of the expertise overlap between the proceduralist and physicians without weighting. The other approaches we consider include one where we use the median expertise overlap and another where we focus on the expertise overlap between the proceduralist and physician with the largest number of visits. In all cases, we find a negative and highly significant effect of expertise overlap on mortality, demonstrating that the exact construction of expertise overlap does not drive our results.

Table B.1: Expertise Overlap and 30-Day Mortality:
Alternative Constructions of Expertise Overlap

	Dependent variable is 30-day mortality			
		Alternative measures		
	Baseline	Unweighted average	Median number of visits	Physician with largest number of visits
	(1)	(2)	(3)	(4)
Expertise overlap	-0.041	-0.0443	-0.0433	-0.0313
	[0.0106]***	[0.0118]***	[0.0116]***	[0.0081]***
Mean of dep. variable	0.056	0.056	0.056	0.056
Observations	176108	176108	176108	176108
Basic controls	✓	✓	✓	✓

Notes. This table presents estimates of the effect of expertise overlap on 30-day mortality considering alternatives ways to measure the expertise overlap. Column 1 repeats the baseline specification for ease of comparison. Column 2 assumes that all physicians have an equal contribution of care and uses simple averages of the expertise overlap between the proceduralist and physicians without weighting. Column 3 uses the median expertise overlap between the proceduralist and physicians. Column 4 focuses on the expertise overlap between the proceduralist and physician with the largest number of visits. All regressions use the baseline set of covariates reported in Table 1, column 5. See notes to Table 1 for details. Standard errors are clustered at the hospital level. * $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

B.3 Alternative Samples

Appendix Table B.2 documents that the results hold in different samples. A possible concern is that the pandemics may have been confounding our results due to the upsurge in mortality, especially among the elderly, and increased hospital capacity constraints. Our results remain very robust when we exclude patients who underwent PCI during the COVID-19 pandemic crisis.

Our analysis focuses on patients in emergency situations. Although we find similar results when we include non-emergency patients in the sample (Appendix Table B.2, column 3), one might be worried about misclassification of a patient's emergency status since it depends on how the health professional who first treat the patient interpret her signs and symptoms. As a robustness test, we limit the sample to patients experiencing a heart attack. This condition requires immediate treatment due to its acute nature and, once diagnosed, it is considered an emergency case by clinical protocols, which plausibly minimizes the risk of misclassification and any degree of selection into teams. The results obtained from using this subsample are displayed in column 4 of Appendix Table B.2. The point estimate is -0.0508 (standard error=0.0156) and thus somewhat larger in magnitude relative to the baseline. We conclude that any bias due to misclassification of a patient's emergency status in the benchmark sample is unlikely to be a major issue.

Finally, we exclude a few proceduralists with many patients in our sample, over 2500 patients, with virtually no impact on the estimated coefficient and its standard error.[1]

[1] Panel B of Appendix Figure A.2 displays the distribution of patients across proceduralists. The average number of patients per proceduralist is 462, with a standard deviation of 541 patients. Of all proceduralists, 25 percent have over 613 patients in total, 10 percent have treated over 1200 patients, and 1 percent have seen over 2500 patients.

Table B.2: Expertise Overlap and 30-Day Mortality:
Alternative Samples

	Dependent variable is 30-day mortality					
		Alternative samples				
	Baseline	Excl. observations with high expertise overlap	Include non-emergency patients	Heart-attack patients	Excl. patients treated in 2020	Excl. proceduralists with many patients (over 2500)
	(1)	(2)	(3)	(4)	(5)	(6)
Expertise overlap	-0.041	-0.0676	-0.0363	-0.0507	-0.044	-0.0407
	[0.0106]***	[0.0128]***	[0.0090]***	[0.0155]***	[0.0105]***	[0.0108]***
Mean of dep. variable	0.056	0.0564	0.0482	0.0754	0.0541	0.0558
Observations	176108	167309	221363	91525	156587	166348
Basic controls	✓	✓	✓	✓	✓	✓

Notes. This table presents estimates of the effect of expertise overlap on 30-day mortality considering alternatives samples. Column 1 repeats the baseline specification for ease of comparison. Column 2 excludes observations in the last vintile of the residualized expertise overlap distribution. Column 3 includes non-emergency patients in the sample. Column 4 focuses on patients who experienced a heart attack. Column 5 excludes patients who underwent PCI in 2020. Column 6 excludes proceduralists who treated 2500 or more patients in total during the entire study period. All regressions use the baseline set of covariates reported in Table 1, column 5. See notes to Table 1 for details. Standard errors are clustered at the hospital level.

* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

B.4 Varying Controls

To evaluate the robustness of the results to control choice, we use a machine learning technique, namely least absolute shrinkage and selection operator (LASSO), to select covariates among the set of patient and physician characteristics. This procedure chooses the subset of covariates that better predicts mortality and has a relatively low penalty in terms of the size of the model (Belloni et al., 2014). In addition to the baseline set of covariates, we consider a full set of interactions between gender, race, age, and comorbidities as well as among all physician covariates. Appendix Table B.3 shows that the results from the specifications identified by LASSO are extremely similar to those reported in Table 1, with the estimated coefficient ranging between -0.0412 and -0.0405. The results from the specifications selected by LASSO tend to be more precisely estimated, which is unsurprising since LASSO penalizes models with poor predictors of mortality and thus focuses on those that significantly reduce sampling variation.

We further assess the robustness of our results to alternative versions of the baseline specification in Appendix Table B.4. The point estimate remains negative and highly significant when we remove all controls and fixed effects, thus depicting the unconditional bivariate relationship, as well as when we use proceduralist fixed effects instead of proceduralist-by-year fixed effects. We also consider more demanding specifications that exploit a narrower source of variation, including hospital-by-day-of-week-by-month fixed effects as well as proceduralist-by-day-of-week and proceduralist-by-month fixed effects. The coefficient remains sizeable and very precisely estimated.

Table B.3: Expertise Overlap and 30-Day Mortality:
Covariates Selected by LASSO

	Dependent variable is 30-day mortality					
		Covariates included in LASSO				
	Baseline	Patient vars	Patient vars + interactions	Physician vars	Physician vars + interactions	All
	(1)	(2)	(3)	(4)	(5)	(6)
Expertise overlap	-0.041	-0.0405	-0.0405	-0.0409	-0.0409	-0.0403
	[0.0106]***	[0.0107]***	[0.0107]***	[0.0106]***	[0.0106]***	[0.0106]***
Mean of dep. variable	0.056	0.056	0.056	0.056	0.056	0.056
Observations	176108	176108	176108	176108	176108	176108

Notes. This table presents estimates of the effect of expertise overlap on 30-day mortality, where the covariates are selected using a LASSO procedure. Column 1 repeats the baseline specification for ease of comparison. In column 2, we start with the patient characteristics displayed in Panel A of Figure 1. In column 3, we consider patient characteristics and a full set of two-way interactions between all of them. In column 4, we include the physician covariates displayed in Panel B of Figure 1. Column 5 considers the same physician characteristics but adds all two-way interactions between them. Column 6 includes all covariates used in columns 3 and 5. All regressions control for hospital-by-day-week, hospital-by-month, proceduralist-year, physician's case-related specialty fixed effects. Since patients often see more than one physician, the physician's case-related specialty indicators are constructed as weighted averages of the physician's case-related specialty indicators, where the weights are the share of hospital visits provided by each physician to that patient. Standard errors are clustered at the hospital level.

* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

Table B.4: Expertise Overlap and 30-Day Mortality: Varying Controls

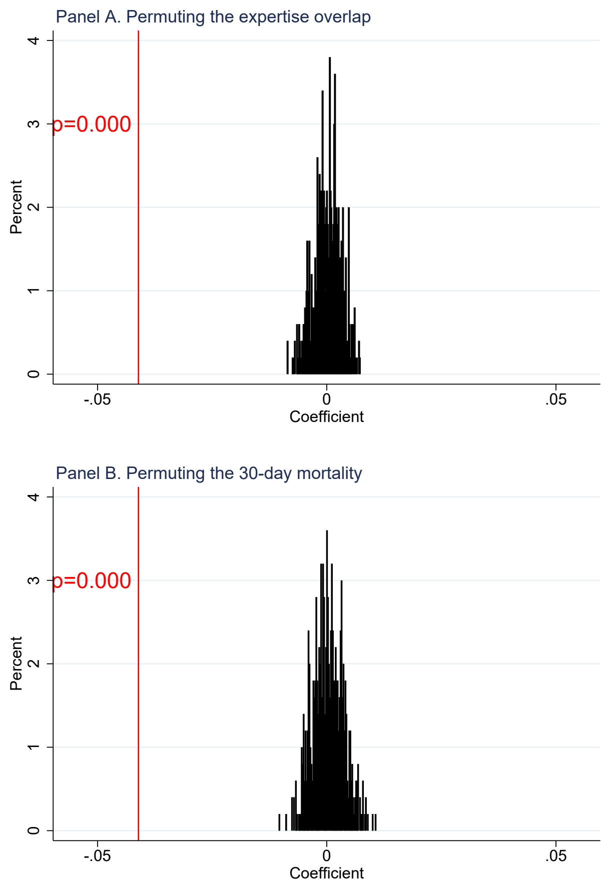
	Dependent variable is 30-day mortality				
	(1)	(2)	(3)	(4)	(5)
Expertise overlap	-0.041	-0.0548	-0.0391	-0.042	-0.0418
	[0.0106]***	[0.0075]***	[0.0084]***	[0.0109]***	[0.0106]***
Mean of dep. variable	0.056	0.056	0.056	0.056	0.056
Observations	176108	176108	176108	176108	176108
Year FE			✓		
Proceduralist FE			✓		
Proceduralist × year FE	✓			✓	✓
Proceduralist × month FE					✓
Proceduralist × day-of-week FE					✓
Physician's case-related specialty FE	✓		✓	✓	✓
Hospital × month FE	✓		✓		✓
Hospital × day-of-week FE	✓		✓		✓
Hospital × day-of-week × month FE				✓	
Patient characteristics	✓		✓	✓	✓
Physician characteristics	✓		✓	✓	✓

Notes. This table presents estimates of the effect of expertise overlap on 30-day mortality considering alternatives specifications. Column 1 repeats the baseline specification for ease of comparison. Column 2 removes all controls and fixed effects. Column 3 includes proceduralist and year fixed effects instead of proceduralist-by-year fixed effects. Column 4 uses hospital-by-day-of-week-by-month fixed effects instead of hospital-by-day-of-week and hospital-by-month fixed effects. Column 5 considers proceduralist-by-day-of-week and proceduralist-by-month fixed effects instead of proceduralist-by-year fixed effects. Standard errors are clustered at the hospital level.
* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

B.5 Robust Inference

Our baseline model uses standard errors clustered at the hospital level, thus allowing for arbitrary correlation in residuals across doctors within the same hospital and for serial correlation at the proceduralist or hospital level. However, because some doctors work for different hospitals within the same health region, a possible concern is that there could be correlation in residuals across hospitals, which could be important in practice. Appendix Table B.5 shows similar results when we use standard errors clustered either at the proceduralist, municipality, or health region as well as when we use standard errors two-way clustered at the hospital and year-month level. In addition, our conclusions are unaffected when we implement permutation tests that randomly assign either expertise overlap or 30-day mortality status to patients (Appendix Figure B.3).

Figure B.3: Permutation Tests



Notes. This presents placebo effects obtained from two permutation tests. Panel A samples the set of observed expertise overlap and reassigns it randomly across patients. We then regress 30-day mortality on the reassigned values controlling for the baseline covariates and repeat this procedure 500 times. Panel B repeats the same exercise but rather permutes 30-day mortality. In both cases, the permutations are within proceduralists.

Table B.5: Expertise Overlap and 30-Day Mortality: Robust Inference

	Dependent variable is 30-day mortality				
	(1)	(2)	(3)	(4)	(5)
Expertise overlap	-0.041	-0.0411	-0.0411	-0.0411	-0.0411
	[0.0106]***	[0.0064]***	[0.0101]***	[0.0100]***	[0.0121]***
Mean of dep. variable	0.056	0.056	0.056	0.056	0.056
Observations	176108	176108	176108	176108	176108
Alternative standard errors:					
Clustered by hospital (baseline)	✓				
Clustered by proceduralist		✓			
Clustered by municipality			✓		
Clustered by health region				✓	
Two-way clustered by hospital and year-month					✓
Basic controls	✓	✓	✓	✓	✓

Notes. This table evaluates the robustness of the baseline results in Table 1 to alternative inference approaches. Column 1 repeats the baseline specification for ease of comparison. Column 2 uses standard errors clustered at the proceduralist level (966 clusters). Column 3 uses standard errors clustered at the municipality level (136 clusters). Column 4 considers standard errors clustered at the health region level (123 clusters). Column 5 uses two-way clustered standard errors at the hospital and year-month level (201 and 143 clusters). All regressions use the baseline set of covariates reported in Table 1, column 5. See notes to Table 1 for details.

* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

C Threats to Identification

C.1 Number of Specialties and Cardiovascular-Related Specialties

Table C.1: Expertise Overlap and 30-Day Mortality:
Controlling for Number of Specialties

	Dependent variable is 30-day mortality			
	(1)	(2)	(3)	(4)
Expertise overlap	-0.041	-0.0414	-0.0421	-0.0392
	[0.0106]***	[0.0106]***	[0.0108]***	[0.0109]***
Mean of dep. variable	0.056	0.056	0.056	0.056
Observations	176108	176108	176108	176108
Number of specialties		✓	✓	✓
Number of cardiovascular-related specialties			✓	✓
Basic controls	✓	✓	✓	✓

Notes. This table evaluates the robustness of the baseline results in Table 1 to controlling for the total and cardiovascular-related number of specialties. Column 1 repeats the baseline for ease of comparison. Column 2 adds the total number of specialties. Column 3 instead includes the number of cardiovascular-related specialties. Column 4 controls simultaneously for the total and cardiovascular-related number of specialties. The basic set of controls include the same used in column 5 of Table 1. See notes to Table 1 for details. Standard errors are clustered at the hospital level.
* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

C.2 Selection into the Procedure

Construction of simulated overlap. As discussed in the paper, a major concern is that patients select into the procedure based on the possible teams available. An approach to investigate this possibility is to look at the relationship between the potential expertise overlap and probability of undergoing PCI. To generate this measure, we calculate the average expertise overlap over all possible proceduralist-physician teams available on the day patients enter the hospital. We identify physicians available on a given date and hospital based on whether she or he provided any care on that date in that hospital. Of course, this approach is imperfect and could be subject to measurement error. However, because the hospital files provide information on all procedures performed by physicians in great detail, we believe that the potential for measurement error is minimal. One might imagine instances where a physician might not see any patient on a given day even if she or he was present that day, but this is unlikely to be the case given that our sample is composed of high-volume hospitals where the demand for physicians is always high.

To sum up, we construct the simulated expertise overlap using the following step-by-step procedure in which we:

- Step 0: Identify the set of proceduralists and physicians available on the day that each patient was admitted to the hospital.
- Step 1: Calculate the degree of overlap in specialties between each possible proceduralist-physician pair.
- Step 2: Take the average over all possible proceduralist-physician expertise overlap for each admission day and hospital.

Note that this measure varies at the admission day and hospital level. Hence, it is not the same as that used in our physician availability design in Section 6.

Table C.2: Probit Model of PCI Treatment:
Estimated Marginal Effects

Independent variable	Marginal effect	Independent variable	Marginal effect
Male	0.0456	Out-state visitor	-0.0706
	[0.0039]***		[0.0338]**
Over age 80	0.029	Primary hypertension	-0.0402
	[0.0108]***		[0.0223]*
Age 75-19	0.0397	Diabetes complicated	-0.0944
	[0.0081]***		[0.0268]***
Age 70-74	0.0397	Diabetes uncomplicated	-0.0182
	[0.0076]***		[0.0272]
Age 65-69	0.0439	Chronic ischemic heart disease	0.049
	[0.0069]***		[0.0776]
Age 60-64	0.0523	Kidney disease	-0.2338
	[0.0062]***		[0.0219]***
Age 55-59	0.0574	HIV/AIDS	-0.1624
	[0.0054]***		[0.0472]***
Age 50-54	0.053	Obesity	-0.0633
	[0.0039]***		[0.0197]***
White	0.0076		
	[0.0115]		
Black	-0.0245		
	[0.0203]		
Race is missing	0.0519		
	[0.0166]***		
Observations		1976501	

Notes. This table contains marginal effects from a probit model of PCI treatment choice. The sample includes patients with a diagnosis of angina pectoris, acute myocardial infarction, other acute ischemic heart disease, or chronic ischemic heart disease. These are the most common diagnoses for PCI, accounting for approximately 98 percent of PCI treatments in our analysis sample. The model includes indicators for state, year, and month. Standard errors are clustered at the hospital level.

* p < 0.1;** p < 0.05;*** p < 0.01.

Table C.3: Potential Expertise Overlap and PCI probability

	Dependent variable is PCI treatment			
	(1)	(2)	(3)	(4)
Simulated expertise overlap	0.0338	0.0587	0.0304	0.0302
	[0.0986]	[0.0969]	[0.0386]	[0.0385]
Mean of dep. variable	0.4507	0.4507	0.4507	0.4507
Observations	1847482	1847482	1847463	1847463
Day-of-admission FE		✓	✓	✓
Hospital $\times$ (day-of-week, month, and year) FE			✓	✓
Patient characteristics				✓

Notes. This table presents estimates of the effect of potential expertise overlap on the probability of PCI treatment. To construct the potential expertise overlap, we first identify all proceduralists and physicians available on the day patients entered the hospital. We then compute the average expertise overlap between each potential proceduralist and physicians for each admission day and hospital. See Appendix Section C.2 for further details. The sample includes patients with a diagnosis either for angina pectoris, acute myocardial infarction, other acute ischemic heart disease, or chronic ischemic heart disease. These are the most common diagnoses for PCI, accounting for approximately 98 percent of PCI treatments in our analysis sample.
* p < 0.1;** p < 0.05;*** p < 0.01.

Table C.4: Actual Expertise Overlap and Predicted PCI Probability

	Dependent variable is predicted PCI treatment				
	(1)	(2)	(3)	(4)	(5)
Expertise overlap	-0.000185	0.000005	-0.0002	-1.3E-05	0.000028
	[0.000587]	[0.000598]	[0.000140]	[0.000607]	[0.000081]
Mean of dep. variable	0.4529	0.4530	0.4529	0.4529	0.4530
Observations	174934	173893	174934	174934	173893
Proceduralist × year FE	✓	✓	✓	✓	✓
Physician's case-related specialty FE	✓	✓	✓	✓	✓
Hospital × time FE		✓			✓
Patient characteristics			✓		✓
Physician characteristics				✓	✓

Notes. This table presents estimates of the effect of expertise overlap on predicted PCI probability. The expertise overlap is defined in equations (1) and (2). See Table 1 for details on the covariates used in the model. Predicted PCI probability is obtained from a probit regression of actual PCI status on patient characteristics that include age category indicators, gender, race, comorbidity indicators, and indicators for state, year, and month. The sample is based on patients with the most common diagnoses for PCI: angina pectoris, acute myocardial infarction, other acute ischemic heart disease, or chronic ischemic heart disease. Standard errors are clustered at the hospital level.

* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

Table C.5: Expertise Overlap and 30-Day Mortality:
Accounting for Selection into PCI

	Dependent variable is 30-day mortality						
		Control for predicted PCI treatment				Control for inverse	Sample limited to high PCI
	Baseline	Linearly	Quadraticly	Cubicly	Quarticly	Mills ratio	prob. patients
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Expertise overlap	-0.041	-0.0409	-0.0409	-0.0409	-0.0409	-0.0408	-0.0407
	[0.0106]***	[0.0106]***	[0.0106]***	[0.0106]***	[0.0106]***	[0.0106]***	[0.0258]
Mean of dep. variable	0.056	0.0555	0.0555	0.0555	0.0555	0.0555	0.0573
Observations	176108	174934	174934	174934	174934	174934	36103
Sample	All	Common diagnosis for PCI					
Basic controls	✓	✓	✓	✓	✓	✓	✓

Notes. This table evaluates the robustness of the baseline results in Table 1 to taking into account selection into PCI. Column 1 repeats the baseline for ease of comparison. Columns 2 through 5 control linearly, quadratically, and cubically for predicted PCI treatment respectively. Column 6 rather controls for the inverse Mill's ratio. Column 7 limits the sample to patients who are in the top 20 percent of the PCI propensity distribution and thus have an *ex-ante* high probability of undergoing PCI. All regressions use the baseline set of covariates reported in Table 1, column 5. Predicted PCI treatment is obtained from a probit regression of actual PCI status on patient characteristics that include age category indicators, gender, race, comorbidity indicators, and indicators for state, year, and month. The sample is based on patients with the most common diagnoses for PCI: angina pectoris, acute myocardial infarction, other acute ischemic heart disease, or chronic ischemic heart disease. The basic set of controls include the same used in column 5 of Table 1. See notes to Table 1 for details. Standard errors are clustered at the hospital level.

* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

C.3 Case Severity and Team-Specific Experience

Table C.6: Expertise Overlap and 30-Day Mortality:
Controlling for Primary Diagnosis

	Dependent variable is 30-day mortality	
	Baseline	Controlling for patient primary diagnosis
	(1)	(2)
Expertise overlap	-0.041	-0.0384
	[0.0106]***	[0.0098]***
Mean of dep. variable	0.056	0.056
Observations	176108	176108
Primary diagnosis FE		✓
Basic controls	✓	✓

Notes. This table evaluates the robustness of the baseline results in Table 1 to controlling for patient primary diagnosis. Column 1 repeats the baseline for ease of comparison. Column 2 adds patient primary diagnosis fixed effects. The basic set of controls include the same used in column 5 of Table 1. See notes to Table 1 for details. Standard errors are clustered at the hospital level.
* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

Table C.7: Results Are Not Driven by Shared Work Experience

	Dependent variable is:		
	Shared work experience (in days)	30-day mortality	
	(1)	(2)	(3)
Expertise overlap	8.8064	-0.041	-0.0364
	[6.3490]	[0.0106]***	[0.0097]***
Mean of dep. variable	51.44	0.056	0.0526
Observations	168238	176108	168238
Shared work experience			✓
Basic controls	✓	✓	✓

Notes. This table investigates the role of shared work experience in driving the baseline results. Shared work experience is defined as in Chen (2021): the number of times that proceduralists and physicians have worked together in the past two years. Since patients often see more than one physician, the shared work experience for a patient is constructed as the weighted average of the shared work experience between the proceduralist and each physicians treating the patient, where the weights are the share of hospital visits provided by each physician to that patient. Column 1 estimates the association between the expertise overlap and shared work experience. Column 2 repeats the baseline results reported in column 5 of Table 1. Column 3 adds shared work experienced as a control variable. The basic controls include the same as the ones used in column 5 of Table 1. See notes to Table 1 for details. Standard errors are clustered at the hospital level.
* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

C.4 The Role of Nurses

Table C.8: Results Are not Driven by Changes in Supply of Nurses

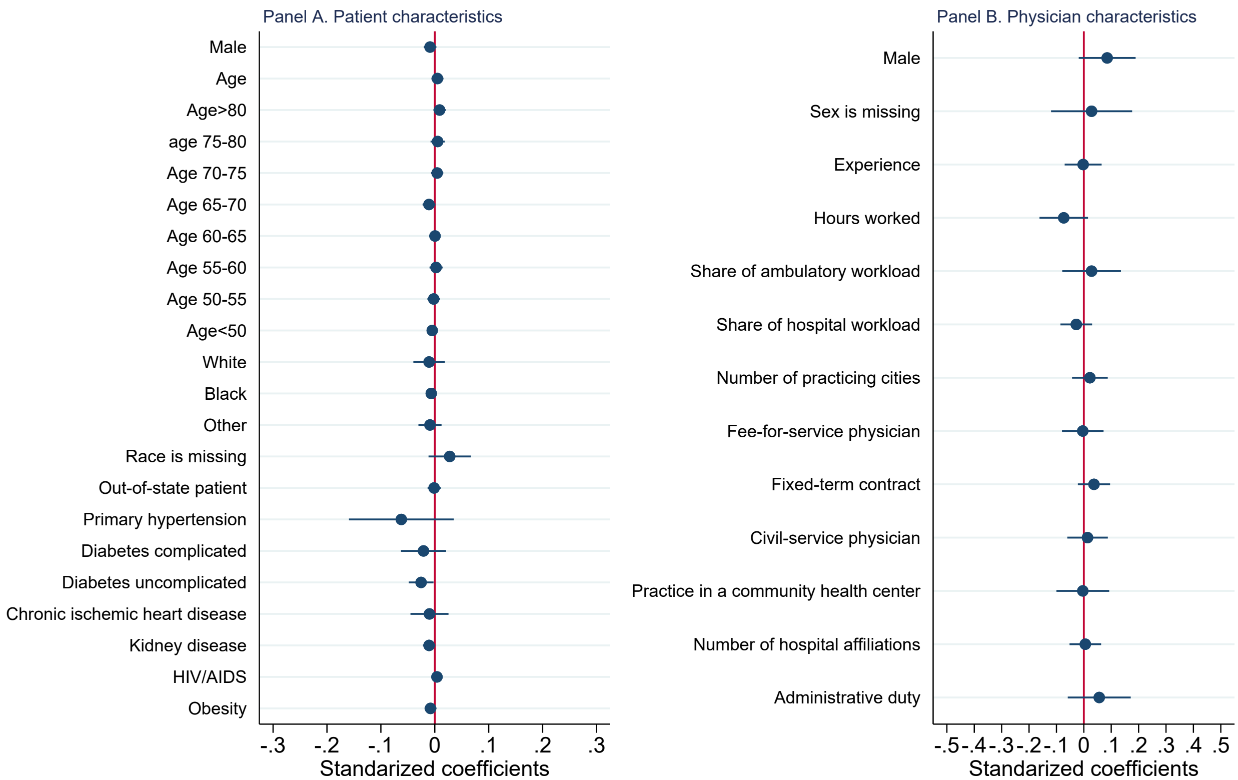
	Dependent variable is:		
	Nurses per ICU bed	30-day mortality	
	(1)	(2)	(3)
Expertise overlap	-0.002	-0.041	-0.0411
	[0.0022]	[0.0106]***	[0.0107]***
Mean of dep. variable	0.0802	0.056	0.0556
Observations	175012	176108	175012
Nurses per ICU bed			✓
Basic controls	✓	✓	✓

Notes. This table investigates the role of changes in nurse rates on the day the patients enter the hospital. Column 1 estimates the association between the expertise overlap and nurse per intensive care unit (ICU) beds. Column 2 repeats the baseline results reported in column 5 of Table 1. Column 3 estimates the effect of expertise overlap but including the nurse rate as a control variable. The basic controls include the same as the ones used in column 5 of Table 1. See notes to Table 1 for details. Standard errors are clustered at the hospital level.
* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

D Physician Availability Design

D.1 Testing for Balance

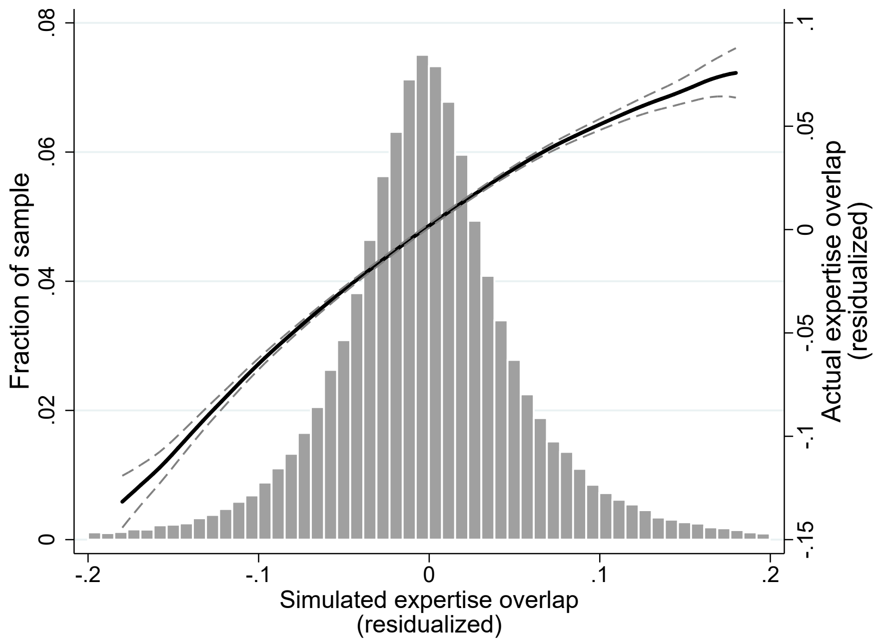
Figure D.1: Covariate Balance:
Physician Availability Design



Notes: Each coefficient is from a different regression where the simulated expertise overlap is the independent variable of interest and the predetermined characteristics are the dependent variables. The regressions include proceduralist, date-of-admission, day-of-week×hospital, and month×hospital fixed effects. All coefficients are standardized for ease of readability. Standard errors are clustered at the hospital level.

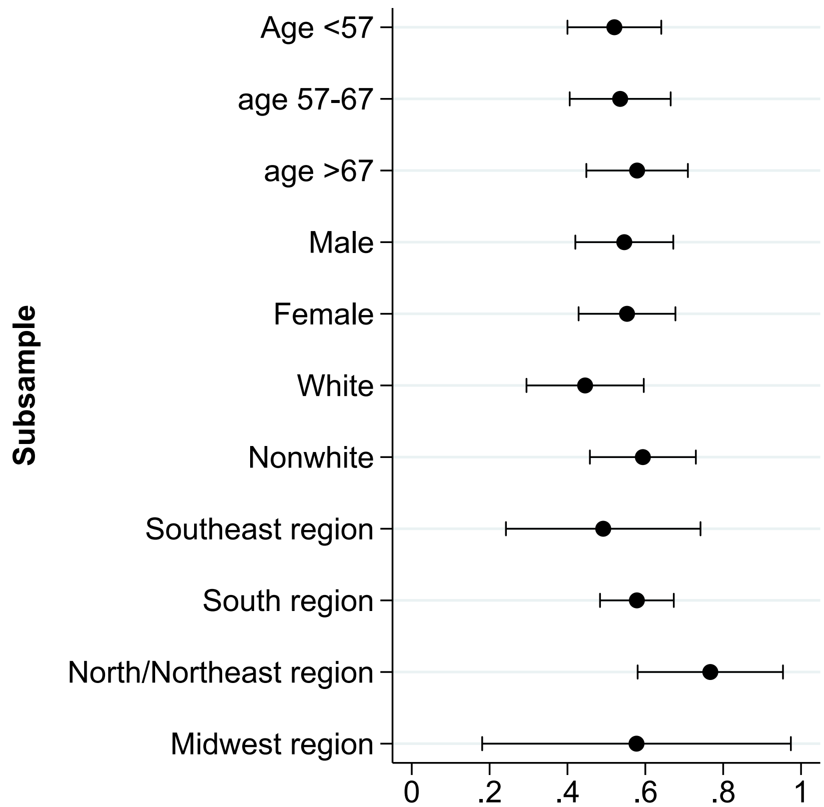
D.2 First Stage, and Testing for the Monotonicity Assumption

Figure D.2: Distribution of Simulated Expertise Overlap and First Stage



Notes. This figure reports the distribution of the simulated expertise overlap that is calculated following the procedure described in Section 6. The actual and simulated expertise overlap are residualized with respect to proceduralist and date-of-admission fixed effects. The relationship between both variables plotted in the figure is estimated a local linear regression.

Figure D.3: Sub-Sample First Stage Estimates



Notes. This table reports heterogeneity in first stage estimates using model (6). Each coefficient corresponds to a different subsample. All models include proceduralist, date-of-admission, day-of-week×hospital, month×hospital fixed effects. 95 confidence intervals are based on standard errors clustered at the hospital level.

D.3 Robustness Checks

Table D.1: First Stage, Reduced Form, and 2SLS Estimates:
Robustness to Hospital-Specific Time Trends

	Dependent variable is:		
	Expertise overlap	30-day mortality	
	(First Stage)	(Reduced-Form)	(2SLS)
	(1)	(2)	(3)
Expertise overlap			-0.0589
			[0.0254]**
Simulated expertise overlap	0.4851	-0.0286	
	[0.0462]***	[0.0126]**	
Kleibergen and Paap (2006) F statistics			107.3502
Mean of dep. variable	0.4288	0.056	0.056
Observations	175349	175349	175349
Hospital-specific linear time trends	✓	✓	✓
Basic controls	✓	✓	✓

Notes. This table reports first-stage, reduced form, and 2SLS coefficient estimates. The instrument for expertise overlap is a simulated measure of expertise overlap based on the availability of physicians at the time patients are admitted to the hospital. See Section 6 for details on the construction of the instrument. Patient and physician characteristics include all those displayed in Figure D.1. Standard errors are clustered at the hospital level.
* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

Table D.2: First Stage, Reduced Form, and 2SLS Estimates: Unweighted Instrument

	Dependent variable is:		
	Expertise overlap	30-day mortality	
	(First Stage)	(Reduced-Form)	(2SLS)
	(1)	(2)	(3)
Expertise overlap			-0.0468
			[0.0200]**
Simulated expertise overlap	0.6258	-0.0293	
	[0.0825]***	[0.0119]**	
Kleibergen and Paap (2006) F statistics			57.6009
Mean of dep. variable	0.4288	0.056	0.056
Observations	175349	175349	175349
Basic controls	✓	✓	✓

Notes. This table reports first-stage, reduced form, and 2SLS coefficient estimates. The instrument for expertise overlap is a simulated measure of expertise overlap based on the availability of physicians at the time patients are admitted to the hospital. This table uses the unweighted version of the instrument described in Section 6. Patient and physician characteristics include all those displayed in Figure D.1. Standard errors are clustered at the hospital level.
* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

Table D.3: Expertise Overlap and 30-Day Mortality:
Robustness of 2SLS Estimates to Sample Restrictions

	Dependent variable is 30-day mortality				
		Alternative samples			
	Baseline (1)	Include non-emergency patients (2)	Heart-attack patients (3)	Excl. patients treated in 2020 (4)	Excl. proceduralists with many patients (over 2500) (5)
Expertise overlap	-0.0447	-0.0454	-0.0442	-0.0604	-0.0495
	[0.0199]**	[0.0165]***	[0.0310]	[0.0208]***	[0.0198]**
Kleibergen and Paap (2006) F statistics	84.1518	81.617	98.241	70.7864	80.4512
Mean of dep. variable	0.056	0.0491	0.0756	0.0545	0.0563
Observations	175349	216316	90100	156843	166666
Basic controls	✓	✓	✓	✓	✓

Notes. This table presents 2SLS estimates of the effect of expertise overlap on 30-day mortality considering alternatives samples. Column 1 repeats the baseline specification for ease of comparison. Column 2 includes non-emergency patients in the sample. Column 3 focuses on patients who experienced a heart attack. Column 4 excludes patients who underwent PCI in 2020. Column 5 excludes patients over 80 years of age. Column 6 excludes proceduralists who treated 2500 or more patients in total during the entire study period. All regressions use the baseline set of covariates reported in Table 2. See notes to Table2 for details. Standard errors are clustered at the hospital level.

* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

Table D.4: Expertise Overlap and 30-Day Mortality:
Robust Inference for 2SLS Estimates

	(1)	(2)	(3)	(4)	(5)	(6)
	Panel A: Second Stage					
Expertise overlap	-0.0447	-0.0447	-0.0447	-0.0447	-0.0447	-0.0447
	[0.0199]**	[0.0169]***	[0.0192]**	[0.0198]**	[0.0200]**	[0.0202]**
	Panel B: First Stage					
Simulated expertise overlap	0.5399	0.5399	0.5399	0.5399	0.5399	0.5399
	[0.0588]***	[0.0513]***	[0.0586]***	[0.0584]***	[0.0591]***	[0.0588]***
Kleibergen and Paap (2006) *F* statistics	84.1518	110.8693	84.7443	85.3709	83.3688	84.1518
Mean of dep. variable	0.056	0.056	0.056	0.056	0.056	0.056
Observations	175349	175349	175349	175349	175349	175349
Alternative standard errors:						
Clustered by hospital (baseline)	✓					
Clustered by proceduralist		✓				
Clustered by municipality			✓			
Clustered by health region				✓		
Two-way clustered by hospital and year-month					✓	
Lee et al. (2022)'s S.E. correction						✓
Basic controls	✓	✓	✓	✓	✓	✓

Notes. This table evaluates the robustness of the 2SLS results in Table 2 to alternative inference approaches. Column 1 repeats the baseline specification for ease of comparison. Column 2 uses standard errors clustered at the proceduralist level (966 clusters). Column 3 uses standard errors clustered at the municipality level (136 clusters). Column 4 considers standard errors clustered at the health region level (123 clusters). Column 5 uses two-way clustered standard errors at the hospital and year-month level (201 and 143 clusters). Column 6 corrects the second-stage standard error using the tF procedure introduced by Lee et al. (2022). All regressions use the baseline set of covariates reported in Table 2. See notes to Table 2 for details.

* $p < 0.1$;** $p < 0.05$;*** $p < 0.01$.

E Details of the Simulation Exercises

In Section 8, we are interested in understanding which types of polices are more cost effective in improving welfare. We compare the consequences of two different policies that alter physicians' work schedules and thus average proceduralist-physician expertise overlap: i) increasing the number of hours worked up to 100 percent of a subset of physicians who have above-average overlap in specialties with proceduralists; and ii) reallocating physicians from days of the week where they have below-average overlap in specialties with proceduralists to other days of the week. Note that these interventions alter the *availability* of physicians, the former by increasing workload intensity and the latter by reallocating physicians across regular workdays. Therefore, we use the results from the physician availability design to calculate the counterfactual mortality reductions. This section provides details on this procedure.

E.1 Increasing Physicians' Workload

This intervention seeks to simulate the consequences of increasing the number of hours that certain physicians work. Physicians who have above-average expertise overlap with proceduralists and work fewer than 40 hours per week are eligible for this policy. The latter restriction accounts for the possibility that increasing the workload intensity of overloaded physicians may be implausible.

Eligible physicians. To identify eligible physicians, we follow a 5-step procedure:

1. Identify all the physicians and proceduralists available for each date on which they provided any services in a given hospital.

2. For physician k in hospital f available on date d, calculate her or his expertise overlap with each proceduralist available on that date and take the average over all the calculated expertise overlap. Let z_{kf} denote this physician-specific average.

3. Compute the average expertise overlap over all the possible proceduralist-physician pairs for each hospital, denoted by z_f.

4. Define eligible physicians as those with $z_{kf} < z_f$ and actual hours worked $<$ 40.

5. Increase up 100 percent the number of hours worked of eligible physicians: $\hat{h} = \min\{40, 2 \times$ actual hours worked$\}$.

Mortality reductions. To compute the counterfactual mortality reductions, we use the physician availability design introduced in Section 6. Specifically, we recalculate the simulated expertise overlap used in that section by replacing the actual number of hours worked with the simulated number of hours worked for each physician:

$$\tilde{z}_{jd} \equiv \text{simulated expertise overlap}_{jd} = \sum_{k \in \mathcal{K}(d)} w_{kd} \times z_{jkd} \quad \text{with} \quad w_{kd} = \frac{\hat{h}_{kd}}{\sum_{k \in \mathcal{K}(d)} \hat{h}_{kd}} \tag{E.1}$$

We then obtain the aggregate mortality reductions using the reduced-form results reported in column 2 of Table 6 as follows:

$$\widehat{\text{mortality reductions}} = N(\mathbb{E}[y_{ijt}] - \mathbb{E}[\tilde{y}_{ijt}]) = N\hat{\pi}_1(\mathbb{E}[\hat{z}_{jd}] - \mathbb{E}[\tilde{z}_{jd}]) = N\hat{\pi}_1(\bar{\hat{z}} - \bar{\tilde{z}}) \tag{E.2}$$

where N represents the number of patients in the sample and $\hat{\pi}_1$ is the reduced-form coefficient that measures the effects of the simulated expertise overlap on mortality. We multiply the differences in the expected mortality status by N to obtain an estimated number of lives saved.

Policy costs. The costs of policy implementation include the monetary compensation paid to physicians. We use information from the Census conducted in 2010 to obtain information on physicians' hourly wages —R$113.74 in 2022 R$. To increase the number of hours that physicians work, the government would need to increase hourly wage such that physicians would be willing to increase their labor supply. This requires that we make assumptions on the labor supply elasticity. Chetty (2012) reviews existing estimates in the literature and finds an approximate intensive margin elasticity of 0.33, which is fairly close to that found in studies focusing on physicians' labor supply decisions (Showalter and Thurston, 1997). Given a labor supply elasticity of 0.33, the percentage change in hourly wage required to increase physicians' labor supply up to 100 percent would be approximately 303 percent:

$$\Delta logh = \eta \Delta logW \quad \text{or} \quad \Delta logW = \Delta logh/\eta \equiv 303\% = 100\%/0.33$$

Hence, the aggregate costs of policy implementation can be obtained as follows:

$$\textstyle\sum_{k \in \mathcal{K}} \sum_{t \in \mathcal{T}} \left(\hat{W}_{kt} \times \hat{h}_{kt} - W_{kt} \times h_{kt}\right)$$

where $\hat{W}$, and W represent the new and old salaries respectively. The subscript t now indexes weeks.

E.2 Reallocation across Days of the Week

This intervention seeks to alter the physicians' regular days of work without any consequence on work intensity.

Eligible physicians. We consider physicians below-average overlap in specialties with proceduralists to be eligible. We follow a 6-step procedure to identify eligible physicians and implement the reallocation:

1. Identify all the physicians and proceduralists available for each date they provided any services in a given hospital.

2. For physician k in hospital f available on day d of week q, calculate her or his expertise overlap with each proceduralist available on that week and take the average over all the calculated expertise overlap. Let z_{kfq} denote this physician-specific average for week q.

3. Compute the average expertise overlap over all the possible proceduralist-physician pairs for each hospital-week cells, denoted by z_{fq}.

4. Define eligible physicians as those with $z_{kfq} < z_{fq}$

5. Reallocate each eligible physician k from day d to day d' within the same week q. Both days d and d' are chosen at random with the restriction that $d \neq d'$.

6. For each eligible physician, choose one corresponding "control" physician from the pool of ineligible ones to be reallocated from day d' to day d within the same week q. This corresponding physician is chosen at random.

Note that step 6 ensures that the number of physicians available on each day of the week is not affected. Since eligible physicians have below-average expertise overlap, this allocation intervention would increase the average expertise overlap in the sample, despite the fact that the reallocation days and corresponding control physicians in step 6 are chosen at random.

Mortality reductions. To compute the counterfactual mortality reductions, we use the physician availability design introduced in Section 6. Specifically, we recalculate the simulated expertise overlap used in that section taking into account the reallocation of physicians across days of the week. In doing so, we note that the simulated expertise overlap increases from 0.32 in the baseline to 0.45 after the reallocation of physicians. This represents an increase of more than 40 percent. We then obtain the aggregate mortality reductions using the reduced-form results reported in column 2 of Table 6:

$$\text{mortality }\hat{\text{reductions}} = N(\mathbb{E}[y_{ijt}] - \mathbb{E}[{y^*}_{ijt}]) = N\hat{\pi}_1(\mathbb{E}[\hat{z}_{jd}] - \mathbb{E}[{z^*}_{jd}]) = N\hat{\pi}_1(\bar{\hat{z}} - \bar{z^*}) \tag{E.3}$$

where N represents the number of patients in the sample, $\hat{\pi}_1$ is the reduced-form coefficient that measures the effects of the simulated expertise overlap on mortality, and z^* denotes the counterfactual simulated expertise overlap. We multiply the differences in the expected mortality status by N to obtain an estimated number of lives saved.

Policy costs. We assume that the cost of reallocating physicians within regular working days (e.g., Tuesday versus Wednesday) is zero. However, we assume that physicians reallocated to weekends or holidays need to be compensated by increasing their income wages by 100 percent on those days. Therefore, in practice, the costs of policy implementation correspond to the additional income wages paid to reallocated physicians on those days. We use information from the Census conducted in 2010 to obtain information on physicians' daily wages —R\$540.82 in 2022 R\$. Hence, the aggregate costs of policy implementation can be obtained as follows:

$$\sum_{k\in\mathcal{K}}\sum_{d\in\mathcal{D}}\Delta W_{kd}\times\mathbf{1}[d=\text{weekend},\text{holiday}]$$

where ΔW represents the additional daily income wage paid on weekends and holidays. The term $\mathbf{1}[\cdot]$ is an indicator variable for whether day d is a weekend day or holiday.

F Data Sources and Construction

This section introduces a new patient-doctor-procedure dataset covering the entire Brazilian public health system, which we assemble by combining information from several administrative registries. Our analysis uses data covering the 2009-2020 period, a choice determined on the front end by the availability of physician data that can be linked to procedures and on the back end by the availability of mortality data. The preparation of death certificate records takes about 2 years, so we did not have any information for 2021 and 2022 at the time of preparation of this manuscript. We next describe these data in detail.

F.1 Provider Data

We have access to background information on providers from the *Cadastro Nacional de Estabelecimentos de Saúde* (CNES). It is a rich source of data collected monthly since 2005 that covers all health professionals in any private or public health facility in Brazil. A major strength of these data is their universal nature and high-frequency observations, with minimal underreporting. They contain key information on doctor characteristics, including specialty, weekly hours of work, and the identification codes of the establishments for which the doctor provides services, among others. The reporting of a specialty in a given month depends on whether the physician provided some service related to that specialty in that month. As a result, the number of specialties associated with a doctor will be underestimated in some periods. To mitigate this issue, we assign doctors a specialty based on whether they reported that specialty in some past period. In practice, this makes little difference in our final dataset: the average number of specialties per doctor before and after the imputation approach is 1.95 and 2.03 respectively. We track doctors over time using the unique health identifier (*Cartão Nacional de Saúde*, CNS).

For each doctor, we generate a panel of data with a record for each month in which they appeared in the CNES database. In each time period, we observe all doctors' specialties and other characteristics. We limit the sample to doctor-month-year cells with positive hours worked in SUS-affiliated hospitals, as we do not observe detailed patient characteristics and outcomes for those patients admitted to non-SUS hospitals.

F.2 Inpatient Procedure Registries

For each patient admitted to a SUS-affiliated hospital, we observe the procedures completed by health care providers the from the *Serviços Profissionais* files of the *Sistema de Informações Hospitalares* (SIH-SP). The SIH-SP covers the universe of hospital procedures provided through the public health system. All procedures have an identifier code created by the Ministry of Health to standardize service reporting, with multiple "layers" describing the procedures at different levels of detail. There exists one main procedure and there could be several secondary procedures per patient. The main procedure corresponds to the basic reason for the treatment, whereas the secondary procedures complement the main one. For example, consider a patient with a heart attack who was treated with PCI. In this case, the PCI treatment would be the main procedure and any tests, exams, or physician visits during the hospital stay would be classified as secondary procedures. We have access to the identity of the health professional who performed a given procedure since January 2008, which allows us to identify proceduralists and physicians who provide pre- and post-procedure care.

The data also contain information on the amount paid per procedure to the providers under the national fee schedule managed by the Ministry of Health. We use this information to create a measure of total hospital spending on each patient, converted into January 2022 *Reais* using the consumer price index. We restrict the sample to patients whose main procedure was a PCI.

F.3 Patient Data

Additional data on patients are drawn from the *Reduzida* files of the *Sistema de Informações Hospitalares* (SIH-RD). These files include information on date of birth, date of admission, date of discharge, diagnosis codes, race, sex, a hospital identifier, and an admission code that allows us to match these data to the inpatient procedure files described above. The diagnosis codes are divided into a primary and up to nine secondary diagnoses. The former is the health condition responsible for the admission and the latter are diseases that coexist at the time of the admission or develop during the hospital stay. We use the secondary diagnoses to create controls for comorbidities that commonly affect patients, such as hypertension, diabetes, chronic ischemic heart disease, and congestive heart failure. Key to our research design is that these data provide information on whether the patient was admitted to the hospital because of emergency health conditions.

F.4 Matching Doctor, Procedure, and Patient Files

Our goal is to generate a dataset of admission cases for which we can observe doctor backgrounds, procedures, as well as basic characteristics of patients. We start by matching the doctor data to the procedure files. A complication with this procedure is that the procedure files do not provide the individual health identifier CNS before June 2012 but only the individual tax identifier (*Cadastro de Pessoas Físicas*, CPF). Note that the doctor data contain only CNS identifiers. To avoid loss of data, we obtain information on the corresponding CNSs associated to each CPF prior to June 2012 using an application programming interface developed by the Ministry of Health. All of the physicians searched in the application were successfully found and hence we were able to obtain their CNSs. With the CNS, it is straightforward to match the procedure files to the doctor data. For each month, we link the doctor data to procedure registries based on the unique individual identifiers. The matching rate is 98 percent. It is not perfect due to the minimal underreporting in the provider database. We exclude patient cases where at least one doctor was not matched to the procedure data.

We then link the resulting doctor-procedure dataset to the extended hospital patient-level files. For this purpose, we use a unique patient case identifier available in both files. We successfully matched 100 percent of all cases. The resulting dataset is at the doctor-by-procedure-by-patient level.

F.5 Matching Hospital and Mortality Records

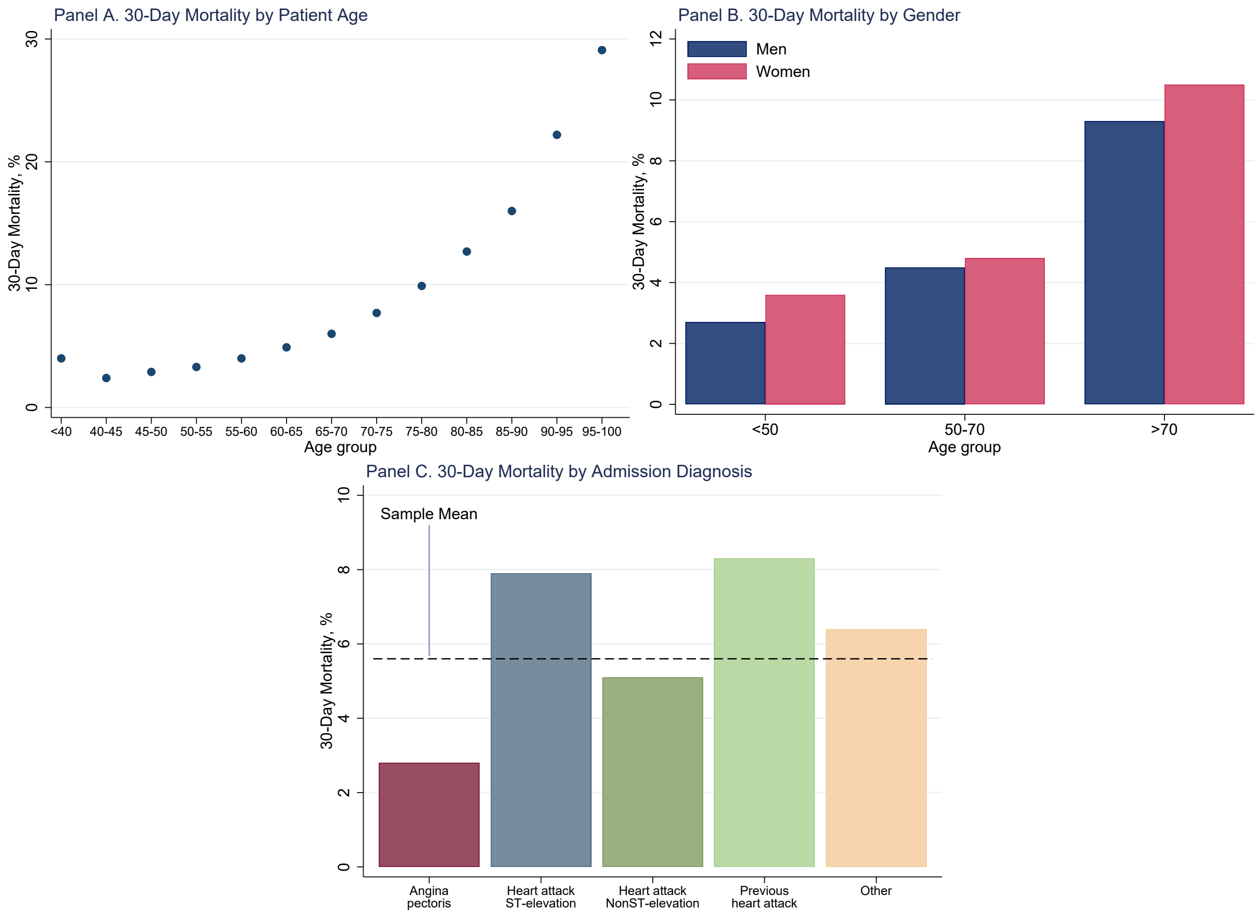
Our main outcome is an indicator for mortality within 30 days after the PCI. Data on mortality come from the *Sistema de Informações de Mortalidade* (SIM). These data contain comprehensive information on all deceased individuals in Brazil, including date of death, date of birth, race, sex, and place of residence. Unfortunately, there are not identifiers that allow us to link mortality and hospital records directly. Therefore, we match these data based on exact date of birth, location of residence, gender, and race. We generate a mortality indicator variable for whether a patient is matched to the mortality data within 30 days after the hospital discharge.

Of course, this matching procedure is imperfect and introduces measurement error in a patient's mortality status. But since there are only a few patients undergoing PCI who were born on the same date, residing within the same location, of the same gender and race, we believe this measurement error should be negligible. Indeed, in the hospital data, we observe unique location-of-residence$\times$date-of-birth$\times$gender$\times$race patient cells in 76 percent of cases. In 20 percent of cases, we observe clusters of exactly 2 patients. The remaining 6 percent include clusters of more than 3 patients. Remarkably, among patients that were matched to a death record in the mortality database, 99 percent corresponds to unique matches. Therefore, we have virtually no death records that match to more than one patient. While mortality status will be measured with error for some patients, there is no reason to expect that it is correlated with the proceduralist-physician expertise overlap. Thus, the extent to which patient's mortality is subject to error, we would be less likely to detect significant effects but without causing bias.

The data suggest that this measurement error is indeed negligible. Most notably, the 30-day mortality rate in our matched data is extremely similar to that estimated by a number of longitudinal studies of patients undergoing PCI in Brazil. For example, Machado et al. (2018) report a 30-day mortality rate of 6 percent among patients undergoing PCI, which is comparable to the 5.6 percent 30-day mortality rate observed in our data. Similarly, D'Avila et al. (2015) look at a sample of patients experiencing myocardial infarction who underwent PCI and find a overall mortality rate of 8.5 percent (the weighted-mean of columns 1 and 2 in row 1 of Table 4). In our sample, patients experiencing a heart attack have a 30-day mortality rate of 7.8 percent, which is fairly comparable to that in D'Avila et al. (2015).

The mortality pattern in the data is also consistent with a number of well-documented facts in the medical literature (see Appendix Figure F.1). At the most basic level, it is characterized by a very steep mortality risk with respect to age: for those over age 90, the 30-day mortality rate is 12 times that of those aged 40-50. In the same line, patients admitted because of myocardial infarction, the most severe condition, have much higher mortality rates than those admitted because of angina pectoris, the least severe condition, with differences on the order of 67 to 120 percent. Among patients experiencing a heart attack, those with a more severe diagnosis experience higher mortality rates than those with a less severe diagnosis. Indeed, patients with a ST elevation myocardial infarction exhibit a 30-day mortality rate of almost 8 percent, which is as much as 40 percent higher than that of patients with non-ST elevation myocardial infarction. The data also reveal that the 30-day mortality rate is systematically higher for women than men, which is consistent with the medical evidence that women have a higher incidence of adverse outcomes following PCI than men (Cowley et al., 1985; Maynard et al., 1997; Kelsey et al., 1993). This is striking given that the mortality rate in the general population is higher for men, suggesting that the gender-specific pattern in our data is unlikely to be an artifact of our merging strategy. Overall, the evidence suggests that our mortality indicator has useful empirical content and that any measurement error is unlikely to be systematic.

Figure F.1: 30-Day Mortality Across Subgroups



Notes. Panel A plots the percentage of patients who died within 30 days since the hospital discharge by age groups. Panel B plots the 30-day mortality rate by age and gender groups. Panel C plots the 30-day mortality rate by admission diagnosis.

References

Bavishi, Chirag, Sripal Bangalore, Dipen Patel, Saurav Chatterjee, Vrinda Trivedi, and Jacqueline E Tamis-Holland, "Short and long-term mortality in women and men undergoing primary angioplasty: A comprehensive meta-analysis," *International Journal of Cardiology*, 2015, *198*, 123–130.

Chen, Yiqun, "Team-specific human capital and team performance: Evidence from doctors," *American Economic Review*, 2021, *111* (12), 3923–62.

Chetty, Raj, "Bounds on elasticities with optimization frictions: A synthesis of micro and macro evidence on labor supply," *Econometrica*, 2012, *80* (3), 969–1018.

Cowley, MICHAEL J, SUZANNE M Mullin, SHERYL F Kelsey, KM Kent, AR Gruentzig, KM Detre, and ER Passamani, "Sex differences in early and long-term results of coronary angioplasty in the NHLBI PTCA Registry.," *Circulation*, 1985, *71* (1), 90–97.

Doll, Jacob A, Colin I O'Donnell, Meg E Plomondon, and Stephen W Waldo, "Contemporary Clinical and Coronary Anatomic Risk Model for 30-Day Mortality After Percutaneous Coronary Intervention," *Circulation: Cardiovascular Interventions*, 2021, *14* (12), e010863.

D'Avila, Alan Castro, Renato Roese Filho, Marcia Moura Schmidt, Karina Melleu, Cristiano de Oliveira Cardoso, Carlos Antonio Mascia Gottschall, Alexandre Schann de Quadros, and Alexandre Damiani Azmus, "Angioplastia coronariana primária em pacientes com mais de 80 anos," *Revista Brasileira de Cardiologia Invasiva*, 2015, *23* (4), 261–265.

Kelsey, Sheryl F, Margaret James, Ann Lu Holubkov, Richard Holubkov, MJ Cowley, and KM Detre, "Results of percutaneous transluminal coronary angioplasty in women. 1985-1986 National Heart, Lung, and Blood Institute's Coronary Angioplasty Registry.," *Circulation*, 1993, *87* (3), 720–727.

Kleibergen, Frank and Richard Paap, "Generalized reduced rank tests using the singular value decomposition," *Journal of econometrics*, 2006, *133* (1), 97–126.

Lee, David S, Justin McCrary, Marcelo J Moreira, and Jack Porter, "Valid t-ratio Inference for IV," *American Economic Review*, 2022, *112* (10), 3260–90.

Machado, Guilherme Pinheiro, Gustavo Neves de Araújo, Stéfani Mariani, Elvis Pellin Cassol, Felipe Homem Valle, Ana Maria Rocha Krepsky, Luiz Carlos Corsetti Bergoli, Sandro Cadaval Gonçalves, Rodrigo Vugman Wainstein, and Marco Vugman Wainstein, "On-vs. off-hours admission of patients with ST-elevation acute myocardial infarction undergoing percutaneous coronary interventions: data from a tertiary university Brazilian hospital," *Clinical and biomedical research. Porto Alegre. Vol. 38, n. 1 (2018), p. 30-34*, 2018.

Maynard, Charles, Nathan R Every, Jenny S Martin, Peter J Kudenchuk, and W Douglas Weaver, "Association of gender and survival in patients with acute myocardial infarction," *Archives of internal medicine*, 1997, *157* (12), 1379–1384.

Showalter, Mark H and Norman K Thurston, "Taxes and labor supply of high-income physicians," *Journal of Public Economics*, 1997, *66* (1), 73–97.