Gender Differences in Medical Evaluations: Evidence from Randomly Assigned Doctors *

Marika Cabral and Marcus Dillender †

April 2023

Abstract

Little is known about what drives gender disparities in health care and related social insurance benefits. Using data and variation from the Texas workers' compensation program, we study the impact of gender match between doctors and patients on medical evaluations and associated disability benefits. Compared to differences among their male patient counterparts, female patients randomly assigned a female doctor rather than a male doctor are 5.2% more likely to be evaluated as disabled and receive 8.6% more subsequent cash benefits on average. There is no analogous gender-match effect for male patients. Our estimates indicate that increasing the share of female patients evaluated by female doctors may substantially shrink gender gaps in medical evaluations and associated outcomes.

*For providing helpful comments, we thank Marcella Alsan, Sandra Black, Amitabh Chandra, Seema Jayachandran, Adriana Lleras-Muney, Heidi Williams, as well as participants of the NBER Summer Institute Health Care Meetings 2021, the Chicago Booth Junior Health Economics Summit 2020, and the ASSA annual meetings 2021. We thank Daniel Jordan Alvarez, Bokyung Kim, Mu Yang Shin, and Jinyeong Son for their excellent research assistance. Cabral gratefully acknowledges financial support from the National Science Foundation CAREER Award (1845190). Cabral and Dillender gratefully acknowledge financial support for this research from the US Social Security Administration. The research reported herein was performed pursuant to grant RDR18000003 from the US Social Security Administration (SSA) funded as part of the Retirement and Disability Research Consortium. The opinions and conclusions expressed are solely those of the author(s) and do not represent the opinions or policy of SSA, any agency of the Federal Government, or NBER. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of the contents of this report. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply endorsement, recommendation or favoring by the United States Government or any agency thereof.

†Cabral: University of Texas at Austin Department of Economics and NBER, marika.cabral@utexas.edu. Dillender: Vanderbilt University and NBER, marcus.o.dillender@vanderbilt.edu.

Gender disparities are observed in many settings ranging from education to labor markets to financial transactions. A growing body of evidence from medical studies suggests these disparities extend to health care settings, where recent work has shown that female patients—relative to male patients—receive less health care for similar medical conditions and are more likely to be told by providers that their symptoms are emotionally driven rather than arising from a physical impairment.[1] Differences in doctors' evaluations of medical issues for male and female patients may be a key factor contributing to observed differences in treatment. Beyond influencing the treatments patients receive, medical evaluations also impact benefit eligibility in social insurance programs. Recent evidence suggests there are large gender disparities in social insurance programs that rely on medical evaluations. For example, Low and Pistaferri (2019) show female applicants for Social Security Disability Insurance are 20 percentage points more likely to be denied benefits than similar male applicants. While there is mounting evidence that gender disparities exist in health care and related outcomes, there is little evidence on what drives these disparities and what policies may affect them.

One frequently discussed policy prescription to address gender disparities in health care is diversifying the physician workforce. Despite gender parity in the training of new physicians, nearly two-thirds of current physicians are men (Kaiser Family Foundation 2019). In specialties that often evaluate physical impairments, like orthopedics, the gender imbalance is even larger (Jagsi et al. 2014). Some have pointed to gender imbalances among physician evaluators as a potential contributor to gender disparities in disability evaluations in social insurance programs. For example, a recent class-action lawsuit filed against the California workers' compensation system alleges that doctors are more likely to classify female occupational injuries as being non-work-related than male occupational injuries and argues that medical examiners being overwhelmingly male likely contributes to male-female disparities in benefit determinations.[2]

The gender of doctors could play a role in medical evaluations if patients and doctors having similar characteristics and backgrounds improves patient-doctor communication or reduces provider bias against a patient or against the types of health issues that the patient is likely to have. However, studying the role of doctor gender in medical evaluations is difficult because both patients and doctors can often influence patient-doctor matches. In this paper, we overcome this key challenge by leveraging random assignment of doctors to patients for medical evaluations within the setting of workers' compensation insurance. To the best of our knowledge, our study provides the first estimates of the impact of the gender of the evaluating doctor on gender disparities in medical evaluations and subsequent social insurance benefit receipt. Further, we leverage additional administrative and survey data to provide evidence on potential mechanisms and to quantify gender homophily in patient-doctor matches in circumstances where patients have the ability to choose their own providers. Finally, we use our estimates to explore the policy implications of our findings.

Our study leverages comprehensive administrative data and random assignment of doctors to patients within the Texas workers' compensation insurance system. Random assignment of doctors to patients occurs in this setting through the dispute resolution process. Insurers and injured workers may request independent medical evaluations to settle disputes over an injured worker's impairment level. Medical doctors (MD), doctors of osteopathy (DO), and doctors of chiropractic (DC) can all apply through the Texas Department of Insurance to be included in the list of independent doctors eligible to perform these evaluations. Conditional on the injured worker's county, assignment of an injured worker to a doctor is random among

[1] For example, see Chen et al. (2008); Hoffmann and Tarzian (2001); Schulman et al. (1999); Tunks, Bellissimo, and Roy (1990).

[2] "Being a Woman Isn't a Pre-Existing Condition, Says This California Workers' Comp Suit" by Christina Cauterucci July 6, 2016, *Slate*.

doctors with the same credentials (i.e., MD, DO, and DC). Thus, after controlling for the injured worker's county and the doctor's credential, pre-determined characteristics of the injured worker should be unrelated to the gender of the evaluating doctor. The random assignment of doctors to patients means that differences in assessments between male and female doctors stem from the doctors themselves rather than from differences in the types of patients assigned to doctors.

An injured worker is eligible for continued cash benefits if the worker has an injury-related disability at the time of the independent exam, so a claimant will receive cash benefits after an exam if and only if the examining doctor assesses the worker as having a continued impairment. Our empirical approach compares the impact of being randomly assigned a female doctor rather than a male doctor on the subsequent benefits received by female claimants to the analogous impact of the gender of the assigned doctor on benefits for male claimants. Our study draws on unique, linked administrative data covering all medical bills and cash benefits paid for workers' compensation claimants in Texas, as well as information on all independent medical exams performed through the dispute resolution system. To validate our reliance on the random assignment mechanism, we verify that doctor gender is uncorrelated with pre-determined patient and injury observable characteristics. We then estimate the differential impact of the gender of the assigned doctor on post-evaluation benefit receipt for exams for female patients versus exams for male patients.

There are a number of strengths of our research design and setting. First, the combination of several features of our setting—the random assignment of doctors, a diverse pool of patients and doctors, the large scale of the randomized evaluations, and the availability of administrative data—allow for a transparent and comprehensive evaluation of the effects of gender match on medical evaluations. Second, while doctors' evaluations of patients have the potential to influence virtually all health care interactions, studying these evaluations is challenging and several key features of this setting allow us to overcome these challenges—the standardized mandate of exams in this setting, the need to formalize assessments, and the data on later outcomes based on these exams. Finally, these exams determine an important outcome for the patient—eligibility for subsequent cash disability benefits—and the availability of broader administrative data allows us to provide context for our findings.

Beyond providing a useful setting to investigate patient-doctor gender match, independent medical evaluations are a central component of workers' compensation programs nationwide. For instance, 39% of claims with cash benefits in Texas in 2013 had a dispute that necessitated an independent medical exam, and these disputed claims accounted for 56% of program-wide claim costs or more than $800 million per injury year.[3] Policies surrounding these exams—including rules for determining the pool of eligible doctors and the doctor assignment processes—are a subject of constant policy debate in workers' compensation programs.[4] Our study provides evidence on how such policies can affect claimant outcomes and gender disparities in these outcomes. In addition, the subject of the exams—evaluation of pain and disability—and gender disparities in these types of evaluations are of interest in other related social insurance programs—including the Veterans Affairs disability program and Social Security Disability Insurance.[5] More broadly, many social insurance and means-tested programs rely on program gatekeepers to determine benefit eligibility, and our findings point to the potential importance of gender match between program gatekeepers and claimants in other government programs.

Our estimates indicate that patient-doctor gender match increases evaluated disability and subsequent

[3] Similar rates of dispute-triggered independent medical exams are observed in other large workers' compensation programs. For example, in 2018, California and New York, which have more workers covered by workers' compensation insurance than all other states excluding Texas, had rates of more than 1 independent medical evaluation for every 3 claims.

[4] For examples, refer to Department of Consumer and Business Services (2004) and Mulcahy et al. (2020).

[5] For examples, refer to Department of Veterans Affairs (2010); Murdoch et al. (2021); Trabanino (2020); Daly et al. (2013).

cash disability benefits for female patients but has little impact on outcomes of male patients.[6] Compared to differences among their male patient counterparts, female patients randomly assigned a female doctor rather than a male doctor are 3.1 percentage points more likely to be evaluated as having an ongoing disability and receive 8.6% more cash benefits on average—or $483 evaluated at the mean of $5,622. There is no analogous gender-match effect for male patients. We note the magnitude of these effects is sizable—the estimated 3.1 percentage point increase in the likelihood of being evaluated as disabled is nearly large enough to offset the entire observed gender gap in this outcome when male doctors evaluate claimants.

We explore potential mechanisms through additional analysis within the setting of these randomized evaluations and through a supplemental survey we conduct. We investigate whether the gender-match effect we document appears to arise from mechanisms inherently linked to gender rather than other patient or doctor characteristics that are correlated with gender. We find no significant differences in outcomes following exams by female and male doctors by other patient characteristics (aside from patient gender), and no differences in outcomes across male and female patients by other doctor characteristics (aside from doctor gender). Heterogeneity analysis reveals the gender-match effect for female patients appears across the board for exams with differing baseline patient characteristics, suggesting that gender match has near universal impacts on the outcomes of female patients. We contextualize our findings using rich baseline data on patients to evaluate whether patient-doctor gender match ameliorates or magnifies differences in the evaluation of male and female patients conditional on observables. Controlling for available baseline patient information, the estimates indicate that female doctors evaluate female and male patients as similarly disabled while male doctors evaluate female patients as less disabled than male patients. While only suggestive, this evidence is consistent with male doctors evaluating female patients against a stricter standard than male patients and female doctors applying similar standards to male and female patients.

To complement our analysis of independent medical exams, we present novel evidence from a survey we conducted which suggests stark gender differences in interactions within the health care system and suggests possible mechanisms behind the importance of patient-doctor gender match in medical evaluations. The survey responses suggest that women—relative to men—more often report having a negative experience where a doctor didn't understand their concerns, had assumed something without asking, talked down to them, made them feel uncomfortable, or didn't believe them. When asked about how a doctor's gender influences the likelihood of having a positive interaction, women were much more likely than men to report an own-gender doctor would be more likely to treat them with respect, understand their concerns, believe them, provide needed testing and treatments, make them feel comfortable, and ask appropriate questions instead of making assumptions. The asymmetry of responses between men and women supports the plausibility of the main finding from the randomized evaluations—that doctor gender is an important factor in the evaluation of female patients more so than in the evaluation of male patients—and suggests that differences in doctor behavior may be an important mechanism behind the main finding.

While our primary analysis focuses on disputed claims with independent medical exams, we use comprehensive data on all workers' compensation claims in Texas to contextualize our findings. Paralleling gender gaps observed among disputed claims, female workers' compensation claimants in the broader population are 3.6 percentage points less likely to receive cash benefits than their male counterparts with the same observable characteristics, which represents a 15.3% reduction relative to the mean rate of benefit receipt

[6]We note that the effects we estimate may run through both gender and sex. Because gender and sex are highly correlated in practice and it is often unclear which is reflected in the data, we cannot distinguish the extent to which the effects we estimate reflect the importance of gender identity as opposed to biological sex. For simplicity and to follow conventions in the economics literature, throughout we use gender to refer to both gender and sex.

among male claimants. When given the choice of doctors, female workers' compensation patients select a female treating doctor 1.4 percentage points—or 5.1%—more often than comparable male patients select a female treating doctor. While these estimates could reflect the preferences of both men and women to see own-gender doctors, evidence from the survey we conducted suggests that women—compared to men—more often express a preference to see own-gender providers, more often select own-gender providers when given the choice, and report being willing to pay more to see an own-gender provider.

Finally, we explore the policy implications of our findings. Our estimates imply that increasing the share of independent medical evaluations performed by female doctors from 17% to 50% would cause a 0.88 percentage point increase in the share of female patients evaluated as disabled, closing approximately 41% of the gender gap conditional on observables among disputed claims. An alternative random assignment mechanism that randomly assigns doctors of the same gender to evaluate disputed claims—rather than the current gender-blind random assignment process—would increase the share of female patients who receive cash benefits by 2.2 percentage points, offsetting the entire gender gap conditional on observables among disputed claims. We also discuss the potential policy implications of our findings for gender gaps in workers' compensation insurance more broadly, beyond disputed claims. Overall, this analysis reveals that a structural feature of the health care system—the under-representation of female doctors—is a major contributor to gender disparities in medical evaluations in this setting and indicates that policies aimed at increasing gender diversity among doctors or relative gender homophily in patient-doctor matches may substantially reduce gender gaps in evaluated disability.[7]

Beyond providing estimates that inform ongoing policy debates, this paper contributes to several literatures. First, our paper is motivated by a recent medical literature documenting gender gaps in health care treatments and related outcomes.[8] Despite growing evidence of gender disparities in health care and related settings, it is unclear what drives these disparities. Our work contributes to this literature by providing the first evidence on the role of doctors—and in particular, the role of patient-doctor gender match—in contributing to gender disparities in medical evaluations and linked social insurance benefits. Second, our work also contributes to an emerging literature considering the impacts of race or gender match between doctors and patients in health care settings (Greenwood, Carnahan, and Huang 2018; Hill, Jones, and Woodworth 2020; Alsan, Garrick, and Graziani 2019). Our paper extends this literature by leveraging random assignment to provide the first estimates of the causal effect of gender match between patients and doctors on medical evaluations—the key first step in most health care interactions. Further, our evidence on gender homophily in patient selections of providers when patients have choice complements prior work illustrating gender homophily in patient selections of urologists (McDevitt and Roberts 2014) and in referrals among doctors (Zeltzer 2020).[9] Finally, our findings contribute to a larger literature investigating the impact of the gender of authority figures on outcomes for women in contexts ranging from education to the labor market. Our results highlight the potential broader importance of gender match in settings where authority figures evaluate women.[10]

[7] By identifying the composition of doctors as a structural factor that is a major contributor to gender disparities in medical evaluations in this setting, our work complements an emerging literature in economics identifying systemic discrimination (Kline, Rose, and Walters, 2022; Bohren, Hull, and Imas, 2022) and a qualitative medical literature discussing structural factors leading to gender differences in health care treatments (Hoffmann and Tarzian, 2001; Samulowitz et al., 2018).

[8] See Cabral and Dillender (2021) for a more comprehensive review of this literature. Beyond work in the medical literature, a few other recent papers in the economics literature explore gender differences in health care settings. For examples, refer to Chandra and Staiger (2010), who document and analyze racial and gender gaps in heart attack treatments, and to Dupas and Jain (2021), who study gender disparities in health care treatments in India.

[9] Further, our evidence on gender disparities in evaluations of patients complements work by Sarsons (2017), which documents gender disparities in how physician ability is interpreted following positive and negative patient outcomes.

[10] For example, see related work in the setting of education (e.g., Carrell, Page, and West 2010; Muralidharan and Sheth 2016; Lim and

I Background and Data

Workers' Compensation Insurance Workers' compensation is state-regulated insurance that provides covered employees cash and medical benefits for work-related injuries. While there is variation across states in the details of workers' compensation insurance, the basic structure is similar across states. Workers' compensation insurance provides cash benefits to those with temporary or permanent disability due to work-related injuries. In addition, workers' compensation insurance covers all injury-related medical spending at no out-of-pocket cost to the claimant, regardless of a claimant's work status or receipt of cash benefits.

A key feature of workers' compensation insurance is that medical providers evaluate impairments and determine eligibility for cash benefits. In Texas, as in many states, the delivery of medical care and the oversight of cash benefit eligibility in workers' compensation insurance follow a "gatekeeper" model. Workers' compensation claimants choose a "treating doctor", who is responsible for overseeing the claimant's medical care, evaluating the claimant's medical improvement, and assessing any permanent impairment the claimant may have. In addition to receiving reimbursement for typical procedures billed by providers, doctors treating workers' compensation claimants receive payments for additional "case management services" that pertain to their particular gatekeeper role in overseeing benefit eligibility of injured workers.

There are two main types of cash benefits: temporary income benefits and permanent impairment benefits. In Texas, injured workers are eligible for temporary income benefits after a waiting period of seven days, and these benefits are terminated at the earliest of the following: (i) the employee decides to return to work, (ii) the employee's doctor determines that the employee has reached "maximum medical improvement", and (iii) the statutory maximum duration is met. The statutory maximum duration in Texas is two years (104 weeks), but workers rarely receive temporary income benefits for two years. Temporary income benefits replace 70% of a claimant's prior average weekly wage, subject to a maximum and minimum weekly benefit level. Once workers have reached maximum medical improvement or have exhausted their temporary income benefits, they are eligible for permanent impairment benefits if they are still impaired. Treating doctors who conclude that a claimant has a permanent impairment assign an impairment rating—the percentage of permanent impairment of the whole body resulting from the injury. The worker is then paid an unconditional cash transfer that is a function of the severity rating of his permanent impairment and his prior average weekly wage. In Texas—as in many states—permanent impairment benefits are calculated as the product of the claimant's permanent impairment severity rating and the eligible wage replacement rate, where the replacement rate is linear in the claimant's pre-injury wage subject to a maximum benefit.

Independent Medical Evaluations While the treating doctor typically serves as the gatekeeper for cash benefits, insurers and/or injured workers may disagree with the treating doctor's assessment of a claimant's impairment level. As with most states, Texas maintains a database of independent doctors—referred to as "designated doctors"—who can resolve these disputes by performing independent medical evaluations. Either an insurer or an injured worker can request an independent evaluation if they disagree with the treating doctor's assessment, though during our analysis period insurers requested 87% of these exams.

We focus on independent exams where the designated doctors have been asked to assess the workers' maximum medical improvement and/or permanent impairments. Nearly all injuries (99%) that go through

Meer 2019), the labor market (e.g., Biasi and Sarsons 2022; Cullen and Perez-Truglia 2019; Casarico and Lattanzio 2019; Bertrand et al. 2018; Maida and Weber 2019), and academic publishing (Abrevaya and Hamermesh 2012; Card et al. 2019). While gender-match effects in educational and professional settings may run through a complex set of mechanisms (e.g., role model effects, differing management or teaching styles, networking, mentoring, discrimination), the authority figures in our setting (doctors) have a relatively narrow mandate—to medically evaluate a patient's ongoing degree of disability—limiting the set of the plausible mechanisms underlying effects in this setting.

the designated doctor process during our period of analysis meet this criteria.[11] Because permanent impairments are assessed after maximum medical improvement has been reached, designated doctors are nearly always ($> 99.8\%$ of the time) asked to assess both dimensions in a single examination if either determination is requested, though the request for the permanent impairment determination is only relevant if the designated doctor assesses the injured worker as having reached maximum medical improvement. In these exams, the designated doctor is charged with assessing the claimant's healing trajectory and degree of ongoing disability, taking as given the diagnosis of the claimant's injury prior to the exam.[12]

Medical doctors (MD), doctors of osteopathy (DO), and doctors of chiropractic (DC) can apply to be designated doctors through the Texas Department of Insurance (TDI). Their reports are given presumptive weight in dispute resolution, and insurance carriers are required to pay income and medical benefits based on the opinion of the designated doctor. Once a designated doctor has been assigned to a claim, that doctor is the designated doctor for the initially requested independent medical exam and for any subsequently requested independent medical exams for that claim should future disputes arise. The role of the designated doctor exam is narrow and well-defined. The designated doctor is called upon by the TDI to serve as an "impartial, objective medical expert" to settle specified claim-related disputes. Designated doctors do not have previous or ongoing relationships with patients outside of independent medical exams and are not allowed to recommend or provide treatment to these patients. To become certified as a designated doctor, doctors need to complete required training and must re-certify every two years with the TDI.

Texas law defines the scope of designated doctor evaluations, the information doctors are allowed to consider, and the requirements for designated doctors when reporting and justifying their decisions. Designated doctors must review a claimant's medical records—which are required to be supplied by the claimant's treating doctor and insurer—prior to the designated doctor exam. The designated doctor must also review any medical records and medical history information supplied by the claimant. At the designated doctor exam, the doctor conducts a physical examination of the claimant—often including range of motion testing and tests specific to the affected body system(s). The designated doctor then compares the information gathered through the physical examination and medical records with specified guidelines for these evaluations.[13] In reporting the results of the exam, the designated doctor is required to justify his or her decision based on the evidence gathered and specified medical guidelines. Importantly, the guidelines for these exams emphasize the inherent subjectivity in assessing whether a worker is continuing to heal from his/her injury (e.g., whether the workers' condition is expected to continue to improve with further passage of time, medical treatment, or home exercises) and the degree of permanent impairment from an injury (e.g., whether and how the worker's activities continue to be constrained from having been injured), and the guidelines suggest doctors should use their discretion in these decisions.[14]

The assignment of designated doctors to claimants is designed to be random within a local area, among

[11]Though rare, designated doctors can be asked to settle disputes not about maximum medical improvement and permanent impairments, such as whether the impairment was caused by a compensable injury and the extent of the worker's compensable injury.

[12]When these designated doctor exams are requested, the diagnosis code describing the claimant's injury is pre-specified on the request form, and the diagnosis of the claimant's injury is not the subject of the dispute. Instead, the designated doctor takes the diagnosis code information as given and is charged with evaluating the content of the dispute—the healing trajectory of the claimant's injury and the degree of ongoing disability of the claimant.

[13]Texas law states that designated doctors will apply when appropriate the American Medical Association Guides for the Evaluation of Permanent Impairment and Texas Department of Workers' Compensation return-to-work guidelines. Designated doctors will also consider Texas Department of Workers' Compensation treatment guidelines and other evidence-based medical guidelines when appropriate (28 TAC §127.200(a)(7)).

[14]Note that there is likely substantial scope for doctor discretion in these exams broadly, regardless of the nature of the injury or the ease in verifying that the injury occurred (or occurred at work). For instance, while it is straightforward to verify if a worker has a torn ligament through MRI imaging, it is subjective to assess whether the worker is continuing to heal from the torn ligament and the impact of the resulting pain, swelling, and mobility limitations on the worker's permanent work capacity.

doctors eligible to perform an exam. Once an independent evaluation has been requested, the claimant is assigned to the next available designated doctor within the claimant's county who has the appropriate credential to evaluate the injury.[15] Any designated doctor (MD, DO, or DC) can evaluate musculoskeletal injuries, while only MDs and DOs are authorized to evaluate certain non-musculoskeletal injuries, like mental and behavioral disorders.[16] As the vast majority of workers' compensation claims and independent exams are for musculoskeletal injuries, all providers are eligible to do most exams.

Prior to 2013, a representative at the TDI assigned doctors to claimants. But to prevent the possibility of manipulation and to ensure that the process approximates random assignment as designed, in 2013 the TDI implemented an automated system to assign designed doctors to independent medical exams. The TDI maintains a confidential list of designated doctors in each county. The automated system cycles through this list, assigning the next available designated doctor in the claimant's county who has the appropriate credentials to perform the assessment and designated doctors move to the bottom of the list upon assignment. This automated assignment process is confidential, and insurers and claimants do not observe the order of—or current position within—the list of certified designated doctors in each county. Since insurers and claimants have no way of knowing who is the next available designated doctor, they cannot time their requests so that they have a specific designated doctor. Thus, after controlling for the claimant's county and the designated doctor's credential, the designated doctor assigned to a claimant is random among designated doctors with that credential in the claimant's county.[17]

Claimants receive cash benefits after a designated doctor exam if and only if the designated doctor assesses the claimant as having an impairment on the day of the exam. Conditional on the designated doctor viewing the claimant as being impaired, the type of benefits that the claimant receives depends on the doctor's assessment of the injury's healing trajectory. If the designated doctor decides that a claimant is impaired but that the claimant has reached maximum medical improvement, the designated doctor assigns the claimant an impairment rating, and the claimant receives additional permanent impairment cash benefits based on that rating. If the designated doctor decides that a claimant is impaired and still healing from an injury, the claimant can receive additional temporary income benefits and can then be assessed for a permanent impairment at a later date, either by the designated doctor or by the claimant's treating doctor.

[15] Beyond having the appropriate credentials, the assigned designated doctor must not have a conflict of interest with the claimant or the claimant's insurer. For example, the designated doctor cannot have treated the patient prior to the independent exam and cannot have a network contract with the insurer associated with the claim. While our baseline empirical strategy does not account for exclusions based on conflicts of interest, supplemental evidence offers reassurance that conflicts of interest exclusions are unlikely to affect our results. For instance, estimates in Table 2 illustrate that baseline observables are balanced by assigned doctor gender, suggesting that the controls in our baseline specification may be sufficient for isolating conditional random variation in the assigned doctor's gender. Further, we obtain estimates similar to our baseline estimates in specifications with richer fixed effects that work toward accounting for exclusions due to conflicts of interest—insurer by county by credential by exam year fixed effects and insurer by county by credential by exam quarter fixed effects. See Appendix Section B for further discussion of this robustness analysis.

[16] As discussed further below, our research design will isolate as good as random assignment of doctors to claimants conditional on claimant county and provider credential (MD, DO, or DC). For this reason, we exclude less than 3% of designated doctor exams for claimants with specified diagnoses that the TDI flags as requiring an evaluation by an MD or DO with a specific sub-specialty. Certain rare complex injuries, such as traumatic brain injuries or spinal cord injuries, require an evaluation by an MD or DO with certain board-certified sub-specialties. Refer to 28 TAC §127.130(b) for more detail. We use diagnosis codes from the form requesting the designated doctor exam to identify and exclude claimants whose exams potentially require specialized expertise.

[17] As we describe further in Section II, our estimation approach controls for the credential of the assigned designated doctor (known ex post) by claimant county fixed effects, as the designated doctor assigned to a claimant is random among designated doctors with that credential in the claimant's county. An alternative approach to isolate conditional random assignment would be to instead control for each injury type by county combination—the ex ante information the regulator uses to decide which credentials are required of the designated doctor who will evaluate a claimant. We prefer the former approach as it is more robust if the available injury type data imperfectly captures the information used by the regulator to determine required designated doctor credentials. Nevertheless, we find very similar estimates in an alternative specification that controls for observable injury type by county fixed effects. See Appendix Section A for further details on approaches to isolate random assignment and Appendix Tables A4 and A5 for estimates from this alternative specification.

For claimants assessed as having a continued impairment, designated doctors can decide the amount of future benefits at the initial visit by assigning a permanent impairment rating or by assigning a future date at which maximum medical improvement will be reached. If designated doctors assess a claimant as having a continued impairment that is still healing and decline to provide a future date at which maximum medical improvement will occur, decisions about the amount of future benefits revert to the claimant's treating doctor, unless and until the insurer or claimant requests another designated doctor evaluation. Since treating doctors' benefit decisions and insurers' (and workers') decisions about requesting further independent medical evaluations may be influenced by the decisions that the designated doctor makes in the initial evaluation, we focus on all benefits received after the initial evaluation. While designated doctors have sole discretion over whether claimants receive any additional cash benefits after the designated doctor exam, employers, insurers, treating doctors, and claimants can all influence the final amount of cash benefits that claimants assessed as having a continued disability eventually receive. However, designated doctors can still influence benefit amounts, both through their initial assessments and because a claimant's original designated doctor is responsible for any future independent medical evaluations.

Data and Sample To consider the role of doctor's gender in claimants' outcomes, we have compiled a unique administrative dataset through open records requests submitted to the TDI (TDI 2018*a*, TDI 2018*b*, TDI 2018*c*, TDI 2018*d*, and TDI 2018*e*). We obtained data on all independent medical exams that occurred from 2005 to 2017. Since a lack of influence on the designated doctor assignment process is crucial to the empirical strategy, we focus on designated doctor evaluations for injuries that occurred from 2013 to 2017 so that claimants are assigned to a designated doctor through the automated system. We focus on benefits received within 12 months of the exam date and choose 2017 as the exam year cutoff so that we have information on claimants' benefit receipt for at least one year after designated doctor exams in the sample. Since subsequent designated doctor exams are endogenous to the first designated doctor exam, we focus on the date of the first independent medical evaluation. In addition to conducting analysis that focuses on the sample of claimants with randomized medical evaluations, we also use the full sample of claims from 2013 to 2017 to study gender differences in benefit receipt and in the choice of treating doctors more broadly.

We combine the information on designated doctor visits with data on the cash and medical benefits that claimants receive. The cash benefit data include claimant-level information on type of cash benefits received, total benefits received, benefit replacement rate, benefit start and end dates, and injury date (month-year).[18] The medical benefit data come from all medical, pharmacy, and case management bills paid for by workers' compensation insurance and include the following information: procedure type (CPT codes), amount paid, diagnoses codes, date, and provider information. The data also include claimant gender, birth date (month-year), injury date (month-year), and zip code.

We merge in data on doctor gender from the Centers for Medicare & Medicaid Services' (CMS) National Provider Identifier (NPI) registry (CMS 2019). We supplement these data with additional data on doctor characteristics from the Medicare Physician Compare File (CMS 2021) and from open records requests to the Texas Board of Chiropractic Examiners (Texas Board of Chiropractic Examiners 2021) and the Texas Medical Board (Texas Medical Board 2014).[19] We adjust all monetary values to 2017 dollars using

[18]For claimants who receive cash benefits, the data also include information on pre-injury average weekly wage and industry. We use this information for some summary statistics comparing our sample to broader populations in the appendix and for some supplemental analysis on potential mechanisms in Section III. While wage and industry information is useful for describing the sample and providing context for our estimates, it is important to note that exam outcomes can influence whether this information is present in the data. Thus, our primary analysis does not use information on pre-injury wages or industry, and supplemental analysis using this information should be interpreted with caution.

[19]We also obtained the list of active DCs in Texas (Texas Board of Chiropractic Examiners, 2020), as well as various crosswalks (CMS

price information from U.S. Bureau of Labor Statistics (2021). We focus on evaluations for which we know the gender of both the provider and the claimant. We exclude the 3.6% of claimants with missing gender information from all analyses. Our final sample has 70,748 independent medical exams performed by 1,298 designated doctors. Refer to Appendix Section A for additional details about the data and setting, including about the construction of insurer and diagnosis fixed effects used as controls in some analyses and about the prevalence and importance of independent medical evaluations in the Texas workers' compensation insurance system.

We focus on two measures of benefit receipt after a designated doctor exam.[20] The first is an indicator variable equal to one if the claimant receives any cash benefits in the year after the exam. As described above, this variable equals one if and only if the doctor assesses an injured worker as being disabled at the time of the exam. To assess the value of benefits received after the exam, we create a second variable—"normalized additional cash benefits"—to reflect the total cash benefits received in the year after the exam valued at the mean benefit rate for each benefit type based on the population-wide distribution of pre-injury earnings.[21] See Appendix Section A for more details on the construction of this measure.

Table 1 Panel A displays means of key variables for male and female claimants, both for the full set of claimants with injuries occurring from 2013 to 2017 and for the subset of claimants who received independent medical evaluations by 2017. In the sample of claimants with independent medical evaluations, male claimants have higher mean first-day and three-month medical spending than females. Male claimants are also more likely to first receive care in the emergency department than female claimants are and receive income and impairment benefits at higher rates than female claimants do. These gender differences are also seen among all workers' compensation claimants and may reflect gender differences in female and male workers more generally. While there are also notable differences for men and women in the types of injuries that receive independent medical exams, similar gender differences also exist among all claimants. While gender differences parallel one another across these populations, it is important to emphasize that appropriate caution should be used in extrapolating from the estimated impact of gender match from independent medical exams to broader contexts. Section VI discusses the interpretation of our evidence further.

Table 1 Panel B displays characteristics of doctors performing evaluations in Texas and of doctors more broadly. Columns 1 describes the 1,298 designated doctors who perform independent medical exams during our analysis period, while column 2 describes the broader sample of doctors treating Texas workers' compensation patients. There are a few notable patterns. The set of designated doctors performing independent medical evaluations makes up more than 23% of all doctors treating workers' compensation patients and more than 2% of all doctors practicing in Texas more broadly. Second, the share of doctors who are female is similar among designated doctors and all doctors treating workers' compensation patients. Third, relative to the population of doctors treating workers' compensation patients, a larger share of designated doctors have a DC credential rather than an MD or DO credential. Finally, among doctors with MD or DO credentials, the distribution of specialties and the share who graduated from a Top 25 medical school are broadly similar among designated doctors and doctors treating workers' compensation patients.

2019; NBER 2021; United States Census Bureau 2010; Dartmouth Atlas 2013).

[20]Our primary focus is on outcomes for the claimant after the exam that are influenced by the designated doctor's evaluation—specifically, cash disability benefits received by the claimant after the exam. We note that the medical claims data for designated doctor exams do not allow us to observe information about doctor effort or behavior during designated doctor exams—e.g., testing conducted or resources expended during these exams.

[21]The normalized additional cash benefits variable is a continuous measure of compensated disability after the exam. Because actual benefits are a function of both disability and pre-injury earnings, we value the benefits at the mean benefit rate based the population-wide distribution of pre-injury earnings which allows us to interpret this as a continuous measure of disability that is not influenced by an individual's own pre-injury earnings.

For comparison, columns 3 and 4 display characteristics of all doctors in Texas and in the United States, respectively. Relative to doctors overall in Texas, a larger share of doctors treating workers' compensation patients hold DC credentials and specialize in orthopedics or internal/family medicine. Among those with an MD or DO credential, doctors treating workers' compensation patients in Texas and doctors overall in Texas are similarly likely to have graduated from a Top 25 medical school. Compared to doctors treating workers' compensation patients, the shares of doctors who are female are also higher in columns 3 and 4, though there are still about twice as many male doctors as female doctors among doctors treating workers' compensation patients. Along the dimensions considered, Texas doctors are similar to doctors nationally.

II Empirical Strategy

Estimating Equation Our empirical strategy takes advantage of random assignment of designated doctors to claimants. Conditional on the designated doctor's credential and the claimant's county, the assignment of designated doctors to claimants is random, and thus a claimant's underlying characteristics should not be related to the gender of the designated doctor. A key advantage of our setting is that we observe evaluations for claimants of both genders. This allows us to estimate and control for any across-the-board differences in evaluations preformed by male and female doctors. Our baseline specification is as follows:

$$
\begin{aligned}
y_i = &\beta_1 female\_claimant_i + \beta_2 female\_doctor_{d(i)} \\
&+ \beta_3 female\_claimant_i * female\_doctor_{d(i)} + \theta_{c(i)r(i)} + f(\mathbf{X_i}) + \epsilon_i,
\end{aligned} \tag{1}
$$

where i indexes the claimant, $c(i)$ indexes the claimant's county, $d(i)$ indexes the doctor assigned to claimant i, and $r(i)$ is the credential of the doctor assigned to claimant i. In this specification, y_i represents the dependent variable, $female\_claimant_i$ is an indicator for the claimant being female, $female\_doctor_{d(i)}$ is an indicator for the designated doctor performing the exam being a female, $\theta_{c(i)r(i)}$ is a vector of fixed effects for a claimant's county by the doctor's credential, and $\mathbf{X_i}$ is a set of additional controls included in some specifications that describe characteristics of the claimant, doctor, and initial evaluation. In our baseline specification, $\mathbf{X_i}$ includes controls for the year of the claimant's injury and the year of the exam. Because the pool of available designated doctors in a county may change over time, we illustrate our results are robust in alternative specifications that control for interactions of our claimant county by doctor credential fixed effects with the exam year or exam quarter. We also illustrate that our results are similar if we include interactions between the county by credential fixed effects and $female\_claimant_i$. See Appendix Section B for further discussion of these results. Standard errors are clustered at the doctor level.[22]

Within this specification, β_2 represents the difference in outcomes among male claimants randomly assigned to a female doctor relative to those assigned to a male doctor, while $\beta_2 + \beta_3$ represents the analogous difference in outcomes among female claimants. Comparing these differentials across claimants, we can interpret β_3—the main coefficient of interest— as representing the differential impact of the assigned doctor's gender on outcomes of female patients relative to male patients.[23] While our discussion largely

[22]We cluster at the doctor level to account for the possibility that residuals in doctors' exam decisions may be correlated across exams. Appendix Tables A1 and A2 verify that the results are robust to other clustering choices and to not clustering standard errors.

[23]Throughout, we discuss β_3 in Equation (1) as representing the impact of having a female doctor on female claimants relative to the impact of having a female doctor on male claimants. We note that β_3 also represents the sum of the gender concordance effects for female claimants and male claimants. To see this, note that the difference-in-difference estimand can be written as:

$$
\begin{aligned}
&\big(E[y_i|claim_i = f, doc_i = f] - E[y_i|claim_i = f, doc_i = m]\big) - \big(E[y_i|claim_i = m, doc_i = f] - E[y_i|claim_i = m, doc_i = m]\big) \\
&\quad = \big(E[y_i|claim_i = f, doc_i = f] - E[y_i|claim_i = f, doc_i = m]\big) + \big(E[y_i|claim_i = m, doc_i = m] - E[y_i|claim_i = m, doc_i = f]\big),
\end{aligned}
$$

where $claim_i$ represents gender of claimant i and doc_i represents the gender of the doctor assigned to claimant i, where both variables take values either f for female or m for male.

focuses on the estimates of β_3, we also discuss estimates of β_2 as these capture whether female doctors are systematically more or less generous in injury assessments of male claimants compared to their male doctor counterparts. We also report β_1 throughout, so we can contextualize the main estimates relative to gender gaps in benefit receipt among claimants evaluated by male designated doctors.

Identifying Variation Given the conditional random assignment of designated doctors to claimants, the gender of the assigned designated doctor should be unrelated to pre-determined claimant characteristics after controlling for the doctor's credential and claimant's county. An advantage of our setting is that we have rich baseline data on pre-determined claimant observable characteristics to verify that the assignment of doctors appears random and orthogonal to baseline claimant characteristics. We estimate Equation (1) replacing the dependent variable with baseline claimant characteristics. In this analysis, we examine the claimant's age as well as baseline claim characteristics that are determined before the designated doctor assignment, including measures that capture the severity of the injury: an indicator for whether the claim originated with an emergency department visit, medical spending on the first day of claim, total medical spending prior to the designated doctor exam, an indicator for whether the claimant received income replacement benefits prior to the designated doctor exam. We also examine two additional claim characteristics: the impairment type and the duration of time between the injury and the designated doctor exam.

Table 2 presents the resulting coefficient estimates and the associated standard errors and p-values. There are a few patterns to note. First, as expected, the assigned designated doctor's gender is uncorrelated with baseline claimant and injury characteristics, with coefficient estimates that are small and statistically indistinguishable from zero for female doctor (column 4) and the interaction of female doctor and female claimant (column 1). This suggests the random assignment of designated doctors was implemented as required by state regulation. Second, the coefficient estimates for female claimant (in column 7) indicate meaningful differences in characteristics comparing female and male claimants with independent medical exams. Many of these differences reflect differences between female and male workers—and workers' compensation claimants—more generally.[24] Throughout our analysis, we include a female claimant control to account for any systematic differences in the propensity of designated doctors to award further benefits to female claimants relative to male claimants.

III Differences in Evaluations by Doctor Gender

A Baseline Results from Randomized Evaluations

Table 3 displays the key coefficient estimates from Equation (1) for the outcomes of interest. Columns 1 through 3 report the results from an OLS regression investigating whether a claimant was assessed as having a continued disability (or equivalently, whether the claimant received any additional cash benefits), while columns 4 through 6 report the results from a Poisson regression for total normalized additional cash benefits received after the exam.[25] We focus first on the results from the baseline specification, reported in columns 1 and 4. Using the coefficient estimates from the baseline specifications, Figure 1 reports the estimated effect of being evaluated by a female doctor (instead of a male doctor) separately for male patients (β_2) and female patients ($\beta_2 + \beta_3$) along with the estimated differential impact for female patients (β_3). The estimates indicate that female patients randomly assigned a female doctor are 3.1 percentage points [95%

[24] See the discussion of Table 1 in Section I for a comparison of female and male workers' compensation claimants, both in the universe of claimants and among the subset of claimants with independent medical evaluations.

[25] Appendix Table A3 illustrates that we obtain similar implied percent effects when estimating an alternative linear specification for normalized subsequent cash benefits and when estimating an alternative specification that considers unadjusted subsequent cash benefits while controlling for pre-injury wage.

C.I.: 1.1 to 5.1 percentage points], or 5%, more likely to be evaluated as having a continued disability relative to female patients assigned a male doctor. In contrast, the estimated effect of being assigned a female doctor (rather than a male doctor) for male patients is small and statistically insignificant, suggesting that male and female doctors make similar disability assessments when evaluating male claimants. The estimated differential effect of being evaluated by a female doctor (instead of a male doctor) for female patients relative to male patients is 3.1 percentage points; this estimate is statistically distinguishable from zero (p-value 0.003) and large enough to nearly close the 3.2 percentage point gender gap seen when male doctors are the evaluators.

We observe similar patterns when investigating the total normalized cash benefits claimants receive after the exam. Relative to female patients evaluated by male doctors, female patients evaluated by female doctors receive 11.8% more in subsequent benefits on average. There is no analogous gender-match effect for male patients. In fact, the point estimates suggest that being assigned a female doctor (rather than a male doctor) may increase the subsequent cash benefits male claimants receive by 3.2% on average, though this estimate is statistically indistinguishable from zero and is small relative to the 11.8% mean increase in subsequent benefits for female patients assigned a female doctor. Relative to differences among male patients, the effect of being assigned a female doctor rather than a male doctor for female patients is sizable at 8.6%, with a 95% confidence interval spanning 2.7% to 14.5%. This translates to an additional $484 dollars evaluated at the mean of $5,622 and represents a 61% closure of the gender gap in cash benefits observed when male doctors evaluate patients.

The magnitudes of the estimates are large relative to observed gender gaps in both outcomes. Overall, the estimates indicate that the gender of the evaluating doctor is an important determinant of gender gaps in this setting and suggest that the under-representation of females among doctors is likely a major explanation behind observed gender gaps in these outcomes. We probe the robustness of these main findings in a number of ways. The remaining columns of Table 3 illustrate the robustness of the results to the inclusion of doctor fixed effects (columns 2 and 5) or the exclusion of male claimants from the estimation data (columns 3 and 6). In Appendix Tables A4 and A5, we demonstrate the robustness of our findings to the inclusion of different combinations of fixed effects.

B Mechanisms and Additional Evidence from Randomized Evaluations

There are several reasons why gender match between doctors and patients could matter for medical evaluations. Differences in evaluations across male and female doctors may reflect provider discrimination against patients of the opposite gender or favoritism towards patients of the same gender. For instance, patient-doctor gender-match effects could arise if doctors disproportionately discount medical conditions of patients of the opposite gender or disproportionately empathize with patients of their own gender. Various forms of discrimination—statistical and taste-based, as well as systemic or structural factors—may affect how doctors evaluate female and male patients. Further, differences in evaluations could arise if patients behave differently when matched with providers of the same gender, perhaps communicating more or different information that may affect their evaluation. We note that it is not possible to conceptually (or empirically) distinguish between these mechanisms in this setting.[26] Nevertheless, it is important to emphasize that doctors are ultimately responsible for outcomes in the medical evaluations we analyze after gathering information from a physical exam, patients' prior medical records, and patients' own accounts of their condition and work limitations. Thus, differences in evaluation outcomes in our setting are driven

[26]For example, patients matched with own-gender doctors may communicate more or different information during an exam if doctors differentially empathize with same-gender patients and ask more questions when evaluating these patients.

by differences in doctor decisions, regardless of whether the reasons for doctors' differential decisions stem from differences in doctor behavior, patient behavior, or a combination of doctor and patient behavior.

More generally, health care decisions typically involve complex interactions between patients and doctors, with health care interactions spanning a spectrum between doctor and patient agency. For example, while patients are ultimately responsible for deciding whether to take up recommended preventive care, doctors are responsible for performing the medical evaluations we examine. The relative agency in most health care interactions lies between these extremes—with the doctor first evaluating the patient's condition, the doctor then recommending treatment based on this evaluation, and the patient deciding whether to follow the doctor's treatment recommendations. Our work provides evidence on the role of demographic concordance between doctors and patients in doctors' evaluations of patients' medical conditions—the critical initial step in most health care interactions (and in disability determinations) and an aspect of health care interactions where doctor agency is central.[27]

In our setting, we find that gender match between doctors and patients increases evaluated disability and subsequent cash benefits for female patients, while gender match does not matter for the evaluation of male patients. This suggests that mechanisms behind the importance of gender match—such as reduced discrimination, increased favoritism, or improved communication—do not operate symmetrically for men and women. Below, we explore potential mechanisms further through additional analysis within the setting of these randomized evaluations. To complement our analysis with the administrative data, Section V presents evidence from a supplemental survey we conducted to further explore plausible mechanisms.

Other Patient and Doctor Characteristics It is possible that the gender-match effect we document arises because of some other patient or doctor characteristic that is correlated with patient or doctor gender rather than mechanisms inherently linked to gender. We consider this possibility by conducting two sets of supplemental analysis. First, we assess whether female and male doctors treat patients differently by observed patient characteristics other than patient gender. Table 4 reports results from estimating Equation (1) replacing the female claimant indicator variable with other claimant characteristics, where the considered characteristics are indicated in the column headings. We consider demographic characteristics including pre-injury weekly wage, age, and marital status; injury characteristics such as whether the injury is associated with a sprain or muscle issue, whether the injury occurred in a dangerous industry, and the claimant's predicted cash benefits based on baseline information; and characteristics of the patient's prior medical and claim experiences such as whether the patient had previously seen a female doctor, whether the patient had requested a designated doctor exam, and the patient's medical spending on the first day of injury.[28] For all characteristics considered in Table 4, the interaction of female doctor and the additional characteristic is statistically indistinguishable from zero. These results suggest that male and female doctors do not systematically differ in their assessments based on claimant characteristics other than claimant gender.

Second, we conduct a parallel analysis looking at whether outcomes for female and male claimants are different by evaluating doctor characteristics other than doctor gender. Table 5 reports results from

[27] In this way, our work complements prior work by Alsan, Garrick, and Graziani (2019), which investigates the role of doctors' race in health care interactions of black men in a setting where patient agency is central—patient decisions to take up effective and broadly recommended preventive care. The mechanisms behind the importance of demographic concordance in health care settings may differ across different types of health care interactions, depending on the relative agency of doctors and patients. While Alsan, Garrick, and Graziani (2019) argue that the most plausible explanation for why black men increase take up of preventive care when paired with a black doctor rather than a white doctor is by improving communication between patients and providers, the plausible mechanisms behind the importance of gender concordance on outcomes in our setting may be different and more complex given the central role of doctor agency in the medical evaluation outcomes we examine.

[28] We classify agriculture, mining, construction, manufacturing, transportation, or warehousing as dangerous industries. We predict cash benefits based on a lasso model. Refer to Appendix A for the full list of regressors included in the lasso model.

estimating Equation (1) replacing the female doctor indicator variable with other doctor characteristics. We consider doctor characteristics including an indicator variable for the designated doctor being an MD or DO rather than a DC, an indicator for the doctor's MD or DO degree being from a top 25 medical school according to U.S. News and World Report, indicators for the doctor's specialty, an indicator for the doctor being born in Texas, and several measures of doctor experience, including years of experience, share of workers' compensation patients seen in the past year who are female, the number of designated doctor exams performed in the past year, and the share of designated doctor exams in the past year that were for females. For all the doctor characteristics considered in Table 5, the interaction of female claimant and the additional characteristic is statistically indistinguishable from zero, suggesting assessments of female and male claimants do not differ based on evaluating doctor characteristics other than doctor gender.[29]

Taken together, the evidence in Tables 4 and 5 suggests that there may be something fundamental or essential about gender—separable from other observed patient and doctor characteristics—that is driving the gender-match effect we document.

Heterogeneity Analyses Heterogeneity analyses may shed light on possible mechanisms. For instance, if gender match is only relevant among particular subgroups of female patients, this may point to potential mechanisms behind the importance of gender match. Figure 2 considers how the effect of female claimants being assigned to female doctors differs by baseline claimant characteristics by showing estimated coefficients on the interaction of female doctor and female claimant from the baseline specification estimated using data from the indicated subgroups defined based on demographic characteristics, injury characteristics, and prior medical experiences. Panel A displays the estimates for an indicator of any additional cash benefits after the exam (or equivalently, being evaluated as disabled at the time of exam), while Panel B displays the estimates for normalized additional cash benefits. Capped horizontal bars indicate the associated 95% confidence intervals.

Examining Figure 2, it is striking that the gender-match effect estimates are positive within all subgroups considered and statistically significant for many subgroups. The estimated gender-match effects are statistically indistinguishable from each other across subgroups for each characteristic. Using a more granular characterization of injury types, Appendix Figure A1 presents additional evidence which illustrates that the gender-match effect estimates are positive and statistically indistinguishable from one another across female claimants with different types of injuries. The similarity of the estimates across subgroups suggests that patient-doctor gender match has near universal impacts on outcomes for female patients and that doctor discretion may play a large role in these evaluations for a broad set of claimants and injuries.

Magnitude Relative to Gender Gap Conditional on Observables Appendix Tables A6 and A7 display results from regressions where we augment our baseline specification by including successively more controls for baseline claimant and injury characteristics. As expected given the conditional random assignment, the estimated coefficients on the female doctor term and the interaction term remain stable with the addition of further controls. Beyond illustrating the robustness of our gender-match estimates, there are two key findings from this analysis. First, the female claimant coefficient remains negative and similar in magnitude when controlling for observable claimant and injury characteristics, indicating that there are significant gen-

[29] In addition to the possibility that fixed characteristics of doctors could relate to evaluations of females relative to males, doctors' experience evaluating patients has the potential to influence evaluations. For example, doctors could potentially become more consistent in evaluating female claimants after having evaluated more female claimants—in the context of designated doctor exams or more generally among workers' compensation patients seen in their standard practice. The small and statistically indistinguishable coefficient estimates for the interaction term in columns 7 and 9 of Table 5 suggest that doctors' recent prior experience evaluating female patients (both within designated doctor exams and among workers' compensation patients more generally) does not seem to be associated with gender differences in evaluation outcomes.

der gaps conditional on observables when male doctors evaluate claimants. The magnitudes indicate that female patients are 2.6 percentage points less likely to receive benefits and receive 10.4% fewer benefits than male patients with similar observables when male doctors evaluate patients. Second, the coefficient on the female claimant term and the interaction term are of nearly the same magnitude and opposite sign regardless of which controls are included, indicating that having female doctors evaluate claimants shrinks (both raw and conditional) gender gaps substantially—almost eliminating these gaps. These results suggest female doctors evaluate female and male patients as similarly disabled when they have similar observable characteristics, while male doctors evaluate female patients as less disabled than male patients with similar observables. Though only suggestive, these results are consistent with male doctors applying a stricter standard when evaluating female patients than when evaluating male patients and female doctors applying similar standards to male and female patients.[30]

IV Broader Evidence: Gender Gap and Gender Homophily

Beyond the data from independent medical exams used for our primary analysis, we have administrative data on the universe of workers' compensation claims in Texas from 2013 to 2017. Next, we use this data on the entire workers' compensation insurance system to provide broader evidence on gender disparities in cash benefit receipt and gender homophily in the selection of doctors when patients have choice.

Gender Gap in Benefit Receipt While all injured workers qualify for workers' compensation medical benefits, cash benefit eligibility is limited to workers with a qualifying disability and this eligibility is determined by a worker's treating doctor (or in the case of disputed claims, the randomly assigned designated doctor). As discussed in Section I, workers are eligible for cash benefits if the doctor determines that the worker has either: (i) a temporary impairment that leaves him/her unable to work for more than seven days, or (ii) a (partial) permanent impairment. Approximately 22% of all workers' compensation claimants qualify for cash benefits, while the remainder have "medical only" claims that only involve reimbursement for injury-related medical expenditures.

We examine how the propensity to receive any cash benefits varies across male and female claimants. Specifically, we estimate the following equation:

$$\text{I(cash benefits}_i > 0) = \beta female\_claimant_i + \Theta \mathbf{X_i} + \epsilon_i, \tag{2}$$

where i denotes claimant, I(cash benefits$_i > 0$) indicates if the claimant receives any cash benefits (temporary income benefits or permanent impairment benefits), and $female\_claimant_i$ indicates the claimant is female. We illustrate how the estimated gender gap varies with the included additional controls, X_i.

Table 6 displays the estimates. Column 1 displays the estimates with no additional controls. Columns 2 through 6 display the estimates with progressively more controls. These additional controls include information about the claim (insurer), about the claimant (county of residence and age), and about the injury including injury timing (month and day-of-the-week), type (diagnoses), and severity (an indicator for first

[30] Because there is no objective measure of the appropriate level of evaluated disability or efficient level of cash benefits conditional on evaluated disability, we cannot determine whether female doctors are too generous or male doctors are too strict in their evaluations of female claimants. Further, note that our analysis does not speak to whether benefits may optimally vary by individual characteristics including gender. Even among workers with the exact same disability and prior earnings, optimal benefit levels may vary across males and females, as the social planner may want to consider other factors in optimal benefit design—such as the magnitude of the consumption drop experienced by workers upon workplace injury or the degree of behavioral responses to benefit generosity. Note, however, that doctors in this setting (and gatekeepers more generally in other disability-related settings) are tasked with solely evaluating disability rather than considering outside factors. Though beyond the scope of this paper, how individual and injury characteristics may influence optimal benefit design is an important area for future work.

medical treatment taking place at an emergency department).

According to the estimates with no controls (from column 1), on average female claimants are 5.5 percentage points less likely to receive benefits than male claimants, which represents a 23.3% reduction relative to the overall mean rate of cash benefit receipt among male claimants (23.6%). Across the range of controls considered, the estimates indicate that females are between 3.3 to 5.5 percentage points less likely to receive cash benefits than males. The estimate from a specification including the full set of claimant and injury controls indicates that female claimants are 3.6 percentage points less likely to receive cash benefits, or 15.3% lower than the mean rate of benefit receipt among male claimants. The main takeaway from this analysis is that female claimants receive cash benefits at a substantially lower rate than male claimants with similar observables. This evidence points to the broader importance of gender disparities in cash benefit receipt among workers' compensation claimants, beyond the subset of claims with randomized evaluations.

Gender Homophily in Choice of Providers More generally, workers' compensation claimants have the ability to select their own "treating doctor," the doctor who typically oversees their medical care and their eligibility for disability benefits. If women understand that they typically fare better when evaluated by a female doctor rather than a male doctor, women may express a preference for female doctors relative to male doctors and select female doctors more frequently when given the choice.

We investigate whether patient selection of providers reflects gender homophily. Given the differing characteristics of female and male doctors (e.g., specialty, availability), we focus on measuring relative gender homophily—do female patients select female treating doctors more often than male patients select female treating doctors? Specifically, we estimate the following equation:

$$\text{I(Chosen Treating Doctor is Female)}_i = \gamma female\_claimant_i + \sigma \mathbf{X_i} + \delta_{m(i)} + \mu_i, \tag{3}$$

where i denotes claimant and $m(i)$ denotes the claimant's medical market. In this equation, I(Chosen Treating Doctor is Female)$_i$ indicates that the chosen treating doctor is female and $female\_claimant_i$ indicates the claimant is female. Most of this analysis focuses on comparing claimants within the same medical market by including patient medical market fixed effects, $\delta_{m(i)}$, where we define medical markets following the Dartmouth Atlas' definition of Hospital Service Areas (HSAs). We also include additional controls in some specifications (X_i), including insurer, injury month-year, injury diagnoses codes, claimant age, and an interaction between injury month-year, diagnoses, and patient medical market. The sample includes all claims for workers in Texas with an identifiable treating doctor.[31]

Table 7 displays the estimates. Without controls for patient medical market, female patients are 2.3 percentage points more likely to select a female treating doctor, which is an 8.4% increase relative to the share of male patients selecting a female treating doctor (27.4%). Once we control for patient medical market, the effect is roughly 78% of the size indicating that some of the female coefficient estimate in the specification without medical market fixed effects reflects a correlation between patient and doctor characteristics across markets: markets with a greater share of female patients (female injured workers) are markets with a greater share of female doctors. Among female and male patients in the same medical market, females are 1.8 percentage points more likely to select a female doctor than male patients, or 6.6%. After controlling for patient medical market, we see that the estimated difference in the rate that females and males select female doctors is very stable as we add more controls. In particular, we obtain a similar estimate—that females are 1.4 percentage points (5.1%) more likely to select a female doctor—in the most saturated specification. We

[31] See Appendix Section A for more detail on the construction of this sample and the identification of treating doctors.

note that this analysis identifies relative gender homophily, and the resulting patterns could be driven by preferences of either men or women to see own-gender providers. In the following section, we present survey evidence that suggests that women may have stronger preferences to see own-gender providers.

V Survey Evidence

Next, we describe the results of a survey we conducted that had two aims. Our first aim is to assess differences in male and female patients' experiences in the broader health care system and to learn about how and why doctors' gender is associated with differential health care experiences for male and female patients. The questions in this section of the survey speak to possible mechanisms behind our main findings. A second aim of the survey is to understand whether patients exhibit preferences for own-gender providers and whether the intensity of these preferences varies across men and women. The questions within this part of the survey can help us understand the generalizability of the gender homophily patterns documented in the prior section and assess potential explanations for these patterns.

We surveyed 1,519 adults between 30 to 64 years of age in May 2021.[32] The survey was created and conducted using the Qualtrics platform, which aims to include diverse set of respondents across age, gender, ethnicity, household income, and census region. To focus on a sample that overlaps with individuals at risk of a workplace injury, we restricted the pool of eligible respondents to people who have worked in the last 12 months, and we targeted an age distribution of respondents to roughly match the age distribution of workers in the American Community Survey (ACS) (Ruggles et al. 2022).[33] We limited the survey to individuals who identify as either male or female, and we had roughly equal numbers of respondents in each group, with 755 male respondents and 764 female respondents. Appendix Table A8 summarizes this basic information about respondents by gender and displays the analogous summary statistics by gender for workers more generally in the US and for workers' compensation claimants in Texas when overlapping information is available. Because there are some small differences in observable characteristics across male and female respondents, we report survey responses in two ways—raw means by gender and regression-adjusted differences by gender which account for variation in observable characteristics.[34]

Experiences with the Health Care System The first set of questions asks respondents whether they have ever experienced various negative interactions with a health care provider. Table 8 reports the results. The main takeaway from this series of questions is that women are much more likely than men to report a negative experience with a health care provider. For example, 39.3% of women and 34.0% of men report having an experience where a doctor talked down to them or didn't treat them with dignity or respect (a greater than 5 percentage point difference). Women are 9.4 percentage points more likely to report encountering a doctor who didn't understand or relate to their experiences and concerns and 4.2 percentage points more likely to report a doctor made them feel uncomfortable discussing their concerns. Slightly more women than men reported a doctor didn't believe they were telling the truth about their symptoms and reported a doctor assumed something about them without asking. Generally, these documented gender differences are, if anything, even larger when accounting for differences in observable characteristics between male

[32]See Cabral and Dillender (2023) for the data.

[33]When interpreting the survey evidence, an important caveat to bear in mind is that the survey sample is not the same as the population reflected in the workers' compensation insurance claims data. The workers' compensation administrative claims data do not include identifying information for claimants, so we could not have surveyed the exact population represented in the claims data. While the survey sample is designed to be representative of the broader population of non-elderly workers, the workers' compensation claims data cover only workers injured at work. Nevertheless, the survey provides suggestive evidence about potential mechanisms behind the main findings regarding the impact of gender concordance on medical evaluations and about the generalizability of—and causes behind—the observed gender homophily patterns in patient-doctor matches.

[34]Appendix Section C reports the full text of the survey questions we analyze.

and female respondents. While most of these negative experiences were more prevalent among women than men, there were two exceptions: women and men were equally likely to report that a doctor refused a test or treatment they thought they needed and men were more likely to report that a doctor had refused to prescribe pain medication they thought they needed. We note this latter difference in responses could potentially stem from underlying differences in the likelihood of needing (or wanting) pain medication. As a summary measure, we also report the share of respondents who experienced at least one of these negative experiences. Women are 5.8 percentage points more likely to report they have experienced any of the negative experiences asked about in the survey.

The next set of survey questions asks respondents about how doctors' gender influences their expectations about the likelihood of having a positive health care interaction along various quality measures based on their past experiences getting health care. Respondents can answer that they would expect a better experience with a male doctor or a female doctor or that doctors' gender is not informative about the likelihood of having a positive interaction.[35] Figure 3 Panel A reports the share of respondents saying a doctor of their own gender would be more likely to have the positive trait, with the analogous estimates reported in Appendix Table A9.[36] Relative to male respondents, a higher share of female respondents believe that gender concordance with their doctor would be more likely to result in a positive health care interaction than seeing a doctor of the opposite gender. Compared to male respondents, women are more likely to answer that a doctor of their own gender is more likely to: treat them with dignity and respect (34.3% vs. 19.3%), understand or relate to their experiences and concerns (51.7% vs. 29.5%), believe they are telling the truth about their symptoms (30.4% vs. 17.5%), provide needed testing and treatments (24.9% vs. 17.6%), make them feel comfortable discussing their concerns (41.6% vs. 24.0%), and ask appropriate questions instead of making assumptions (32.6% vs. 17.7%). These differences across male and female respondents are large in magnitude, statistically distinguishable from zero, and, if anything, larger when controlling for respondent characteristics. The final two positive attributes asked about pertain to doctors' proximity and qualifications. No meaningful differences between men and women are observed in the share that respond that an own-gender doctor is more likely to be available nearby or to be the most qualified.

Overall, the survey evidence suggests that women—relative to men—more often have their concerns and symptoms discounted by doctors, and women more often report that positive interactions are more likely with an own-gender doctor. The asymmetry of responses between men and women supports the plausibility of the main finding: that doctor gender is an important factor in the evaluation of female patients more so than in the evaluation of male patients.

Preferences over Doctors and Homophily The next set of questions collects information about the propensity of respondents to select male and female doctors and about preferences regarding doctor gender, with the results summarized in Table 8. First, respondents were asked whether they had ever received care from a male doctor and from a female doctor. There are notable differences by respondent gender, with 90.1% of women and 96.8% of men reporting they have received care from a male doctor while 94.1% of women and 76.7% of men report they have received care from a female doctor.

The next several questions ask about preferences related to doctor gender. The survey asks whether respondents prefer to see a doctor who is female, prefer to see a doctor who is male, or it does not make much difference. Women were more likely to express a preference for an own-gender doctor, with 46.3% of

[35] To ensure that the ordering of options did not bias the responses, we randomized at the respondent level which choice—a male doctor or a female doctor—was listed first in these questions.

[36] Appendix Figure A2 illustrates that patterns are very similar if we instead focus on the share of respondents who indicated a doctor of their own gender is at least as likely to have the positive trait.

women preferring a female doctor while only 29.0% of men prefer a male doctor. The results also suggest that women are much less likely to report a preference for an opposite-gender doctor, with only 5.8% of women preferring a male doctor while 13.6% of men prefer to see a female doctor. The share of respondents expressing that doctor gender does not make much difference was smaller among women than men—47.9% among female respondents and 57.4% among male respondents.

To better understand how the intensity of preferences over doctor gender varies across men and women, we next ask respondents to make a hypothetical choice over health care options. Specifically, respondents were told to suppose they needed to see a doctor for medical evaluation and treatment. They were then presented with two doctors who were identical in all other ways except for the listed characteristics, where the listed characteristics described the doctor's gender and the out-of-pocket cost for the visit. The respondents were presented with a randomly assigned choice set among the following four choice sets: [37]

- Female doctor, $30 out-of-pocket cost or Male doctor, $5 out-of-pocket cost
- Female doctor, $10 out-of-pocket cost or Male doctor, $5 out-of-pocket cost
- Female doctor, $5 out-of-pocket cost or Male doctor, $10 out-of-pocket cost
- Female doctor, $5 out-of-pocket cost or Male doctor, $30 out-of-pocket cost.

Figure 3 Panel B plots the share of respondents selecting an own-gender doctor by respondent gender, while Appendix Table A10 displays the underlying estimates. At every co-pay differential, we see that women are at least as likely as men—and typically strictly more likely than men—to select an own-gender doctor. Based on the responses, 48.5% of women are willing to pay an additional $5 copay to see an own-gender provider compared to only 29.3% of men—a 19.2 percentage point difference. When it is cheaper to see an own-gender provider, nearly all women select to see an own-gender provider: 92.9% at an out-of-pocket cost differential of -$5 and 98.8% at an out-of-pocket cost differential of -$25. In contrast, a lower share of men would select to see an own-gender provider when a visit costs less: 83.3% at an out-of-pocket cost differential of -$5 and 85.4% at an out-of-pocket cost differential of -$25. These gender gaps are statistically distinguishable from zero and robust to controlling for respondent characteristics.

Finally, the survey asks respondents to rate the importance of various doctor attributes—including a doctor's gender—when they select a health care provider.[38] As reported in Table 8, doctor gender was rated as at least moderately important by 41.4% of women compared to 33.6% of men—a more than 7 percentage point gap (p-value 0.002). This difference stands in contrast to the similarity of men's and women's responses for many other doctor attributes, such as out-of-pocket cost for a visit, travel time to get to doctor's office, wait time at the doctor's office, and doctor's age.[39]

Overall, this survey evidence helps us interpret the homophily evidence documented in the prior section. The survey responses indicate gender homophily in patient-doctor matches, pointing to the broader relevance of the homophily evidence presented in the prior section. Further, the survey evidence suggests a stark asymmetry in preferences to interact with own-gender providers—women are more likely to express a preference to see own-gender providers, women more often select own-gender providers when given the

[37] Beyond randomizing which choice set was presented, we also randomized which option was presented first within each choice set to ensure primacy bias did not affect the results.

[38] The answer options for this question were: not at all important, slightly important, moderately important, very important, and extremely important.

[39] Aside from doctor's gender, we see gender differences in the stated importance of doctor reviews, with 84.7% of women indicating doctor reviews are at least moderately important compared to 80.7% of men, a 4.0 percentage point difference (p-value 0.038). Women placing more weight on doctor reviews could be in response to their greater propensity to have negative experiences with health care providers in the past.

choice, and women express greater intensity of preferences to see own-gender providers. These gender differences in preferences to see own-gender providers may reflect gender differences in patient expectations of having positive interactions with own-gender versus opposite-gender providers.

VI Discussion

We turn to the potential policy implications of our findings. Our findings indicate that policies that increase the likelihood that female doctors evaluate female patients may increase evaluated disability and benefits for women and shrink documented gender gaps conditional on observables in this setting. We consider the impacts of two types of policies that would increase the prevalence of female doctors evaluating female patients: (i) increasing the share of doctors who are female and (ii) increasing gender homophily in patient-doctor matches.

First, consider both types of policies within our empirical setting—disputed claims with independent medical evaluations in the Texas workers' compensation system. Our estimates imply that increasing the share of independent medical exams performed by female doctors from 17% to 50% would cause a 0.88 percentage point increase in female patients evaluated as disabled, closing 41.3% of the overall gender gap conditional on observables among disputed claims. An alternative—and more direct—way to increase the share of female patients evaluated by female doctors would be to change the doctor assignment process. For instance, patients with disputed claims could be randomly assigned to a doctor of the same gender, rather than the current gender-blind random assignment process in which only 17% of female claimants are assigned female designated doctors. Such a policy would increase the share of female patients who receive cash benefits by 2.2 percentage points, closing the observed gender gap in benefit receipt.

More generally, we may be interested in the broader policy implications for gender gaps in workers' compensation insurance, beyond the subset of disputed claims with randomly assigned evaluators. We conduct back-of-the-envelope counterfactual analysis evaluating the potential impact of two types of policies on the gender gap: increasing the share of female doctors and increasing gender homophily patient-doctor matching.[40] Appendix Section D presents the details of this analysis. Holding all else fixed, this analysis suggests that increasing the share of female treating doctors by 22.1 percentage points—moving from 27.9% to parity—would lead to a 33.7% decrease in the observed gender gap. Alternatively, if we hold fixed the overall gender composition of treating doctors, the analysis suggests that an increase in relative gender homophily from the observed level to a level where female patients select female doctors at twice the rate of male patients would lead to a 17.3% decrease in the gender gap conditional on observables. More broadly, we use our estimates to characterize the mix of gender diversity among doctors and relative gender homophily in patient-doctor matches that would lead to a given decrease in the gender gap conditional on observables. Overall, this analysis reveals that these types of policies may have substantial impacts and complement one another in closing gender gaps in evaluated disability.

VII Conclusion

This paper investigates how male and female doctors differ in their medical evaluations of male and female patients. Leveraging random assignment of doctors to claimants for medical evaluations in the Texas

[40] There are several ways policy may influence the share of treating doctors who are female. For instance, some policies may encourage women to enter relevant medical specialties (e.g., residency hours caps (Wasserman 2019)) or encourage female doctors to accept workers' compensation insurance patients (e.g., through outreach efforts or payment policy). One can also imagine policies that increase gender homophily. For instance, gender homophily may increase if there were a reduction in barriers to sorting (e.g., providing patients a directory of workers' compensation treating doctors) or direct provision of information to claimants about the expected gains from sorting to own-gender doctors (e.g., informing female claimants of the results of this study).

workers' compensation insurance program, we show that being evaluated by a female doctor rather than a male doctor increases female claimants' likelihood of being evaluated as disabled and their subsequent cash disability benefits. In contrast, we show the gender of the evaluating doctor does not impact evaluated disability or subsequent cash benefits for male claimants. These patterns hold for various subgroups and do not seem to be explained by other characteristics of claimants or the evaluating doctors. Our findings indicate that gender of evaluating doctors is an important determinant of gender gaps in this setting, where the estimates imply that having female doctors evaluate patients—rather than male doctors—eliminates the observed gender gap in the likelihood of being evaluated as disabled conditional on observables.

To place these estimates within the broader context, we use data on all workers' compensation claims to assess male-female benefit differences more broadly and to examine patient-doctor sorting when patients have the ability to select their treating doctors. We find that female claimants are 15.3% less likely to receive cash disability benefits than male claimants are after controlling for a rich set of baseline claimant and injury characteristics. Relative to comparable male claimants, female claimants are about 5% more likely to choose a female treating doctor when they are able to choose their own doctor, which is consistent with female patients having a preference for female doctors on average relative to male patients' preference for female doctors. Results from a supplemental survey we conduct suggest that women—relative to men—more often have negative health care experiences, expect better treatment from own-gender doctors, and have stronger preferences to see own-gender doctors.

One implication of our findings is that increasing the share of doctors who are female may increase assessed disability and cash benefits for women, while having little impact on outcomes for men. Extrapolating from our estimates, we find that increasing the share of independent medical evaluations performed by female doctors from 17% to 50% would increase the share of female patients evaluated as disabled, closing approximately 41% of the overall gender gap conditional on observables among claimants with independent medical evaluations. To analyze the broader implications of our findings, we conduct back-of-the-envelope counterfactuals that reveal that policies aimed at either increasing the share of doctors who are female or increasing gender homophily in patient-doctor matches can substantially shrink gender gaps in evaluated disability and work to complement one another in doing so.

More broadly, our findings highlight the importance of gatekeeper discretion in public programs and inequities that can arise from this discretion. Our findings illustrate that gender match between program gatekeepers and claimants can have important implications for the distribution of program benefits. While our discussion of policy implications focuses on policies that shift the composition of evaluators, policymakers and future research may also want to consider broader policies aimed at reducing the impact of discretion on evaluations—such as requiring multiple assessments for high-stakes evaluations, mandating evaluators receive de-biasing training, or encouraging the use of claimant advocates.

Our results also inform broader literatures that study the differential treatment of male and female patients within the health care system and the impact of the gender of authority figures and evaluators on females more broadly. Despite a growing body of evidence documenting that male and female patients receive different treatment in the health care system for reasons that appear unrelated to underlying health needs, determining if structural factors of the health care system play a role in these differences is difficult. Our study suggests that differential evaluations of female patients by male and female doctors likely play a role in differential outcomes for male and female patients. A natural corollary of these findings is that the under-representation of female doctors may be an important contributing factor behind observed gender disparities in health care settings. More broadly, our findings highlight the potential importance of gender match in settings where authority figures evaluate women.

References

Abrevaya, Jason and Daniel S. Hamermesh. 2012. "Charity and Favoritism in the Field: Are Female Economists Nicer (to Each Other)?" *The Review of Economics and Statistics*, 94(1): 202–207.

Alsan, Marcella, Owen Garrick, and Grant Graziani. 2019. "Does Diversity Matter for Health? Experimental Evidence from Oakland." *American Economic Review*, 109(12): 4071–4111.

Bertrand, Marianne, Sandra E Black, Sissel Jensen, and Adriana Lleras-Muney. 2018. "Breaking the Glass Ceiling? The Effect of Board Quotas on Female Labour Market Outcomes in Norway." *The Review of Economic Studies*, 86(1): 191–239.

Biasi, Barbara and Heather Sarsons. 2022. "Flexible Wages, Bargaining, and the Gender Gap." *The Quarterly Journal of Economics*, 137(1): 215–266.

Bohren, J. Aislinn, Peter Hull, and Alex Imas. 2022. "Systemic Discrimination: Theory and Measurement." National Bureau of Economic Research Working Paper 29820.

Cabral, Marika and Marcus Dillender. 2021. "Disparities in Health Care and Medical Evaluations by Gender: A Review of Evidence and Mechanisms." *AEA Papers and Proceedings*, 111: 159–63.

Cabral, Marika and Marcus Dillender. 2023. "Replication Data for Gender Differences in Medical Evaluations: Evidence from Randomly Assigned Doctors." American Economic Association [publisher], Inter-university Consortium for Political and Social Research [distributor]. openicpsr-188361.

Card, David, Stefano DellaVigna, Patricia Funk, and Nagore Iriberri. 2019. "Are Referees and Editors in Economics Gender Neutral?" *The Quarterly Journal of Economics*, 135(1): 269–327.

Carrell, Scott E., Marianne E. Page, and James E. West. 2010. "Sex and Science: How Professor Gender Perpetuates the Gender Gap." *The Quarterly Journal of Economics*, 125(3): 1101–1144.

Casarico, Alessandra and Salvatore Lattanzio. 2019. "What Firms Do: Gender Inequality in Linked Employer-Employee Data." Cambridge-INET Working Paper Series Working Paper 2019/15.

Centers for Medicare & Medicaid Services (CMS). 2019*a*. "ICD-9-CM to and from ICD-10-CM and ICD-10-PCS Crosswalk or General Equivalence Mapping." Version available through the National Bureau for Economic Research. https://www.nber.org/research/data/icd-9-cm-and-icd-10-cm-and-icd-10-pcs-crosswalk-or-general-equivalence-mappings (Date accessed: August 2019).

Centers for Medicare & Medicaid Services (CMS). 2019*b*. "National Plan and Provider Enumeration System (NPPES) Downloadable File." https://www.cms.gov/Regulations-and-Guidance/Administrative-Simplification/NationalProvIdentStand/DataDissemination (Date accessed: July 2019).

Centers for Medicare & Medicaid Services (CMS). 2021. "Doctors and Clinicians National Downloadable File: Performance Year 2019." https://data.cms.gov/provider-data/dataset/mj5m-pzi6 (Date accessed: August 2021).

Chandra, Amitabh and Douglas O Staiger. 2010. "Identifying Provider Prejudice in Healthcare." National Bureau of Economic Research Working Paper 16382.

Chen, Esther H, Frances S Shofer, Anthony J Dean, Judd E Hollander, William G Baxt, Jennifer L Robey, Keara L Sease, and Angela M Mills. 2008. "Gender Disparity in Analgesic Treatment of Emergency Department Patients with Acute Abdominal Pain." *Academic Emergency Medicine*, 15(5): 414–418.

Cullen, Zoë B. and Ricardo Perez-Truglia. 2019. "The Old Boys' Club: Schmoozing and the Gender Gap." National Bureau of Economic Research Working Paper 26530.

Daly, Mary C, Brian Lucking, Jonathan A Schwabish, et al. 2013. "The Future of Social Security Disability Insurance." *FRBSF Economic Letter*, 17.

Dartmouth Atlas. 2013. "ZIP Code to Health Service Area (HSA) Crosswalk (2013)." https://data.dartmouthatlas.org/supplemental/#crosswalks (Date accessed: August 2020).

Department of Consumer and Business Services. 2004. "Workers' Compensation Insurer Medical Examination Study." https://www.oregon.gov/dcbs/reports/Documents/archive/arc-medical/ime/ime-study-rpt.pdf (Date Accessed: February 2022).

Department of Veterans Affairs. 2010. "Review of Combat Stress in Women Veterans Receiving VA Health Care and Disability Benefits." Office of Inspector General.

Dupas, Pascaline and Radhika Jain. 2021. "Women Left Behind: Gender Disparities in Utilization of Government Health Insurance in India." National Bureau of Economic Research Working Paper 28972.

Greenwood, Brad N., Seth Carnahan, and Laura Huang. 2018. "Patient–Physician Gender Concordance and Increased Mortality among Female Heart Attack Patients." *Proceedings of the National Academy of Sciences - PNAS*, 115(34): 8569–8574.

Hill, Andrew, Daniel Jones, and Lindsey Woodworth. 2020. "Physician-Patient Race-Match Reduces Patient Mortality." SSRN Scholarly Paper 3211276, Rochester, NY.

Hoffmann, Diane E. and Anita J. Tarzian. 2001. "The Girl Who Cried Pain: A Bias against Women in the Treatment of Pain." *The Journal of Law, Medicine & Ethics*, 28(4_suppl): 13–27.

Jagsi, Reshma, Kent A Griffith, Rochelle A DeCastro, and Peter Ubel. 2014. "Sex, Role Models, and Specialty Choices among Graduates of US Medical Schools in 2006–2008." *Journal of the American College of Surgeons*, 218(3): 345–352.

Kaiser Family Foundation. 2019. "Professionally Active Physicians by Gender." https://www.kff.org/other/state-indicator/physicians-by-gender (Date accessed: Februrary 2019).

Kline, Patrick, Evan K Rose, and Christopher R Walters. 2022. "Systemic Discrimination Among Large U.S. Employers." *The Quarterly Journal of Economics*, 137(4): 1963–2036.

Lim, Jaegeum and Jonathan Meer. 2019. "Persistent Effects of Teacher–Student Gender Matches." *The Journal of Human Resources*, 55(3): 809–835.

Low, Hamish and Luigi Pistaferri. 2019. "Disability Insurance: Error Rates and Gender Differences." National Bureau of Economic Research Working Paper 26513.

Maida, Agata and Andrea Weber. 2019. "Female Leadership and Gender Gap within Firms: Evidence from an Italian Board Reform." *ILR Review*, 75(2): 488–515.

McDevitt, Ryan C. and James W. Roberts. 2014. "Market Structure and Gender Disparity in Health Care: Preferences, Competition, and Quality of Care." *The RAND Journal of Economics*, 45(1): 116–139.

Mulcahy, Andrew, Barbara Wynn, Kandice Kapinos, Preethi Rao, Rosalie Malsberger, Brian Phillips, and Spencer Case. 2020. "Evaluation of California Senate Bill 863 Medical Care Reforms." https://www.rand.org/content/dam/rand/pubs/research_reports/RR2000/RR2045/RAND_RR2045.pdf (Date Accessed: February 2022).

Muralidharan, Karthik and Ketki Sheth. 2016. "Bridging Education Gender Gaps in Developing Countries: The Role of Female Teachers." *The Journal of Human Resources*, 51(2): 269–297.

Murdoch, Maureen, Michele Roxanne Spoont, Nina Aileen Sayer, Shannon Marie Kehle-Forbes, and Siamak Noorbaloochi. 2021. "Reversals in Initially Denied Department of Veterans Affairs' PTSD Disability Claims after 17 Years: A Cohort Study of Gender Differences." *BMC Women's Health*, 21(70): 1–8.

National Bureau of Economic Research (NBER). 2021. "National Provider Identifier (NPI) to State License Crosswalk." https://www.nber.org/research/data/national-provider-identifier-npi-state-license-crosswalk (Date accessed: August 2021).

Ruggles, Steven, Sarah Flood, Ronald Goeken, Megan Schouweiler, and Matthew Sobek. 2022. "IPUMS USA: Version 12.0 [Dataset]. Minneapolis, MN: IPUMS." https://doi.org/10.18128/D010.V12.

Samulowitz, Anke, Ida Gremyr, Erik Eriksson, and Gunnel Hensing. 2018. ""Brave Men" and "Emotional Women": A Theory-Guided Literature Review on Gender Bias in Health Care and Gendered Norms towards Patients with Chronic Pain." *Pain research & management*, 2018: 1–14.

Sarsons, Heather. 2017. "Interpreting Signals in the Labor Market: Evidence from Medical Referrals." https://scholar.harvard.edu/files/sarsons/files/sarsons_jmp_01.pdf. Accessed on August 2021.

Schnell, Molly and Janet Currie. 2018. "Addressing the Opioid Epidemic: Is There a Role for Physician Education?" *American Journal of Health Economics*, 4(3): 383–410. Appendix B, Table 1.

Schulman, Kevin A, Jesse A Berlin, William Harless, Jon F Kerner, Shyrl Sistrunk, Bernard J Gersh, Ross Dube, Christopher K Taleghani, Jennifer E Burke, Sankey Williams, et al. 1999. "The Effect of Race and Sex on Physicians' Recommendations for Cardiac Catheterization." *New England Journal of Medicine*, 340(8): 618–626.

Texas Board of Chiropractic Examiners. 2020. "Information on Registered Texas DCs." Obtained through open records request to Texas Board of Chiropractic Examiners.

Texas Board of Chiropractic Examiners. 2021. "Texas Board of Chiropractic Examiners Doctors of Chiropractic License Data." https://www.tbce.state.tx.us/verify.html (Data accessed: August 2021).

Texas Department of Insurance (TDI). 2018*a*. "DWC Form-032 Data (Injury Years 2013-2017)." Obtained through open records request to Texas Department of Insurance.

Texas Department of Insurance (TDI). 2018*b*. "DWC Form-069 Data (Exam Years 2012-2017)." Obtained through open records request to Texas Department of Insurance.

Texas Department of Insurance (TDI). 2018*c*. "Workers' Compensation Benefit Data (Injury Years 2013-2017)." Obtained through open records request to Texas Department of Insurance.

Texas Department of Insurance (TDI). 2018*d*. "Workers' Compensation Incident Data (Injury Years 2013-2017)." Obtained through open records request to Texas Department of Insurance.

Texas Department of Insurance (TDI). 2018*e*. "Workers' Compensation Medical Bill Data, SV1, SV2, and SV4 Header and Detail Files (Service Years 2012-2018)." Obtained through open records request to Texas Department of Insurance.

Texas Medical Board. 2014. "Information on Registered Texas MDs and DOs." Obtained through open records request to Texas Medical Board.

Trabanino, Shawn. 2020. "Health, Law, and Ethnicity: The Disability Administrative Law Judge and Health Disparities for Disadvantaged Populations Note." *California Law Review*, 108(6): 2079–2103.

Tunks, Eldon Ed, Anthony Ed Bellissimo, and Ranjan Ed Roy. 1990. *Chronic Pain: Psychosocial Factors in Rehabilitation.* Robert E Krieger Publishing Co.

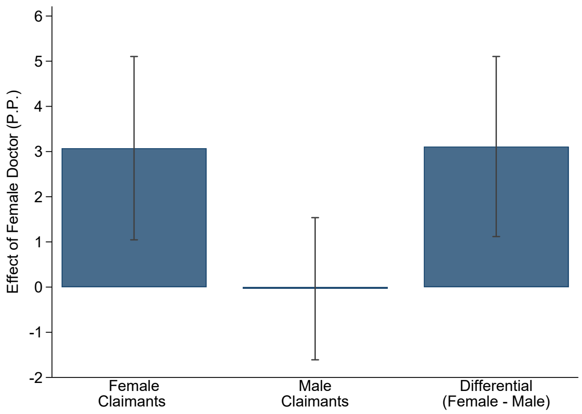
United States Census Bureau. 2010. "ZIP Code Tabulation Area (ZCTA) Relationship File (2010)." https://www.census.gov/geographies/reference-files/time-series/geo/relationship-files.html. (Date accessed: October 2018).

U.S. Bureau of Labor Statistics. 2021. "Consumer Price Index for Urban Wage Earners and Clerical Workers [CWUR0000SA0]." Version available through Social Security Administrative website. https://www.ssa.gov/oact/STATS/cpiw.html. (Date accessed: March 2021).

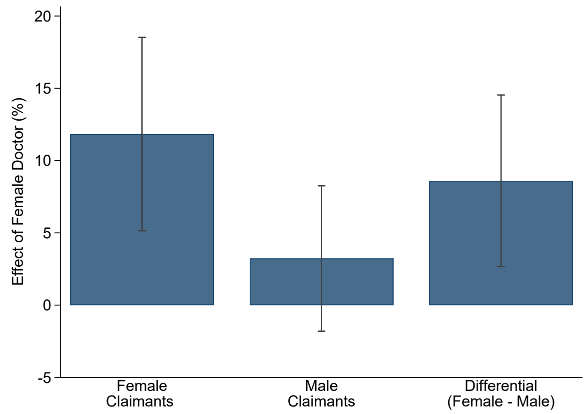
Wasserman, Melanie. 2019. "Hours Constraints, Occupational Choice, and Gender: Evidence from Medical Residents." SSRN Scholarly Paper 3371100, Rochester, NY.

Zeltzer, Dan. 2020. "Gender Homophily in Referral Networks: Consequences for the Medicare Physician Earnings Gap." *American Economic Journal: Applied Economics*, 12(2): 169–97.

Figure 1: Effect of Female Doctor on Benefit Receipt



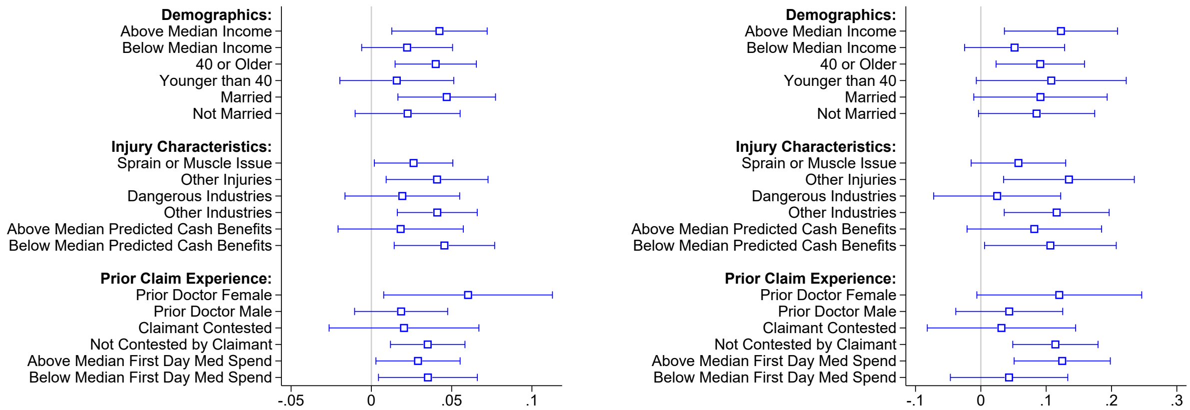
(a) OLS: I(Additional Cash Benefits>0)



(b) Poisson Regression: Additional Cash Benefits

Notes: Each graph displays estimates from a separate regression of Equation (1) that controls for a female doctor indicator variable, a female claimant indicator variable, the interaction of the female doctor and female claimant indicator variables, credential-by-county fixed effects, exam year fixed effects, and injury year fixed effects. The coefficient on the female doctor indicator variable is the effect of a female doctor on male patients. The sum of the coefficients on the female doctor indicator variable and the interaction of the female doctor and female claimant indicator variables is the effect of a female doctor on female patients. The coefficient on the interaction of the female doctor and female claimant indicator variables is the differential effect of a female doctor for female patients. The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017. The 95% confidence intervals displayed along with the coefficient estimates are calculated using standard errors clustered at the doctor level.

Figure 2: Heterogeneity: Estimate on Interaction of Female Doctor and Female Claimant by Subgroup

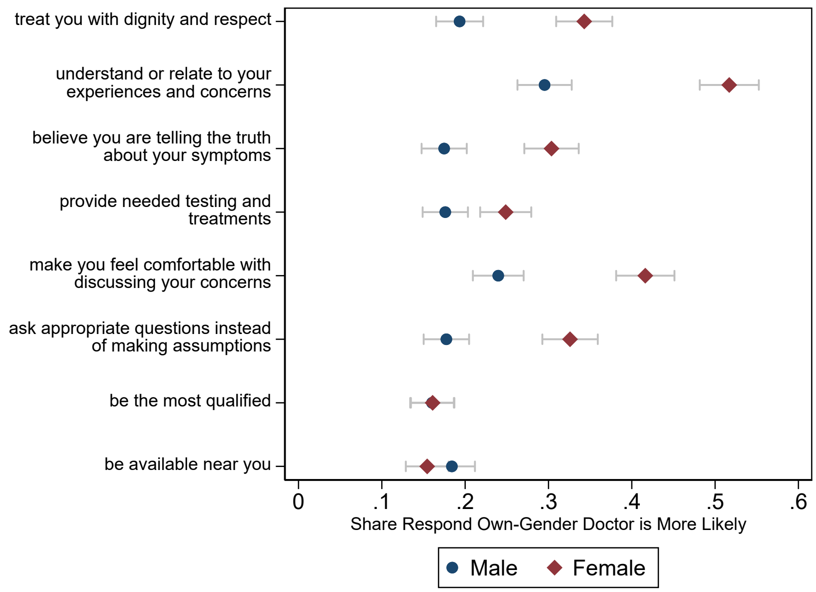


(a) OLS: I(Additional Cash Benefits>0)

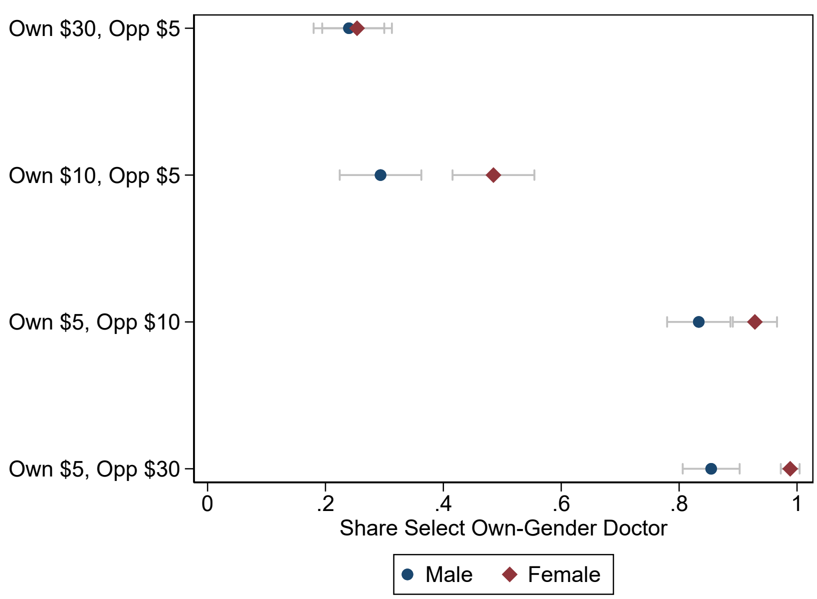
(b) Poisson Regression: Additional Cash Benefits

Notes: Each marker is the coefficient on the interaction of the female doctor and female claimant indicator variables from separate regressions of Equation (1) for the specified sample of claimants. All regressions control for a female doctor indicator variable, a female claimant indicator variable, credential-by-county fixed effects, exam year fixed effects, and injury year fixed effects. The dependent variables are as indicated in the figure: an indicator for receiving any additional cash benefits (Panel A) and (normalized) additional cash benefits received (Panel B). As described in Section I, the presence of wage and industry information is related to the receipt of cash benefits, and thus these variables are only available for a subset of claimants. The 95% confidence intervals displayed along with the coefficient estimates are calculated using standard errors clustered at the doctor level. The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017 and that have non-missing values for the specified characteristic.

Figure 3: Survey Results: Expectations about Health Care Experiences and Hypothetical Choice Questions



(a) "Thinking about your experiences getting health care for yourself, which doctor—male or female—would be more likely to..." (answer options: male doctor, female doctor, male and female doctors are equally likely)



(b) Share Selecting an Own-Gender Doctor in Initial Hypothetical Choice Question by Co-Pay Differential

Notes: The figure above shows the means and associated 95% confidence intervals for survey responses to the indicated question by respondent gender. The survey was conducted by Qualtrics and included 1,519 adults ages 30 to 64. Eligibility was restricted to individuals who reported working at some point in the last 12 months and individuals who self-identified as either male (755 respondents) or female (764 respondents). See Section V for more detail on the survey. Appendix Tables A9 and A10 report raw and regression-adjusted differences in means for these survey questions and the associated p-values.

Table 1: Summary Statistics

Panel A. Workers' Compensation Claimants	All Claimants		Baseline Sample Claimants with Randomized Evaluations	
	Male	Female	Male	Female
	(1)	(2)	(3)	(4)
Age	40.0	43.0	44.4	47.0
Claim Initiated with ED Visit	28%	23%	37%	31%
First-Day Medical Spending	639	442	1,301	806
Three-Month Medical Spending	2,011	1,496	6,266	4,974
Receives Income-Replacement Benefits within 1 Year of Injury	22%	16%	82%	78%
Receives Impairment Benefits within 1 Year of Injury	8%	6%	47%	45%
Injury Type:				
Contusion	12%	20%	8%	12%
Fracture	7%	5%	17%	12%
Muscle Issue	18%	19%	28%	29%
Sprain	27%	31%	34%	38%
Other	36%	26%	13%	8%
Has Female Designated Doctor	-	-	17%	18%
Receives Additional Cash Benefits within 1 Year of Exam	-	-	63%	60%
Amount of Additional Cash Benefits within 1 Year of Exam	-	-	6,563	5,726
Percent of Sample	62%	38%	66%	34%

Panel B. Doctors	Designated Doctors	Texas Doctors Treating Injured Workers	All Texas Doctors	All U.S. Doctors
	(1)	(2)	(3)	(4)
MDs/DOs	57%	85%	91%	92%
Female	20%	22%	34%	34%
Specialty:				
Internal or Family Medicine	40%	48%	37%	38%
Orthopedics	17%	22%	3%	3%
Top 25 Medical School	15%	17%	19%	19%
DCs	43%	15%	9%	8%
Female	19%	19%	28%	28%
N	1,298	5,441	64,554	1,022,068

Notes: Panel A displays means for all 1,076,759 claims occurring from 2013 to 2017 and for the 70,748 claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017. Panel B compares characteristics of doctors in the sample to characteristics of broader populations of doctors. Column 1 contains information on doctors who performed designated doctor exams from 2013 to 2017 in doctor exam data from the TDI. Column 2 contains information on doctors who treated injured workers from 2013 to 2017 in the Texas workers' compensation insurance medical data. Columns 3 and 4 contain information on all doctors in the National Plan and Provider Enumeration System (NPPES) data through 2018. Doctors are identified using a unique identifier from the TDI in column 1 and using NPI in other columns. Information on age and specialty comes from the NPPES data. Information on medical schools attended comes from the Medicare Physician Compare File. Information on medical school rankings comes from schools' average research ranking in the U.S. News and World Report from 2010 to 2017 as reported in Schnell and Currie (2018). If an observation is missing information for a characteristic, that observation is excluded from the percent calculation so total percents sum to 100.

Table 2: Balance

	Female Doctor X Female Claimant (1)	Female Doctor (2)	Female Claimant (3)
Age	-0.173	0.110	2.670
	(0.272)	(0.148)	(0.109)
	[0.525]	[0.459]	[<0.001]
ED Claim	-0.007	0.004	-0.038
	(0.010)	(0.007)	(0.004)
	[0.517]	[0.497]	[<0.001]
Log(First-Day Medical Spending)	0.031	-0.022	-0.178
	(0.028)	(0.018)	(0.012)
	[0.264]	[0.220]	[<0.001]
Log(Med Spending Prior to Exam)	-0.001	-0.003	-0.157
	(0.021)	(0.014)	(0.010)
	[0.948]	[0.821]	[<0.001]
Receives Income-Replacement Benefits Prior to Exam	0.011	0.001	-0.040
	(0.008)	(0.005)	(0.004)
	[0.164]	[0.777]	[<0.001]
Log(Weeks from Injury to Exam)	-0.013	0.004	0.002
	(0.009)	(0.006)	(0.004)
	[0.143]	[0.458]	[0.617]
Injury Type:			
Contusion	0.001	0.005	0.044
	(0.007)	(0.004)	(0.003)
	[0.821]	[0.210]	[<0.001]
Fracture	0.012	-0.010	-0.038
	(0.009)	(0.005)	(0.004)
	[0.147]	[0.051]	[<0.001]
Muscle Issue	0.007	-0.007	0.006
	(0.010)	(0.006)	(0.004)
	[0.529]	[0.249]	[0.183]
Sprain	-0.011	0.007	0.024
	(0.010)	(0.007)	(0.004)
	[0.280]	[0.277]	[<0.001]
Other	-0.009	0.005	-0.035
	(0.007)	(0.005)	(0.003)
	[0.170]	[0.303]	[<0.001]

Notes: This table displays estimates of the coefficients on the female doctor indicator variable, the female claimant indicator variable, and the interaction of the female doctor and female claimant indicator variables from OLS regressions of Equation (1) that control for credential-by-county fixed effects, exam year fixed effects, and injury year fixed effects. Each row represents a separate regression with the dependent variable as indicated in the table. Columns 1, 4, and 7 display the coefficient estimates, columns 2, 5, and 8 display standard errors clustered at the doctor level, and columns 3, 6, and 9 display p-values. In each specification, the sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017 and that have non-missing values for the given dependent variable.

Table 3: Effect on Benefit Receipt

	OLS: I(Additional Cash Benefits > 0)			Poisson Regression: Additional Cash Benefits		
	(1)	(2)	(3)	(4)	(5)	(6)
Female Doctor X Female Claimant	0.031	0.030		0.086	0.080	
	(0.010)	(0.010)		(0.030)	(0.031)	
	[0.002]	[0.004]		[0.004]	[0.010]	
Female Doctor	-0.000		0.030	0.032		0.117
	(0.008)		(0.011)	(0.026)		(0.034)
	[0.964]		[0.004]	[0.208]		[0.001]
Female Claimant	-0.032	-0.030		-0.140	-0.135	
	(0.004)	(0.004)		(0.013)	(0.013)	
	[<0.001]	[<0.001]		[<0.001]	[<0.001]	
Sample	All Claimants	All Claimants	Female Claimants	All Claimants	All Claimants	Female Claimants
Doctor Fixed Effects		x			x	
Mean of Dep. Var.	0.619	0.619	0.600	6,281	6,281	5,726
Mean of Dep. Var. for Females Evaluated by Male Doctors	0.596	0.596	0.596	5,622	5,622	5,622
N	70,748	70,748	23,867	70,748	70,748	23,867

Notes: This table displays estimates of the coefficients on the female doctor indicator variable, the female claimant indicator variable, and the interaction of the female doctor and female claimant indicator variables from regressions of Equation (1) that control for credential-by-county fixed effects, exam year fixed effects, and injury year fixed effects. Each column represents a separate regression with the dependent variable as indicated in the table. The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017. Standard errors clustered at the doctor level are reported in parentheses, and p-values are reported in brackets.

Table 4: Other Patient Characteristics

Dependent Variable: I(Additional Cash Benefits > 0)			
	Demographics		
	Log(Wage)	Age	Married
	(1)	(2)	(3)
X * Female Doctor	-0.005	0.000	0.012
	(0.008)	(0.000)	(0.012)
	[0.526]	[0.341]	[0.301]
Female Doctor	0.039	-0.007	0.002
	(0.050)	(0.019)	(0.010)
	[0.438]	[0.708]	[0.825]
X	0.003	0.002	-0.005
	(0.003)	(0.000)	(0.005)
	[0.419]	[<0.001]	[0.321]
	Injury Characteristics		
	Sprain or Muscle Issue	Dangerous Industry	Predicted Log(Cash Benefits)
	(4)	(5)	(6)
X * Female Doctor	0.017	-0.003	-0.097
	(0.011)	(0.009)	(0.078)
	[0.124]	[0.755]	[0.211]
Female Doctor	0.000	0.011	0.954
	(0.010)	(0.009)	(0.758)
	[0.974]	[0.207]	[0.209]
X	0.040	0.046	0.626
	(0.005)	(0.004)	(0.037)
	[<0.001]	[<0.001]	[<0.001]
	Prior Medical Experience		
	Prior Doctor Female	Claimant Contested	Log(First Day Med Spend)
	(7)	(8)	(9)
X * Female Doctor	0.007	0.001	-0.008
	(0.014)	(0.012)	(0.004)
	[0.626]	[0.901]	[0.084]
Female Doctor	0.013	0.010	0.056
	(0.009)	(0.008)	(0.028)
	[0.140]	[0.205]	[0.048]
X	0.004	0.042	0.016
	(0.006)	(0.006)	(0.002)
	[0.508]	[<0.001]	[<0.001]

Notes: This table displays estimates of the coefficients on the female doctor indicator variable, the indicated claim or claimant characteristic, and the interaction of the female doctor indicator and the indicated claim or claimant characteristic—estimates from regressions of Equation (1) replacing the female claimant indicator with the characteristic indicated in the column. These regressions additionally control for credential-by-county fixed effects, exam year fixed effects, and injury year fixed effects. Each column represents a separate regression with the dependent variable being I(Additional Cash Benefits $>$ 0). The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017 with non-missing information for the indicated characteristic. Standard errors clustered at the doctor level are reported in parentheses, and p-values are reported in brackets.

Table 5: Other Doctor Characteristics

Dependent Variable: I(Additional Cash Benefits > 0)

	Doctor Credential (MD/DO)	Doctor Degree from Top 25 Medical School	Doctor Specialty: Internal or Family Medicine	Doctor Specialty: Orthopedics	Doctor State of Birth Texas
	(1)	(2)	(3)	(4)	(5)
X * Female Claimant	0.003	0.008	0.014	-0.015	-0.002
	(0.009)	(0.023)	(0.012)	(0.015)	(0.009)
	[0.704]	[0.746]	[0.232]	[0.315]	[0.852]
Female Claimant	-0.027	-0.026	-0.028	-0.025	-0.026
	(0.005)	(0.004)	(0.004)	(0.004)	(0.006)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[0.000]
X		-0.017	-0.018	0.037	-0.013
		(0.013)	(0.010)	(0.012)	(0.007)
		[0.179]	[0.072]	[0.003]	[0.062]

	General Experience		Designated Doctor Experience	
	Years of Experience	Share Female Patients in Past Year	Number of Exams in Past Year	Share Exams for Females in Past Year
	(6)	(7)	(8)	(9)
X * Female Claimant	0.000	0.014	0.000	0.023
	(0.000)	(0.017)	(0.000)	(0.032)
	[0.809]	[0.408]	[0.517]	[0.481]
Female Claimant	-0.027	-0.032	-0.030	-0.033
	(0.009)	(0.007)	(0.006)	(0.012)
	[0.002]	[<0.001]	[<0.001]	[0.005]
X	0.000	-0.009	0.000	-0.006
	(0.000)	(0.011)	(0.000)	(0.022)
	[0.315]	[0.412]	[0.712]	[0.795]

Notes: This table displays estimates of the coefficients on the female claimant indicator variable, the indicated doctor characteristics, and the interaction of the female claimant indicator and the indicated doctor characteristic—estimates from regressions of Equation (1) replacing the female doctor indicator with the characteristic indicated in the column. These regressions additionally control for credential-by-county fixed effects, exam year fixed effects, and injury year fixed effects. Each column represents a separate regression with the dependent variable being I(Additional Cash Benefits $>$ 0). The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017 by a doctor with non-missing information for the indicated characteristic. For columns 7 and 9, doctors with no other patients in the previous year in the data are assigned a share of female patients of 0, and the regressions include an indicator variable for doctors with no experience in the previous year. Standard errors clustered at the doctor level are reported in parentheses, and p-values are reported in brackets.

Table 6: Broader Evidence: Gender Differences in Cash Benefit Receipt

	Dependent Variable: I(Cash Benefits > 0)					
	(1)	(2)	(3)	(4)	(5)	(6)
Female Claimant	-0.055	-0.036	-0.033	-0.040	-0.037	-0.036
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Controls						
Insurer Fixed Effects		x	x	x	x	x
Injury Month-Year Fixed Effects			x	x	x	x
Injury Day-of-the-Week Fixed Effects			x	x	x	x
Claimant County Fixed Effects			x	x	x	x
Age Fixed Effects				x	x	x
Diagnosis Code Fixed Effects					x	x
Indicator for First Medical Treatment at ED						x
Mean of Dep. Var.	0.215	0.215	0.215	0.215	0.215	0.215
Mean of Dep. Var. for Males	0.236	0.236	0.236	0.236	0.236	0.236
N	1,076,759	1,076,759	1,076,759	1,076,759	1,076,759	1,076,759

Notes: This table displays coefficients on a female claimant indicator variable from separate OLS regressions of Equation (2). The specifications include controls as noted in the table. The diagnosis code fixed effects are based on the first three digits of the ICD-9 code from the earliest treatment. The sample contains 1,076,759 claims occurring from 2013 to 2017. Standard errors are reported in parentheses, and p-values are reported in brackets.

Table 7: Broader Evidence: Gender Homophily in Choice of Providers

Dependent Variable: I(Chosen "Treating Doctor" is Female)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female Claimant	0.023	0.018	0.017	0.015	0.015	0.015	0.014
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Controls							
Patient Medical Market (HSA) Fixed Effects		x	x	x	x	x	
Insurer Fixed Effects			x	x	x	x	
Injury Month-Year Fixed Effects				x	x	x	
Diagnosis Code Fixed Effects					x	x	x
Age Fixed Effects						x	x
Patient Medical Market (HSA) X Injury Month-Year X Diagnosis Code Fixed Effects							x
Mean of Dep. Var.	0.283	0.283	0.283	0.283	0.283	0.283	0.283
Mean of Dep. Var. for Males	0.274	0.274	0.274	0.274	0.274	0.274	0.274
N	528,538	528,538	528,538	528,538	528,538	528,538	528,538

Notes: This table displays coefficients on a female claimant indicator variable from separate OLS regressions of Equation (3). The sample contains the 528,538 claims occurring from 2013 to 2017 for which we are able to identify the claimants' treating doctor and the gender of the treating doctor. See Appendix Section A for more details. The specifications include controls as noted in the table. The diagnosis code fixed effects are based on the first three digits of the ICD-9 code from the earliest treatment. Standard errors are reported in parentheses, and p-values are reported in brackets.

Table 8: Survey: Health Care Experiences and Preferences

	Female	Male	Difference	p-value	Regression - Female Coefficient Est	Std Error	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Thinking about your experiences with health care visits in the past, have you ever felt that a doctor did any of the following? - Share answered yes							
Talked down to you or didn't treat you with dignity or respect	0.393	0.340	0.052	[0.035]	0.087	(0.026)	[0.001]
Didn't understand or relate to your experiences and concerns	0.450	0.356	0.094	[<0.001]	0.104	(0.027)	[<0.001]
Didn't believe you were telling the truth about your symptoms or concerns	0.317	0.286	0.031	[0.193]	0.045	(0.025)	[0.068]
Refused to order a test or treatment you thought you needed	0.223	0.232	-0.009	[0.666]	0.022	(0.022)	[0.310]
Made you feel uncomfortable discussing your concerns	0.292	0.250	0.042	[0.069]	0.059	(0.024)	[0.014]
Assumed something about you without asking	0.357	0.328	0.029	[0.236]	0.050	(0.026)	[0.054]
Refused to prescribe pain medication you thought you needed	0.175	0.249	-0.074	[<0.001]	-0.055	(0.021)	[0.008]
Share answered yes to any of the above	0.596	0.538	0.058	[0.023]	0.103	(0.026)	[<0.001]
Have you ever received care from a…. - Share answered yes							
Male doctor?	0.901	0.968	-0.068	[<0.001]	-0.062	(0.013)	[<0.001]
Female doctor?	0.941	0.767	0.174	[<0.001]	0.162	(0.019)	[<0.001]
Given the choice, would you prefer to see a doctor who is male or female, or does it not make							
Prefer to see a doctor who is female	0.463	0.136	0.327	[<0.001]	0.315	(0.024)	[<0.001]
Prefer to see a doctor who is male	0.058	0.290	-0.232	[<0.001]	-0.207	(0.020)	[<0.001]
Doesn't make much difference	0.479	0.574	-0.094	[<0.001]	-0.108	(0.028)	[<0.001]
Please indicate how important each of the following characteristics is if you were choosing a doctor - Share indicating at least moderately important							
Out-of-pocket cost for a visit	0.829	0.824	0.005	[0.809]	0.005	(0.022)	[0.807]
Doctor reviews (e.g., on websites like Healthgrades or Google or from friends or	0.847	0.807	0.040	[0.038]	0.048	(0.022)	[0.027]
Travel time to get to doctor's office	0.831	0.834	-0.003	[0.864]	0.002	(0.021)	[0.914]
Wait time at the doctor's office	0.863	0.848	0.015	[0.410]	0.018	(0.020)	[0.361]
Doctor's sex	0.414	0.336	0.077	[0.002]	0.112	(0.026)	[<0.001]
Doctor's age	0.349	0.370	-0.020	[0.416]	0.021	(0.026)	[0.415]

Notes: This table summarizes responses from several questions from the survey described in Section V. The table above reports the raw means for each of the indicated variables, along with the difference by respondent gender and the associated p-value testing whether the difference is zero. The table also includes the estimated coefficient on female (and associated standard error and p-value) from a regression of the indicated response on a female indicator and all the respondent characteristics reported in Appendix Table A8. The survey was conducted by Qualtrics and included 1,519 adults ages 30 to 64. Eligibility was restricted to individuals who reported working at some point in the last 12 months and individuals who self-identified as either male (755 respondents) or female (764 respondents).

APPENDIX

A Additional Details on Setting and Data

Prevalence of Independent Medical Evaluations Note that while 6.7% of claims for injuries occurring from 2013 to 2017 had an independent medical exam by the end of 2017, a higher share of these claims will have an exam at some point, since claims are more likely to have an independent medical exam the longer they are open. For example, for claims for 2013 injuries, 9% have had an independent medical exam by the end of the sample period. Because independent medical exams typically settle disputes about the worker's continued inability to work, they predominately occur for claims with cash disability benefits, which have higher stakes for both employers and injured workers than claims involving only medical benefits. Among claims for 2013 injuries, the 22% of claims receiving any cash benefits accounted for 86% of total medical and cash benefits paid out through the end of the sample period and 39% of these claims had an independent medical exam by the end of the sample period.

Information for Insurer, Diagnosis, and Injury Day of Week Fixed Effects In some specifications, we control for insurer fixed effects. There are 466 distinct insurers represented in the disputed claims sample. Because certified self-insured employers have their own insurer codes, insurer fixed effects are equivalent to controlling for an employer fixed effect among claimants working for some large employers.

We classify claims as sprains, fractures, muscle issues, or contusions using ICD-9 diagnosis codes from medical care received on the first day of injuries. For bills that identify ICD-10 codes, we convert ICD-10 codes to ICD-9 codes using a crosswalk from the Centers for Medicare & Medicaid Services so that we have consistent definitions of injuries over time. We classify claims based on their most frequently listed ICD-9 codes from among the above injury types with ties broken randomly. The 11% of claims without an ICD-9 code on the first day for sprains, fractures, muscle issues, or contusions are classified as claims for Other Injuries. We use this classification when testing for heterogeneity by injury type and for the descriptive statistics shown in Table 1. In addition, we also control for fixed effects for initial diagnoses in certain specifications. These fixed effects are based on the first three digits of the ICD-9 code used to classify the injury type for claims classified as sprains, fractures, muscle issues, or contusions and for the most frequently listed ICD-9 code for Other Injuries. We also use the first treatment to determine the day of the week that injuries occurred.

Identifying Gender of Designated Doctors We identify the gender of approximately 94% of doctors performing designated doctor exams in the sample using information from the CMS NPI registry. For doctors with missing NPI information in the workers' compensation insurance data, we classify their gender based on their first name if at least 99% of the providers in Texas in the CMS registry with that first name have the same gender.

Isolating Conditional Random Assignment As we describe in Section II, our estimation approach controls for the credential of the assigned designated doctor (known ex post) by claimant county fixed effects, as the assignment process means the designated doctor assigned to a claimant is random among designated doctors with that credential in the claimant's county. An alternative approach to isolate conditional random assignment would be to instead control for each injury type by county combination—the ex ante information the regulator uses to decide which subset of designated doctors within a county are eligible to evaluate a claimant. A potential concern with basing the estimation on a control for injury type is that we can only imperfectly observe the injury type information used by the regulator in selecting which doctor credentials are required to evaluate a claimant. Our measure of injury type, which is based on the claimant's diagnosis on the first day of treatment for that injury, is imperfect as it could in principle miss aspects of a claimant's injury that the regulator considers when deciding which types of doctors are eligible to evaluate the claimant (MDs, DOs, and/or DCs). For example, while it is straightforward to identify musculoskeletal injuries, our measure of injury type may miss secondary non-musculoskeletal conditions (e.g., subsequent secondary diagnoses of mental health conditions, secondarily affected body systems) that may lead the regulator to require an evaluation by an MD or DO, rather than a DC. Thus, to be conservative, our baseline approach is to control for credential of the assigned doctor by county, as the assignment mechanism means that it is as good as random which doctor was assigned to the claimant among designated doctors in that county with that credential. In practice, we obtain very similar estimates if we control for injury type by county fixed effects instead of credential by county fixed effects in our analysis. See Appendix Tables A4 and A5 for estimates from this alternative specification. It is not surprising that we obtain similar results using either approach. The measurement error in our injury type variable may be limited in practice, and the vast majority of exams are for claimants with musculoskeletal injuries, whose exams can be performed by any designated doctor regardless of credential.

Construction of Normalized Additional Benefits Measure To assess the value of benefits received after the exam, we create a second variable—"normalized additional cash benefits"—to reflect the total cash benefits received in the year after the exam valued at the mean benefit rate for each benefit type based on the population-wide distribution of pre-injury wages. To create this measure, we first calculate the mean benefit rate for each benefit type—the population-wide mean inflation-adjusted benefits paid per week receiving benefits based on all claims with nonzero benefits.[1] We then calculate the normalized additional benefits received as the sum of the post-evaluation temporary income benefits and permanent impairment benefits valued at the mean benefit rate for that benefit type. Specifically, we multiply the mean benefit rate for each type of benefit by the number of weeks the claimant received benefits and then add these two amounts to calculate the normalized additional cash benefits. A feature of this measure is that it depends on only the claimant's evaluated degree of disability rather than the claimant's pre-injury wages.

Identifying Treating Doctors We identify a claimant's treating doctor as the doctor who submits a workers' compensation report describing a claimant's ability to work. If a claimant does not have a bill that explicitly states that a work status exam was performed, we identify the claimant's treating doctor as the doctor who bills for case management services. If multiple doctors bill for case management services for a claim, we identify the treating doctor as the first doctor to bill for case management services. We can identify the treating doctor and the treating doctor's gender for about half of claims in our sample using this approach. For the other half of claims, either no case management services were billed, which often happens for medical-only claims (i.e., claims without cash benefits), or the case management services were billed to a health care organization rather than to a specific provider (in which case, we cannot determine the gender of the treating doctor). We obtain similar results if we instead adopt a treating doctor definition based on first office visit, which likely misidentifies treating doctors more often than our baseline approach does but leads to fewer claimants having missing treating doctor gender.

Predicting Cash Benefits with Lasso Model To create the measure of predicted cash benefits, we first fit a lasso model of normalized cash benefits where benefits are measured through one year after the designated doctor exam. We then predict claimants' cash benefits based on this lasso model. For the lasso model, we include indicator variables for ten-year age bins, wage deciles, day of the week of first medical treatment, industry, and injury type. We also include the cost of first-day medical treatment and indicator variables for first treatment occurring in the emergency department and for marital status.

B Additional Results

Robustness to Included Controls Appendix Tables A4 and A5 further probe the robustness of our findings to the inclusion of different combinations of fixed effects. The estimates from these alternative specifications are similar to our baseline estimates. For instance, we obtain similar estimates when excluding injury year and exam year fixed effects. We also obtain similar estimates when replacing separate county-credential and exam year fixed effects with county-credential-year fixed effects or county-credential-quarter fixed effects. This rules out alternative explanations related to the pool of designated doctors changing over time. We also obtain similar estimates in a specification including county-doctor-year fixed effects, in which the coefficient on the female doctor and female claimant interaction term is identified using differences in outcomes between male and female claimants who are assigned the same doctor from the same pool of

[1] As described in the text, temporary income benefits are paid for the weeks that individuals miss work while healing from temporary impairments. Permanent impairment severity—as rated by a claimant's treating doctor or designated doctor—determines the number of weeks permanent impairment benefits are paid, though these benefits are paid regardless of whether a claimant remains out of work or returns to work. Among individuals who receive non-zero benefits for a particular benefit type, we calculate an individual's inflation-adjusted benefit rate for that benefit type as the ratio of the inflation-adjusted dollars of benefits received to the weeks receiving benefits. We then calculate the population-wide mean weekly rate for a particular benefit type by taking the mean weekly rate among individuals who receive non-zero benefits for that benefit type.

potential designated doctors. Further, the estimates are similar in specifications with interactions between claimant gender and time, which validates that the documented differences are not driven by unobserved factors related to female claimant evaluations that vary across time or space. We also obtain similar estimates when we replace the county-by-credential fixed effects with county-by-diagnosis fixed effects, an alternative approach to isolating conditional random assignment.[2] We obtain similar estimates if we supplement the baseline specification with insurer-by-year fixed effects, which rules out changes in insurer practices across time as an explanation of the results. Further, we obtain similar estimates when we include fixed effects for insurer-county-credential-year or insurer-county-credential-quarter, providing reassurance that exclusions due to conflicts of interest with insurers do not impact the estimates. Finally, we obtain similar estimates in specifications where we interact all our baseline controls with claimant gender. Overall, the results are very similar regardless of the combination of fixed effects included.

C Survey Questions: Full Text

A Experiences with Health Care System (summarized in Appendix Table A9)

- Thinking about your experiences with health care visits in the past, have you ever felt that a doctor did any of the following? *Answer options: Yes, has happened; No, has not happened.*
  - Talked down to you or didn't treat you with dignity or respect
  - Didn't understand or relate to your experiences and concerns
  - Didn't believe you were telling the truth about your symptoms or concerns
  - Refused to order a test or treatment you thought you needed
  - Made you feel uncomfortable discussing your concerns
  - Assumed something about you without asking
  - Refused to prescribe pain medication you thought you needed
- Thinking about your experiences getting health care for yourself, which doctor—male or female—would be more likely to *Answer options: Male doctor; Female doctor; Male and Female doctors are equally likely.*[3]
  - treat you with dignity and respect?
  - understand or relate to your experiences and concerns?
  - believe you are telling the truth about your symptoms or concerns?
  - provide needed testing and treatments?
  - make you feel comfortable with discussing your concerns?
  - ask appropriate questions instead of making assumptions?
  - be the most qualified?
  - be available near you?

B Preferences and Homophily (summarized in Appendix Table A10)

Hypothetical Choice Questions The hypothetical choice questions are designed to measure willingness to pay to see an own-gender provider. Before this set of questions, we provided the following framing:

> **Now we will move to a set of questions that will help us learn the importance you place on doctor attributes and out-of-pocket cost per visit. For these questions, suppose you need to see a doctor for medical evaluation and treatment.**
>
> Please provide the answers that best reflect your preferences.
>
> Please click on "Next" to continue...

[2]See Section I and Appendix Section A for a detailed description of the designated doctor assignment mechanism and alternative approaches to empirically isolate conditional random assignment.

[3]To ensure primacy bias did not impact our results, we randomized which option—Male doctor or Female doctor—appeared first for this set of questions.

Respondents were then presented with a hypothetical choice question. Specifically, the respondents were presented with one of four choice sets:

- Female doctor, \$30 out-of-pocket cost or Male doctor, \$5 out-of-pocket cost
- Female doctor, \$10 out-of-pocket cost or Male doctor, \$5 out-of-pocket cost
- Female doctor, \$5 out-of-pocket cost or Male doctor, \$10 out-of-pocket cost
- Female doctor, \$5 out-of-pocket cost or Male doctor, \$30 out-of-pocket cost.

The choice set presented to each respondent was randomly assigned among those above. An example of the presentation of this hypothetical choice question is copied below:

If the two doctors presented below are identical in every other way except for the listed characteristics, which one would you prefer to see for evaluation and treatment?

Note: *There is no right or wrong answer. You should select the option that best reflects your personal preferences.*

(Please select one option below)

Doctor 1	**Doctor 2**
Doctor sex: Female	Doctor sex: Male
Out-of-pocket cost for visit: \$30	Out-of-pocket cost for visit: \$5
○	○

For completeness, we asked sequential follow-up questions—moving either upward or downward from the initial randomly assigned choice set among the choice sets listed above—based on the respondent's answer to prior questions to discern the maximum the respondent is willing to pay for a doctor of a particular gender in terms of out-of-pocket cost differentials. Thus, each respondent is asked at most four hypothetical choice questions, though the median respondent was only asked 2 questions (mean 2.5 questions). To ensure that primacy bias did not lead to a higher percentage of respondents choosing either the male or female doctor option, we randomly assigned which option–the male doctor or the female doctor option—appeared first in each of the four choice sets above, both in the initially assigned hypothetical choice question and all follow-up hypothetical choice questions. While the discussion in the text focuses on the results from the initial randomly assigned choice set, Appendix Figure A5 illustrates that the patterns in selecting own-gender doctors are very similar when based on estimates from the full set of hypothetical choice questions asked of respondents.

Other Questions

- Have you ever received care from a. . . .
  - Male doctor? *Answer options: Yes; No.*
  - Female doctor? *Answer options: Yes; No.*
- Given the choice, would you prefer to see a doctor who is male or female, or does it not make much difference to you? *Answer options: Prefer to see a doctor who is female; Prefer to see a doctor who is male; Doesn't make much difference.*[4]
- Please indicate how important each of the following characteristics is if you were choosing a doctor. *Answer options: Not at all important; Slightly important; Moderately important; Very important; Extremely important.*
  - Out-of-pocket cost for a visit
  - Doctor reviews (e.g., on websites like Healthgrades or Google or from friends or family)

[4]We randomized which option was presented first—Prefer to see a doctor who is female or Prefer to see a doctor who is male—to ensure primacy bias did not affect the estimates.

  - Travel time to get to doctor's office
  - Wait time at the doctor's office
  - Doctor's sex
  - Doctor's age

C Respondent Characteristics (summarized in Appendix Table A8)

- Are you male or female? *Answer options: Male; Female; Other.*
- What is your age? *Answer options: under 30 years; 30 to 39 years; 40 to 49 years; 50 to 59 years; 60 to 64 years; 65 years or over.*
- Have you worked in the last 12 months? *Answer options: Yes; No.*
- In the last year, have you ever been unemployed or out of the labor force? *Answer options: Yes; No.*
- Are you currently married, living with a partner, widowed, divorced, separated, or have you never been married? *Answer options: Married; Living with partner; Not married and not living with a partner.*
- Are you of Hispanic, Latino, or Spanish origin? *Answer options: Yes; No; Decline to state.*
- Do you or anyone in your household work in a health care delivery setting, such as a doctor's office, clinic, hospital, nursing home, or dentist's office? *Answer options: Yes, I work in health care delivery setting; Yes, both myself and someone else in household; Yes, someone else in household works in health care delivery setting; No.*
- Are you, yourself, now covered by any form of health insurance or health plan? *Answer options: Yes; No.*
- In the past five years, have you interacted with a doctor to get health care for yourself? *Answer options: Yes; No.*
- In the last five years, have you experienced chronic physical pain that has interfered with your daily activities? *Answer options: Yes; No.*
- What is your race? *Answer options: White; Black or African American; Asian; Other; Decline to state.*
- What is the highest level of school you have completed or the highest degree you have received? *Answer options: Less than High School; High School Graduate or GED; Some College; Bachelor's Degree; Postgraduate Degree.*
- About how much do you earn in a year through your job(s)? *Answer options: Less than $10,000; $10,000 to less than $20,000; $20,000 to less than $30,000; $30,000 to less than $40,000; $40,000 to less than $50,000; $50,000 to less than $75,000; $75,000 to less than $90,000; $90,000 to less than $100,000; $100,000 or more.*
- What is the industry of the main job you have held in the past year? (Please select the answer option that fits best.) *Answer options: Agriculture/Forestry/Fishing/Hunting; Arts/Entertainment/Accommodation/Food Services; Information/Finance/Real Estate/Professional Services; Health Care/Educational Services; Manufacturing; Mining/Utilities/Construction; Public Administration/Other Services; Wholesale Trade/Retail Trade/Transportation.*

D Policy Counterfactuals

Below, we outline the setup for the back-of-the-envelope policy counterfactuals. In the discussion below, references to the "gender gap" pertain to the gender gap conditional on all available observable characteristics. To conduct this analysis, we combine our estimates on the effect of gender match from the randomized evaluations with our broader evidence on the gender gap and gender homophily in patient selection of treating doctors within workers' compensation insurance more generally. We note that this broader back-of-the-envelope counterfactual analysis involves extrapolation beyond the randomized evaluations, and thus the quantitative findings from this analysis should be interpreted with the appropriate caution.

A Setup

We define some notation to describe this analysis. Let r_j be the share of patients of gender j who choose a female treating doctor, and let s denote the share of patients who are female. In this notation, the share of overall treating doctors who are female can be expressed as: $r \equiv r_m(1-s) + r_f(s)$. We define relative gender homophily as the ratio of the share of female patients selecting female doctors to the share of male patients selecting female doctors, $Z \equiv \frac{r_f}{r_m}$. Let Gap_j denote the expected gender gap in benefit receipt— the percent reduction in the rate of benefit receipt for females relative to the analogous rate for males—when female claimants are evaluated by doctors of gender j. Using this notation, the overall gender gap in benefit receipt can be expressed as: $Gap \equiv Gap_m(1-r_f) + Gap_f(r_f)$. We define the gender-match effect as the ratio of the gender gap when patients are evaluated by female doctors relative to the gender gap when patients are evaluated by male doctors: $X \equiv \frac{Gap_f}{Gap_m}$.

It is straightforward to show that we can express the gender gap as a function of the overall share of treating doctors who are female (r) and the degree of homophily in the market (Z):

$$Gap(r, Z) = Gap_m \times (1 - \frac{r}{\frac{1-s}{Z} + s}) + Gap_f \times (\frac{r}{\frac{1-s}{Z} + s}), \tag{1}$$

where $Gap_f \equiv \frac{XGap^0}{r_f^0(X-1)+1}$, $Gap_m \equiv \frac{Gap^0}{r_f^0(X-1)+1}$, and Gap^0 and r_f^0 denote the overall gender gap and share of female patients selecting female doctors in the status quo, respectively. Appendix Section C below presents this derivation in more detail. In the calculations that follow, we take the share of female patients (s) as given at the observed share in the workers' compensation insurance population: 38%.

We make two simplifying assumptions. First, we assume that patient-doctor gender match causes the same percent reduction in gender gaps in the likelihood of cash benefit receipt in broader workers' compensation insurance as it does within the set of claims with randomized evaluations. In other words, we assume the gender match effect (X) we estimate among the set of randomized evaluations applies more broadly within workers' compensation insurance. Second, we assume relative gender homophily in patient selection of treating doctors (Z) is constant with respect to the overall share of female treating doctors. For instance, female patients may select female doctors at Z times the rate that male patients select female doctors, but Z is constant as the share of doctors who are female varies.[5]

B Results

Using the relationship defined in Equation (1), Figure A3 illustrates the effect of varying one dimension holding the other fixed at the observed values in the status quo. In both panels, the vertical axis displays the expected gender gap in benefit receipt—the percent reduction in the rate at which females receive cash benefits relative to analogous rate for males with the same observable characteristics, while the horizontal axis displays either the share of female treating doctors (Panel A) or the relative gender homophily in patient-doctor matches (Panel B). For reference, each figure displays an "x" representing the observed combination in the status quo: a gender gap is 15.3% (i.e., females are 15.3% less likely to receive cash benefits than males with the same observables), the share of treating doctors who are female is 27.9%, and the degree of relative gender homophily is 1.051 (i.e., female patients select female doctors at 1.051 times the rate that male patients select female doctors).

Holding fixed the observed degree of gender homophily, Figure A3 Panel A plots the relationship between the gender gap and the share of female treating doctors. This figure indicates that increasing the share of female treating doctors by 22.1 percentage points—moving from 27.9% to parity—-would lead to a 33.7% decrease in the observed gender gap. The observed degree of gender homophily in patient-doctor matches works to reinforce the effects of increasing gender diversity among doctors, with a 22.1 percentage point increase in the share of female treating doctors translating to a 22.8 percentage point increase in the share of female claimants who see female doctors. Additionally, Figure A3 Panel A can be used to determine the increase in the share of female doctors necessary to offset a given amount of the gender gap in

[5] While it would be straightforward to extend these calculations to allow gender homophily to vary with the share of doctors who are female, we avoid doing this for a few reasons. First, whether Z and r are correlated—and the direction of any correlation—is ex ante unclear. Second, we cannot credibly estimate how Z varies with r. While Section IV presents estimates of relative gender homophily across the workers' compensation insurance system in Texas, there is no plausibly exogenous variation in the availability of female doctors across geography to identify how relative gender homophily varies with the share of doctors who are female.

cash benefit receipt. To offset a quarter of the gender gap in benefit receipt observed in the status quo, it would take a 16.4 percentage point increase in the share of treating doctors who are female.

Figure A3 Panel B plots the relationship between the gender gap and the degree of gender homophily, fixing the gender composition of doctors as observed in the status quo. An increase in relative gender homophily from the observed level to two (i.e., where female patients choose female doctors at twice the rate that male patients choose female doctors), would lead to a 17.3% decrease in the observed gender gap. Holding all else equal, female patients would need to choose female doctors 2.7 times as often as male patients choose female doctors to offset a quarter of the observed gender gap.

More broadly, we can characterize the trade-offs between policies that may increase gender diversity among doctors and policies that increase sorting of patients to own-gender doctors. Figure A4 illustrates level curves of Equation (1), which characterize the combinations of gender composition in the doctor workforce and gender homophily patient-doctor matches that would result in a given value of the gender gap. The combination of conditions in the status quo is represented as “x” in this figure. There are at least two important qualitative properties worth noting. First, the level curves are negatively sloped. This indicates that a given reduction in the gender gap can be achieved by trading-off increases in the share of doctors who are female and increases in relative gender homophily. Second, the level curves are convex relative to the origin. This convexity reflects the fact that the inputs—the share of doctors who are female and relative gender homophily—are complementary in reducing the gender gap.

Suppose we are interested in understanding the change in conditions needed to reduce the gender gap from 15.2% to 12.5%. This reduction in the gender gap could be accomplished by increasing the share of doctors female to 39.8%, holding homophily fixed at 1.051. Alternatively, it could be accomplished by increasing homophily to 2.06, holding the share of female doctors fixed at 27.9%. More generally, the combinations of conditions that would lead to the same closure of the gap are depicted by the level curve corresponding to 0.125, where the difference between this curve and the observed conditions “x” indicates the changes in conditions necessary to achieve this closure. Reducing the gender gap to 12.5% could be accomplished by a convex combinations of changes in the gender composition of treating doctors and relative gender homophily—for example, by increasing gender homophily to 1.75 and the share of treating doctors who are female to 30.1%, or increasing gender homophily to 1.25 and the share of treating doctors who are female to 35.9%.

C More Detail on Policy Counterfactuals

The goal of this analysis is to understand how gender gaps conditional on observables are impacted by: (i) the share of treating doctors who are female (r) and (ii) the degree of gender homophily in patient-doctor matches (Z). In other words, we would like to express the gender gap (conditional on observables) as a function of the overall share of treating doctors who are female and the degree of homophily in the market: $Gap(r, Z)$. We derive this function in two steps. First, following directly from the definitions above, we can express the share of female claimants seeing female doctors (r_f) as a function of the share of female treating doctors overall (r) and the degree of relative gender homophily (Z):

$$r_f = \frac{r}{\frac{1-s}{Z} + s}. \tag{2}$$

Second, we can use our estimate of the gender-match effect (X) and the definitions to infer the values of Gap_m and Gap_f:

$$Gap_m \equiv \frac{Gap^0}{r_f^0(X-1)+1} \tag{3}$$

$$Gap_f \equiv \frac{XGap^0}{r_f^0(X-1)+1}, \tag{4}$$

where Gap^0 and r_f^0, respectively, represent the observed values of the gender gap and the share of female patients selecting female treating doctors in the status quo.

Putting these together, we obtain:

$$Gap(r, Z) = Gap_m \times (1 - \frac{r}{\frac{1-s}{Z} + s}) + Gap_f \times (\frac{r}{\frac{1-s}{Z} + s}), \tag{5}$$

where $Gap_f \equiv \frac{XGap^0}{r_f^0(X-1)+1}$ and $Gap_m \equiv \frac{Gap^0}{r_f^0(X-1)+1}$.

Table A1: Effect on Indicator for Receiving Additional Benefits, Sensitivity to Clustering Level

	Dependent Variable: I(Additional Cash Benefits > 0)					
	(1)	(2)	(3)	(4)	(5)	(6)
Female Doctor X Female Claimant	0.031	0.031	0.031	0.031	0.031	0.031
	(0.010)	(0.010)	(0.009)	(0.010)	(0.009)	(0.010)
	[0.002]	[0.002]	[0.001]	[0.001]	[<0.001]	[0.001]
Female Doctor	-0.000	-0.000	-0.000	-0.000	-0.000	-0.000
	(0.008)	(0.006)	(0.006)	(0.006)	(0.008)	(0.008)
	[0.964]	[0.952]	[0.954]	[0.951]	[0.964]	[0.963]
Female Claimant	-0.032	-0.032	-0.032	-0.032	-0.032	-0.032
	(0.004)	(0.004)	(0.007)	(0.005)	(0.007)	(0.005)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Level of Clustering						
Doctor	x				x	x
None		x				
County			x		x	
County by Quarter				x		x
Mean of Dep. Var.	0.619	0.619	0.619	0.619	0.619	0.619
N	70,748	70,748	70,748	70,748	70,748	70,748

Notes: This table displays estimates of the coefficients on the female doctor indicator variable, the female claimant indicator variable, and the interaction of the female doctor and female claimant indicator variables. Each column represents a separate regression with the dependent variable being I(Additional Cash Benefits > 0). The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017. Standard errors clustered at the indicated level are reported in parentheses, and p-values are reported in brackets.

Table A2: Effect on Amount of Additional Benefits, Sensitivity to Clustering Level

	Dependent Variable: Additional Cash Benefits					
	(1)	(2)	(3)	(4)	(5)	(6)
Female Doctor X Female Claimant	0.086	0.086	0.086	0.086	0.086	0.086
	(0.030)	(0.029)	(0.026)	(0.029)	(0.027)	(0.031)
	[0.004]	[0.003]	[0.001]	[0.003]	[0.002]	[0.005]
Female Doctor	0.032	0.032	0.032	0.032	0.032	0.032
	(0.026)	(0.016)	(0.015)	(0.016)	(0.024)	(0.025)
	[0.208]	[0.050]	[0.031]	[0.046]	[0.178]	[0.202]
Female Claimant	-0.140	-0.140	-0.140	-0.140	-0.140	-0.140
	(0.013)	(0.013)	(0.016)	(0.014)	(0.017)	(0.014)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Level of Clustering						
Doctor	x				x	x
None		x				
County			x		x	
County by Quarter				x		x
Mean of Dep. Var.	6,281	6,281	6,281	6,281	6,281	6,281
N	70,748	70,748	70,748	70,748	70,748	70,748

Notes: This table displays estimates of the coefficients on the female doctor indicator variable, the female claimant indicator variable, and the interaction of the female doctor and female claimant indicator variables. Each column represents a separate Poisson regression with the dependent variable being the amount of additional normalized benefits received after the exam. The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017. Standard errors clustered at the indicated level are reported in parentheses, and p-values are reported in brackets.

Table A3: Effect on Amount of Additional Benefits, Alternative Specifications

	(1)	(2)	(3)	(4)	(5)	(6)
Female Doctor X Female Claimant	0.086	498.237	0.109	494.014	0.101	513.389
	(0.030)	(189.407)	(0.033)	(184.872)	(0.034)	(189.974)
	[0.004]	[0.009]	[0.001]	[0.008]	[0.003]	[0.007]
Female Doctor	0.032	206.351	0.027	174.474	0.020	143.214
	(0.026)	(170.932)	(0.027)	(184.661)	(0.026)	(187.043)
	[0.208]	[0.228]	[0.308]	[0.345]	[0.441]	[0.444]
Female Claimant	-0.140	-844.925	-0.405	-2,245.698	-0.233	-1,261.951
	(0.013)	(79.111)	(0.015)	(80.123)	(0.015)	(82.918)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Estimation						
Poisson	x		x		x	
OLS		x		x		x
Dependent Variable						
Additional Cash Benefits	x	x				
Additional Cash Benefits, without Normalization			x	x	x	x
Control for log(wage)					x	x
Mean of Dep. Var. for Females Evaluated by Male Doctors	5,622	5,622	4,478	4,478	4,752	4,752
N	70,748	70,748	70,748	70,748	67,233	67,233
Implied % Effect of Coefficient on Female Doctor X Female Claimant		0.089		0.110		0.108
Implied $ Effect of Coefficient on Female Doctor X Female Claimant	483.731		487.017		478.240	

Notes: This table displays estimates of the coefficients on the female doctor indicator variable, the female claimant indicator variable, and the interaction of the female doctor and female claimant indicator variables. The estimation and dependent variables are indicated in the table. The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017. Standard errors clustered at the doctor level are reported in parentheses, and p-values are reported in brackets.

Table A4: Effect on Indicator for Receiving Additional Benefits, Robustness to Varying Fixed Effects

	Dependent Variable: I(Additional Cash Benefits > 0)											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Female Doctor X Female Claimant	0.031	0.031	0.033	0.037	0.040	0.030	0.030	0.030	0.028	0.026	0.033	0.030
	(0.010)	(0.010)	(0.010)	(0.011)	(0.014)	(0.010)	(0.010)	(0.011)	(0.010)	(0.013)	(0.016)	(0.011)
	[0.002]	[0.002]	[0.002]	[0.001]	[0.005]	[0.004]	[0.004]	[0.006]	[0.006]	[0.050]	[0.036]	[0.004]
Female Doctor	-0.000	-0.002	-0.001	-0.002		0.000	-0.000	0.000	-0.000	0.004	-0.009	0.000
	(0.008)	(0.008)	(0.008)	(0.008)		(0.008)	(0.008)	(0.009)	(0.008)	(0.010)	(0.010)	(0.008)
	[0.964]	[0.822]	[0.917]	[0.849]		[0.972]	[0.967]	[0.956]	[1.000]	[0.656]	[0.407]	[0.980]
Female Claimant	-0.032	-0.031	-0.031	-0.031	-0.020			-0.032	-0.021	-0.017	-0.019	
	(0.004)	(0.004)	(0.005)	(0.005)	(0.006)			(0.005)	(0.005)	(0.006)	(0.007)	
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[0.001]			[<0.001]	[<0.001]	[0.006]	[0.007]	
Additional Controls												
County by Credential Fixed Effects	x	x				x	x		x			x
Exam Year Fixed Effects	x							x				x
Injury Year Fixed Effects	x		x	x	x	x	x	x	x	x	x	x
County by Credential by Exam Year Fixed Effects			x									
County by Credential by Exam Quarter Fixed Effects				x								
County by Provider by Exam Year Fixed Effects					x							
Gender by County and Gender by Exam Year Fixed Effects						x						
Gender by County by Exam Year Fixed Effects							x					
County by Diagnosis Fixed Effects								x				
Insurer by Exam Year Fixed Effects									x			
Insurer by County by Credential by Exam Year Fixed Effects										x		
Insurer by County by Credential by Exam Quarter Fixed Effects											x	
Baseline Controls by Gender												x
Mean of Dep. Var.	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619	0.619
N	70,748	70,748	70,748	70,748	70,748	70,748	70,748	70,748	70,748	70,748	70,748	70,748

Notes: This table displays estimates of the coefficients on the female doctor indicator variable, the female claimant indicator variable, and the interaction of the female doctor and female claimant indicator variables. Each column represents a separate regression with the dependent variable being I(Additional Cash Benefits $>$ 0). The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017. Standard errors clustered at the doctor level are reported in parentheses, and p-values are reported in brackets.

Table A5: Effect on Amount of Additional Benefits, Robustness to Varying Fixed Effects

	Dependent Variable: Additional Cash Benefits											
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Female Doctor X Female Claimant	0.086	0.084	0.087	0.096	0.103	0.082	0.079	0.090	0.081	0.059	0.065	0.084
	(0.030)	(0.031)	(0.031)	(0.031)	(0.038)	(0.031)	(0.031)	(0.031)	(0.030)	(0.033)	(0.036)	(0.031)
	[0.004]	[0.006]	[0.005]	[0.002]	[0.007]	[0.007]	[0.010]	[0.004]	[0.007]	[0.074]	[0.072]	[0.006]
Female Doctor	0.032	0.025	0.040	0.038		0.033	0.036	0.035	0.032	0.058	0.030	0.033
	(0.026)	(0.026)	(0.025)	(0.024)		(0.026)	(0.025)	(0.026)	(0.026)	(0.025)	(0.026)	(0.026)
	[0.208]	[0.332]	[0.105]	[0.123]		[0.200]	[0.150]	[0.176]	[0.217]	[0.021]	[0.251]	[0.194]
Female Claimant	-0.140	-0.141	-0.137	-0.138	-0.111			-0.140	-0.091	-0.082	-0.085	
	(0.013)	(0.013)	(0.013)	(0.013)	(0.018)			(0.014)	(0.014)	(0.016)	(0.018)	
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]			[<0.001]	[<0.001]	[<0.001]	[<0.001]	
Additional Controls												
County by Credential Fixed Effects	x	x				x	x		x			x
Exam Year Fixed Effects	x							x				x
Injury Year Fixed Effects	x		x	x	x	x	x	x	x	x	x	x
County by Credential by Exam Year Fixed Effects			x									
County by Credential by Exam Quarter Fixed Effects				x								
County by Provider by Exam Year Fixed Effects					x							
Gender by County and Gender by Exam Year Fixed Effects						x						
Gender by County by Exam Year Fixed Effects							x					
County by Diagnosis Fixed Effects								x				
Insurer by Exam Year Fixed Effects									x			
Insurer by County by Credential by Exam Year Fixed Effects										x		
Insurer by County by Credential by Exam Quarter Year Fixed Effects											x	
Baseline Controls by Gender												x
Mean of Dep. Var.	6,281	6,281	6,281	6,281	6,281	6,281	6,281	6,281	6,281	6,281	6,281	6,281
N	70,748	70,748	70,748	70,748	70,748	70,748	70,748	70,748	70,748	70,748	70,748	70,748

Notes: This table displays estimates of the coefficients on the female doctor indicator variable, the female claimant indicator variable, and the interaction of the female doctor and female claimant indicator variables. Each column represents a separate Poisson regression with the dependent variable being the amount of additional normalized benefits received after the exam. The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017. Standard errors clustered at the doctor level are reported in parentheses, and p-values are reported in brackets.

Table A6: Effect on Indicator for Receiving Additional Benefits, Controlling for Additional Claimant and Injury Characteristics

	Dependent Variable: I(Additional Cash Benefits > 0)						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female Doctor X Female Claimant	0.031	0.027	0.027	0.028	0.027	0.028	0.027
	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)	(0.010)
	[0.002]	[0.008]	[0.008]	[0.006]	[0.008]	[0.007]	[0.008]
Female Doctor	-0.000	0.001	0.001	0.000	0.002	0.002	0.002
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
	[0.964]	[0.903]	[0.900]	[0.961]	[0.825]	[0.842]	[0.835]
Female Claimant	-0.032	-0.020	-0.020	-0.026	-0.026	-0.026	-0.026
	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Additional Controls							
Insurer Fixed Effects		x	x	x	x	x	x
Injury Day-of-the-Week Fixed Effects			x	x	x	x	x
Age Fixed Effects				x	x	x	x
Diagnosis Code Fixed Effects					x	x	x
Indicator for First Medical Treatment at ED						x	x
Medical Spending on First Treatment Date							x
Mean of Dep. Var.	0.619	0.619	0.619	0.619	0.619	0.619	0.619
N	70,748	70,748	70,748	70,748	70,748	70,748	70,748

Notes: This table displays estimates of the coefficients on the female doctor indicator variable, the female claimant indicator variable, and the interaction of the female doctor and female claimant indicator variables from OLS regressions of Equation (1) that control for credential-by-county fixed effects, exam year fixed effects, and injury year fixed effects. Each column represents a separate regression with the dependent variable being I(Additional Cash Benefits $>$ 0). The specifications include additional controls as noted in the table. The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017. Standard errors clustered at the doctor level are reported in parentheses, and p-values are reported in brackets.

Table A7: Effect on Amount of Additional Benefits, Controlling for Additional Claimant and Injury Characteristics

	Dependent Variable: Additional Cash Benefits						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Female Doctor X Female Claimant	0.086	0.078	0.078	0.080	0.076	0.076	0.076
	(0.030)	(0.031)	(0.031)	(0.031)	(0.031)	(0.030)	(0.030)
	[0.004]	[0.012]	[0.012]	[0.009]	[0.013]	[0.012]	[0.012]
Female Doctor	0.032	0.033	0.033	0.031	0.035	0.035	0.035
	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)	(0.026)
	[0.208]	[0.199]	[0.200]	[0.227]	[0.169]	[0.179]	[0.178]
Female Claimant	-0.140	-0.087	-0.087	-0.103	-0.106	-0.105	-0.104
	(0.013)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)	(0.014)
	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]	[<0.001]
Additional Controls							
Insurer Fixed Effects		x	x	x	x	x	x
Injury Day-of-the-Week Fixed Effects			x	x	x	x	x
Age Fixed Effects				x	x	x	x
Diagnosis Code Fixed Effects					x	x	x
Indicator for First Medical Treatment at ED						x	x
Medical Spending on First Treatment Date							x
Mean of Dep. Var.	6,281	6,281	6,281	6,281	6,281	6,281	6,281
N	70,748	70,748	70,748	70,748	70,748	70,748	70,748

Notes: This table displays estimates of the coefficients on the female doctor indicator variable, the female claimant indicator variable, and the interaction of the female doctor and female claimant indicator variables from Poisson regressions of Equation (1) that control for credential-by-county fixed effects, exam year fixed effects, and injury year fixed effects. Each column represents a separate regression with the dependent variable being the amount of additional normalized benefits received after the exam. The specifications include additional controls as noted in the table. The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017. Standard errors clustered at the doctor level are reported in parentheses, and p-values are reported in brackets.

Table A8: Survey: Respondent Characteristics

	Survey Respondents				Workers in ACS		Workers' Comp Claimants	
	Female (1)	Male (2)	Difference	p-value	Female (3)	Male (4)	Female (5)	Male (6)
Ever unemployed in last 12 months	0.270	0.225	0.044	[0.045]	-	-	-	-
Married	0.516	0.640	-0.124	[<0.001]	0.590	0.639	0.454	0.558
Hispanic	0.093	0.074	0.019	[0.187]	0.158	0.184	-	-
Works in healthcare	0.179	0.117	0.063	[0.001]	0.193	0.052	-	-
Has health insurance	0.878	0.891	-0.013	[0.424]	0.908	0.867	-	-
Interacted with a doctor to get health care in last five years	0.932	0.918	0.014	[0.299]	-	-	-	-
Experienced chronic physical pain that interfered with daily activities in last five years	0.442	0.491	-0.049	[0.056]	-	-	-	-
Age:								
30-39 years	0.355	0.270	0.085	[<0.001]	0.314	0.322	0.262	0.339
40-49 years	0.260	0.351	-0.091	[<0.001]	0.292	0.292	0.315	0.311
50-55 years	0.296	0.309	-0.013	[0.587]	0.282	0.276	0.296	0.244
60-64 years	0.089	0.070	0.019	[0.176]	0.112	0.110	0.092	0.076
Race:								
White	0.822	0.837	-0.015	[0.434]	0.721	0.744	-	-
Black or African American	0.085	0.073	0.012	[0.377]	0.135	0.106	-	-
Asian	0.043	0.036	0.007	[0.457]	0.069	0.066	-	-
Other	0.038	0.045	-0.007	[0.490]	0.074	0.085	-	-
Decline to State	0.012	0.009	0.003	[0.632]	-	-	-	-
Highest Level of Education:								
< High School	0.009	0.011	-0.001	[0.778]	0.068	0.103	-	-
High School Graduate or GED	0.181	0.146	0.035	[0.065]	0.208	0.266	-	-
Some College	0.344	0.225	0.119	[<0.001]	0.309	0.277	-	-
Bachelor's Degree	0.291	0.350	-0.059	[0.014]	0.245	0.218	-	-
Postgraduate Degree	0.175	0.269	-0.093	[<0.001]	0.171	0.136	-	-
Annual Earnings:								
<10K	0.064	0.029	0.035	[0.001]	0.103	0.054	0.030	0.010
10-20K	0.084	0.034	0.049	[<0.001]	0.129	0.071	0.168	0.054
20-30K	0.134	0.056	0.078	[<0.001]	0.146	0.103	0.292	0.155
30-40K	0.140	0.090	0.050	[0.002]	0.139	0.119	0.166	0.180
40-50K	0.115	0.095	0.020	[0.208]	0.109	0.109	0.104	0.162
50-75K	0.208	0.179	0.029	[0.148]	0.186	0.215	0.140	0.263
75-100K	0.135	0.225	-0.090	[<0.001]	0.086	0.117	0.073	0.117
>100K	0.120	0.291	-0.171	[<0.001]	0.102	0.213	0.028	0.059
Industry:								
Agriculture/Forestry/Fishing/Hunting	0.014	0.026	-0.012	[0.096]	0.007	0.018	0.006	0.014
Arts/Entertainment/Accommodation/Food Services	0.064	0.061	0.003	[0.796]	0.075	0.065	0.065	0.033
Information/Finance/Real Estate/Professional Services	0.170	0.274	-0.104	[<0.001]	0.208	0.217	0.071	0.065
Health Care/Educational Services	0.297	0.131	0.166	[<0.001]	0.376	0.113	0.432	0.086
Manufacturing	0.063	0.136	-0.074	[<0.001]	0.068	0.145	0.069	0.147
Mining/Utilities/Construction	0.037	0.077	-0.040	[0.001]	0.023	0.153	0.018	0.198
Public Administration/Other Services	0.190	0.142	0.048	[0.012]	0.103	0.094	0.193	0.259
Wholesale Trade/Retail Trade/Transportation	0.165	0.152	0.013	[0.502]	0.140	0.196	0.146	0.199
Region:								
West	0.174	0.208	-0.034	[0.093]	0.232	0.246	-	-
Midwest	0.246	0.197	0.049	[0.022]	0.212	0.210	-	-
Northeast	0.212	0.217	-0.005	[0.806]	0.181	0.176	-	-
South	0.368	0.377	-0.010	[0.697]	0.375	0.369	-	-

Notes: Columns 1 and 2 of this table display respondent characteristics from the survey described in Section V. The survey was conducted by Qualtrics and included 1,519 adults ages 30 to 64. Eligibility was restricted to individuals who reported working at some point in the last 12 months and individuals who self-identified as either male (755 respondents) or female (764 respondents). For comparison, columns 3 through 6 of the table also display characteristics of workers ages 30 to 64 from the 2019 American Community Survey and characteristics of workers' compensation claimants in Texas ages 30 to 64 injured between 2013 and 2017. As described in Section I, the presence of wage and industry information is related to the receipt of cash benefits in the workers' compensation data. All wages are in 2020 dollars, and the American Community Survey numbers are weighted using IPUMS weights.

Table A9: Survey: Experiences with the Health Care System

	Female	Male	Difference	p-value	Regression - Female Coefficient Est	Std Error	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Thinking about your experiences with health care visits in the past, have you ever felt that a doctor did any of the following? - Share answered yes							
Talked down to you or didn't treat you with dignity or respect	0.393	0.340	0.052	[0.035]	0.087	(0.026)	[0.001]
Didn't understand or relate to your experiences and concerns	0.450	0.356	0.094	[<0.001]	0.104	(0.027)	[<0.001]
Didn't believe you were telling the truth about your symptoms or concerns	0.317	0.286	0.031	[0.193]	0.045	(0.025)	[0.068]
Refused to order a test or treatment you thought you needed	0.223	0.232	-0.009	[0.666]	0.022	(0.022)	[0.310]
Made you feel uncomfortable discussing your concerns	0.292	0.250	0.042	[0.069]	0.059	(0.024)	[0.014]
Assumed something about you without asking	0.357	0.328	0.029	[0.236]	0.050	(0.026)	[0.054]
Refused to prescribe pain medication you thought you needed	0.175	0.249	-0.074	[<0.001]	-0.055	(0.021)	[0.008]
Share answered yes to any of the above	0.596	0.538	0.058	[0.023]	0.103	(0.026)	[<0.001]
Thinking about your experiences getting health care for yourself, which doctor---male or female---would be more likely to [Answer options: male doctor, female doctor, male and female doctors are equally likely]							
Share selected doctor of own-gender							
treat you with dignity and respect?	0.343	0.193	0.150	[<0.001]	0.178	(0.024)	[<0.001]
understand or relate to your experiences and concerns?	0.517	0.295	0.222	[<0.001]	0.232	(0.027)	[<0.001]
believe you are telling the truth about your symptoms or concerns?	0.304	0.175	0.129	[<0.001]	0.159	(0.023)	[<0.001]
provide needed testing and treatments?	0.249	0.176	0.073	[0.001]	0.103	(0.022)	[<0.001]
make you feel comfortable with discussing your concerns?	0.416	0.240	0.176	[<0.001]	0.179	(0.026)	[<0.001]
ask appropriate questions instead of making assumptions?	0.326	0.177	0.148	[<0.001]	0.177	(0.023)	[<0.001]
be the most qualified?	0.161	0.160	0.001	[0.969]	0.016	(0.020)	[0.406]
be available near you?	0.154	0.184	-0.030	[0.124]	-0.007	(0.020)	[0.733]
Share did not select doctor of opposite gender							
treat you with dignity and respect?	0.931	0.820	0.111	[<0.001]	0.108	(0.018)	[<0.001]
understand or relate to your experiences and concerns?	0.932	0.826	0.105	[<0.001]	0.088	(0.018)	[<0.001]
believe you are telling the truth about your symptoms or concerns?	0.928	0.816	0.112	[<0.001]	0.114	(0.018)	[<0.001]
provide needed testing and treatments?	0.932	0.853	0.079	[<0.001]	0.079	(0.017)	[<0.001]
make you feel comfortable with discussing your concerns?	0.932	0.837	0.095	[<0.001]	0.090	(0.017)	[<0.001]
ask appropriate questions instead of making assumptions?	0.933	0.813	0.120	[<0.001]	0.120	(0.018)	[<0.001]
be the most qualified?	0.958	0.890	0.068	[<0.001]	0.053	(0.014)	[<0.001]
be available near you?	0.914	0.898	0.016	[0.298]	0.011	(0.016)	[0.501]

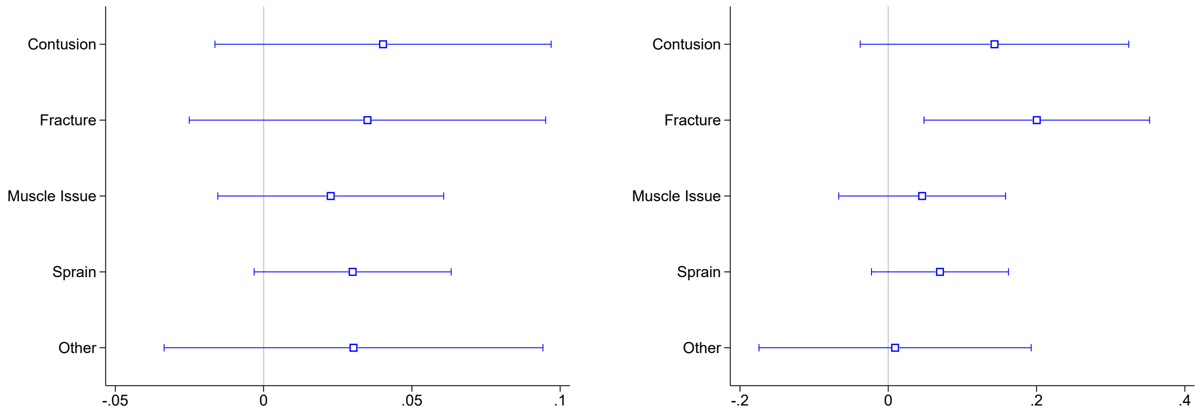
Notes: This table summarizes responses from several questions from the survey described in Section V. The table above reports the raw means for each of the indicated variables, along with the difference by respondent gender and the associated p-value testing whether the difference is zero. The table also includes the estimated coefficient on female (and associated standard error and p-value) from a regression of the indicated response on a female indicator and all the respondent characteristics reported in Appendix Table A8. The survey was conducted by Qualtrics and included 1,519 adults ages 30 to 64. Eligibility was restricted to individuals who reported working at some point in the last 12 months and individuals who self-identified as either male (755 respondents) or female (764 respondents).

Table A10: Survey: Preferences over Providers and Homophily

	Female	Male	Difference	p-value	Regression - Female Coefficient Est	Std Error	p-value
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Have you ever received care from a…. - Share answered yes							
Male doctor?	0.901	0.968	-0.068	[<0.001]	-0.062	(0.013)	[<0.001]
Female doctor?	0.941	0.767	0.174	[<0.001]	0.162	(0.019)	[<0.001]
Given the choice, would you prefer to see a doctor who is male or female, or does it not make much difference to you? - Share select each option below							
Prefer to see a doctor who is female	0.463	0.136	0.327	[<0.001]	0.315	(0.024)	[<0.001]
Prefer to see a doctor who is male	0.058	0.290	-0.232	[<0.001]	-0.207	(0.020)	[<0.001]
Doesn't make much difference	0.479	0.574	-0.094	[<0.001]	-0.108	(0.028)	[<0.001]
Choices based on initial hypothetical choice question: doctor gender and out-of-pocket cost - Share select own-gender doctor							
Own Gender $30 vs. Opp Gender $5	0.254	0.240	0.014	[0.748]	0.029	(0.048)	[0.544]
Own Gender $10 vs. Opp Gender $5	0.485	0.293	0.192	[<0.001]	0.200	(0.054)	[<0.001]
Own Gender $5 vs. Opp Gender $10	0.929	0.833	0.095	[0.005]	0.067	(0.036)	[0.064]
Own Gender $5 vs. Opp Gender $30	0.988	0.854	0.134	[<0.001]	0.129	(0.031)	[<0.001]
Choices based on full set of hypothetical choice questions: doctor gender and out-of-pocket cost - Share select own-gender doctor							
Own Gender $30 vs. Opp Gender $5	0.229	0.147	0.082	[<0.001]	0.091	(0.021)	[<0.001]
Own Gender $10 vs. Opp Gender $5	0.427	0.273	0.154	[<0.001]	0.159	(0.026)	[<0.001]
Own Gender $5 vs. Opp Gender $10	0.929	0.828	0.102	[<0.001]	0.088	(0.018)	[<0.001]
Own Gender $5 vs. Opp Gender $30	0.970	0.910	0.060	[<0.001]	0.050	(0.014)	[<0.001]
Please indicate how important each of the following characteristics is if you were choosing a doctor - Share indicating at least moderately important							
Out-of-pocket cost for a visit	0.829	0.824	0.005	[0.809]	0.005	(0.022)	[0.807]
Doctor reviews (e.g., on websites like Healthgrades or Google or from friends or family)	0.847	0.807	0.040	[0.038]	0.048	(0.022)	[0.027]
Travel time to get to doctor's office	0.831	0.834	-0.003	[0.864]	0.002	(0.021)	[0.914]
Wait time at the doctor's office	0.863	0.848	0.015	[0.410]	0.018	(0.020)	[0.361]
Doctor's sex	0.414	0.336	0.077	[0.002]	0.112	(0.026)	[<0.001]
Doctor's age	0.349	0.370	-0.020	[0.416]	0.021	(0.026)	[0.415]

Notes: This table summarizes responses from several questions from the survey described in Section V. The table above reports the raw means for each of the indicated variables, along with the difference by respondent gender and the associated p-value testing whether the difference is zero. The table also includes the estimated coefficient on female (and associated standard error and p-value) from a regression of the indicated response on a female indicator and all the respondent characteristics reported in Appendix Table A8. The survey was conducted by Qualtrics and included 1,519 adults ages 30 to 64. Eligibility was restricted to individuals who reported working at some point in the last 12 months and individuals who self-identified as either male (755 respondents) or female (764 respondents).

Figure A1: Heterogeneity: Estimate on Interaction of Female Doctor and Female Claimant by Injury Type

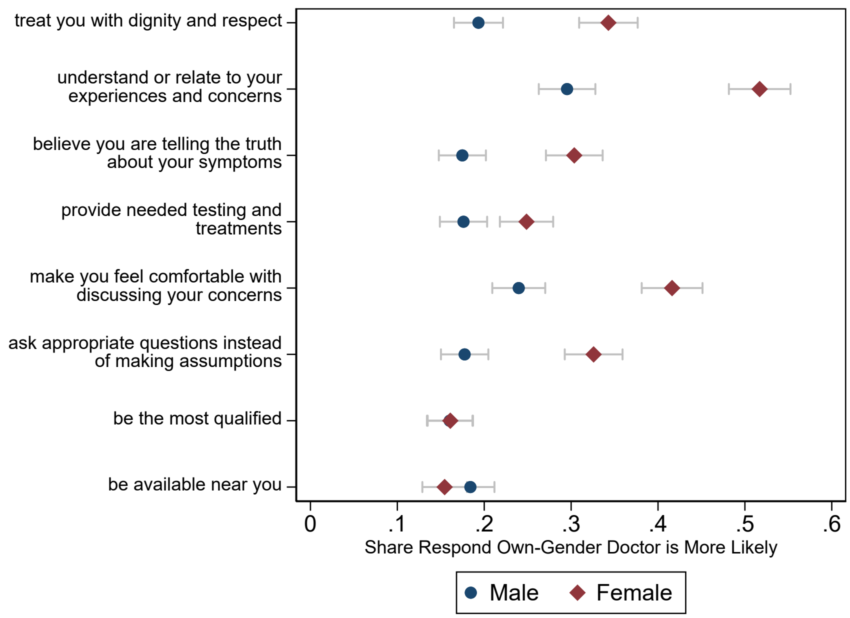


(a) OLS: I(Additional Cash Benefits>0)

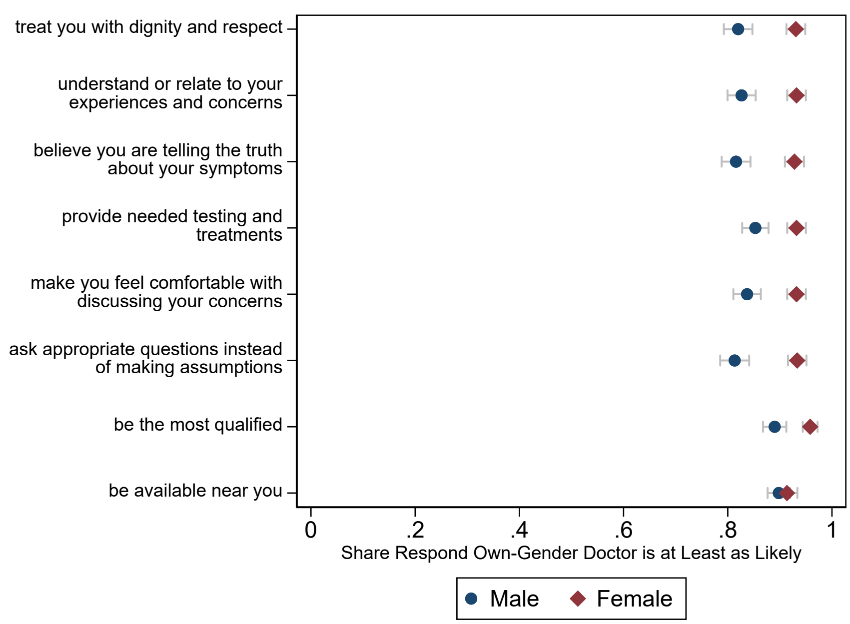
(b) Poisson Regression: Additional Cash Benefits

Notes: Each marker is the coefficient on the interaction of the female doctor and female claimant indicator variables from separate regressions of Equation (1) for the specified sample of claimants. All regressions control for a female doctor indicator variable, a female claimant indicator variable, credential-by-county fixed effects, exam year fixed effects, and injury year fixed effects. The dependent variables are as indicated in the figure: an indicator for receiving any additional cash benefits (Panel A) and (normalized) additional cash benefits received (Panel B). The 95% confidence intervals displayed along with the coefficient estimates are calculated using standard errors clustered at the doctor level. The sample includes claims occurring from 2013 to 2017 that had an independent medical exam by the end of 2017 and that have non-missing values for the specified characteristic.

Figure A2: Additional Results from Survey: "Thinking about your experiences getting health care for yourself, which doctor—male or female—would be more likely to..." [answer options: male doctor, female doctor, male and female doctors are equally likely]



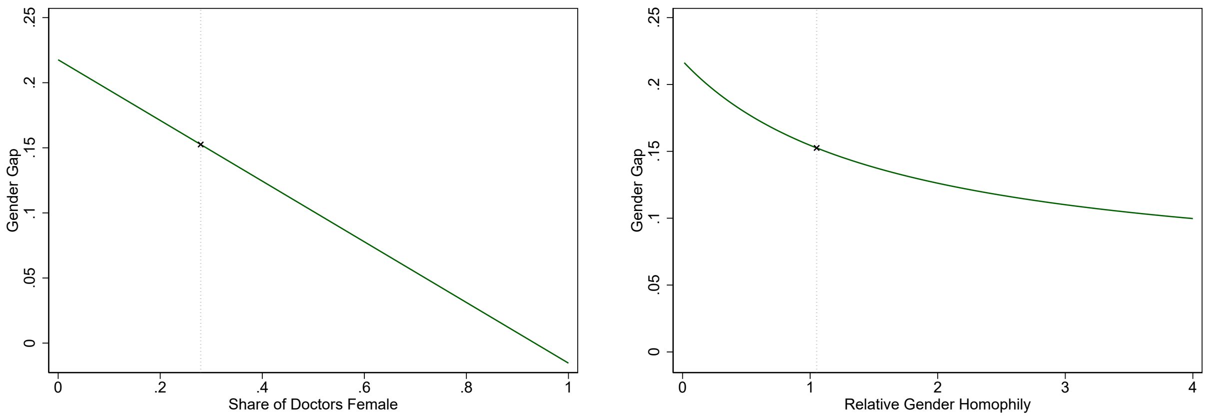
(a) Share that Select Own-Gender Doctor



(b) Share that Do Not Select Opposite-Gender Doctor

Notes: The figure above shows the means and associated 95% confidence intervals for survey responses to the indicated question by respondent gender. The survey was conducted by Qualtrics and included 1,519 adults ages 30 to 64. Eligibility was restricted to individuals who reported working at some point in the last 12 months and individuals who self-identified as either male (755 respondents) or female (764 respondents). See Section V for more detail on the survey. Table A9 reports raw and regression-adjusted differences in means for these survey questions and the associated p-values.

Figure A3: Counterfactual Policy Analysis: Partial Effects of Varying Share Doctors Female or Gender Homophily on Gender Gaps

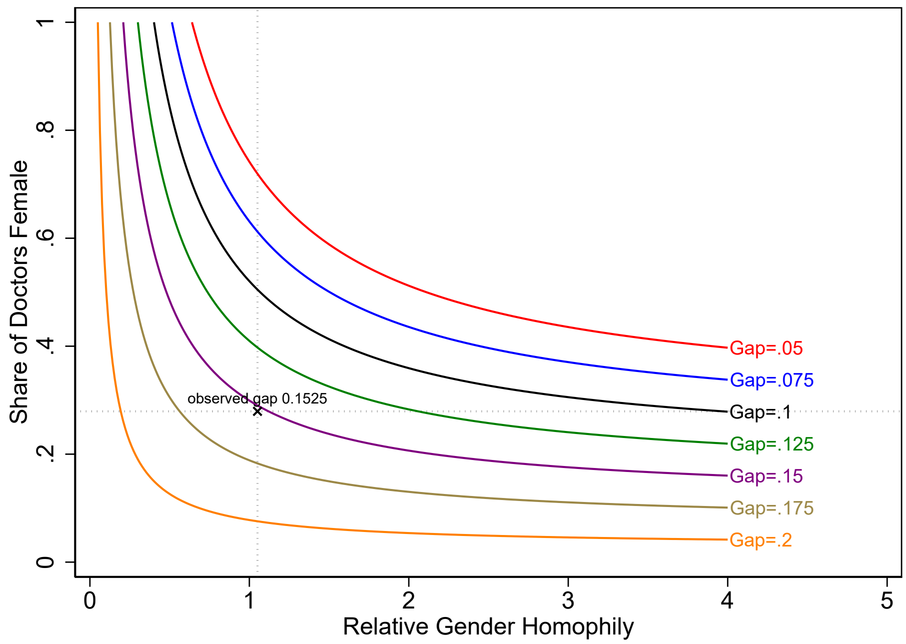


(a) Share of Doctors who are Female

(b) Relative Gender Homophily

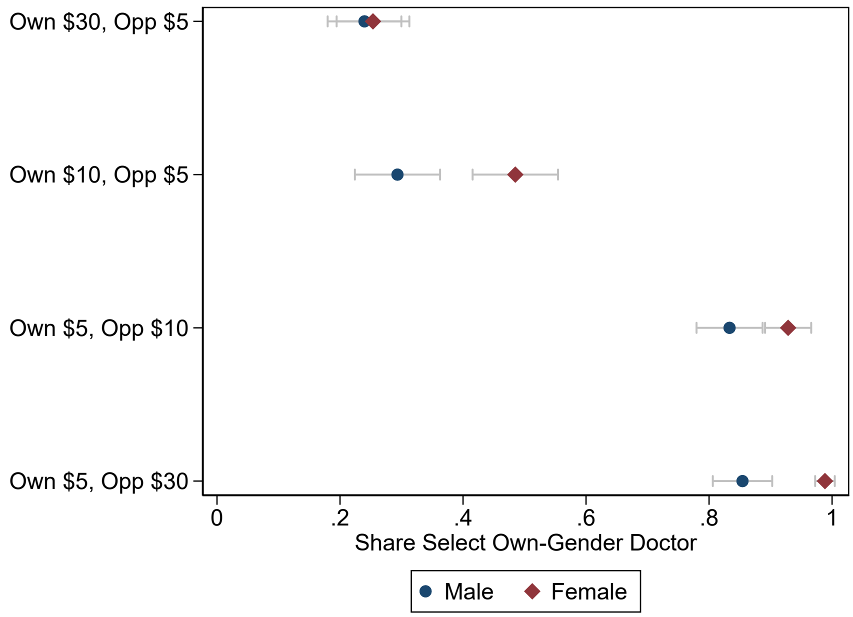
Notes: This figure displays the results from policy counterfactual analysis described in Section VI and Appendix Section D. This analysis combines the broader findings regarding the gender gap and relative gender homophily in workers' compensation more generally, with the estimated effects of gender match among claims with randomized evaluations. This analysis draws on Equation (1), which relates the gender gap (conditional on observable characteristics) to the share of doctors who are female and relative gender homophily in patient-doctor matches. For the purposes of this figure, the gender gap represents the percent reduction in the likelihood of cash benefits for female patients relative to male patients with the same observable characteristics. The point indicating the observed values in the status quo is indicated with an "x" in each panel, where the gender gap is 15.3% (i.e., females are 15.3% less likely to receive benefits than males with the same observables), the share of treating doctors who are female is 0.279, and the degree of relative gender homophily is 1.051 (i.e., female patients select female doctors at 1.051 times the rate that male patients select female doctors). Panel A displays the partial effects of varying the share of doctors who are female, holding relative gender homophily fixed. Panel B displays the partial effects of varying the degree of relative gender homophily, holding the share of doctors who are female fixed.

Figure A4: Counterfactual Policy Analysis: Effect of Share Doctors Female and Gender Homophily on Gender Gaps

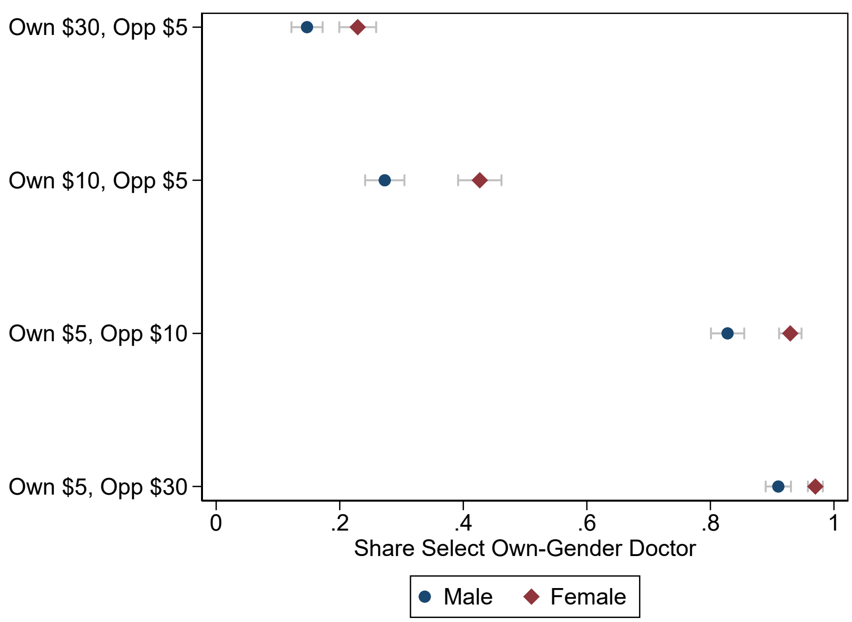


Notes: This figure displays the results from policy counterfactual analysis described in Section VI and Appendix Section D. This analysis combines the broader findings regarding the gender gap and relative gender homophily in workers' compensation more generally, with the estimated effects of gender match among claims with randomized evaluations. This figure shows the level curves of Equation (1), which relates the gender gap (conditional on observables) to the share of doctors who are female and relative gender homophily in patient-doctor matches. For the purposes of this figure, the gender gap represents the percent reduction in the likelihood of cash benefits for female patients relative to male patients with the same observable characteristics. The point indicating the observed values in the status quo is indicated with an "x", where the gender gap is 15.3% (i.e., females are 15.3% less likely to receive benefits than males with the same observables), the share of treating doctors who are female is 0.279, and the degree of relative gender homophily is 1.051 (i.e., female patients select female doctors at 1.051 times the rate that male patients select female doctors).

Figure A5: Additional Results from Survey: Share Selecting an Own-Gender Doctor in Hypothetical Choice Questions When Varying Co-Pay Differential



(a) Based on Initial Hypothetical Choice Question



(b) Based on Full Set of Hypothetical Choice Questions

Notes: The figure above shows the share selecting an own-gender doctor for each co-pay differential in hypothetical choice questions and the 95% confidence intervals by respondent gender. The survey was conducted by Qualtrics and included 1,519 adults ages 30 to 64. Eligibility was restricted to individuals who reported working at some point in the last 12 months and individuals who self-identified as either male (755 respondents) or female (764 respondents). See Section V for more detail on the survey. Table A10 reports raw and regression-adjusted differences in means for these survey questions and the associated p-values.