Worth the Wait? Willingness to Pay for Gender-Concordant Physicians Across Medical Specialties

Helen Kissel,* Tamri Matiashvili,† and Helena Roy‡§

September 11, 2026

Abstract

Patients increasingly select physicians through online "find a doctor" tools, which display information such as physician gender and patient ratings, rather than relying on physician referral, word-of-mouth, or soonest availability. Medical literature documents female patients' stated preferences for female physicians in certain specialties, but whether and how patients act on these preferences when given the tools to do so remains an open question. Using a randomized hypothetical choice experiment that mirrors the information presented in typical online physician-search platforms, we find that female patients have strong preferences for gender-concordant physicians. Women are willing to wait 1.6 additional weeks for a female urologist and 2.2 additional weeks for a female psychiatrist, two specialties with more gender-sensitive components, but show no statistically significant concordance preference in orthopedics, a more gender-neutral specialty. Men exhibit no significant concordance preference in any specialty. Among women, the gender concordance preference is stronger among those who report prior negative healthcare experiences due to gender. We complement these experimental findings with evidence from Medicare claims data on bladder cancer patients. We find that female patients are more likely than male patients to see female urologists, and women who match with a female urologist wait longer than women who match with a male urologist. Male patients exhibit no corresponding wait-time difference by physician gender. These findings suggest that patient preferences over physician characteristics, and patients' growing capacity to act on those preferences, mean that the availability of female healthcare providers may affect patterns of health-seeking behavior such as provider choice and timing of care.

*Kissel: Simon Fraser University – helen_kissel@sfu.ca

†Matiashvili: Stanford University — tamrim@stanford.edu

‡Roy: Hasso Plattner Institute – helena.roy@hpi.de

§We are grateful to Marcella Alsan, Liran Einav, Harsh Gupta, Petra Persson, Maria Polyakova, Maya Rossin-Slater, Paul Sniderman, Ariel Stern, Michael Tomz, and Heidi Williams for feedback and support. Participants of workshops at Stanford University and the CSWEP Health Economics Meetings provided helpful comments. This research was funded by grants from the Institute for Research in the Social Sciences (IRiSS) and the Stanford Institute for Economic Policy Research (SIEPR) at Stanford University. This study was approved by the Stanford University Institutional Review Board (Protocol IDs 84149 and 84962).

A note on language: In this paper, we follow the American Medical Association style guide when reporting race and ethnicity (capitalizing the names of all races and ethnicities).

1. Introduction

Physician demographic attributes can have profound effects on demand and take-up of medical care (Alsan, Garrick and Graziani, 2019). The medical literature provides ample examples of patients expressing a preference for a physician of the same gender, stretching across a range of specialties including gender-sensitive emergency department visits (Nolen et al., 2016), endoscopy procedures (Varadarajulu, Petruff and Ramsey, 2002), urology (McDevitt and Roberts, 2014), and gynecology (Reyes, 2008). These preferences may be consequential if the physician workforce is not representative of the population, and if patients have the ability to act on these preferences. Using national administrative data on roughly 1.19 million physicians, we document that patients' ability to match with a gender-concordant provider varies widely across both specialty and location. For example, 17% and 10% of Hospital Referral Regions have no female urologists and orthopedists, respectively.

The strength of the expressed gender concordance preferences and whether they translate into differences in care-seeking behavior and downstream health outcomes remains largely unknown. This is particularly true following the advent of online "find a doctor" tools, which have enabled patients to review demographic and other physician characteristics online, before scheduling an appointment. Whereas previously patients may have matched with a physician based on a referring physician's recommendation, scheduling considerations, or word-of-mouth, they are now able to review multiple profiles online, often seeing physician photographs and reading other patients' reviews. Thus, patients' ability to act on any preferences they may have has increased in recent years. If patients are willing to delay care to match with a concordant provider, these preferences could have meaningful implications for the timing of treatment and disease progression; something patients may be unable to account for, given the asymmetric information that characterizes healthcare (Arrow, 1963). These preferences may also vary systematically across medical specialties, particularly with the degree of gender sensitivity in care. Understanding whether and how these preferences are acted upon is important for assessing their consequences for both patient health and the allocation of the physician workforce across specialties and geographic areas, where the gender composition of providers may constrain patients' ability to match.

This paper addresses this knowledge gap by measuring the willingness-to-pay (WTP) for gender concordance preferences across different types of specialties in terms of appointment wait time using a hypothetical choice experiment, and by documenting how these preferences manifest in real-world patient behavior when access to concordant providers may be limited. The hypothetical choice experiment allows us to go one step beyond direct questions about preferences for same-sex physicians and to quantify these preferences in terms of an appointment attribute that can have important consequences for health: waiting longer to see a physician. The real-world administrative data allow us to confirm that these preferences are reflected in actual patient behavior and to measure their health implications.

To quantify how patients may alter their care-seeking behavior to satisfy their gender concordance preference, we use a randomized hypothetical choice experiment. Respondents on an online survey platform were randomized into one of three medical specialties: urology, psychiatry, or orthopedics. These specialties were chosen to reflect variation in the salience and potential importance of gender and other social factors, a feature we call gender sensitivity.

Urology is a gender-sensitive specialty in that it relates to an intimate physical area, and can involve physical examination (something that is made clear to respondents in the survey). Psychiatry could include gender-sensitive discussions, depending on patient history, and involves sensitive discussions that can make a patient feel emotionally vulnerable, but no physical exam. Orthopedics, by comparison, is relatively neutral in relating to non-gender-specific physical areas and involving less sensitive communication than psychiatry. Respondents were then asked to repeatedly choose between two physician appointments with randomized appointment attributes, including physician gender, wait time, travel time, cost, and years of physician experience. Respondents also saw patient star ratings (ratings out of five stars), which proxy for physician quality. Such ratings are now commonplace, and often displayed by healthcare systems, who collected reviews from verified patients. For example, Stanford Healthcare sends patients a link to a survey after their appointment, and displays the results and associated star ratings on physician profiles. Search engines such as Google will privilege these ratings and display them in search results, given that they are from verified patients, over public sites such as Yelp or Healthgrades.

To measure how patients trade off gender concordance against other attributes, we employ a mixed logit model estimated separately by respondent gender and specialty, allowing preference heterogeneity within each group by treating physician-appointment attribute coefficients as draws from a population distribution. For each group, we calculate willingness-to-pay (WTP) for gender concordance as the marginal rate of substitution between the gender concordance indicator and appointment cost attributes, specifically wait time (weeks), co-pay (dollars), and travel time (hours). We focus primarily on wait time as our numeraire, given its particular salience, the potential for differing understanding of its cost arising from the asymmetry of information in healthcare, and subsequent downstream health implications.

We find that a significant gender concordance premium exists for female physicians in some specialties, but by no means all. The strength of the premium varies according to the gender sensitivity of the specialty. Female respondents exhibit meaningful preferences for gender concordance in the gender-sensitive context of urology, where they are willing to wait up to two additional weeks to see a female physician in our hypothetical scenario. We find that women are, on average, willing to wait even longer for gender-concordant psychiatrists, but that in this specialty, variance is higher, a finding we decompose later. Women randomized into the orthopedics scenario and men in all scenarios exhibit little to no preference for gender concordance.

We also show that the strength of gender concordance preferences differs by respondents' reported prior experiences. Women who report having been treated negatively or unfairly due to gender in a healthcare setting have a higher willingness-to-pay for concordance than those who do not. However, even within the group with no such experiences, the willingness-to-pay for gender concordance in the gender-sensitive specialty, urology, is statistically significantly different than zero. The contrast is sharpest in psychiatry, where women who report prior negative experiences are willing to wait an average of 2.3 additional weeks for a female psychiatrist than those without such experiences, while those without such experiences exhibit no significant preference.

We present suggestive evidence on how perceived urgency moderates the willingness to delay care in pursuit of gender concordance. Among female respondents in the urology arm, those who perceived their symptoms as non-urgent were willing to wait approximately twice

as long for a gender-concordant physician than those who perceived their symptoms as urgent. Yet even the latter group's willingness-to-pay remained close to 1.6 weeks, suggesting that patients who understood that the presented symptoms required prompt attention were still willing to tolerate delay.

Finally, we examine whether the preferences observed in the experiment and heterogeneity in access to gender-concordant physicians translate into real-world care-seeking behavior. Using a 20% sample of Medicare claims data from 2006–2018, we study how gender concordance preferences affect patient-physician matching and wait times in the gender-sensitive specialty of urology. We focus on patients who ultimately receive a bladder cancer diagnosis, for whom delays in care may have particularly high stakes. Consistent with the survey findings, the Medicare appointments data show female patients are more likely to see female urologists than their male counterparts. As the supply of female urologists increases in a Hospital Referral Region (HRR), all patients are more likely to see female urologists, but the gradient is far steeper for women, suggesting their ability to act on gender-concordant preferences is constrained by the limited supply of female urologists. The claims data allow us to estimate appointment wait time by the time elapsed between the first urology visit before the bladder cancer diagnosis and the immediately preceding primary care physician visit. We find that female patients who see a female urologist wait about four days longer than those who see a male urologist, with no corresponding gap for male patients. These patterns suggest that preferences for gender-concordant physicians among female bladder cancer patients are leading to delays in access to care.

This paper contributes to several strands of literature. The most directly related one concerns the economics of gender-concordant matching in healthcare markets. Reyes (2006; 2008) shows that excess demand for female OB/GYNs clears through both monetary and non-monetary prices: female physicians in that specialty charge higher fees and face longer wait times, with the relative adjustment depending on the rigidity of price-setting institutions. McDevitt and Roberts (2014) study urology and find that markets without a female urologist have significantly higher bladder cancer mortality among female patients - a finding we expand on using both behavioral and granular appointment-level evidence. Walker et al. (2024) combine survey and field experiment data to show that female primary care physicians face excess demand, with female patients preferring them by over 50 percentage points, resulting in longer wait times among female primary care physicians. Lamiraud et al. (2026) find that in France, in areas with a relative scarcity of gender-concordant GPs of either sex, increasing the supply of gender-discordant physicians brings little benefit to the under-served gender. Our paper contributes to this literature by providing the first direct experimental estimates of willingness to pay for concordance across multiple specialties that vary in gender sensitivity, linking those estimates to real-world care-seeking patterns, describing a key dimension of variation in the preferences (gender sensitivity of specialty), and tracing the downstream consequences for diagnosis timing and severity.

A second strand examines how patient–physician demographic concordance affects health outcomes and the mechanisms behind it. Alsan, Garrick and Graziani (2019) use a randomized field experiment to show that Black men are more likely to accept preventive care when assigned a Black physician, with effects concentrated among patients with the highest initial distrust of medicine. Greenwood, Carnahan and Huang (2018) find higher in-hospital mortality among female heart attack patients treated by male physicians, Wallis et al. (2022) document

worse surgical outcomes for female patients with male surgeons, and Tsugawa et al. (2017) and Miyawaki et al. (2024) find lower mortality among elderly Medicare patients of female internists, with the benefit being largest for female patients. Our paper suggests that if patients actively seek concordant providers and are willing to wait to do so, then the health benefits of concordance documented in this literature will be unevenly distributed depending on local access to concordant physicians.

A third strand of literature examines how the proliferation of online physician-search platforms has fundamentally altered the market for physician selection (Chen and Lee, 2024; Chan, 2026).[1] Prior to the emergence of platforms such as Healthgrades, Zocdoc, and Yelp, patients relied on a narrow set of information channels, such as physician referrals, word of mouth, and advertising, that offered limited coverage and little ability to compare providers on observable attributes. Online platforms have since expanded both the choice set available to patients and the richness of information at their disposal, enabling comparisons across physician experience, credentials, patient ratings, and narrative reviews at low search cost. Some of these platforms allow users to preemptively filter on physician characteristics, such as gender or language. Therefore, while in the past, patients may have stated that they preferred a physician with certain demographic characteristics, they had limited ability to act on these preferences. Chan (2026) shows that when patients select physicians through online platforms, providing quality signals substantially reduces discrimination against Black and Asian doctors, suggesting that the informational design of these platforms, and particularly provision of a quality signal, affects how patients choose care. In this paper, we show how these online tools alter patient-physician matching by enabling patients to act directly on their preferences.

The remainder of the paper proceeds as follows. Section 2 documents variation in the gender composition of physicians across specialties and regions. Section 3 describes our hypothetical choice experiment design and empirical methodology. Section 4 describes the results of the experiment. Section 5 links our findings to real-world behavior using Medicare data. Section 6 concludes.

2. Background

In this section, we document two descriptive statistics that relate to patient preferences over physician characteristics, and how those might impact downstream health choices and outcomes. The first is substantial geographic heterogeneity in physician gender composition and in the availability of female physicians across specialties; for example, 17% of all US Hospital Referral Regions have no female urologists. When combined with gender concordance preferences, this heterogeneity could lead patients to wait longer to see a physician, particularly in areas with limited or no access to gender-concordant providers. Relatedly, we also document wait times across specialties and over time, to provide a real-world baseline against which to compare our willingness-to-pay estimates by wait time.

[1] Similar changes have occurred in the market for other healthcare providers, such as nursing homes (Werner et al., 2012) or hospitals (McCarthy, Sanbower and Sánchez Aragón, 2024).

2.1 Physician Sex by Specialty and Area

To better understand the accessibility of male and female physicians by area and clinical specialty, we make use of the publicly available National Plan and Provider Enumeration System (NPPES) data file for September 2025, obtained from the National Bureau of Economic Research (NBER). The NPPES is maintained by the Centers for Medicare and Medicaid Services (CMS) and contains records for every healthcare provider issued a National Provider Identifier (NPI) in the US, including their gender, practice location, and one or more Healthcare Provider Taxonomy codes that encode their type and specialty. We use the ZIP from practice location to infer Hospital Referral Region (HRR), regional markets defined by referral patterns for tertiary medical care, via a ZIP-to-HRR crosswalk published by the Dartmouth Atlas.

Starting from 9,129,558 total NPI records, we restrict the sample to individual providers (dropping organizational entities), remove deactivated NPIs, and then retain only records whose primary taxonomy code maps to an allopathic or osteopathic physician specialty, which excludes nurses, therapists, pharmacists, and other non-physician providers.[2] A further 330 records are dropped for missing gender, and 14,907 for practice ZIP codes that do not match the HRR crosswalk (primarily APO/FPO military addresses and US territories). The final analytic file contains 1,189,697 active physicians with known gender and HRR assignment. A summary of these sample restrictions can be seen in Table A1.

There is substantial variation in the share of female physicians across medical specialty, as shown in Figure 1. Relatively few specialties have a majority female workforce; only pediatrics, obstetrics and gynecology, and dermatology, and none are above 70% female. In other specialties, such as urology and orthopedics, less than 20% of physicians are female.

This creates variation in access to female and male physicians across the US. We map the share of physicians who are female across HRRs for four specialties in Figure 2.[3] For urology and orthopedics, the median share of female physicians by HRR is around one in ten; for obstetrics and gynecology and psychiatry, it is four to six in ten.

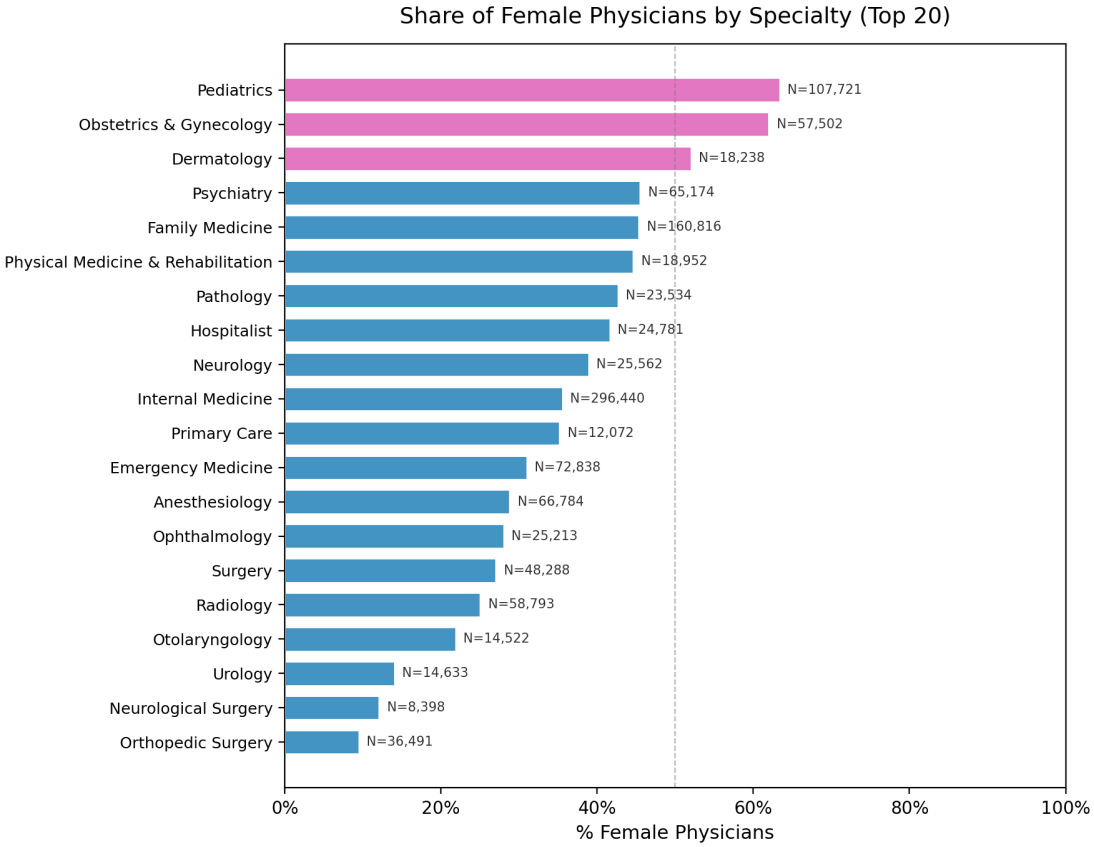
Specialties with a lower share of female physicians result in some HRRs having no female physicians for that specialty. Figure A1 demonstrates that is a relatively common occurrence in urology and orthopedics. Out of 306 HRRs, 53 (17.3%) have no female urologist, and 30 (9.8%) have no female orthopedist. There is also significant geographic dispersion of these HRRs across the US. Other specialties with no female physicians in some HRRs include neurological surgery (102 HRRs, 33.4%) and primary care (24 HRRs, 7.8%). The four largest

[2]Each NPI record in the NPPES file may list up to 15 Healthcare Provider Taxonomy codes, reflecting the range of services a provider is qualified to deliver. One of these codes can be designated as the provider's "primary" taxonomy via a flag field (pprimtax). We assign each physician a single specialty using the first taxonomy code flagged as primary (checking slots 1 through 3 in order); if no primary flag is set, we default to the first-listed taxonomy code. Among the 1,204,934 active physicians in our sample, 37.5% list two or more taxonomy codes. Of these physicians, 68.6% have their primary flag on the first taxonomy slot, 21.9% on the second, and 6.3% on the third. The remaining 3.2% (38,982 physicians) have no primary flag set on any of their first three codes and are assigned the specialty corresponding to their first-listed taxonomy code. We map each taxonomy code to a broad specialty category (e.g. "Internal Medicine", "Psychiatry", "Orthopedic Surgery") using the NUCC Healthcare Provider Taxonomy Code Set classification level, which groups subspecialties under their parent discipline. Each physician is counted exactly once under a single specialty; secondary taxonomy codes are not used.

[3]We show which HRRs lack access to any female physician (by the same specialties) in Figure A1.

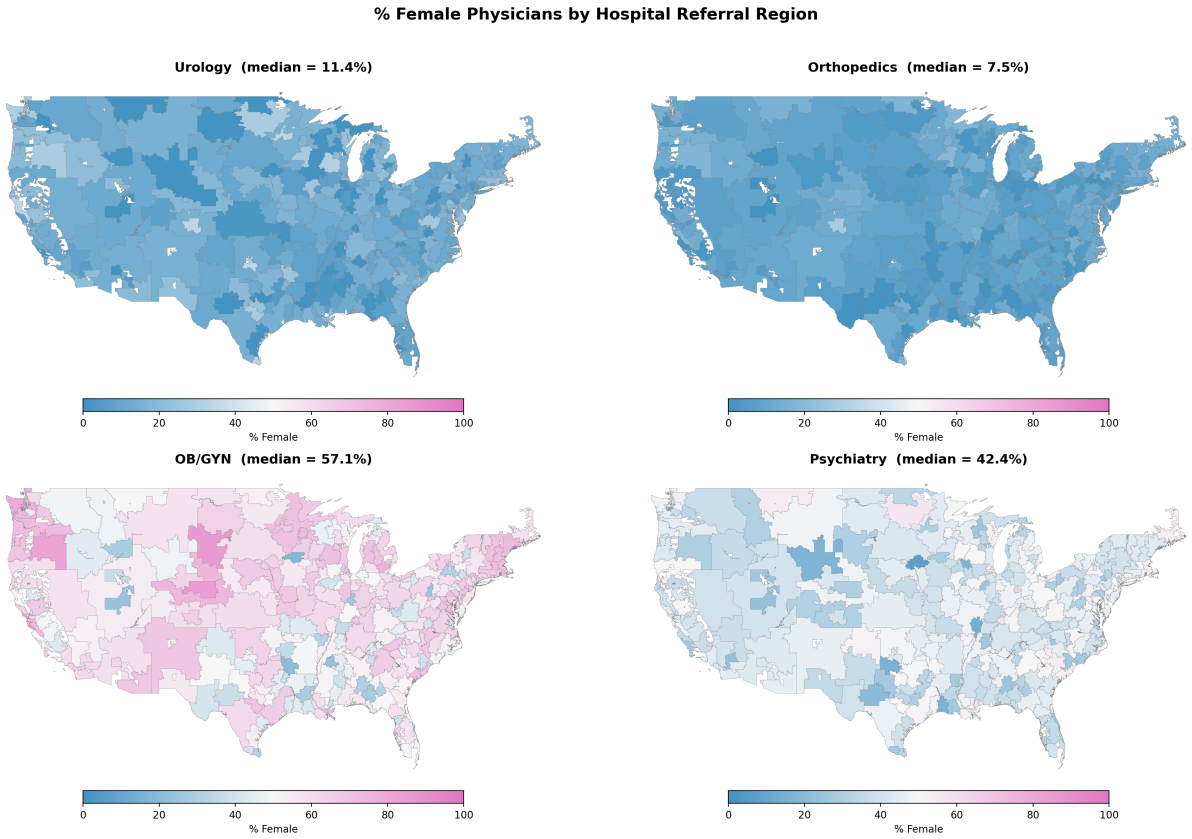
specialties (Internal Medicine, Family Medicine, Pediatrics, Emergency Medicine) all have at least one female physician in every HRR in the country.

Figure 1: Gender by Medical Specialty



Notes: Data from NPPES NPI file September 2025, downloaded from NBER. Analysis restricted to top twenty specialties by number of physicians. (Internal medicine includes gastroenterology, cardiology, nephrology, pulmonology, rheumatology, and other subspecialties.) Specialty inferred from provider taxonomy codes.

Figure 2: Share of Female Physicians by Hospital Referral Region and Specialty



Notes: Data from NPPES NPI file September 2025, downloaded from NBER. Analysis restricted to top twenty specialties by number of physicians. Specialty inferred from provider taxonomy codes. Practice location linked to HRR via ZIP.

2.2 Wait Time by Specialty

Average wait times for six major specialties across 15 major US metro areas stood at 31 days in 2025, up 48% since 2004 (AMN Healthcare, 2025). Wait times have increased in most specialties over this period. In obstetrics and gynecology, for example, average wait times grew from 23.3 days in 2004 to 41.8 days in 2025, an 80% increase (Figure A2). These baseline wait times ground our experimental estimates: for example, a female patient willing to wait an additional two weeks to see a female physician is willing to wait, on average, 50% longer than she otherwise would. This average also conceals substantial geographic variation. Across the 15 metro areas surveyed, average wait times ranged from 12 days in Atlanta to 65 days in Boston.

3. Experimental Design & Methodology

In this section, we outline our experimental design and associated methodology. Our experimental design is, in essence, a hypothetical choice (or conjoint) experiment, a method of

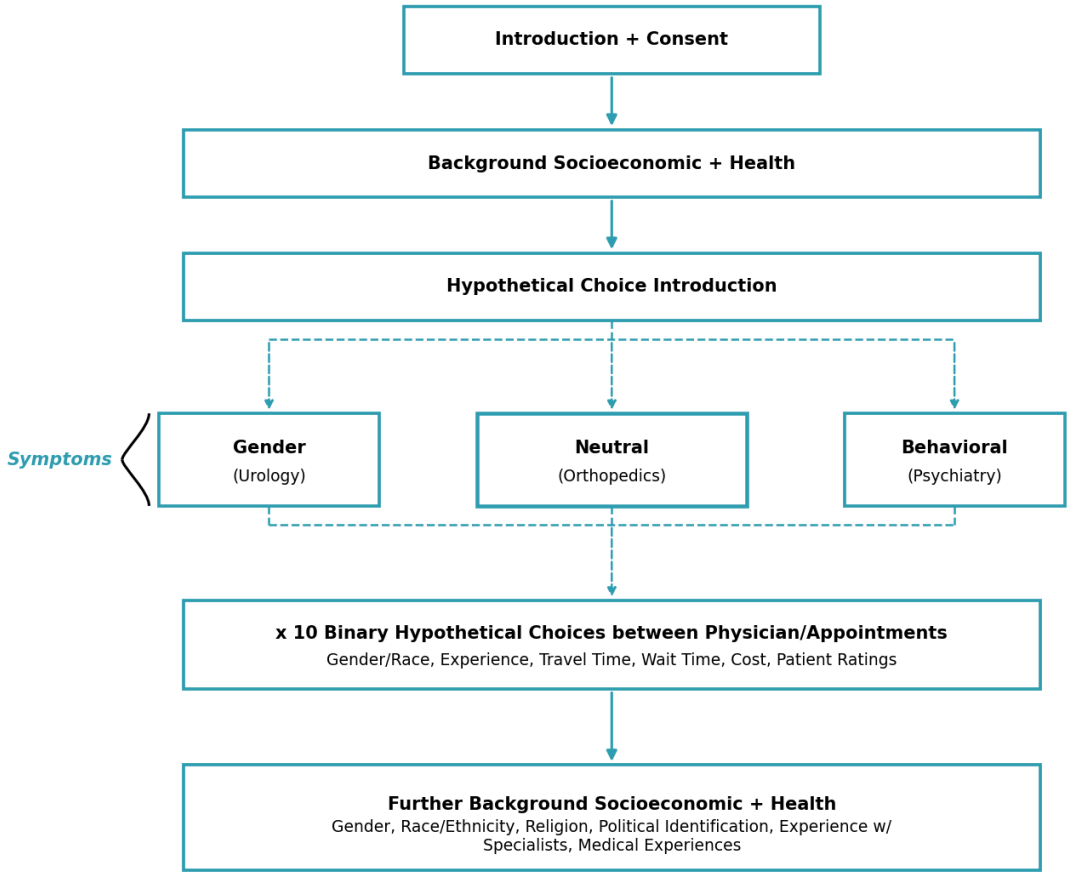
growing use in economics (Wiswall and Zafar, 2018; Gallen and Wasserman, 2023; Andrew and Adams, 2025; Chan, 2026). Though this tool relies on responses collected outside of a real healthcare context, it provides numerous advantages over observational data. First, by designing the choice environment within the survey, we have control over the respondent's information and choice sets – two key objects usually unobservable in observational healthcare data. Second, this design control means that respondents can be presented with gender-concordant physicians as options, something that does not exist in many healthcare markets, as documented above. Finally, the tool affords us the opportunity to examine the drivers of preferences (such as perceived urgency or prior experience in healthcare) in more detail than would be possible in an active healthcare environment.

The experimental design uses randomized variation in physician attributes to identify patient willingness-to-pay in appointment wait time for gender concordance across specialties. The data combine a large sample of participant choices between physician-appointments across three different medical specialties, enabling us to estimate willingness-to-pay for gender concordance via a discrete choice model.

3.1 Experimental Design

We begin with a survey intended to estimate willingness-to-pay for gender concordance, relative to other physician-appointment attributes. A diagram of the survey flow is available in Figure 3.

Figure 3: Willingness-to-Pay Survey Flow



Respondents enter the survey, consent, and pass an instructional manipulation check to assess attentiveness. We then ask a range of background socioeconomic questions (e.g. age, household income) and health questions (e.g. chronic conditions, prescriptions) to gauge their experience interacting with the healthcare system.

Respondents are then randomized into referral scenarios spanning more- to less-gender-sensitive specialties and choose between appointments with randomized characteristics. We explain that we are going to present them with a scenario in which they have developed some symptoms and are seeking medical advice. We tell them that we will then ask them to choose between potential physicians and appointments. We outline three to four symptoms and provide a time horizon over which the symptoms have been occurring. Then, we explain that in the scenario the patient's primary care physician has referred them to see a specialist. We explain what the specialist's role is and what will take place at the appointment. We provide more detail on the information the respondents are offered in Appendix B.

The hypothetical choices are then made. We present respondents with ten choices, each between two physician-appointments. The respondents are able to choose neither appointment, which serves as our outside option. We finish by eliciting further background and health information. Critically, we only ask for respondent gender *after* the hypothetical choices have

been made, to avoid making this characteristic more salient to respondents. We also ask explicitly if respondents have ever had a negative medical experience, and if they perceive that as being due to their gender. Respondents are also offered the chance to give any feedback on the survey, and we ask for their guess as to what the purpose of the survey was.

Our experiment has three key aspects of variation to help us estimate a range of relevant willingness-to-pay metrics.

(1) Variation in medical setting – Respondents are randomly assigned to one of three medical areas with associated symptoms for an underlying condition: urology, orthopedics, or psychiatry. We chose to use three different areas to vary contextual factors that might be relevant to preferences for physician demographics, such as gender, and to thereby maximize the external validity of our results. In particular, the medical areas we chose vary substantially on three important characteristics: the influence of sensitive gender norms; the influence of sensitive social/behavioral norms; and the share of female specialists.

(2) Variation in physician gender – We vary physician gender randomly for each physician profile (20 total, over 10 choices). We decided the probability of each gender with two goals: first, approximate the share of female physicians in the true distribution and so have a plausible share of both genders; second, generate a sufficient number of choices where both options were available. We chose a 60% chance of a male physician and a 40% chance of a female physician. This is close to the true distribution over all specialties, while also allowing for a higher number of observed choices between physicians of different genders.

(3) Variation in numéraires – Patients can pay for concordance via various numéraires: money, travel time, experience, patient ratings, or wait time. We included all of these in each physician profile, but we will focus on wait time as our primary numéraire.

Designing the hypothetical choice Throughout the design of the survey experiment, our goal has been to mimic the real-world process of choosing a physician as closely as possible. We reviewed online search tools for physicians – both independent and those built internally to several major healthcare systems (academic and non-academic). The information offered in these search tools heavily influenced the information we chose to include (and exclude) in our hypothetical choice. Generally, we found that most healthcare systems presented a physician's photo and name, as well as their specialty, clinic/location, and often a patient star rating. Age could often be roughly inferred from photos, and often more detail on experience and education was available by clicking on physician profiles (but not immediately visible in a side-by-side comparison on physicians). Some search tools presented appointment availability and distance from the user's location (if the user entered e.g. a ZIP code). Others allowed users to filter results by physician gender, giving preferences for gender concordance an opportunity to impact not only the choice of physician, but the choice set overall. We provide screenshots of some of the online search tools we looked at in Appendix B.3.

Influenced by these examples, we chose photos and names as our preferred method of conveying gender.[4] This has the benefit of providing greater realism to physician search online, as well as avoiding highlighting gender in the physician choice, potentially creating

[4]We use stock images that return for the search term "physician" or "doctor". Many of these images are from photoshoots designed to generate stock images of people as physicians, and we select images such that there is greater uniformity in photo quality and composition.

experimental demand effects. Following the experimental literature, we chose to blur these photos (Fong and Luttmer, 2009; Chan, 2026). Doing so largely preserves the ability to infer gender, but creates enough noise in the image that other hard-to-observe factors influencing physician choice (such as "attractiveness", professionalism, or just the quality of the photo) are made somewhat uniform across images.

After the information treatments, we gave respondents a preview of the hypothetical choice. We told respondents the exact number of choices we would ask them to make and emphasized that they should choose as they would for themselves. We then outlined the characteristics of the physician-appointments that they would be shown, and explained each one. Critically, we told respondents to assume that the physicians and appointments were otherwise identical in other aspects; implicitly, choices should be made based *only* on variation in the characteristics we outline. Respondents then made ten choices across physician-appointments. Figure 4 provides an example of the type of choice they were offered within the survey environment.

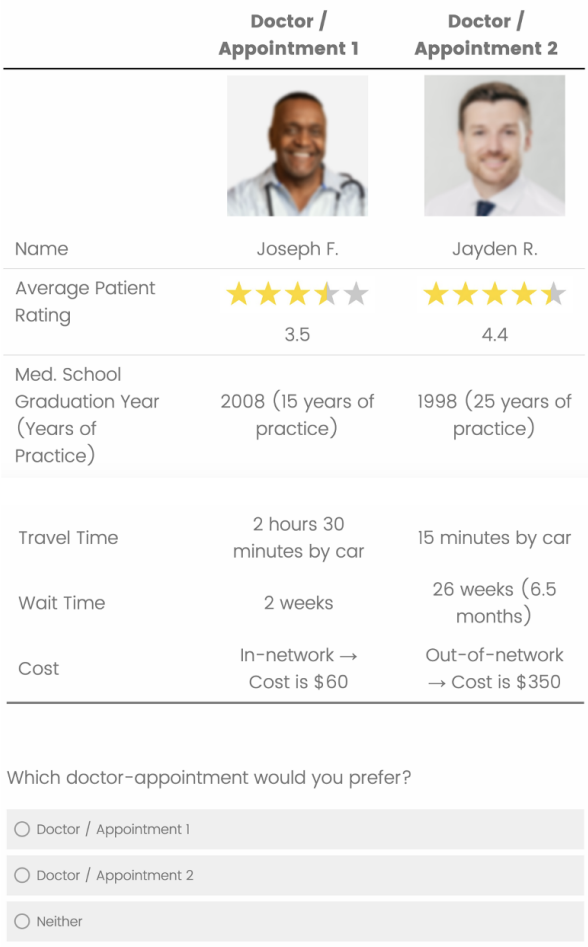


Figure 4: Example of Hypothetical Choice Between Physician-Appointments

We surveyed 2,800 respondents. In our analysis, we include all respondents who passed the initial attention check of the survey, were not flagged as potentially using LLMs or bots

by the survey platform, and did not complete the survey exceptionally fast, as defined by the survey platform. Demographic characteristics by treatment groups are presented in Appendix Table C2. The three groups are balanced across different demographic characteristics, such as age, gender, race, education, marital status, employment, and health status, among others.

3.2 Methodology

We estimate how respondents value physician-patient gender concordance using a discrete choice framework. Let i index respondents, j index physician-appointment options within choice task t, and $g \in \{\text{men}, \text{women}\}$ denote the respondent gender group. We estimate models separately for each group g and specialty, pooling choices across respondents within each group.

The utility that respondent i in group g derives from choosing appointment option j in task t is given by:

$$U_{ijt}^{(g)} = X_{ijt}' \beta_i^{(g)} + \epsilon_{ijt} \tag{1}$$

where X_{ijt} is a vector of observed characteristics of the physician and appointment, including an indicator for the outside option, $\beta_i^{(g)}$ is respondent i's vector of marginal utilities over these characteristics, and ϵ_{ijt} is an i.i.d. Type I Extreme Value error term denoting choice-specific utility shocks. The outside option (choosing neither physician) enters the model as a binary variable in X_{ijt} rather than as a separate intercept, and the model is estimated without a constant.

The vector X_{ijt} includes: cost (co-pay, in dollars), wait time (in weeks), travel time (in hours), physician experience (in years), patient ratings (in stars), an indicator for the outside option, and $\text{GenderConcordant}_{ijt} = 1$ if the physician and respondent share the same gender.

Our estimation strategy is a mixed logit model, which allows preference heterogeneity within each group g by treating a subset of coefficients as draws from a population distribution rather than as fixed group-level parameters. Specifically, we assume:

$$\beta_i^{(g)} = \mu^{(g)} + \eta_i^{(g)}, \qquad \eta_i^{(g)} \sim \mathcal{N}\left(0,\, \Sigma^{(g)}\right) \tag{2}$$

where $\mu^{(g)}$ is the vector of population mean utilities for group g and $\Sigma^{(g)}$ is a diagonal covariance matrix governing within-group preference dispersion. Following Chan (2026), we treat all preference-relevant attribute coefficients as random: particularly, we allow the coefficients on gender concordance, wait time, cost, travel time, and patient ratings to be random. Preference heterogeneity is most plausible and most consequential for these attributes: for example, how much patients value gender concordance varies with their reported prior experiences and perceived urgency, as we document below. The coefficients on physician experience and the outside option are held fixed.[5] The choice probability

[5]Fixing the outside option coefficient is standard practice, as it is a binary indicator for opting out rather than a continuous attribute with meaningful individual-level preference variation. Physician experience is fixed for parsimony; results are robust to allowing it to be random.

for respondent i in group g is then the integral of the standard logit probability over the distribution of $\beta_i^{(g)}$:

$$P_{ijt}^{(g)} = \int \frac{\exp(X_{ijt}'\beta_i^{(g)})}{\sum_{k\in\mathcal{C}_{it}} \exp(X_{ikt}'\beta_i^{(g)})} f\left(\beta_i^{(g)} \mid \mu^{(g)}, \Sigma^{(g)}\right) d\beta_i^{(g)} \tag{3}$$

This integral is approximated by simulation using 500 Halton draws (Train, 2009). We estimate the model by maximum simulated likelihood, maximizing:

$$\mathcal{L}^{(g)} = \sum_i \sum_{t=1}^{10} \sum_{j\in\mathcal{C}_{it}} \mathbb{1}(Y_{it} = j) \cdot \log \hat{P}_{ijt}^{(g)} \tag{4}$$

where $\hat{P}_{ijt}^{(g)}$ is the simulated choice probability. The population mean coefficients $\mu^{(g)}$ are interpreted as average marginal utilities of appointment attributes within group g. The estimated standard deviations $\sqrt{\Sigma_{kk}^{(g)}}$ provide a direct measure of within-group preference heterogeneity for each attribute k: a large estimated standard deviation on the gender concordance coefficient, relative to its mean, indicates that concordance preferences vary substantially across respondents within a group, over and above the heterogeneity captured by our observed subgroup splits.

We calculate willingness-to-pay (WTP) for gender concordance as the marginal rate of substitution between the gender concordance indicator and each of three cost-of-access numéraires: wait time (weeks), co-pay (dollars), and travel time (hours):

$$\text{WTP}_{\text{Gender, Wait}}^{(g)} = -\frac{\mu_{\text{GenderConcordant}}^{(g)}}{\mu_{\text{WaitTime}}^{(g)}} \tag{5}$$

$$\text{WTP}_{\text{Gender, Cost}}^{(g)} = -\frac{\mu_{\text{GenderConcordant}}^{(g)}}{\mu_{\text{CoPay}}^{(g)}} \tag{6}$$

$$\text{WTP}_{\text{Gender, Travel}}^{(g)} = -\frac{\mu_{\text{GenderConcordant}}^{(g)}}{\mu_{\text{TravelTime}}^{(g)}} \tag{7}$$

WTP is computed as the ratio of population mean coefficients, which is a consistent estimator of mean WTP in the population.[6] Standard errors are obtained via the delta method. These estimates represent the average amount respondents in group g are willing to forgo — in weeks of wait, dollars, or hours of travel — to see a gender-concordant physician relative to a non-concordant one, holding all other attributes constant. Throughout, we focus on WTP expressed in wait time (weeks), as wait time is a salient and clinically meaningful measure of access costs, though results expressed in cost and travel time are reported alongside for completeness.

[6]The ratio of mean coefficients is not identical to the mean of the individual-level WTP ratios $-\beta_{i,\text{GenderConcordant}}^{(g)}/\beta_{i,\text{WaitTime}}^{(g)}$, which can have heavy tails or undefined moments when both coefficients are normally distributed (Train, 2009). The ratio of means is the standard estimator in applied work and is consistent for mean WTP under mild regularity conditions.

4. Experimental Results

We present the experimental results in four parts. We first examine stated preferences for same-gender physicians, which broadly confirm that women report stronger concordance preferences than men. We then estimate WTP for gender concordance from the hypothetical choice experiment, finding that women are willing to wait longer to see a female physician in more gender-sensitive specialties, while men exhibit little to no concordance preference across specialties. We show that these preferences are stronger among women who report prior negative healthcare experiences due to gender, and that perceived urgency attenuates but does not eliminate concordance preferences.

4.1 Stated Preferences

Before turning to the results of the hypothetical choice experiment, we first examine whether respondents report preferences for same-gender physicians, and whether these differ by respondent gender. These questions were asked after the choice tasks had concluded, so they do not contaminate the experiment itself, though responses may be influenced by the experiment, which could have elicited these preferences.

We find that 34.3% of respondents report a preference for same-gender physicians, compared to only 6.5% who report a preference for different-gender physicians. There are pronounced gender differences, which we outline in Table 1. While 51.2% of women report preferring a same-gender physician, only 17.7% of men do, a difference of 33.4 percentage points that is statistically significant. At the same time, only 2.9% of women report preferring different-gender physicians, compared to 10.1% of men, suggesting that female physicians may be preferred overall.

Table 1: Stated preferences and beliefs by gender ($N = 2793$)

	Male ($n = 1392$)		Female ($n = 1368$)		Diff.	
	n	%	n	%	(pp)	
Gender concordance preference						
Prefer own gender	247	17.7	700	51.2	+33.4	***
Prefer other gender	140	10.1	39	2.9	-7.2	***
No preference	1005	72.2	629	46.0	-26.2	***
Negative experience (gender)						
Yes	101	7.3	413	30.2	+22.9	***
No	1291	92.7	955	69.8	-22.9	***
Trust in medical professionals						
Most can be trusted	1275	91.6	1201	87.8	-3.8	**
Most cannot be trusted	117	8.4	165	12.1	+3.7	**

Two-proportion z-test: $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$

These questions also reveal substantial differences in trust in medicine and reported gender-

related experiences. Women are substantially more likely to report negative experiences in healthcare due to gender (30.2% versus 7.3%, a 22.9 percentage point difference) and report lower trust in medical professionals than men (87.8% versus 91.6%, a 3.8 percentage point gap).

The overall takeaway from this end-of-survey questionnaire is that, if patient behavior maps to stated preferences, we would expect, on average, weak to no concordance preferences among men, and substantial preferences for concordance among women.

4.2 Hypothetical Choice Experiment

We next analyze the results from the hypothetical choice experiment. Table 2 presents willingness-to-pay (WTP) estimates across different denominators (wait time, cost, travel time) and for different appointment attributes (wait time, cost, travel time, experience, patient ratings, gender concordance). We find, as expected, that respondents value lower wait time, lower cost, lower travel time, more experience, and higher patient ratings.

A striking pattern emerges when examining how the three access-cost attributes — wait time, cost, and travel time — are valued relative to one another. The dollar value of travel time is essentially constant across specialties and sexes, hovering around \$71–\$91 per hour (Panel B), and symmetrically, the travel-time cost of a dollar is equally uniform at roughly 0.011–0.014 hours per dollar (Panel C). This consistency is intuitive, as the opportunity cost of time spent traveling reflects stable individual time preferences that do not vary with the clinical reason for the visit.

Wait time, by contrast, is priced very differently depending on specialty and sex. In dollar terms, men in urology are willing to pay \$29 to avoid a week's wait, compared to \$12 in psychiatry — and the same divergence appears when wait time is expressed in travel-time terms (Panel C), ranging from 0.35 hours per week for male urology patients to 0.1 hours for female psychiatry patients. This heterogeneity is consistent with wait time carrying specialty-specific costs beyond the pure opportunity cost of time: perceived urgency and consequences of diagnostic delay plausibly vary across clinical settings in ways that the cost of a car ride does not.

An extra star in patient ratings generates substantial willingness-to-pay across groups, but its valuation varies in interesting ways across numéraires and specialties. Men in the urology arm are willing to pay \$209 for an extra star in patient ratings, but would wait 7.3 additional weeks for it — whereas in psychiatry, men would wait 12.2 weeks for the same improvement in ratings but pay only \$146. This pattern suggests that the implicit exchange rate between dollars and wait time itself varies across specialties, consistent with wait time carrying greater urgency costs in different specialties and presented conditions.

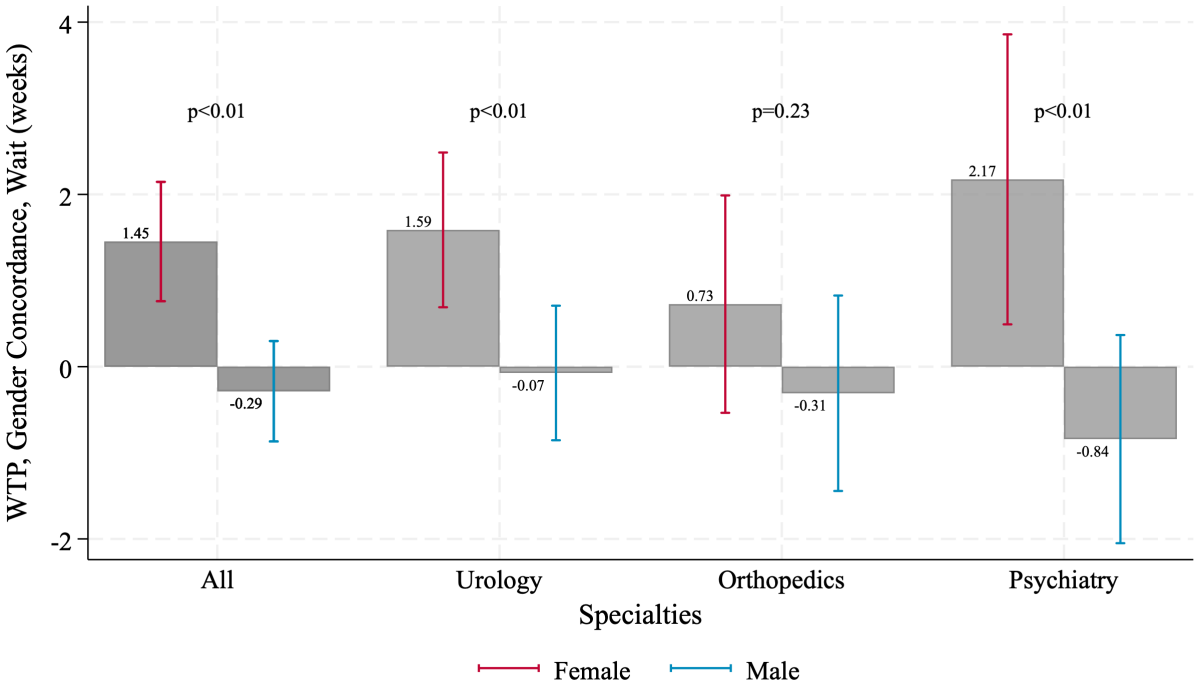
Next we turn to our primary outcome: willingness-to-pay (WTP) for gender concordance. As discussed above, we focus on wait time as our primary WTP measure, given its particular relevance for downstream health outcomes. For better exposition, Figure 5 plots these WTPs by gender and specialty. Pooling the three specialties, we find that women are willing to wait an extra 1.4 weeks to see a female physician, and this estimate is statistically significantly different from zero and the male estimate, which is very close to zero.

Table 2: Willingness-to-Pay for Different Attributes by Specialty, Gender, and Numéraire

	Urology		Orthopedics		Psychiatry	
	Men	Women	Men	Women	Men	Women
Panel A: WTP in Wait Time (Weeks)						
Cost	-0.035	-0.049	-0.076	-0.081	-0.084	-0.115
	(0.003)	(0.003)	(0.005)	(0.005)	(0.006)	(0.009)
Travel Time	-2.879	-3.976	-5.387	-7.357	-6.748	-9.694
	(0.238)	(0.290)	(0.429)	(0.532)	(0.537)	(0.818)
Experience	0.060	0.026	0.106	0.112	0.174	0.166
	(0.033)	(0.034)	(0.047)	(0.048)	(0.050)	(0.065)
Patient Ratings	7.304	6.080	12.687	12.101	12.225	14.308
	(0.501)	(0.479)	(0.819)	(0.770)	(0.843)	(1.167)
Gender Concordant	-0.074	1.588	-0.309	0.726	-0.841	2.175
	(0.399)	(0.458)	(0.579)	(0.644)	(0.616)	(0.859)
Panel B: WTP in Cost (Dollars)						
Wait Time	-28.648	-20.329	-13.223	-12.348	-11.952	-8.687
	(2.314)	(1.326)	(0.914)	(0.786)	(0.836)	(0.672)
Travel Time	-82.464	-80.831	-71.224	-90.842	-80.655	-84.207
	(8.081)	(6.067)	(5.631)	(6.308)	(6.354)	(6.078)
Experience	1.729	0.523	1.396	1.384	2.077	1.441
	(0.939)	(0.682)	(0.624)	(0.594)	(0.595)	(0.563)
Patient Ratings	209.242	123.595	167.756	149.422	146.115	124.294
	(18.314)	(10.098)	(10.761)	(9.062)	(9.855)	(8.368)
Gender Concordant	-2.109	32.281	-4.086	8.961	-10.056	18.890
	(11.419)	(9.366)	(7.651)	(7.953)	(7.363)	(7.425)
Panel C: WTP in Travel Time (Hours)						
Cost	-0.012	-0.012	-0.014	-0.011	-0.012	-0.012
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Wait Time	-0.347	-0.251	-0.186	-0.136	-0.148	-0.103
	(0.029)	(0.018)	(0.015)	(0.010)	(0.012)	(0.009)
Experience	0.021	0.006	0.020	0.015	0.026	0.017
	(0.011)	(0.008)	(0.009)	(0.007)	(0.007)	(0.007)
Patient Ratings	2.537	1.529	2.355	1.645	1.812	1.476
	(0.230)	(0.132)	(0.181)	(0.115)	(0.143)	(0.110)
Gender Concordant	-0.026	0.399	-0.057	0.099	-0.125	0.224
	(0.139)	(0.117)	(0.107)	(0.088)	(0.092)	(0.088)
Panel D: SD of Random Coefficients						
SD(Gender Concordant)	0.116	0.621	0.077	0.565	0.136	0.492
SD(Wait Time)	0.118	0.115	0.080	0.074	0.086	0.076
SD(Cost)	0.006	0.005	0.006	0.006	0.007	0.007
SD(Travel Time)	0.385	0.422	0.491	0.674	0.764	0.685
SD(Patient Ratings)	0.926	0.853	0.639	0.618	0.727	0.708
N	14190	14100	13890	14160	14100	13350

Notes: Willingness-to-pay (WTP) estimates are derived from a mixed logit model estimated separately by respondent sex and medical specialty (Urology, Orthopedics, and Psychiatry). The coefficients on gender concordance, wait time, cost, travel time, and patient ratings are treated as random, drawn from a normal distribution; the coefficients on physician experience and the outside option are fixed. The model is estimated by maximum simulated likelihood using 500 Halton draws. Respondents made 10 discrete choices each between two physician profiles varying in cost, wait time, travel time, physician experience, patient ratings, and gender concordance. Panel A reports WTP expressed in weeks of wait time, Panel B in dollars of out-of-pocket cost, and Panel C in hours of travel time. Panel D reports the estimated standard deviations of the random coefficients, which capture within-group preference heterogeneity. In each panel, WTP for a given attribute is computed as the negative ratio of that attribute's mean coefficient to the mean numeraire coefficient, with standard errors obtained via the delta method.

Figure 5: WTP in Wait Time for Gender Concordance



Notes: N = 2,793. Sample construction and estimation same as in Table 2. Bars represent point estimates and capped lines represent 95% confidence intervals.

We next analyze the results for the gender-sensitive specialty: urology. We find that women are willing to wait, on average, 1.6 extra weeks to see a female urologist, and this difference is statistically significant and statistically significantly different from the male estimate. We find that men do not exhibit a preference for urologists of either gender.

A different pattern emerges in orthopedics, a specialty we chose for its relative neutrality on social and gender norms. On average, women are willing to wait 0.7 additional weeks to see a female orthopedist, but this estimate is not statistically significant and not different from a male estimate, which is also close to zero.

The female gender concordance preference is strongest in psychiatry, a specialty we chose for its sensitivity to social norms (including gender), but lacking physical examination. Women are willing to wait an average of 2.17 additional weeks to see a gender-concordant specialist and this estimate is statistically significantly different from zero. Notably, the variance of this estimate is much higher than that of urology. Men exhibit no statistically significant preference, and, if anything, the coefficient is negative. The difference between men and women for *concordant* physicians is statistically significant, but is not if we instead evaluate it through the lens of preference for *female* physicians.

Beyond mean WTP estimates, our mixed logit framework also identifies the degree of preference heterogeneity within each group. Panel D of Table 2 reports the estimated standard deviations of the random coefficients, which capture within-group preference heterogeneity not explained by observed characteristics. Two patterns are worth noting. First, the estimated standard deviation of the gender concordance coefficient is substantially larger for women

than for men across all three specialties — for example, 0.621 for women versus 0.116 for men in urology — indicating that women's concordance preferences are considerably more dispersed around their group mean. This dispersion is consistent with the presence of distinct subgroups of women with very different underlying preferences, and motivates the heterogeneity analyses we present below. Second, the standard deviation of the wait time coefficient is small and similar across all groups (ranging from 0.074 to 0.118), indicating that sensitivity to wait time is approximately homogeneous across respondents. This is reassuring for the interpretation of our WTP estimates: because the denominator of the WTP ratio is precisely and homogeneously estimated, variation in WTP for gender concordance across subgroups reflects true heterogeneity in concordance preferences rather than heterogeneity in the valuation of time.

4.3 The Role of Reported Prior Negative Healthcare Experiences Due to Gender

As we observed in Section 4.1, women are substantially more likely than men to report having been treated negatively or unfairly by a medical professional due to their gender. Those who have had such experiences may place a higher value on gender concordance, anticipating that a female physician is less likely to perpetuate them. Alternatively, some other unobserved characteristic may drive both higher likelihood of perceiving mistreatment due to gender and higher gender concordance preferences. Even if the latter is the primary factor behind group differences, the differences can still be informative about heterogeneity in concordance preferences. We present results separately for women who report prior negative gender-related medical experiences and those who do not.

The results are shown in Figure 6. Pooling across all three specialties, women with prior perceived negative experiences are willing to wait more than twice as long for a gender-concordant physician than those without — on average, approximately one additional week, although this difference is not statistically significant. This pattern holds across all three specialties, and with similar magnitudes. It is important to mention here that while the differences are not statistically significant, they all consistently point in the same direction: those women who report having had negative experiences in healthcare due to gender are willing to wait longer to see a female physician.

In urology, women who report prior negative experiences are willing to wait roughly twice as long (2.25 weeks) as those without (1.3 weeks), and both estimates are statistically significantly different from zero. In orthopedics, the estimate for women who report prior negative experiences is almost three times larger, but neither group's estimate is statistically distinguishable from zero, consistent with the weak concordance preferences we observe in this neutral specialty overall.

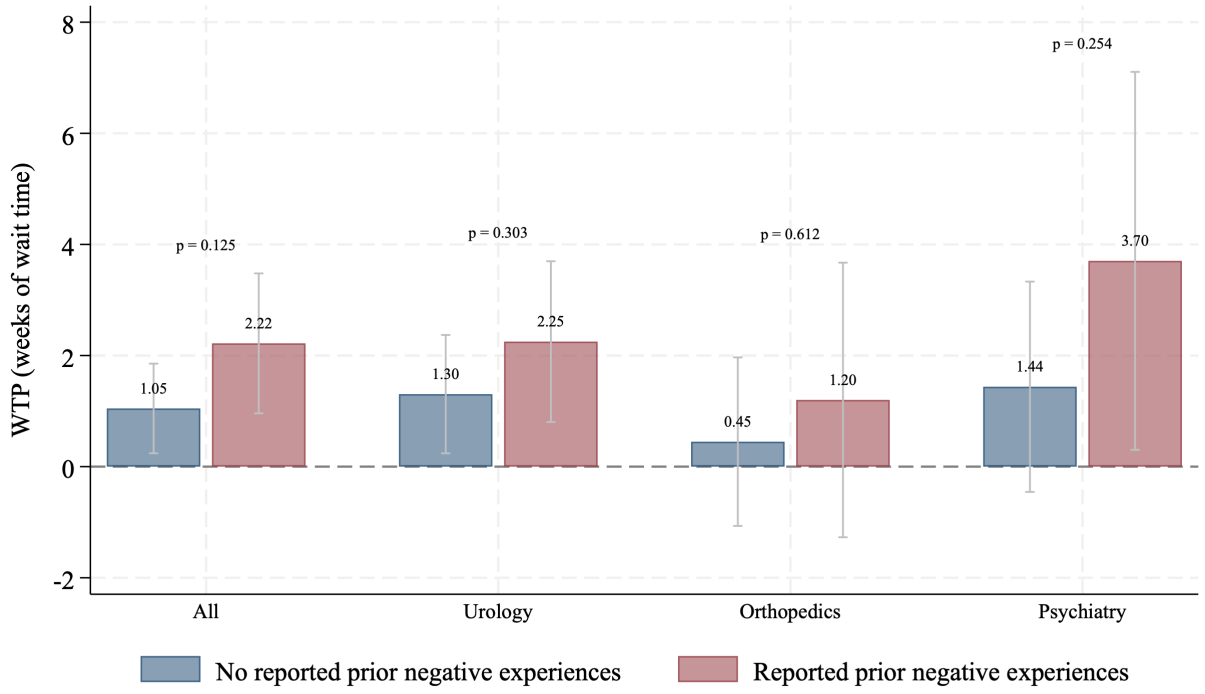
In psychiatry, women who report prior negative experiences are willing to wait an average of 3.7 additional weeks for a female psychiatrist, an estimate that is statistically significant, while the estimate for those without such experiences is 1.4 weeks and is statistically indistinguishable from zero. This demonstrates one potential reason why, while there appears to be an overall preference for gender concordance with a psychiatrist among female respondents, there is substantial variance. Psychiatry was chosen as a specialty

because it may be sensitive to social norms, including gender, given patients discuss private and often sensitive topics with a psychiatrist. If negative experiences that are perceived to be based on gender fall into that category, this could explain the varying but strong preferences in this clinical specialty.

These results suggest that gender concordance preferences are magnified by prior negative healthcare experiences. Women who report having encountered gender-based mistreatment in medical settings are willing to bear substantially higher costs to see a female physician, and a substantial share (30%) of our female respondents do report such mistreatment.

This has two implications. First, it suggests that concordance preferences partly reflect a response to past discrimination: women who believe their gender has negatively impacted past medical experiences are those most willing to wait for a female physician. Second, it implies that the welfare costs of a non-representative physician workforce fall disproportionately on those who may have already been negatively affected by it.

Figure 6: Female WTP in Wait Time for Gender Concordance by Prior Negative Experiences



Notes: Sample construction and estimation same as in Table 2, limited to female respondents, who are further broken down by whether they reported having been treated negatively or unfairly by a medical professional due to their gender. Male respondents are excluded due to a very small share reporting such negative experiences. Bars represent point estimates and capped lines represent 95% confidence intervals. P-values test the null hypothesis of no difference in gender concordance WTP between those with and without reported negative experiences due to gender.

4.4 The Role of Urgency Perception

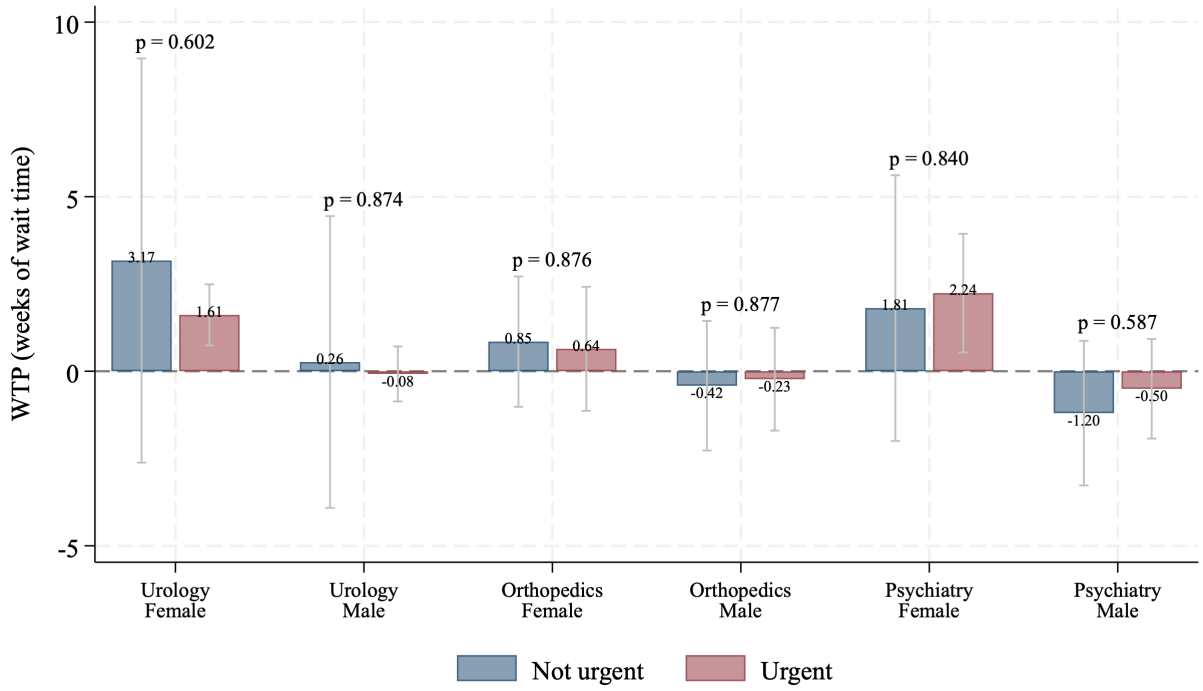
Willingness to wait may depend on the perceived urgency of symptoms. Within our experiment, respondents varied in how urgent they perceived the presented medical issue to be: 93% of

participants assigned to the urology scenario viewed their symptoms as "quite" or "very" urgent, compared to 54% in the orthopedics scenario and 61% in the psychiatry scenario.

Perceived urgency is, in turn, correlated with willingness to pay for gender concordance. Figure 7 shows that women who did not perceive the urological symptoms as urgent were willing to wait an average of 3.2 additional weeks to see a female urologist, compared to 1.6 weeks among those who perceived the symptoms as urgent, a twofold difference. The difference is not statistically significant, however, due to the small sample of respondents who did not perceive the issue as urgent.

Notably, even among respondents who perceived the issue as urgent, willingness to wait remained statistically significantly different from zero. This suggests that gender concordance preferences may lead to delayed care even when patients recognize that their symptoms are time-sensitive. These findings provide direct experimental support for the real-world patterns we document in Section 5, where women who are ultimately diagnosed with bladder cancer in areas with limited access to female urologists wait longer to see a concordant provider and thus delay diagnosis.

Figure 7: WTP in Wait Time for Gender Concordance by Urgency Perception



*Notes:*Sample construction and estimation same as in Table 2. Groups are broken down by whether they perceive the presented symptoms as "very" or "quite" urgent, or "not" or "not at all" urgent. Bars represent point estimates and capped lines represent 95% confidence intervals. P-values test the null hypothesis of no difference in gender concordance WTP between urgent and non-urgent respondents.

5. Real-World Application Using Medicare Data

The experimental results show that female patients prefer appointments with female physicians and are willing to wait longer to see them—particularly in specialties that are gender-sensitive and involve physical examinations, such as urology. As discussed in Section 2, female physicians are underrepresented in urology, with some areas of the country having no female urologist at all. If female patients are indeed waiting longer to access the limited supply of female urologists, the combination of gender concordance preferences and the undersupply of female urologists could carry meaningful real-world consequences for timely care.

Bladder cancer provides a useful setting to examine this question. The main symptom of bladder cancer is blood in the urine, often accompanied by pain with urination and frequent urination (NHS, 2025). Because these symptoms overlap with more common conditions such as urinary tract infections, patients may not recognize the urgency and may delay care in favor of other appointment characteristics. While incidence is higher among men, women are diagnosed at later stages on average and face a worse prognosis (Pignot, Barthélémy and Borchiellini, 2024). McDevitt and Roberts (2014) find that markets without a female urologist have higher mortality rates for female bladder cancer patients, all else equal. Our experimental results offer a mechanism: female patients reporting symptoms consistent with bladder cancer indicated a preference for waiting longer in order to see a female urologist. In this section, we show that patterns of realized patient-physician matching in Medicare claims are consistent with the experimental findings.

5.1 Summary Statistics in Medicare Claims

We use Medicare Fee-for-Service (FFS) claims from 2006–2018, which comprise a random 20% sample of all Medicare FFS beneficiaries. Medicare is a useful setting to study bladder cancers because the disease disproportionately affects older adults, with roughly nine in ten diagnoses occurring in patients over age 55 (American Cancer Society, 2024). We focus on the carrier file, which includes all claims submitted by individual providers. Each claim includes the date of service, provider NPI, diagnosis codes, and procedure codes, allowing us to reconstruct the full sequence of outpatient encounters in the period leading up to a bladder cancer diagnosis. We restrict the sample to beneficiaries with a first observed diagnosis of bladder cancer starting at the age of 66.[7]

There is no direct measure of wait times in Medicare claims, and Medicare fee-for-service does not typically require a referral to book an appointment with a specialist. However, in practice, patients often present first to a primary care physician before scheduling with a urologist. We define wait time as the number of days between a patient's last primary care claim and their first urology contact, capturing the delay between initial symptom presentation and specialist evaluation.

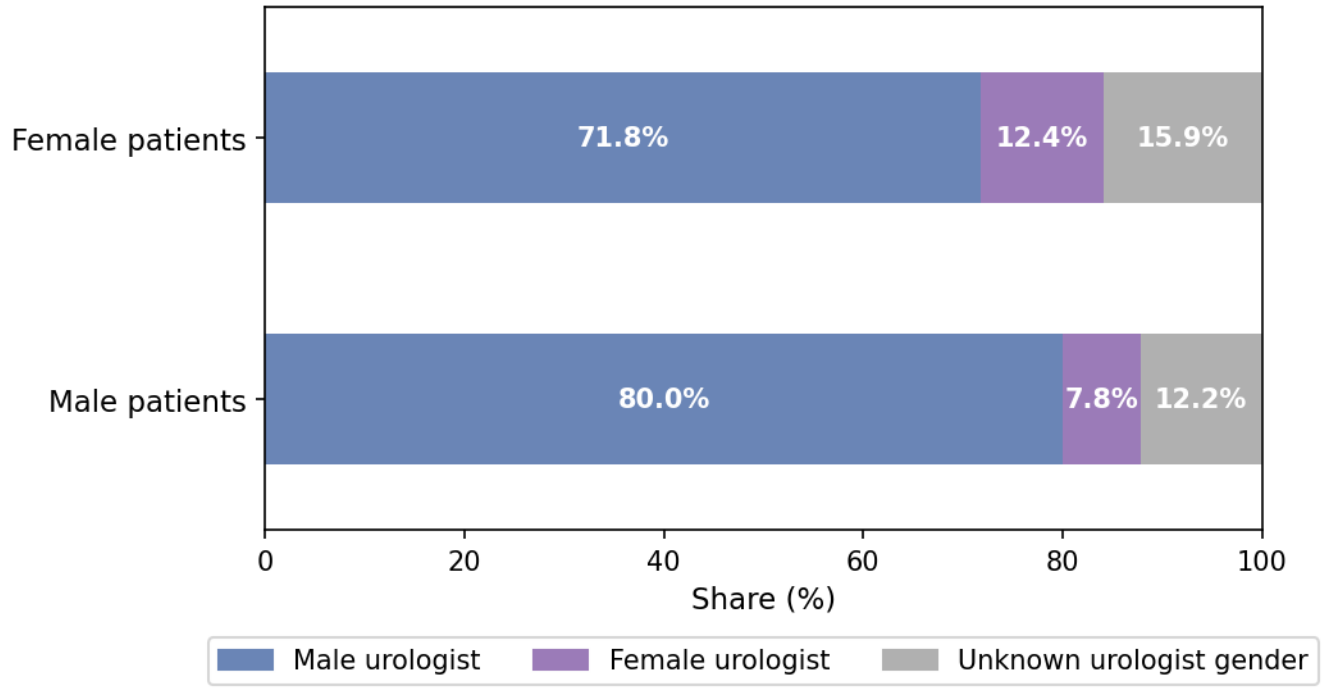
We construct this measure working backward from the diagnosis date. Because the diagnosis date is not necessarily the date of first urology contact, many patients have related claims from laboratories and other providers in the days or weeks preceding a formal diagnosis,

[7]We exclude beneficiaries with a first observed diagnosis of bladder cancer at age 65, because these patients are more likely to have received a diagnosis before enrolling in Medicare insurance.

we identify the first urologist seen prior to diagnosis as our anchor. In cases where no urology claim is observed before diagnosis but the diagnosing provider is a urologist, we treat that provider as the first urologist contact. Wait time is then defined as the number of days elapsed between the last primary care claim preceding this first urology visit and the urology visit itself. While the inferred wait time measure is subject to measurement error, there is no reason to expect mismeasurement to differ systematically between male and female patients.

Of the approximately 150,000 beneficiaries identified to have a new bladder cancer diagnosis after the age of 65, 29% are female. We are able to identify at least one urology claim for 114,698 beneficiaries, and are able to calculate an estimated wait time for 90,373 beneficiaries. As shown in Figure 8, female patients are more likely to see female urologists, which is consistent with our survey results. We find that, on average, female patients have shorter wait times compared with male patients (32.4 days vs. 35.3 days) but are diagnosed at slightly older ages (77.8 vs 77.2).

Figure 8: Gender of Urologist by Patient Gender



Notes: Figure shows the share of patients who see a male vs. female urologist, by patient gender. If a patient has a claim with more than one urologist, we use the gender of the first urologist in the year prior to bladder cancer diagnosis. Our primary source of physician gender is Medicare Data on Provider Practice and Specialty (MD-PPAS). For those with missing gender information, we supplement with the National Plan and Provider Enumeration System (NPPES) data. N = 114,698 bladder cancer patients with at least one urologist claim.

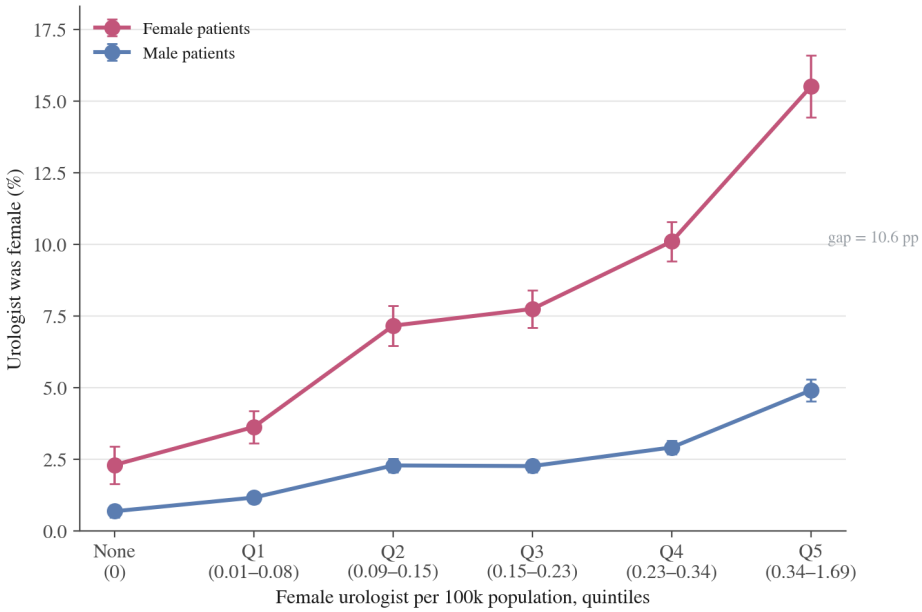
5.2 Wait times and realized gender concordance

Our key question of interest is whether women with bladder cancer are waiting longer for urology appointments in order to see female urologists. We first estimate how patient-physician matching responds to the supply of female urologists. We focus on hospital referral

regions (HRRs), and calculate the number of female urologists using NPPES records. We count the total number of urologists with a practice location address in each HRR from the 2019 NPPES data, excluding those with an enumeration year after 2018 or who deactivated their NPI prior to 2006. This gives an estimate of the cumulative count of female urologists in each HRR between 2006-2018.

Figure 9 plots the share of patients matched to a female urologist against the local supply of female urologists per capita, by patient gender. For both male and female patients the match rate with female physicians rises with supply, but the gradient is far steeper for female patients. This pattern is consistent with the gender-concordance preferences for urology documented among female respondents in the survey, and suggests that the limited supply of female urologists constrains the share of female patients who are able to match with one.

Figure 9: Share of Patients Matching with Female Urologists, By Patient Gender

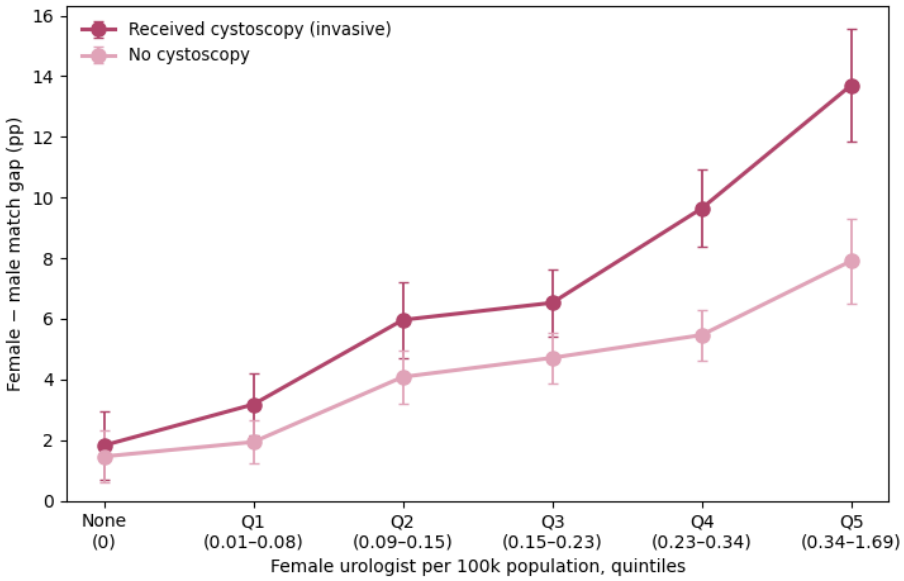


Notes: Figure shows the share of patients whose urologist was female, by patient gender, across quintiles of female urologist supply in the patient's Hospital Referral Region (HRR). Supply is measured as the number of female urologists per 100,000 HRR residents, constructed from 2019 NPPES practice location. HRRs with no female urologist are shown separately; the remaining HRRs are grouped into quintiles, with bin cutoffs reported below the axis. If a patient has a claim with more than one urologist, we use the gender of the first urologist seen. Capped lines denote 95 percent confidence intervals for the share of patients seeing a female urologist.

In our survey scenario, patients in the urology arm were told that during the hypothetical appointment the doctor would perform a cystoscopy, a procedure in which a thin tube with a camera and light is passed through the urethra to examine the inside of the bladder. The exact wording of the described appointment is shown in Figure B6. We chose to describe this procedure in order to emphasize that the appointment would involve an intimate physical examination. We would expect an anticipated intimate or invasive procedure to drive stronger preferences for gender concordance.

We examine whether this pattern appears in real appointments in the Medicare claims data by splitting patients by whether they received a cystoscopy, a standard procedure in the diagnosis of bladder cancer, during their first urology visit. In Figure 10 we plot the gap between female and male patients in the probability of matching with a female urologist, across quintiles of female urologists per capita in each patient's HRR. The darker pink line shows the gap for patients who received a cystoscopy at their first urology visit; the lighter pink line shows those who did not. The gap is positive in all quintiles, meaning female patients are more likely to be matched with female urologists, and it increases with supply of female urologists for both groups. The gradient is steeper for patients who received a cystoscopy, suggesting that those selecting into appointments involving a more intimate procedure are more responsive to the supply of female physicians. This is consistent with the hypothesis that concordance preferences are stronger for women when the appointment involves a gender-sensitive examination. However this figure is purely descriptive, as we do not observe what procedures patients anticipated when booking the visit and the two groups may differ in other ways.

Figure 10: Sorting into Female Urologists, by Cystoscopy Receipt



Notes: Figure plots the gap between female and male patients in the share matched to a female urologist (in percentage points). The x-axis shows quintile of female urologists per 100,000 HRR residents; HRRs with no female urologists are shown separately. If a patient has a claim with more than one urologist, we use the gender of the first urologist seen. The sample is split by whether the initial urology visit involved a cystoscopy. Capped lines denote 95 percent confidence intervals for the female–male gap.

If female urologists are in excess demand by female patients, then we would expect female patients waiting for a female urologist to wait longer. To test this we estimate the following equation separately for male and female patients.

$$WaitTime_{iht} = \alpha + \beta \mathbb{1}[\text{Female urologist}]i + \gamma_t + \mu_h + \varepsilon_{iht} \tag{8}$$

where $\mathbb{1}[\text{Female urologist}]_i$ indicates that the first urologist patient i saw was female, γ_t are year of diagnosis fixed effects, and μ_h are HRR fixed effects. $WaitTime_{iht}$ is days between the last primary care visit and first urology contact for patient i in HRR h in year t. Standard errors are clustered at the HRR level. Estimating the specification on male patients serves as a placebo. If female urologists practice in settings where wait times are longer for reasons unrelated to patient gender, the coefficient should be positive for men as well.

Results are shown in Table 3. Female patients who see a female urologist wait 3.97 days longer than female patients who see a male urologist, against a mean wait of 31.0 days. Adding year fixed effects raises the estimate to 4.7 days, and adding HRR fixed effects leaves it essentially unchanged. For male patients, the corresponding estimates are small, negative, and statistically indistinguishable from zero in every specification. The null for male patients suggests the increased wait time to see a female urologist is specific to female patients matching on gender, rather than a general feature of female urologists' practices. Male patients face similar wait times regardless of the gender of the urologist they see, suggesting they are not trading off wait time against physician gender in the way female patients appear to be.

Table 3: Wait Times to See Female Urologists, By Patient Gender

	Wait time PCP to urologist (days)					
	Women			Men (placebo)		
	(1)	(2)	(3)	(4)	(5)	(6)
Female urologist	3.97**	4.73***	4.65***	-1.54	-1.05	-1.24
	(1.33)	(1.33)	(1.32)	(1.27)	(1.26)	(1.32)
Mean wait (days)	31.0	31.0	31.0	33.9	33.9	33.9
Year fixed effects		✓	✓		✓	✓
HRR fixed effects			✓			✓
Observations	21,495	21,495	21,495	55,596	55,596	55,596
R^2	0.001	0.004	0.022	0.000	0.001	0.010

Notes: OLS estimates of Equation (8). The outcome is the number of days between the last primary care visit and first urology contact prior to bladder cancer diagnosis. *Female urologist* indicates that the first urologist seen prior to diagnosis was female. Columns 1–3 estimate the specification on female patients; columns 4–6 repeat it on male patients. *Mean wait* reports the sample mean of the outcome for the relevant patient group. The sample is restricted to patients with an observed wait time and an identified first urologist of known gender. Standard errors clustered at the HRR level in parentheses. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

6. Discussion

Many medical specialties have large gender gaps in the physician workforce, and prior literature has documented that some patients prefer to see physicians of the same gender. Yet evidence on whether patients act on these preferences, and what the health consequences

of this are, are limited. This is especially important given the increasing agency patients are exercising in their healthcare choices, often using online search and information tools. While in the past physicians may have recommended a particular specialist, or patients may have felt unable to ask for a specific demographic of physician, patients now increasingly have the option to choose a physician and an appointment themselves. Given the asymmetry of information in healthcare, they may do so without fully considering potential health consequences, particularly with respect to the progression of disease with delay.

This paper provides direct experimental evidence on the willingness to pay for gender-concordant physicians across specialties that differ in gender sensitivity, and links those estimates to real-world care-seeking behavior. Our results confirm that gender concordance preferences impact preferences over appointments, but they are concentrated among women and in gender-sensitive specialties. We find women are willing to wait approximately 1.6 additional weeks to see a female urologist, while men exhibit little to no concordance preference across all three specialties studied. In orthopedics, a more gender-neutral specialty, neither men nor women display statistically significant concordance preferences. In psychiatry, where gender can play an important role, we find that women's estimate is large in magnitude and statistically significant, but the variance is also high. The strength of concordance preferences in the more gender-sensitive settings, and their absence in orthopedics, suggests that these preferences are not a general aversion to opposite-sex providers, but are instead tied to the intimate or sensitive nature of specific clinical interactions.

These preferences are heterogeneous in important ways. Preference for gender concordance is stronger among women who report having previously experienced negative or unfair treatment in medical settings due to their gender. Approximately 30% of female respondents report such experiences, and these women are willing to wait roughly two additional weeks for a female physician. Additionally, the perceived urgency of symptoms moderates but does not eliminate concordance preferences.

Using Medicare claims data, we find suggestive evidence that these preferences matter for real-world access to care. We focus on the specialty with the strongest experimental concordance preferences, urology, and examine a subset of patients with particularly high stakes: those ultimately diagnosed with bladder cancer. Consistent with the experimental results, female bladder cancer patients are substantially more likely to see a female urologist than their male counterparts. As the supply of female urologists in an HRR increases, all patients are more likely to match with a female urologist, but the gradient is far steeper for female patients. Taken together, these patterns suggest that women's ability to act on the concordance preference documented in our experiment is constrained by the local supply of female urologists. Among female patients, we find that those who see a female urologist wait on average four days longer than those who see a male urologist. Male patients show no corresponding gap, consistent with the delay reflecting excess demand for female urologists by female patients, rather than differences in how female physicians practice. Overall, the patterns in urology appointments within Medicare claims data suggest that preferences for gender-concordant physicians among women with bladder cancer are obtained at the cost of delays in care.

The findings of the survey, in conjunction with the Medicare application, have meaningful policy implications. To the extent that delayed care worsens health outcomes, the under-representation of women in certain specialties is a public health concern. Policies aimed at

diversifying the specialty workforce, or mitigating the concerns of female patients seeing male physicians in these areas (e.g., via chaperone policies), may have downstream benefits for patient health. Notably, we do not find strong concordance preferences in non-gender-sensitive specialties like orthopedics, suggesting that female patients do not have an across-the-board preference for female physicians. We also find that concordance preferences are in part driven by past negative experiences with the healthcare system, suggesting that interventions targeting the quality of patient-physician interactions could also play an important role. Additionally, we find that perceived urgency attenuates concordance preferences, though does not eliminate them. In future extensions of this work, we will test interventions that target methods of increasing patient comfort in gender-discordant clinical encounters. We will also expand the Medicare analysis to additional specialties and examine the downstream effects of appointment delays on diagnosis timing and disease progression more directly.

References

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

American Cancer Society. 2024. "Bladder Cancer Risk Factors."

AMN Healthcare. 2025. "2025 Survey of Physician Appointment Wait Times and Medicare and Medicaid Acceptance Rates." AMN Healthcare Physician Solutions Survey Report. Survey conducted January 15–February 24, 2025, across 15 major metropolitan areas and six medical specialties.

Andrew, Alison, and Abi Adams. 2025. "Revealed Beliefs and the Marriage Market Return to Education." *The Quarterly Journal of Economics*, 140(3): 2107–2162.

Arrow, Kenneth J. 1963. "Uncertainty and the Welfare Economics of Medical Care." *American Economic Review*, 53(5): 941–973.

Chan, Alex. 2026. "Discrimination Against Doctors: A Field Experiment." Conditionally accepted, *American Economic Review*. Available at http://tinyurl.com/chandiscrimination.

Chen, Yiwei, and Stephanie Lee. 2024. "User-Generated Physician Ratings and Their Effects on Patients' Physician Choices: Evidence from Yelp." *Journal of Marketing*, 88(1): 77–96.

Fong, Christina M., and Erzo F. P Luttmer. 2009. "What Determines Giving to Hurricane Katrina Victims? Experimental Evidence on Racial Group Loyalty." *American Economic Journal: Applied Economics*, 1(2): 64–87.

Gallen, Yana, and Melanie Wasserman. 2023. "Does information affect homophily?" *Journal of Public Economics*, 222: 104876.

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*, 115(34): 8569–8574.

Lamiraud, Karine, Morgane Le Guern, Michael Rockinger, and Christine Sevilla-Dedieu. 2026. "Preference for Gender Concordant GP and Medical Deserts." *Review of Economics of the Household*.

McCarthy, Ian, Kaylyn R. Sanbower, and Leonardo Sánchez Aragón. 2024. "Online Reviews and Hospital Choice." Working paper.

McDevitt, Ryan, 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): 68–93.

Miyawaki, Atsushi, Anupam B. Jena, Lisa S. Rotenstein, and Yusuke Tsugawa. 2024. "Comparison of Hospital Mortality and Readmission Rates by Physician and Patient Sex." *Annals of Internal Medicine*, 177(5): 586–594.

NHS. 2025. "Symptoms of Bladder Cancer."

Nolen, Haley A., Justin Xavier Moore, Joel B. Rodgers, Henry E. Wang, and Lauren A. Walter. 2016. "Patient Preference for Physician Gender in the Emergency Department." *The Yale Journal of Biology and Medicine*, 89(2): 131–42.

Pignot, Géraldine, Philippe Barthélémy, and Delphine Borchiellini. 2024. "Sex Disparities in Bladder Cancer Diagnosis and Treatment." *Cancers*, 16(23): 4100.

Reyes, Jessica Wolpaw. 2006. "Do Female Physicians Capture Their Scarcity Value? The Case of OB/GYNs." National Bureau of Economic Research Working Paper 12528.

Reyes, Jessica Wolpaw. 2008. "Gender Preference and Equilibrium in the Imperfectly Competitive Market for Physician Services." *Eastern Economic Journal*, 34(4): 431–448.

Train, Kenneth E. 2009. *Discrete Choice Methods with Simulation.* . 2nd ed., Cambridge:Cambridge University Press.

Tsugawa, Yusuke, Anupam B. Jena, José F. Figueroa, E. John Orav, David M. Blumenthal, and Ashish K. Jha. 2017. "Comparison of Hospital Mortality and Readmission Rates for Medicare Patients Treated by Male vs Female Physicians." *JAMA Internal Medicine*, 177(2): 206–213.

Varadarajulu, Shyam, Carol Petruff, and William H. Ramsey. 2002. "Patient Preferences for Gender of Endoscopists." *Gastrointestinal Endoscopy*, 56(2): 170–173.

Walker, Brigham, Janna Wisniewski, Sarah Tinkler, Jillian Torres, and Rajiv Sharma. 2024. "Identity and Access: Gender-Based Preferences and Physician Availability in Primary Care." *Journal of Economic Behavior & Organization*, 225: 707–722.

Wallis, Christopher J.D., Angela Jerath, Natalie Coburn, Zachary Klaassen, Amy N. Luckenbaugh, David E. Magee, Amanda E. Hird, Kate Armstrong, Damian Broe, Bhavya Ravi, Girish Kulkarni, and Raj Satkunasivam. 2022. "Association of Surgeon–Patient Sex Concordance With Postoperative Outcomes." *JAMA Surgery*, 157(2): 146–156.

Werner, Rachel M., Edward C. Norton, R. Tamara Konetzka, and Daniel Polsky. 2012. "Do Consumers Respond to Publicly Reported Quality Information? Evidence from Nursing Homes." *Journal of Health Economics*, 31(1): 50–61.

Wiswall, Matthew, and Basit Zafar. 2018. "Preference for the Workplace, Investment in Human Capital, and Gender*." *The Quarterly Journal of Economics*, 133(1): 457–507.

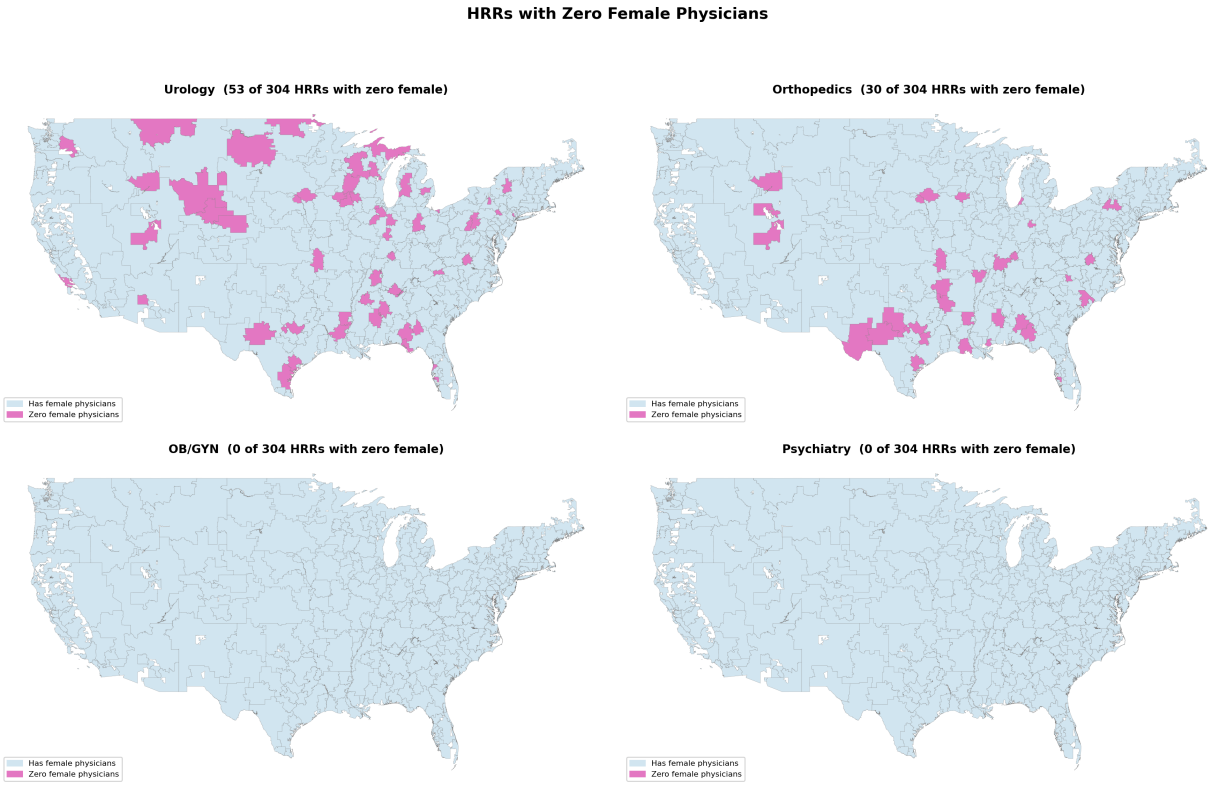
A. Background

Table A1: Sample Construction and Restrictions

Step	N remaining	N dropped	% of total
Total NPI records	9,129,558	0	100.0%
Restrict to individuals (Entity Type = 1)	6,962,229	2,167,329	76.3%
Drop deactivated NPIs	6,947,618	14,611	76.1%
Restrict to physician taxonomy codes	1,204,934	5,742,684	13.2%
Require non-missing gender (M/F)	1,204,604	330	13.2%
Match to HRR via practice ZIP code	1,189,697	14,907	13.0%

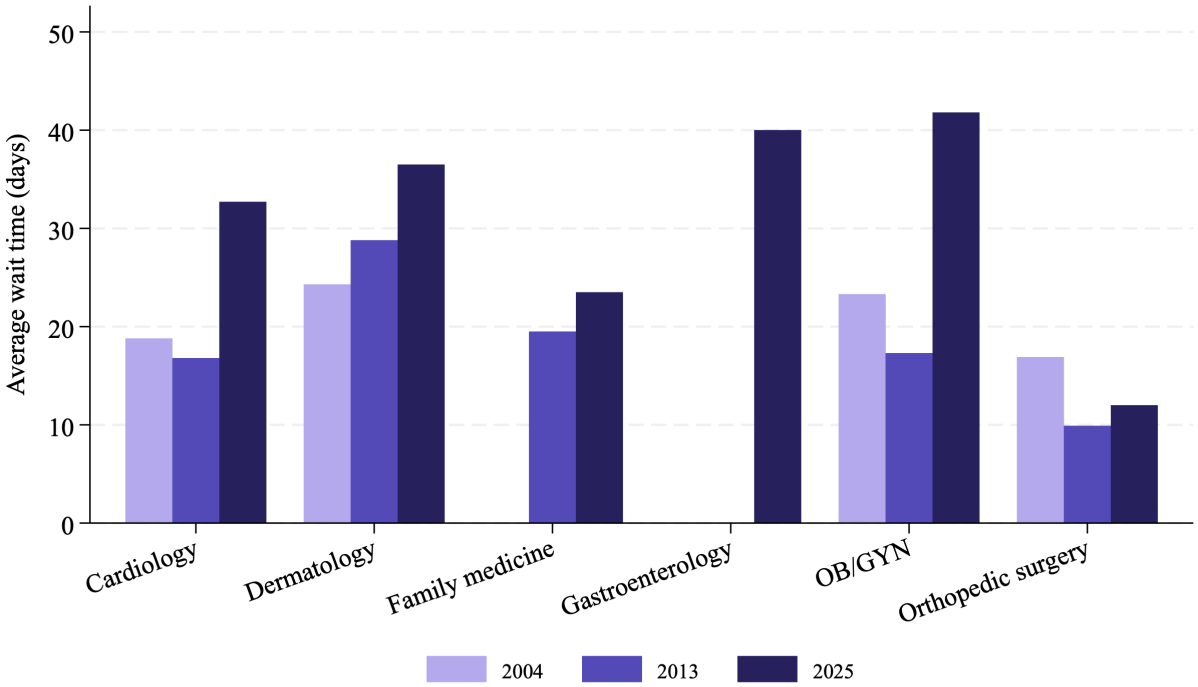
Notes: NPPES NPI file for September 2025, downloaded from the NBER.

Figure A1: Hospital Referral Regions with Zero Female Physicians by Specialty



Notes: Data from NPPES NPI file September 2025, downloaded from NBER. Analysis restricted to top twenty specialties by number of physicians. Specialty inferred from provider taxonomy codes. Practice location linked to HRR via ZIP. 304 HRRs displayed (Hawaii and Alaska excluded).

Figure A2: Wait Times by Specialty over Time



Notes: Source: AMN Healthcare Survey of Physician Appointment Wait Times. Average across 15 major US metro areas. Family medicine not surveyed in 2004; gastroenterology first included in 2025.

B. Detail on Survey Design

In this appendix, we outline details in our survey design and the motivation behind them.

B.1 Explanation of Hypothetical Choice

After consenting to the survey, passing an instructional manipulation check, and answering some background questions, respondents are introduced to the symptom scenarios and hypothetical choice component of the survey. We begin by showing them the following text.

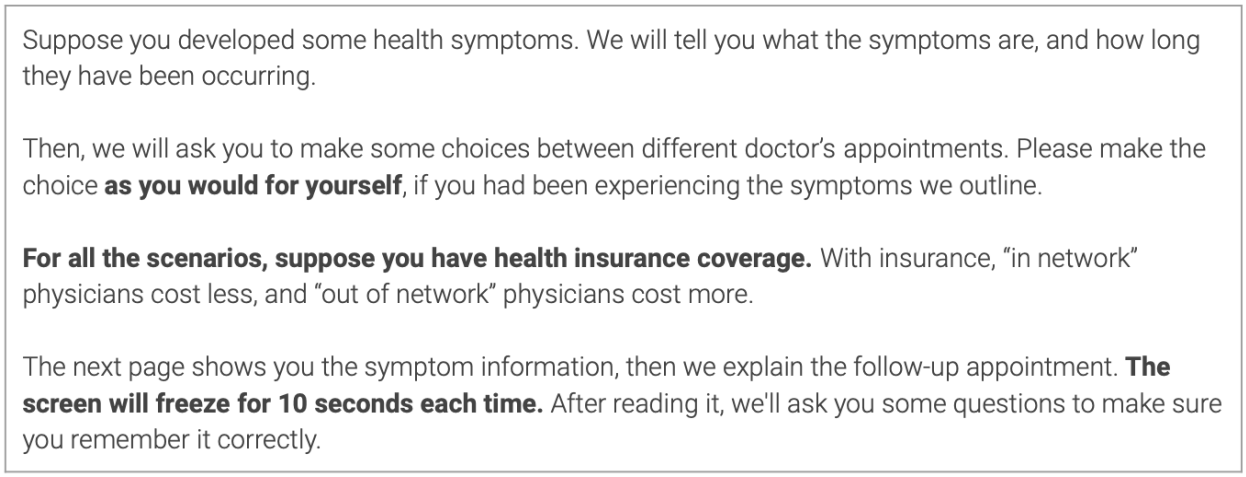


Figure B3: Introduction to Symptom Scenario and Hypothetical Choice

We made several important choices in this introduction:

- We emphasize that the hypothetical choices should be made from the perspective of the respondent, as if he/she were making the choice in reality ("Please make the choice **as you would for yourself**...").
- We asked respondents to imagine they had health insurance coverage when making their choice. This is because we wanted to include a monetary cost for each physician-appointment, to allow for calculation of a willingness-to-pay in dollar terms for concordance. Critically, we did not want this cost to be inferred as a measure of physician quality (e.g. more expensive physicians were better), so sought to find an explanation that was independent on underlying physician characteristics. We described physicians as in- or out-of-network to motivate differences in cost.
- We told respondents what information they would see, and that each information screen would be frozen for 10 seconds. Freezing screens can increase attrition in surveys, so we kept the time interval short and forewarned respondents to avoid confusion and motivate engagement with the material.

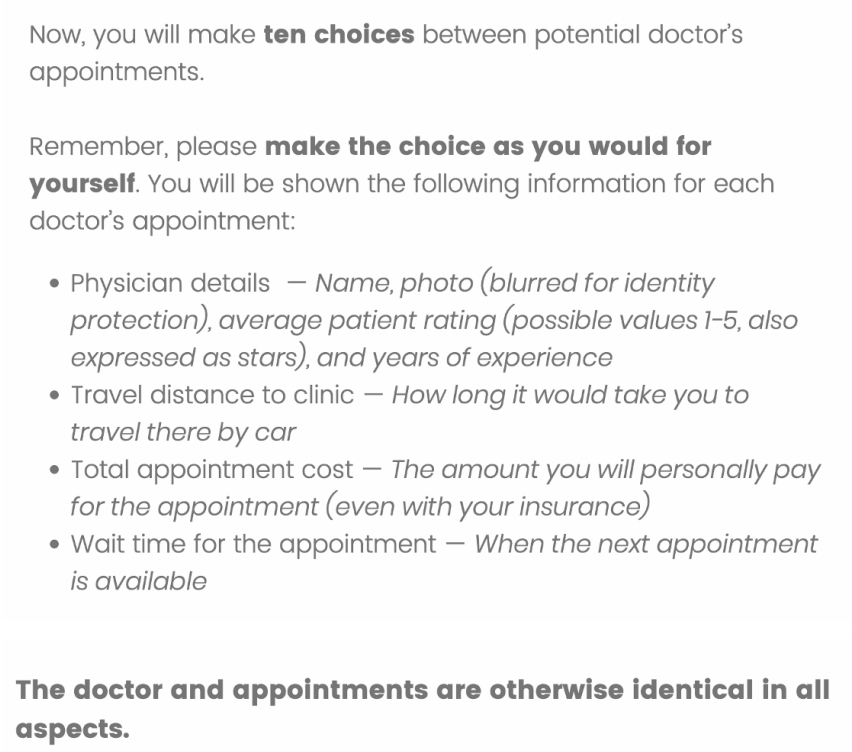


Figure B4: Preview to Hypothetical Choices

B.2 Information Treatments

After previewing the scenarios and hypothetical choice exercise, respondents were randomized into a speciality/symptom scenario. First, they were asked to suppose that they had been experiencing a collection of symptoms. We outline those symptoms and specialty/disease areas in Figure B5.

Neutral *Orthopedics* *Osteoarthritis Symptoms*	Over the past month, suppose you developed these symptoms: • Stiffness in your knee when getting up from your desk • A "scraping" sensation in your knee when you walk • A dull ache in your knee after walking for longer periods
Gender *Urology* *Bladder Cancer Symptoms*	Over the past month, suppose you developed these symptoms: • You have soreness in your hips and lower back • You need to wake up several times at night to urinate • You have noticed some blood in your urine on occasion
Social/Behavioral *Psychiatry* *Depressive Symptoms*	Over the past month, suppose you developed these symptoms: • You have been struggling to get to sleep at night • You do not feel as interested in your regular leisure activities • You have found it difficult to concentrate when at work • You have been feeling sad most days

Figure B5: Symptom Scenarios

We made several important choices in these symptom outlines:

- We kept the information load similar across scenarios. In all cases we said symptom development had been occurring for a month and detailed 3-4 symptoms.
- We tried to maintain a non-emotive tone with no symptoms that presented an obvious emergency that might warrant a visit to the emergency department. For example, in the case of depressive symptoms, we avoided any mention of self-harm. This was because we wanted the choice of organizing a specialist appointment, potentially several weeks in advance, to be a plausible one.
- We chose gender-neutral symptoms such that male and female respondents could experience the same symptom scenarios, to allow for more precise comparisons of male versus female preferences for concordance.
- We detailed symptoms found on reputable websites for specific diseases, such as the Centers for Disease Control and Prevention (CDC), Mayo Clinic, and National Health Service (of the United Kingdom).

Respondents were then told that they had been referred to see a specialist by a primary care physician. We made this choice to make the choice of seeing a specialist more plausible, and prevent respondents from reneging on the choice pathway we had laid out. For example, if we had asked them to choose between specialists, a respondent might have chosen visiting their usual primary care physician as an outside option. We then detailed exactly what type of specialist they had been referred to, as well as what would happen at the appointment. We outline this content in Appendix Figure B6.

Neutral *Orthopedics* *Osteoarthritis Symptoms*	A primary care doctor referred you to see an **orthopaedic surgeon**. An orthopaedic surgeon is a doctor who diagnoses and treats conditions related to the bones, joints, and muscles. At the appointment, the doctor will examine your knee, ask you about your symptoms, and might order some imaging tests (e.g. X-ray, ultrasound) and lab tests (e.g. blood tests). The doctor will also discuss possible diagnoses and treatment options.
Gender *Urology* *Bladder Cancer Symptoms*	A primary care doctor has referred you to see a **urologist**. A urologist is a doctor who diagnoses and treats conditions related to the urinary tract and reproductive system. At the appointment, the doctor will examine your abdomen and ask about your symptoms. The doctor will also examine the inside of your bladder by passing a thin tube with a camera and light at the end through your urethra (the tube through which you urinate).
Social/Behavioral *Psychiatry* *Depressive Symptoms*	A primary care doctor has referred you to see a **psychiatrist**. A psychiatrist is a doctor who diagnoses and treats conditions related to mental health. At the appointment, your psychiatrist will ask you about you and your family's medical history, your daily life, and your mental health. The appointment will be 1-1.5 hours long, to give you time to discuss your feelings in depth.

Figure B6: Specialist Scenarios

We made several important choices in these symptom outlines:

- We kept the format and length of the specialist introduction similar across treatments, to facilitate comparisons across the specialties.
- We detailed exactly what would happen at each appointment, in case the respondent did not know of the specialty or what typically occurred at such visits. Providing this information was also essential for realism, given it is in the realm of what a primary care physician would tell a patient in an initial visit.
- We highlighted elements of the appointment where the norms we were interested in came into play. So, in urology we explicitly stated that the appointment would include an intimate physical exam, given we are interested in how preferences for concordance might vary across specialties with different levels of physical engagement and gender norms. In psychiatry, we explicitly stated that the appointment would be 1-1.5 hours long, and would involve a detail review of the respondent's personal life.

After these information treatments, the respondents were asked to answer several questions about the information. As we detailed above, the respondents had been forewarned these questions would occur, as an incentive for reading the information carefully. Though their answers did not impact their compensation at all, answering the questions incorrectly extended the survey time; those who did not correctly answer all the questions were routed to the information for a second time.

The last section before the hypothetical choice asked respondents to react to the information they had been shown. These questions were designed to help us understand how

medical knowledge might impact willingness-to-pay for concordance with, for example, time to appointment. For example, we asked respondents how urgent they thought the issue was, as well as if they had any predicted diagnoses.

B.3 Examples of Physician Search Tools

Below we provide some examples of physician search tools that we drew on while designing our hypothetical choice experiment.

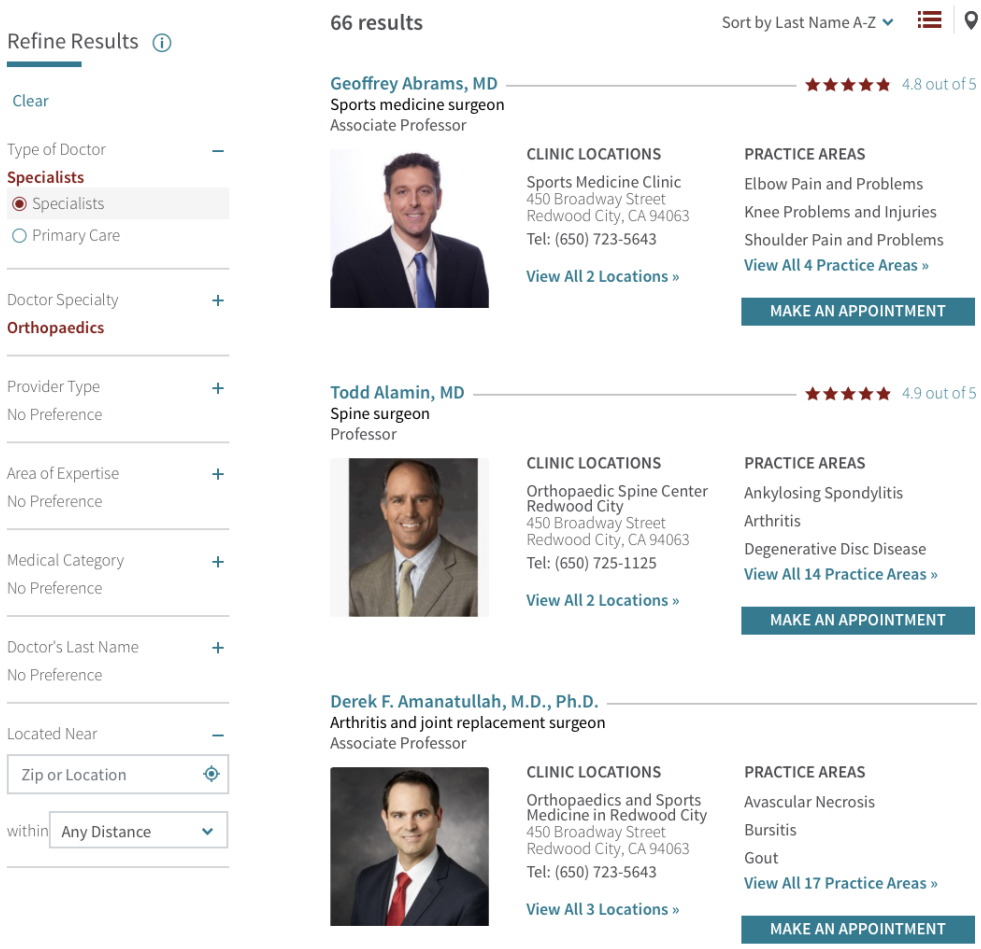


Figure B7: Search results for Stanford Orthopedics Specialists

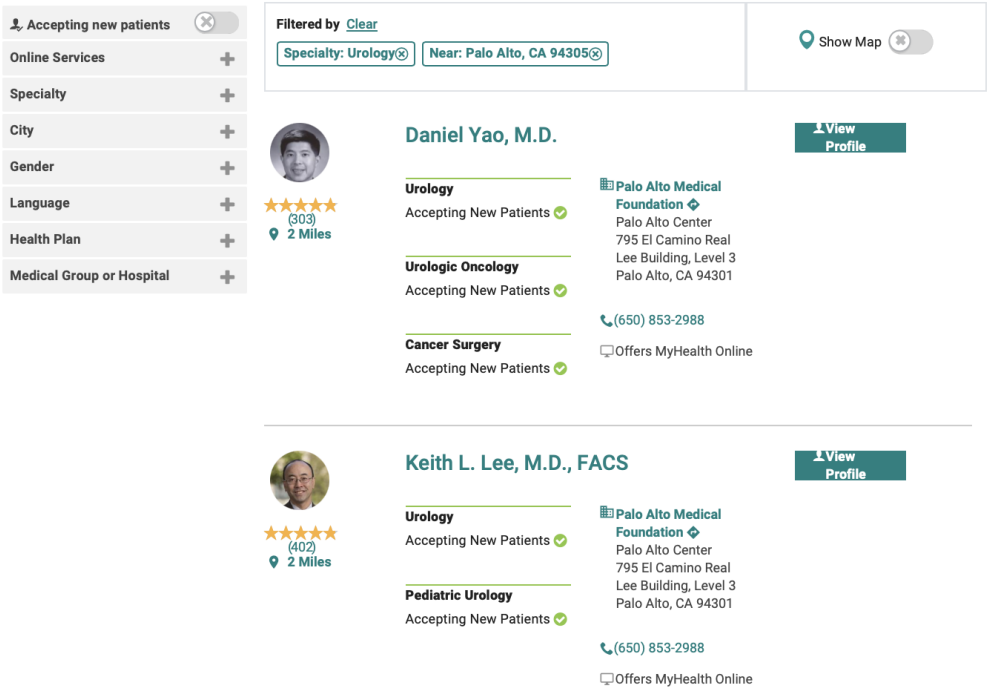


Figure B8: Search results for Sutter Health Urologists

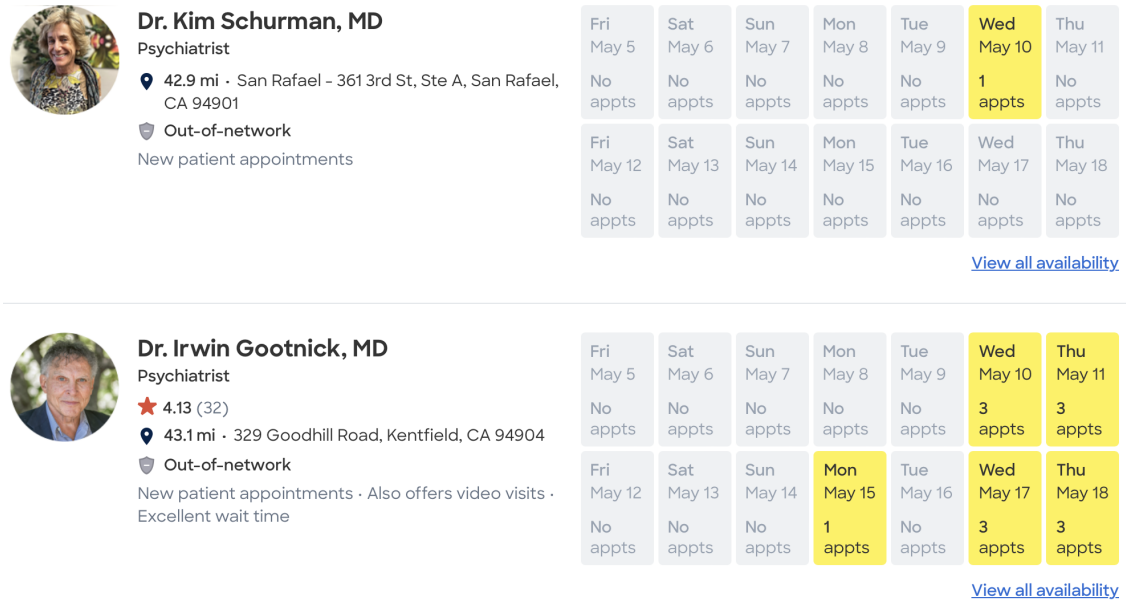


Figure B9: Search results for Zocdoc Psychiatrists

C. Experimental Results

C.1 Experimental Checks

Table C2: Demographic characteristics by treatment group (completes, $n = 2793$)

	Orthopedics ($n = 935$)	Urology ($n = 943$)	Psychiatry ($n = 915$)	Test	
Age, mean (SD)	42.6 (13.2)	42.9 (13.4)	42.6 (14.0)	$F = 0.14$	
Gender					
Male	463 (49.5%)	473 (50.2%)	470 (51.4%)	$\chi^2 = 0.65$	
Female	472 (50.5%)	470 (49.8%)	445 (48.6%)		
Race					
White	468 (50.1%)	469 (49.7%)	477 (52.1%)	$\chi^2 = 1.25$	
Black	467 (49.9%)	474 (50.3%)	438 (47.9%)		
Education					
Less than HS	6 (0.6%)	7 (0.7%)	9 (1.0%)	$\chi^2 = 2.45$	
High school	119 (12.7%)	126 (13.4%)	111 (12.1%)		
Some college	192 (20.5%)	185 (19.6%)	171 (18.7%)		
College+	618 (66.1%)	625 (66.3%)	624 (68.2%)		
Marital status					
Single	446 (47.7%)	458 (48.6%)	421 (46.0%)	$\chi^2 = 1.00$	
Married	385 (41.2%)	380 (40.3%)	384 (42.0%)		
Separated/divorced	88 (9.4%)	85 (9.0%)	84 (9.2%)		
Employment					
Employed	726 (77.6%)	731 (77.5%)	728 (79.6%)	$\chi^2 = 1.42$	
Unemployed	209 (22.4%)	212 (22.5%)	187 (20.4%)		
Insurance					
Insured	926 (99.0%)	933 (98.9%)	901 (98.5%)	$\chi^2 = 1.04$	
Uninsured	8 (0.9%)	9 (1.0%)	12 (1.3%)		
Health status					
Chronic condition	437 (46.7%)	439 (46.6%)	434 (47.4%)	$\chi^2 = 0.16$	
Prescription medications	434 (46.4%)	421 (44.6%)	436 (47.7%)	$\chi^2 = 1.71$	
Emergency room use	750 (80.2%)	782 (82.9%)	718 (78.5%)	$\chi^2 = 6.00$	*
Religion					
Christian	503 (53.8%)	538 (57.1%)	511 (55.8%)	$\chi^2 = 6.35$	
Jewish	10 (1.1%)	9 (1.0%)	13 (1.4%)		
Muslim	9 (1.0%)	15 (1.6%)	16 (1.7%)		
Atheist/Agnostic/None	384 (41.1%)	350 (37.1%)	346 (37.8%)		
Other	29 (3.1%)	31 (3.3%)	28 (3.1%)		
Political affiliation					
Republican	198 (21.2%)	195 (20.7%)	211 (23.1%)	$\chi^2 = 8.40$	
Democrat	432 (46.2%)	435 (46.1%)	417 (45.6%)		
Independent	254 (27.2%)	265 (28.1%)	259 (28.3%)		
Other/no preference	51 (5.5%)	48 (5.1%)	28 (3.1%)		

χ^2 test for categorical; one-way ANOVA F-test for continuous.

$^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$