Gender and the Market for Expertise: What can we Learn from Second Opinions?
Draft. Not for Broader Distribution

Mathijs de Vaan and Toby Stuart

Introduction

The market for expertise is enormous. The US legal market and the management consulting industry, for example, each approach 100 billion dollars in annual revenue. The market for medical services dwarfs both; it exceeds one trillion dollars.

Across markets for expertise, the literature provides striking evidence of gender differences in compensation and rates of career advancement. These differences exist against a backdrop of shifting occupational choices of women, with a vast uptick in the female share of graduates in law, medical, and business schools (Blau et al., 2013). Despite near gender equality in rates of entry into the advice-based professions, there remain persistent earnings disparities between male and female professionals. Studies have documented persistent gender gaps in earnings, promotion rates, and other career outcomes for attorneys (Beckman and Phillips 2005; Azmat and Ferrer 2017), physicians and surgeons (Zeltzer 2020; Altieri et al 2018), and scientists (Galotti and De Domenico 2019).

As contemporary scholars of the sociology of professions observe, expertise always unfold in relational contexts that are imbued with core, social processes (Azocar and Fereee 2017). Experts perform roles that entail creating and/or delivering guidance, which is embedded in specific social relationships and subject to the evaluative processes that pervade markets for uncertain goods. Within the division of labor in these markets, most practitioners serve as intermediaries between the production of knowledge and its situational application. Experts are the service providers and clients make consumption choices amidst evaluations of the (uncertain) quality of advice given. Therefore, the marketability of an expert's service is fundamentally reliant on the trust and confidence he or she engenders among current and prospective clients (Stehr and Grundmann 2011). The quintessence of trust in the quality of guidance in expert-client dyads is at the center of the argument and evidence in this paper.

We posit that an overarching reason for gender differences in career outcomes in the professions is that "buyers" in markets for expertise may be less trusting of guidance from female experts. If this supposition is correct–if some or all clients are more prone to question the advice of female experts–it directly undermines the market value of female-supplied expertise. This is because a central property in markets for expertise is that as a commodity, the value of professional advice declines in situations in which its reliability is judged to be in doubt, because the buyer must then hire additional experts to formulate a course of action. Another way to state this assertion is that questioning expertise will cause a downstream wage disparity because it necessitates purchases of second opinions. A significant body of theory suggest that his may be especially true if the buyers are male and the sellers are female.

In this article, we empirically examine gendered patterns in clients' acceptance of expert advice by studying the frequency of second opinions among patients that visit specialist providers in healthcare. Physician visits are an ideal context to examine these questions because we can truly observe population-level data with essentially no ambiguity about the gender of clients (patients) and experts (medical specialists). We first show that the probability of obtaining a second opinion decisively varies with the gender of the patient. We find that *female patients*, on average, seek second opinions at a higher rate than do male patients. However, this baseline difference masks the true dynamic in the data: the likelihood of securing a second opinion conditional on seeing a *female specialist* is much larger for male patients than for female patients. Moreover, compared to the decisions of female patients to obtain a second opinion, the decision of male patients is strongly guided by the gender of the specialist. Women seek second opinions at unconditionally higher rates, but male patients seeing female specialists are most likely to consult an additional physician.

Next, we examine the probability of switching to a specialist of the opposite gender, conditional on pursuing a second opinion. We find a similar pattern here: when the first specialist a patient sees is the opposite gender, male patients are *substantially* more likely to switch to a same-gender specialist than are female patients. Finally, we conduct a back-of-the-envelope calculation to assess the implications of these patterns on earnings for specialists. If we assume that patient spending that shifted to the second specialist represents the missed earnings by the first specialist and if we express these missed earnings in total earnings of the first specialist (because women earn much less than men), we find that the missed earnings for female physicians as a fraction of their total earnings are 15% higher than the missed earnings for male physicians because their expertise is questioned.

Theory

When making decisions that require specialized knowledge, actors often consult experts in an attempt to improve decision-making or comply with complex legal or regulatory requirements. For example, organizations hire management consultants and bankers for strategic and financial advice [2], legal matters are handled with counsel from attorneys [9], and patients consult physicians to monitor and treat health conditions[3].

The locus of the study of expertise in sociology has been work on the professions. Within this literature, expertise itself generally has been viewed as an attribution that arises from professional credentialing. Expertise is a quality that is sometimes assumed to exist when an expert is certified by a legitimated credentialing authority (Collins and Evans 2007). This focus has led to work on the professions as a type of organizational form, with jurisdictional boundaries demarcated by inter-profession jostling for the monopolization of esoteric knowledge (Abbott 1988). In consequence, some have argued that the sociology of the professions has placed the analysis of organizational and institutional forms ahead the content and context of what experts actually do (Azocar and Fereee 2017).

Against this backdrop, a second approach to studying expertise derives from the intellectual tradition of symbolic interactionism. In this view, expertise is less objective and codified. Rather, because expertise is created and shared in embedded market contexts characterized by myriad evaluative uncertainties, it is part and parcel to all of the nuance attached to social relationships. This contrasts to a view of expertise as an objective resource that is created and controlled by the professions (Hughes, 1994; Barley, 1989; Becker, 1970). It also connects the study of expertise to the enormous bodies of scholarship in the sociology of work, occupations, and economic markets. For instance, we know that in markets, evaluations of products of uncertain quality often default to assessments of the identity and characteristics of producers, because the latter is far easier to assess.

Given the market value placed on their advice, experts often garner compensation that places them in the upper strata of the income distribution [4]. Despite being well-compensated *on average*, however, there is substantial variation in the pay of male and female experts[5, 1, 10]. Prior research has identified a variety of mechanisms through which this pay gap has emerged including, but not limited to, outright gender bias [7], childcare expectations [8], and differential selection into higher versus lower paying job roles within professional service firms and occupations [6].

In this paper we build on prior work but focus on a general, under-emphasized mechanism for the gender pay gap in high-skill occupations. Specifically, we study gender patterns in the incidence of second opinions among medical specialists. Although not true in all instances, many second opinion arise in expert engagements because the client has less than complete trust in the counsel of the original advisor. Insofar as there are predictable and consistent gendered patterns in the incidence of second opinions, we believe that this phenomenon reveals one of the essential mechanisms that explains gender disparities in occupational outcomes for experts. In particular, we present evidence that strongly suggests that male clients are less likely to place full confidence in the expertise of female providers.

Context

Second opinions occur in healthcare for a number of reasons. First, patients are sometimes miss-assigned to specialists. In that case, they are re-referred (generally by the first specialist) to a second specialist that is a better match to the underlying health problem. Second, SOs also may shift patients from lower quality specialists to higher quality specialists, potentially improving health outcomes. Third, in complex cases, specialists themselves may prefer a second opinion, when a diagnosis or treatment protocol is ambiguous. But regardless of its impact on health outcomes, second opinions may also reveal preferences held by patients and implicitly, the trust that patients place on the recommendations of their physician. Patients often have discretion in choosing whether to accept the recommendations from the first-seen specialist or to pursue a second opinion. As these individual patient decisions aggregate, they may widen or narrow the gender pay gap in medicine.

In this paper, we describe a strategy to identify second opinions from claims data and detail several analyses to validate cases as second opinions. We then document gender differences in second opinion rates in a large sample of speciality visits.

Data

We examine the Massachusetts All Payers Claims Database (MA APCD). The MA APCD is collected and maintained by the Center for Health Information and Analysis (CHIA) and contains remarkably comprehensive information derived from the medical and pharmacy claims of virtually every resident in Massachusetts between January 1, 2010, and December 31, 2015. The data encompass in the vicinity of 1 billion distinct medical claims. Massachusetts requires all health insurers in the state to report detailed information on every medical and pharmacy claim they receive. CHIA collects these data and prepares them for use in research. For instance, CHIA processes the data to create a hashed identifier to link records of individuals that change insurance plans over time. The MA APCD contains multiple data files, but we mainly draw from the medical claims file. Data in the medical claims file include a physician identifier, a patient identifier, diagnosis codes, dates and locations of provider visits, identification of medical procedures performed, charged dollar amounts, and, importantly, referral information. The latter includes an indicator for whether the patient was referred to a given specialist and an identifier for the specific, referring physician.

Sampling

Data as comprehensive as the the state All Payers Claims Database have only become available recently, and there is no established method for identifying second opinions in medical claims.[1] We define the risk set for second opinion to the set of all first-time visits to specialist for a new medical condition. There are the set of index visits to specialists that have the potential to result in a second opinion.

From the medical claims data, we leverage the referral indicator to sample all instances in which a patient is referred to a specialist in specialty i for the first time[2] and label these the index visits.[3] We then limit these referrals to office visits to specialists (CPT codes 992**) because our goal is to identify cases in which a patient is sent to a specialist to obtain a medical diagnosis or treatment recommendation. In constructing the sample, we exclude three clinical scenarios. First, we eliminate all index cases that begin with emergency department visits where there is limited scope for patients to exercise discretion. Second, we further limit this sample to referrals originating from primary care physicians. This step is important because it removes most instances in which a patient is referred to a specialist that is a mismatch because the provider does not treat the focal diagnosis. In

[1] There is a survey-based literature on second opinions, although it is small relative to the importance of the phenomenon. To our knowledge, this is the first article to propose and validate a method for identifying second opinions in population claims data.

[2] For example, if a patient is referred to an orthopedist at time t, but the patient had seen an orthopedist in the past for the same or another health issue, we exclude that visit.

[3] The referral indicator is used in HMO and POS plans, but it also is populated in a few PPO plans. The indicator identifies the physician that provided the referral.

these instances, the first-seen specialist generally refers the patient to a second physician that has a clinical focus that matches the diagnosis. Third, we limit the sample to patients ages 18 years and older. The resulting sample includes almost 1.6 million referrals from PCPs to specialists that resulted in an office visit of the patient to a specialist. The sample by gender of the specialist in the index visit is shown in table 1.

Table 1: Descriptive statistics by gender of the specialist in the index visit

	Male specialist	Female Specialist
Patient age (Mean)	48.94	46.51
Female patient (%)	53.93	72.00
Specialist graduation year (mean)	1986	1993
Female PCP (%)	39.47	53.66
Charlson index score (Mean)	0.62	0.53
Insurance type (%)		
Health Maintainance Organization (HMO)	72.19	72.75
Point of Service (POS)	7.79	8.63
Preferred Provider Organization (PPO)	6.99	7.95
Medicaid	3.04	2.66
Exclusive Provider Organization (EPO)	1.87	1.81
Other insurance type	8.12	6.19
Provider specialty (%)		
Dermatology	16.05	34.56
Orthopaedic Surgery	16.20	2.96
Otolaryngology	7.51	5.25
Urology	8.10	1.68
Surgery	6.53	7.00
Gastroenterology	7.58	4.79
Ophthalmology	6.07	5.22
Neurology	5.07	5.84
Cardiovascular Disease	6.15	2.73
Obstetrics and Gynecology	1.67	9.47
Other specialty	19.07	20.50
Observations	1,210,491	385,778

Notes: This sample includes only index visits by adult patients who were referred by their PCP and had not previously seen a specialist in the focal specialty.

For each index visit, we record whether the patient is referred by their PCP to another specialist in the same speciality within 180 days of the index visit.[4] To be precise, we define a second opinion to occur when, conditional on having not consulted a specialist in field k in the past: a patient consults two physicians in the same medical specialty k in a 180-day window, and both appointments were established by a referral from a PCP. The

[4] Setting the window to 180 days was motivated by the fact that in some specialties it takes months before a patient can be seen by a specialist. In sensitivity analyses, we re-estimate the effects using different time windows and find comparable patterns.

resulting data are shown in table 2. The table shows that female patients and female specialists are over-represented in SO cases compared to no SO cases, which hints at a differential gender pattern in pursuing second opinions. Table 2 also shows that some specialties are over-represented in second opinion cases while others are under-represented. Finally, the table shows that about 4% of all visits result in a second opinion, but that there is variation by specialty.

Table 2: Descriptive statistics by whether the index visit results in a second opinion

	Second Opinion	No Second Opinion
Patient age (Mean)	47.77	48.37
Female patient (%)	59.31	58.25
Specialist graduation year (mean)	1988	1987
Female specialist (%)	25.55	24.11
Female PCP (%)	43.92	42.86
Charlson index score (Mean)	0.67	0.59
Insurance type (%)		
Health Maintainance Organization (HMO)	73.13	72.29
Point of Service (POS)	8.60	7.97
Preferred Provider Organization (PPO)	6.83	7.24
Medicaid	2.84	2.95
Exclusive Provider Organization (EPO)	1.48	1.87
Other insurance type	7.11	7.68
Provider specialty (%)		
Dermatology	19.91	20.55
Orthopaedic Surgery	17.87	12.81
Otolaryngology	6.75	6.97
Urology	7.94	6.49
Surgery	5.59	6.68
Gastroenterology	4.46	7.00
Ophthalmology	9.16	5.73
Neurology	5.14	5.26
Cardiovascular Disease	5.83	5.30
Obstetrics and Gynecology	5.56	3.48
Other specialty	11.78	19.72
Observations	60,557	1,535,712

Notes: This sample includes only index visits by adult patients who were referred by their PCP and had not previously seen a specialist in the focal specialty.

Validating the second opinion label

Because we do not have an externally validated ground truth about whether a second opinion as defined in our data is in fact a second opinion (versus a first-time visit for a second medical condition), we perform several tests to establish that the second opinions we have identified have a set of characteristics that distinguish them from other specialist visits. First, we assume that second opinions increase in likelihood with the severity of the health condition. To assess this, we examine average 1-year medical spending and the 1-year surgery probability associated with the the diagnosis listed by the specialist in the index visit. The intuition is that diagnoses that are expensive to treat or have a high probability of requiring surgery should see higher rates of second opinions.

To compute expenditure statistics, we identified all instances in which a Massachusetts resident j was diagnosed with diagnosis i for the first time and aggregated all allowed amounts on medical claims in the following year. We then computed the mean and the median of this sum, by diagnosis. The 1-year surgery probability was computed using the same strategy but rather than aggregating spending, we identified whether patients had surgery in the year following the diagnosis.[5] We then regress the mean 1-year spending, the median 1-year spending, and the 1-year surgery probability on a second opinion indicator. The results are shown in table 3. The estimates suggest that spending and surgery probability associated with the diagnosis of the patient are between 5% and 7% higher for cases we label as second opinions.

Table 3: Second opinions and the severity of a diagnosis

	Dependent variable:		
	Mean 1-year spending	Median 1-year spending	1-year surgery probability
	(1)	(2)	(3)
Second opinion	644.701***	284.634***	0.010***
	(39.079)	(14.731)	(0.001)
Constant	12,597.660***	4,109.473***	0.206***
	(6.853)	(2.519)	(0.0001)
Observations	1,596,269	1,596,269	1,596,269

Notes: The estimates in model 1 and model 2 can be interpreted as changes in dollar amounts.

An unknown fraction of second opinions occur because there is a mismatch between the expertise of the first-seen physician and a patient's health condition. There are many instances of sub-specialization within the major medical fields, and patients occasionally are mis-assigned in the referral process. Note that in constructing the data, this is exactly why we excluded all second opinions that result from a referral that was made by a specialist provider, rather than a PCP. However, it remains possible that the PCP is the referrer of record for insurance purposes, even if the suggestion to see a different physician originated from the specialist in the first visit.

To evaluate the possibility of such expertise mismatches, we compare the exact treatment histories of the specialist in the first visit and the specialist in the second visit. The intuition for this analysis is that, if patients are frequently referred to specialists who do

[5]We used the 'narrow' surgery definition developed by HCUP to identify CPT codes that indicate surgery: https://www.hcup-us.ahrq.gov/toolssoftware/surgeryflags_svcproc/surgeryflagssvc_proc.jsp.

not have the expertise to treat their conditions, the experience distribution of the first- and second-opinion physicians should be more dissimilar than two, same-specialty physicians chosen at random. Conversely, if the pair of physicians handle similar cases (and therefore represent suitable alternatives for first and second opinions for a given diagnosis), we should observe that their treatment histories are more similar than two, randomly selected providers.

We conduct this analysis by extracting the 1-year treatment history of the first and second specialist. This yields two vectors of frequencies of procedure codes performed by the specialists. We then compute the cosine similarity between these two vectors. To benchmark the resulting distribution of similarities we also compute distances between the 1-year treatment history of the first specialist and two, alternatively sampled specialists. In one approach, we randomly sample a specialist in the same speciality who also treated patients on the same day as the index visit. In a second approach, we match specialists on the diagnosis that they most commonly treat and then randomly sample a specialist in the same speciality who also treated patients on the same day as the index visit.

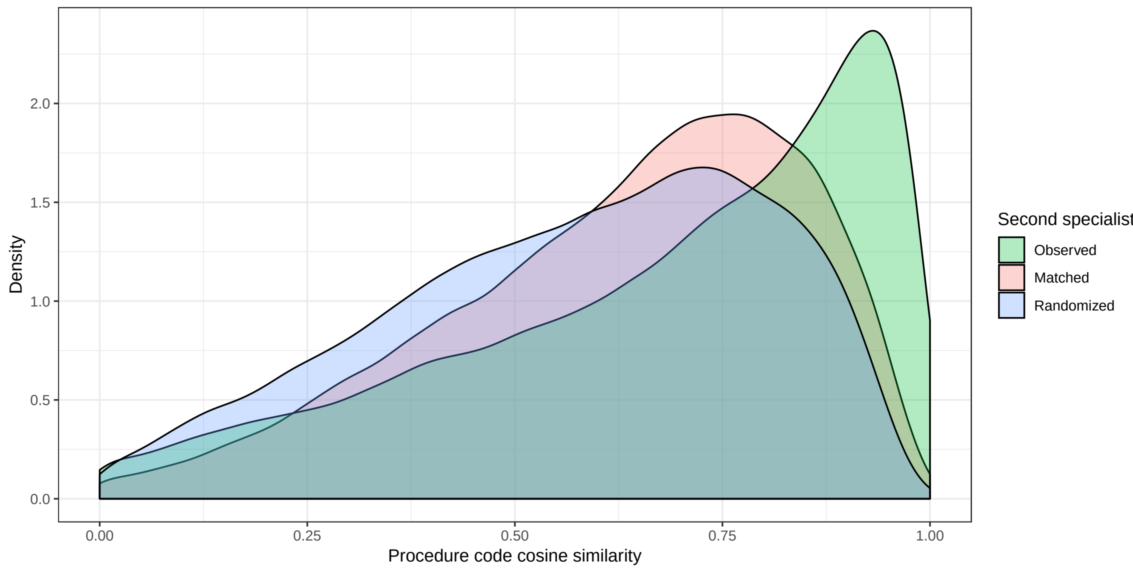
The three resulting distributions of similarity scores are shown in figure 1. Here, the results are remarkable. We find that on average, the actual pair of first- and second-opinion specialists is much more similar than the two pairs with counterfactually assigned specialists in the second visit. This shows us that mismatches between expertise and health condition in the original PCP referral to first specialists is unlikely to be common.[6]

The third analysis we conducted to evaluate the sample of second opinions aims to identify cases where new health issues emerge between the first and second specialty visit. If this new health issue requires expertise of the same specialty as the provider in the first visit, the second visit may not represent a second opinion. We therefore examine the rate at which, in cases we label as second opinions, the patient visits their PCP between visit 1 and 2 and is assigned a different diagnosis code than the one from the first specialist visit. We restrict these diagnoses to those that are reasonably likely to relate to a health issue that can be treated in the focal speciality. Specifically, the diagnosis needs to occur at least once in every 1,000 visits to a physician in that specialty. We find that about 7% of all second opinion cases occur in this circumstance. Note that this includes cases where the PCP simply inputs a diagnosis code for the same health issue but is less specific than the one assigned by the specialist. In other words, this final analysis also suggests that the majority of cases we identify as a second opinion are, in fact, those.

In the analyses that follow we will be using the sample of second opinion cases described in table 2, but we also conduct multiple sensitivity checks that exclude the cases that, based on our previous two analyses, are possible false positives. Specifically, we exclude cases we label as second opinions but could be new health issues, we exclude cases in which the imputed pair of specialists is more similar than the actual pair and we reduce the window size from 180 days to 3 months and 1 month.

[6] In the Robustness Checks section of the paper we add another strategy to evaluate the impact of the potential expertise mismatch issue that is based on Physician Compare's specialty taxonomy.

Figure 1: Distribution of cosine similarities between observed, matched and randomly assigned pairs of specialists.



Note: This graph shows the distribution of similarities between the treatment histories of the first and second specialist in green. The two counterfactual scenarios are shown in red and blue. Compared to the counterfactual distributions, the observed distribution is shifted to the right and clearly peaks close to 1, suggesting high levels of similarity.

Evaluating baseline homophily

We use the sample described in table 1 to estamate a choice model aimed at examining gender homophily in the first visit of the patient to a specialist. Both patient and PCP are likely to drive the choice for a specific specialist we account for similarity in gender between both the PCP and specialist and the patient and specialist. We estimate gendered preferences of both patients and PCPs using a conditional logit model for the probability of PCP j or patient i selecting specialist k. The identifying variation comes from differences within each patient's and PCP's choice set. As a result, any patient or PCP-level attributes are differenced out.

The data consist of an observation for each dyad (ij, k), with specialist and dyad (pairwise) characteristics X_{ik}, X_{jk}, and a binary outcome standing for whether the dyad is linked. In other words, each referral from a PCP to a specialist generates a realized PCP-specialist and patient-specialist dyad and our aim is to establish – for each referral – a set of counterfactuals consisting of specialists the PCP could have referred to and the patient could have visited. To establish the risk set of specialists, we compare each chosen specialist k with unchosen specialists k from the same Hospital Referral Region (HRR), from the same specialty of k and that practiced in the same month as the specialist to which j actually referred and i was referred to. This choice makes a conservative (weak) assumption about substitutability. Specifically, specialists in the same city and medical specialty are not assumed to be perfect substitutes. Rather, it only assumes that specialists from different markets or from different medical specialties are not substitutes. Variation in other characteristics that are included as controls (e.g., distance) capture the actual substitutability within those cells.

The results are included in table 4. Estimating the conditional logit is computationally expensive, so we break up the sample by year.[7] The results are consistent across years. Each model suggests that, all else being equal, patients are about 20% more likely to visit a specialist of the same gender. PCPs exhibit weaker preferences but are still 12 to 15% more likely to refer to a specialist of the same gender. When jointly interpreting the Same gender (Patient) coefficients and the Female specialist coefficients, the models suggest that while female patients are slightly more likely to select a male specialist over a female specialist, the difference is substantially more pronounced for male patients. Finally, experience and distance are important determinants of referrals too – the experience gradient suggests an inverted U shape with mid career specialists being in highest demand. In sum, these results suggest that, on average, male patients exhibit strong homophilous preferences in their selection of a specialist. If anything, these results stack the deck against finding any gender differences in our SO analyses, because male patients with stronger beliefs about male versus female specialists have selected into male specialists in the first stage of the referral process.

Table 4: Baseline homophily first visit

Dependent Variable:	Patient visited specialist/PCP referred to specialist				
	2010	2011	2012	2013	2014
	(1)	(2)	(3)	(4)	(5)
Variables					
Female specialist	-0.2813***	-0.2674***	-0.2357***	-0.2555***	-0.2384***
	(0.0043)	(0.0053)	(0.0056)	(0.0058)	(0.0065)
Same gender (Patient)	0.2355***	0.2318***	0.1966***	0.2019***	0.1894***
	(0.0045)	(0.0055)	(0.0058)	(0.0061)	(0.0068)
Same gender (PCP)	0.1529***	0.1228***	0.1291***	0.1255***	0.1437***
	(0.0042)	(0.0052)	(0.0055)	(0.0058)	(0.0065)
Grad. year pre 80	0.2395***	0.0671***	-0.0028	-0.0538***	-0.1031***
	(0.0056)	(0.0062)	(0.0065)	(0.0067)	(0.0075)
Grad. year 80-00	0.4113***	0.2453***	0.1933***	0.1488***	0.1065***
	(0.0051)	(0.0055)	(0.0056)	(0.0056)	(0.0060)
Distance (Km)	-0.0866***	-0.0862***	-0.0829***	-0.0819***	-0.0821***
	(0.0003)	(0.0004)	(0.0004)	(0.0004)	(0.0004)
Fit statistics					
# of visits	468,259	312,457	268,130	241,147	190,830
Observations	94,729,369	60,668,246	49,669,119	43,812,135	33,301,615
Pseudo R^2	0.15167	0.14842	0.14696	0.14613	0.14664

Notes: All models are estimated using conditional logits.
$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$ (Two-tailed tests).

[7] Sample size is smaller each year because we restrict our sample to only include patients who visit a specialist in a specialty for the first time.

Research design

In our main specifications, we estimate the dependence of the second opinion rate on the gender of the specialist and patient. One of the main concerns in such a specification is that specialist or pateint gender is correlated with unobserved variables that are also correlated with the error term (i.e. $corr(X_i, u_i) \neq 0$). Obvious candidates for such a variable include the quality of the specialist and health status of the patient. While we are able to include a rich set of controls based on the track record of the specialist and the health history of each patient, it is unlikely that conditioning on these variables fully satisfies the assumption that $corr(X_i, u_i) = 0$.

There are, however, strategies to address this concern. Most importantly, we are able to estimate models using physician fixed effects (FEs). We run two separate regressions, one for female and one for male specialists, to show that even if we hold quality of the specialist constant, men and women request second opinions at different rates.[8] In addition to the FE specification, we also leverage the idea that if second opinion rates were fully explained by differences in health status between male and female patients, then, all else being equal, the decisions to request a second opinion should be independent of the gender of the specialist seen first.

Results

In table 5 we show the regression results of six Ordinary Least Squares (OLS) regressions in which the outcome of interest is the decision to request a second opinion. Note that the regressions shown in table 5 as well as all other regressions in the remainder of the paper include several control variables: 1) age dummies to capture patient age, 2) graduation year dummies to capture specialist experience, 3) the Charlson score based on the 1-year medical history of the patient to capture health status, and 4) insurance plan dummies to capture plan specific incentives. The first model includes only the gender of the specialist and shows that female specialists see their patients request a second opinion more frequently than male specialists. The direction of the effect remains the same after including specialty and year fixed effects, but grows substantially in size after including the specialty fixed effects. This indicates that female providers are more likely to sort into specialties with lower second opinion rates. Models 4 and 5 include the gender of the patient and an interaction between gender of the patient and gender of the specialist. The results suggest that female patients, on average, request second opinions more frequently than male patients but that the decision-making of male patients is much more sensitive to the gender of the specialist. In model 6, we include PCP fixed effects, which do not alter the results.

In figure 2 we visualize the point estimates for each of the four possible gender pairings based on model 5 in table 5. The graph shows that the second opinion rate for male patients varies much more, depending on the gender of the specialist. Note that the male patient - male physician combination is used as the reference group in the regression and in figure 2.

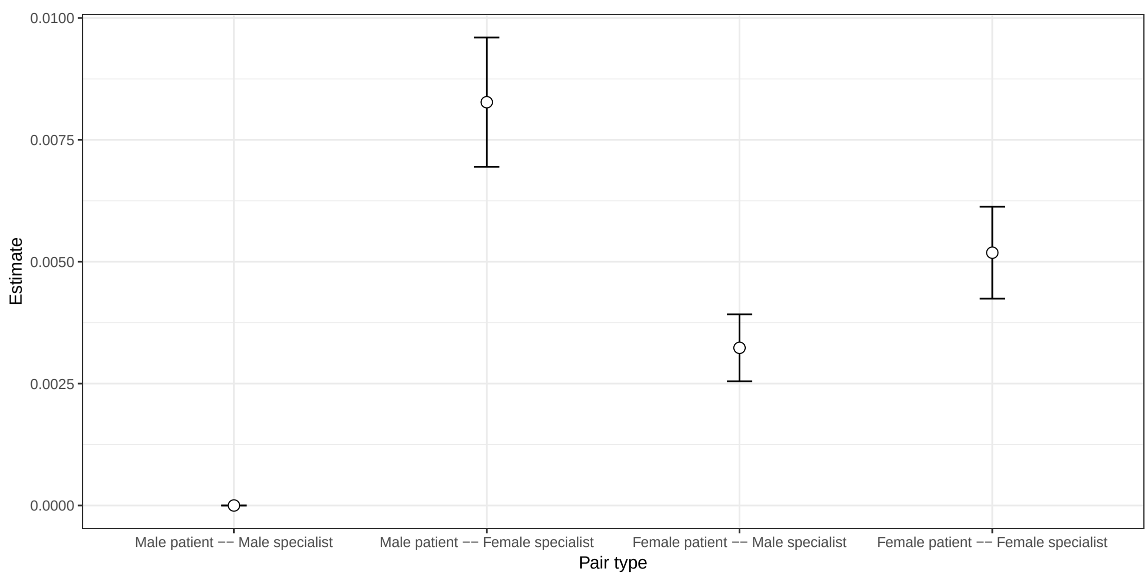
[8] Most patients are observed only once so including patient FEs is not feasible.

Table 5: Second opinion rate and gender

	Dependent variable:					
	Probability of second opinion after first consult					
	(1)	(2)	(3)	(4)	(5)	(6)
Female specialist	0.002***	0.004***	0.004***	0.004***	0.008***	0.008***
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.001)	(0.001)
Female patient				0.002***	0.003***	0.003***
				(0.0003)	(0.0004)	(0.0004)
Female specialist*Female patient					−0.006***	−0.006***
					(0.001)	(0.001)
Constant	0.043***	0.011*	0.012*	0.011*	0.010	
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)	
Specialty FEs	No	Yes	Yes	Yes	Yes	Yes
Year FEs	No	No	Yes	Yes	Yes	Yes
PCP FEs	No	No	No	No	No	Yes
Observations	1,596,269	1,596,269	1,596,269	1,596,269	1,596,269	1,596,269

Notes: All models are estimated using OLS. Logistic regressions yield similar results.
*p<0.05; **p<0.01; ***p<0.001 (Two-tailed tests).

Figure 2: Estimated effect sizes – extensive margin



Note: This graph shows the change in second opinion rate by pair type. The four types included are: Female specialist - female patient, Female specialist - male patient, Male specialist - female patient and Male specialist - male patient where the latter served as the reference group. The graph clearly shows that the behavior of male patients is substantially more affected by the gender of the specialist.

As mentioned earlier, one may argue that the patterns shown in table 5 may be driven by differences in specialist quality. To address this issue we estimate two regressions, one for male and one for female specialists, that includes specialist fixed effects. The results are shown in table 6. Although the coefficient estimates are similar, the effect sizes are different because the probability of a patient getting a second opinion is substantially more likely when the specialist in the first visit is female. Specifically, if the specialist is female,

female patients are 7.1% less likely than male patients to get a second opinion. If the specialist is male, male patients are 9.1% less likely than female patients to get a second opinion. Again, we find that the decision of male patients to pursue a second opinion is more strongly guided by the gender of the specialist than the decision of female patients to obtain a second opinion.

Table 6: Second opinion rate and gender

	Dependent variable:	
	Probability of second opinion after first consult by specialist gender	
	Female specialist	Male specialist
	(1)	(2)
Female patient	-0.003^{**}	
	(0.001)	
Male patient		-0.003^{***}
		(0.0004)
Specialist FEs	Yes	Yes
Year FEs	Yes	Yes
Observations	385,778	1,210,491

Notes: Both models are estimated using OLS. Logistic regressions yield similar results.
*p<0.05; **p<0.01; ***p<0.001 (Two-tailed tests).

Next, we examine whether, conditional on seeing a specialist for a second opinion, patients make different choices about the gender of the specialist. We are particularly interested in the following thought experiment: conditional on seeing a male or female specialist first, do male and female patients switch to the other gender at different rates? Table 7 shows the regression coefficients of a model predicting the probability that patients getting a second opinion see two specialists with different genders in the two visits. The table shows that there is less gender switching when the first specialist is male compared to female. It also shows that men are much more likely to switch gender than women if the specialist in the first visit was female. And although female patients are more likely to switch gender when the first specialist is female, the discrepancy is much less pronounced compared to male patients. These differences are visualized in figure 3. In sum, this evidence is consistent with patient decisions to request a second opinion being less driven by gender when the specialist is male. It is also consistent with men holding much stronger preferences for male specialists than women hold for either male or female specialists.

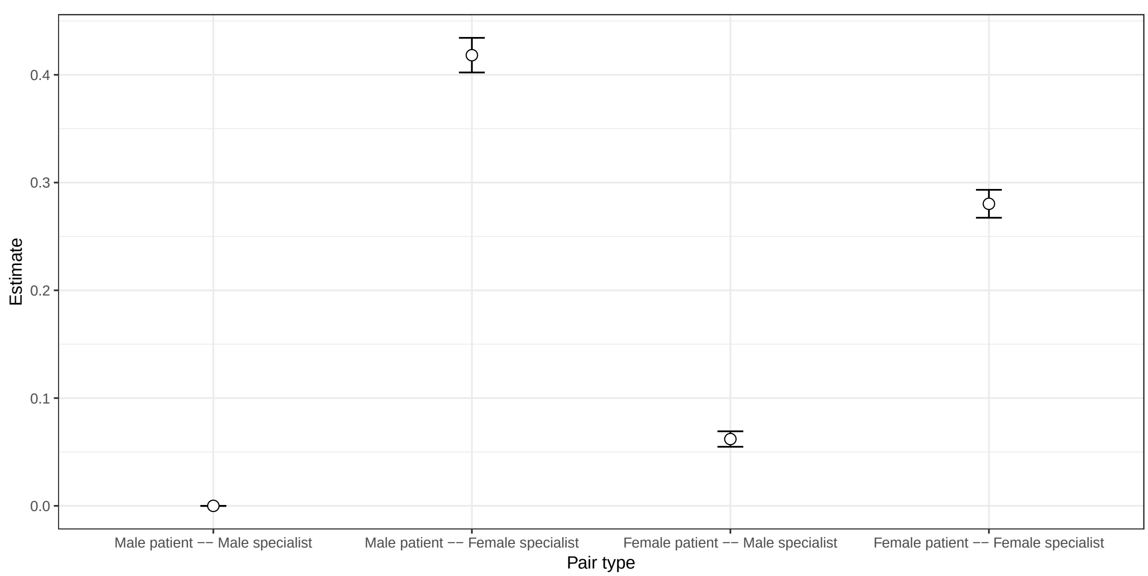
Finally, we conduct a back-of-the-envelope calculation to assess the implications of these patterns on earnings for specialists. First note that when a patient obtains a SO the patient obviously may choose to return to the original specialist for treatment. The data suggest that this is rare: on average, patient spending on services from the second specialist is 80% higher than spending on services of the first specialist if the patient goes in

Table 7: Gender switching in second opinions

	Dependent variable:	
	Probability of switching specialist gender	
	(1)	(2)
Female specialist	0.287***	0.418***
	(0.005)	(0.008)
Female patient		0.062***
		(0.004)
Female specialist*Female patient		−0.200***
		(0.010)
Constant	−0.006	−0.052
	(0.065)	(0.065)
Specialty FEs	Yes	Yes
Year FEs	Yes	Yes
Observations	60,557	60,557

Notes: Both models are estimated using OLS. Logistic regressions yield similar results.
*p<0.05; **p<0.01; ***p<0.001 (Two-tailed tests).

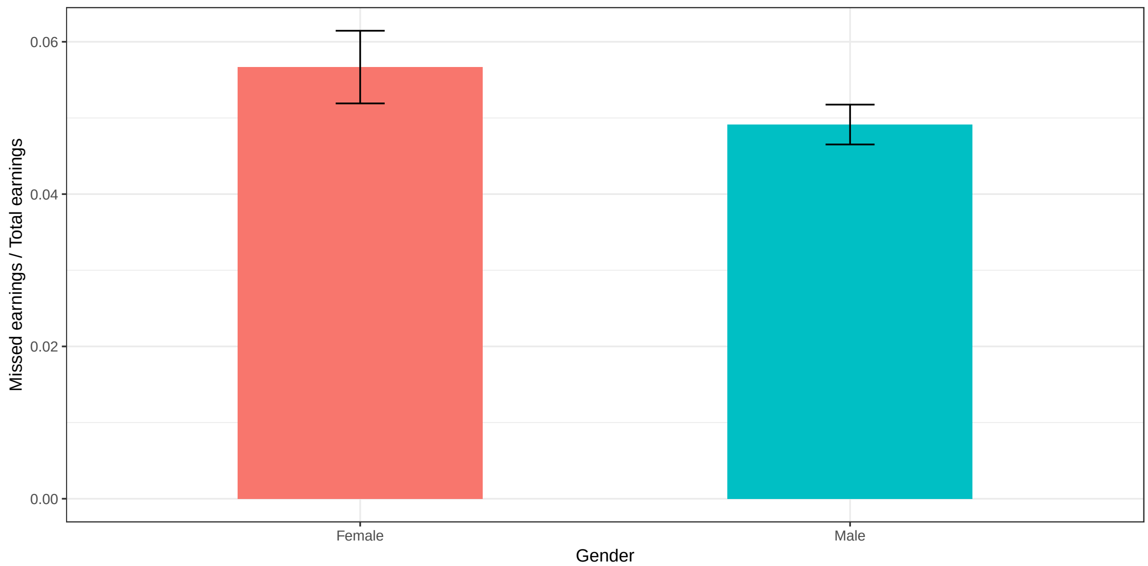
Figure 3: Estimated effect sizes – intensive margin



Note: This graph shows gender switching for the sample of second opinions.

for an SO. Also, if we compare spending on services from specialists in the first visit between SO and non-SO cases, specialists in non-SO cases bill 42% more to the patient than

Figure 4: Differences in missed earnings



Note: This graph shows the difference in "missed earnings" expressed as a fraction of total earnings for female specialists compared to males. The 15% difference in the mean is statistically significant (t = 2.72).

specialists in SO cases. This suggests that one may think of the spending on the second specialist as the "missed earnings" for the first specialist. Therefore, if we assume that patient spending that shifted to the second specialist represents the missed earnings by the first specialist and if we express these missed earnings in total earnings of the first specialist (because women earn much less than men), we find that the missed earnings for female physicians as a fraction of their total earnings are 15% higher than the missed earnings for male physicians because their expertise is questioned. These differences are shown in figure 4.

Readmissions

In table 8, we estimate the difference in the 30-day readmission rates between male and female specialists, for both male and female patients. Readmission rates have been targeted by policy interventions in an attempt to reduce costs of health care and improve health outcomes ody2019decreases. Readmissions represent an undesirable outcome that has been shown to result from substandard care and has been used by health researchers to evaluate quality of physicians, hospitals and health systems. Here, we examine differences in adjusted 30-day readmission rates between male and female specialists. We record a readmission for every index visit by identifying patients who were admitted to a hospital in the year following the index visit and were readmitted to the hospital within 30 days of the initial admission date. We adjust our estimates by patient age and the Charlson comorbidity score and we include specialty fixed effects. In model 1, we simply compare readmission rates between male and female specialists and show that patients of female specialists are

9% less likely (the base rate is 1.5%) to experience a readmission in the year following the index visit. Female patients are 8% less likely than male patients to be readmitted in the year following the hospitalization. In model 2 we include an interaction between the gender of the specialist and the gender of the patient. The results show that female patients treated by female specialists (compared to female patients seen by male specialists) are less likely to experience readmissions. Male patients, however, experience indistinguishable readmission rates when treated by a female specialist compared to treatment from a male specialist. In model 3 and 4, we examine robustness of these results by 1) limiting the sample to only index visits that did not lead to an SO and 2) limiting the sample to only specialists that had at least one SO event in the sample. The results are consistent with the results presented in model 2. In sum, this evidence is inconsistent with the idea that male specialists have male specific qualities that would benefit male patients that they treat.

Table 8: Readmission rates by specialist gender

	Dependent variable:			
	Probability of 30-day readmission in year following specialist visit			
	Full sample		Non-SO visits	Specialists w/ SOs
	(1)	(2)	(3)	(4)
Female specialist	-0.001^{***}	0.0001	-0.00003	0.0001
	(0.0002)	(0.0004)	(0.0004)	(0.0004)
Female patient	-0.001^{***}	-0.001^{***}	-0.001^{***}	-0.001^{**}
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Female specialist * Female patient		-0.002^{***}	-0.002^{***}	-0.002^{***}
		(0.0005)	(0.0005)	(0.0005)
Observations	1,596,269	1,596,269	1,535,712	1,533,424

Notes: All models are estimated using OLS. Logistic regressions yield similar results. Model 4 uses gender of the firsty specialist, model 5 uses gender of the second specialist.
*p<0.05; **p<0.01; ***p<0.001 (Two-tailed tests).

Robustness checks

In this section we show the results for various robustness checks. In table 9 we show the results for the same regression presented in table 5, model 5, but we exclude cases we labeled as second opinions but may not be second opinions. Specifically, in model 1 we exclude cases where the observed specialists are less similar than the imputed pair. In model 2, we leverage differences between two specialty taxonomies. The specialty taxonomy used throughout the paper comes from the Centers for Medicare and Medicaid Services (CMS) and by design the specialist in the first and second visit have the same specialty. Physician Compare, which is another data source that provides specialty data, gives physicians the opportunity to list more fine-grained specialty data. We therefore exclude those cases where the Physician Compare specialties are not the same for the first and second specialist. The results are shown in model 2 of table 5. In model 3, we exclude cases where patients see their PCP in between the first and second specialist visit and are assigned a different diagnosis that is plausibly treated by a specialist in the specialty. In models 4 and 5, we reduce the window after the first visit to 3 months and to 1 month, respectively. While the effect sizes change slightly, the patterns described earlier in this paper remain stable.

A second robustness check we conducted addresses the issue of left censoring. Our data start on January 1, 2010. For individuals who have their 'first' specialist visit in 2010, we do not have a full year of data to establish that they have not seen a specialist in the focal specialty and we lack the data to compile a 1-year medical history to capture health status. We therefore re-estimated table 5, model 5 but we excluded all cases in which the first visit took place before January 1, 2011. Note that this removes a large number of cases from our sample. This is by design since our sampling restrictions (i.e. only first visits to a specialty), frontload many of our cases in the sample. However, despite the drastic reduction in sample size, the effect sizes remain remarkably stable and in the same direction.

Table 9: Second opinion rate and gender

	Dependent variable:				
	Probability of second opinion after first consult				
	(1)	(2)	(3)	(4)	(5)
Female specialist	0.006***	0.008***	0.006***	0.006***	0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.0004)
Female patient	0.003***	0.003***	0.002***	0.002***	0.0004
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0002)
Female specialist*Female patient	−0.004***	−0.005***	−0.005***	−0.005***	−0.003***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Constant	0.002	0.012*	0.0002	0.005	0.001
	(0.004)	(0.005)	(0.004)	(0.005)	(0.003)
Specialty FEs	Yes	Yes	Yes	Yes	Yes
Year FEs	Yes	Yes	Yes	Yes	Yes
Observations	1,571,075	1,585,755	1,581,769	1,580,248	1,559,836

Notes: All models are estimated using OLS. Logistic regressions yield similar results.
*p<0.05; **p<0.01; ***p<0.001 (Two-tailed tests).

Conclusion

In this paper we demonstrate that male and female patients differ substantially when digesting medical advice and determining a course of action to address their health issues.

Table 10: Second opinion rate and gender

	Dependent variable:
	Probability of second opinion after first consult
Female specialist	0.007^{***}
	(0.001)
Female patient	0.003^{***}
	(0.0004)
Female specialist*Female patient	-0.007^{***}
	(0.001)
Constant	-0.014^{***}
	(0.002)
Specialty FEs	Yes
Year FEs	Yes
Observations	1,093,631

Notes: All models are estimated using OLS. Logistic regressions yield similar results.
*p<0.05; **p<0.01; ***p<0.001 (Two-tailed tests).

Specifically, while female patients choose to have their medical condition evaluated by a second specialist more frequently than male patients, the decision making process of male patients seem to be driven more by the gender of the first specialist than the decision making process of female patients. Preliminary results highlight the potential consequences of these gendered preferences: if second opinions shift patients from the first specialist to the second, female specialists face relatively high missed earnings compared to male specialists.

References

[1] Marianne Bertrand, Claudia Goldin, and Lawrence F Katz. Dynamics of the gender gap for young professionals in the financial and corporate sectors. *American economic journal: applied economics*, 2(3):228–55, 2010.

[2] Nicholas Bloom, Aprajit Mahajan, David McKenzie, John Roberts, et al. Do management interventions last? evidence from india. *American Economic Journal: Applied Economics*, 12(2):198–219, 2020.

[3] Gil Eyal. For a Sociology of Expertise: The Social Origins of the Autism Epidemic. *American Journal of Sociology*, 118(4):863–907, January 2013. Publisher: The University of Chicago Press.

[4] Joshua D Gottlieb, Maria Polyakova, Kevin Rinz, Hugh Shiplett, and Victoria Udalova. Who values human capitalists' human capital? healthcare spending and physician earnings. Technical report, US Census Bureau, Center for Economic Studies, 2020.

[5] Mary C Noonan, Mary E Corcoran, and Paul N Courant. Pay differences among the highly trained: Cohort differences in the sex gap in lawyers' earnings. *Social forces*, 84(2):853–872, 2005.

[6] Susan P Phillips and Emily B Austin. The feminization of medicine and population health. *Jama*, 301(8):863–864, 2009.

[7] Heather Sarsons. Interpreting signals in the labor market: Evidence from medical referrals. 2017.

[8] Eliot L Sherman. Discretionary remote working helps mothers without harming non-mothers: Evidence from a field experiment. *Management Science*, 66(3):1351–1374, 2020.

[9] Deepak Somaya, Ian O Williamson, and Xiaomeng Zhang. Combining patent law expertise with r&d for patenting performance. *Organization Science*, 18(6):922–937, 2007.

[10] Dan Zeltzer. Gender homophily in referral networks: Consequences for the medicare physician earnings gap. *American Economic Journal: Applied Economics*, 12(2):169–97, 2020.