Gender Homophily in Referral Networks:
Consequences for the Medicare Physician Pay Gap

Dan Zeltzer (Princeton)

24 December 2015, 11:15-12:30, bld. 72, room 465

Abstract: Female physicians—now a quarter of active U.S. doctors—still work puzzlingly less than their male counterparts. This paper suggests an explanation: gender homophily in referral networks (more same-gender referrals). I model the referral networks forming as doctors choose specialists, showing that homophily is explained by both biased preferences and sorting. I then suggest new ways to identify both types of biases in directed networks. Analyzing administrative data on more than 100 million Medicare physician referrals from 2008–2012, I find significant gender homophily, due mostly to preferences, not sorting. As most referrals are still made by men, homophily lowers demand for female specialists, explaining 10% of the average within-specialty workload gap, and contributing to the absence of women from lucrative specialties that rely on referrals from men. Results imply that increased participation of female physicians facilitates further integration into related specialties, and more broadly, that gender- biased interactions contribute to the propagation of occupational inequalities.

Gender Homophily in Referral Networks: Consequences for the Medicare Physician Pay Gap

Dan Zeltzer[*]

October 2015

Current Version Here

Abstract

Female physicians—now a quarter of active U.S. doctors—still work puzzlingly less than their male counterparts. This paper suggests an explanation: gender homophily in referral networks (more same-gender referrals). I model the referral networks forming as doctors choose specialists, showing that homophily is explained by both biased preferences and sorting. I then suggest new ways to identify both types of biases in directed networks. Analyzing administrative data on more than 100 million Medicare physician referrals from 2008–2012, I find significant gender homophily, due mostly to preferences, not sorting. As most referrals are still made by men, homophily lowers demand for female specialists, explaining 10% of the average within-specialty workload gap, and contributing to the absence of women from lucrative specialties that rely on referrals from men. Results imply that increased participation of female physicians facilitates further integration into related specialties, and more broadly, that gender-biased interactions contribute to the propagation of occupational inequalities.

JEL classification: J72; I11; L14

Keywords: homophily, referrals, networks, gender, discrimination, physician markets

1 Introduction

Despite the successful and rapid entry of women into medicine, female physicians, like females in other occupations, still have substantially lower earnings: women work less than male physicians, and practice lower paying specialties. Why would they so choose, after long and costly investment in education and training as doctors? This question reflects a broader concern that despite the great convergence in education and earnings females still fall behind, for reasons not well understood (Goldin, 2014). Some argue for supply side explanations: women have fundamentally different preferences (Croson and Gneezy, 2009), opt for less competitive jobs (McGee et al., 2015), and are held back by family-related career interruptions and their lasting consequences (Bertrand et al., 2010). Others contend that not all differences are voluntary, as quasi-experimental evidence exists for discrimination by gender (e.g. Goldin and Rouse, 2000). Yet it is difficult to support, let alone quantify, demand-side explanations for the earning gap because data typically reveal only individual characteristics, but not the external differences in opportunities individuals might face because of their gender.

This paper utilizes unique data on the network of referrals between over half a million U.S. Medicare[1] physicians and their payments for 2008–2012, to propose a new demand-side explanation for the pay gap: female physicians receive fewer referrals due to gender homophily (i.e., physicians referring to their same gender). I find evidence for significant gender homophily in Medicare referrals, beyond what is expected from random sorting. Modeling the formation of referral networks, I show that homophily sustains inequality by reducing demand for the minority. And I find that homophily indeed contributes to the Medicare physician pay gap through two channels: first, by reducing demand for female specialists, and second, by discouraging their entry to lucrative specialties relying on referrals from men. Insights extend to disparities on dimensions other then gender, and to any domain where referrals (or connections) are important, such as hiring, where employee referrals are frequently used by firms as an informal way to screen and monitor new employees (Kugler, 2003; Antoninis, 2006; Dustmann et al., 2011; Burks et al., 2014). More broadly, this paper demonstrates how interactions between individuals in professional networks (their "networking") provide a key to the understanding of propagation and perpetuation of inequalities.

Since Medicare payment schedule is common to all physicians, Medicare data are well suited for the study of pay differences due to difference in specialization and in workload, free from the difficulties involved with wage differentials in other settings. Broadly, one can

---

[1]Medicare is the U.S. federal health insurance program for the elderly patient with certain disabilities, and patient with kidney failure.

decompose gender differences in pay to three components: differences in the type of work performed, differences in quantity of work, and, differences in the per-unit pay. And while the popular debate of gender pay disparities is focused on "cents per dollar", per-unit differences, much of the great convergence in earning between genders over time, as well as much of still existing gaps come from the other two factors: women working more, and women entering previously male occupations. Understanding gender workload and specialization is therefore important to understand what path forward would facilitate further pay convergence. Decomposing the Medicare physician earnings gap shows that it confirms earlier findings that female physicians earn much less than male physicians (Weeks et al., 2009; Lo Sasso et al., 2011; Esteves-Sorenson et al., 2012; Seabury et al., 2013). Unconditional on any attributes, a female physician annual payments are 48% lower than those of a male physician. Accounting for specialty differences and career interruptions (Bertrand et al., 2010) explain a significant part, but as much as half of this gap remains unexplained.

Confidential data on more than half a billion Medicare physician services reveal how gender-biases in the work flows between physicians, through referrals, contribute to earnings gaps. Data allow for scrupulous empirical analysis of homophily: a large directed and weighted network spanning many local markets is observed over time. This setup is quite exceptional in that it offers more than just detailed individual data—it records a vast web of micro-level interactions of *between* individuals, with clear economic outcomes. Such data reveal how disparities arise from biased interpersonal interactions. Such interactions may indeed be present in many other settings, but are rarely so systematically recorded[2].

To study homophily, I model local referral networks as forming when doctors (upstream) choose specialists (downstream) to refer to. In these networks, a link exists between a doctor and specialist if one referred patients to the other in a given year, and its weight depends on the volume of work referred. The model builds on Currarini et al. (2009), but extends the analysis from social to referral networks, which are directed. In the model homophily is a combination of both preferences and sorting: doctors may prefer to work with same-gender specialists but may also face biased opportunity pools[3]. It yields testable predictions that distinguish between these mechanisms, and it illuminates their implications for the pay gap. The model has two main predictions: First, the minority receives fewer referrals. Second, unlike sorting, preference biases generate stronger homophily in larger markets: In large markets, extremely biased gender pools—which mask preferences—are less probable.

---

[2]The use of administrative Medicare data to study referral networks has been validated by Barnett et al. (2011), who showed that recorded referrals coincides well with self-reported relationships.

[3]Sorting is causing systematic differences in the opportunity pools to which men and women are exposed. Sorting may be reflecting segregation, or clustering of physicians by gender (see Massey and Denton, 1988; Echenique and Fryer, 2007).

As one doctor's referrals are another's income, the link between gender homophily in referrals and the gender pay gap seems intuitive: with men handling most of outgoing referrals, homophily implies more patients are referred to male specialists and fewer to female specialists. The model helps refine this intuition: it holds only if observed homophily is due to gender-biased preferences, but not if it is due to sorting, as sorting only affects which referrals women get, not how many. Moreover, with gender-biased preference the demand for female specialists depends on the gender composition of both upstream and downstream markets. Female specialists' demand suffers twice: first, there are fewer female doctors, resulting in fewer referrals to females, and second, there are fewer female specialists, making it easier to opt for a male one.

To identify homophily and gender-bias in referrals one needs to account for heterogeneity in the propensity to refer or receive referrals, as discussed by Graham (2014)[4]. I propose a new measure of homophily that compares the fractions of referrals male and female doctors made to males. Unlike previous measures that compared referral rates with the fraction of male specialists, this measure (*Directed Homophily*) is robust to heterogeneity in individual physicians' propensity to refer or receive referrals. Tight control for heterogeneity is further facilitate by comprehensive and standardized reporting of physician characteristics.

A second identification concern when estimating homophily is to distinguish the role of gender-biased preferences from homophily driven by sorting, namely from physicians being over-exposed to their same gender. To rule out sorting I take three complementary approaches: I estimate link formation with controls for rich pairwise characteristics, estimate homophily by comparing referrals of physicians who face similar opportunity pools, and directly test theoretical predictions that are unique to preference-driven homophily. Comparison of referral rates by doctors affiliated with the same hospital provides one way to estimate the presence of gender bias in referrals, as such doctors face similar opportunity pools. Another way is to structurally estimate the impact of gender on referral probability by comparing realized and unrealized referrals, while controlling for pairwise characteristics, including: distance, shared hospital affiliation, similar experience, and shared medical school[5]. Finally, two of the model's predictions: that homophily increases with market size and that it generates demand effects, are shown to be driven by biased-preferences, not sorting, and thus provide another way to test for their existence.

---

[4]Graham (2014) also suggests a method to account for heterogeneity, yet his method is restricted to dense networks (where node degree is proportional to network size), while physician referral networks are sparse

[5]To deal with the technical challenge that referral networks are large and sparse, I sample and match realized and unrealized links. This design is known as *choice-based*, *endogenous stratified*, or *case-control* sampling (Manski and Lerman, 1977; King and Zeng, 2001). Sampling is not required for the estimation of homophily, as it only considers gender correlation within the population of realized links.

I find significant gender homophily in referrals, mostly due to preferences: All else equal, doctors are 10% more likely to refer within- rather than between-gender. Across all markets, female doctors refer to female specialists a third more than males do (19% women-to-women compared with 15% men-to-women), i.e. the average difference between the genders is 4%. This average masks heterogeneity between markets in the availability of female specialists: as predicted by the model, homophily is greater in larger markets. Gender-homophily is stronger among older doctors, and same-gender relationships are also relatively more persistent than between-gender ones. Some sorting does exist, but it only explains a quarter of the overall gender homophily. I also find age-homophily: physicians are 1% more likely to refer to a physician one year closer to them in age. Hence there is a comparable effect between being of a different gender and having a ten-year age difference.

Data on physician first names provide an opportunity to test for prejudice, using a quasi-audit strategy: I check whether homophily is weaker when specialists' names are gender ambiguous (e.g., Alex or Robin) compared with names that reveal gender (e.g., Robert or Jennifer). I find homophily is unchanged, regardless of how revealing are names. This rules out strong-form prejudice, one based on names alone. The emerging picture is not one of ill-disposed individual biases, but rather one in which a widespread tendency to prefer similar others leads to an overall unfavorable environment for the female minority.

Next, I turn to directly testing the predicted effects of homophily, using reduced form specifications with longitudinal data. I find homophily has significant impact on physician pay: more male primary care physicians in a market lead to higher pay for male specialists and lower pay for female specialists. This result is obtained using a monthly panel of individual payments for 2008–2012, with specialist fixed-effects, so it is robust to unobserved heterogeneity in specialists' labor supply, in particular due to other known differences between the genders. The impact of homophily on the workload gap is fairly large, and comparable to that of gender differences in career interruptions.

The overall impact of homophily on the pay gap is in fact larger, since demand disparities discourage female entry to higher-paying specialties, which rely more heavily on referrals: a "boys' club" effect. To test this effect, I leverage the interdependence of medical specialties to construct demand shifters: with gender homophily, female entry to a medical specialty impacts females in other specialties differently, in proportion to the how many referrals flow between them. Using retrospective cohort data from 1965–2005, I find evidence that female entry has indeed been discouraged by the paucity of female referrals. Notwithstanding its adverse effects on the physician gender pay-gap, I find no effects of homophily on patient mortality or cost, suggesting the clinical appropriateness of referrals is not compromised by homophily.

The findings that women in upstream specialties induce positive externalities for women in downstream specialties have two implications: First, helping female doctors further integrate into primary care specialties would facilitate, through referrals, female integration into higher-paying specialties. In particular, it would promote female entry to specialties where much of the work depends on referrals, such as most surgical specialties, where there are currently still very few women. Second, in the longer run, as recent female entrants gradually transforms the gender composition of the physician labor force, the part of the pay gap due to homophily is expected to vanish, or even reverse. Referrals originating in a gender-balanced doctor population would eliminate on average 10% of the within-specialty gender pay-gap.

This paper proceeds as follows: I discuss related literature (Section 1.1); describe the data (Section 2); decompose the gender pay-gap in Medicare (Section 3); define, model, and estimate gender homophily in Medicare referrals (Section 4); estimate homophily's impacts on pay disparities (Section 5); consider prejudice and patient outcomes (Section 6); and conclude (Section 7).

1.1   Related Literature and Background

Homophily, the tendency of individuals to connect more with similar others, is a robust phenomena, which has been documented by numerous studies in sociology, in many different contexts: friendship networks, organizations, weaker acquaintances, and on-line networks; it has been documented on different dimensions including: gender, age, race, religion, and educational level (for a comprehensive survey see McPherson et al., 2001). This paper is part of a nascent economic literature studying homophily, its causes and implications, and the challenges to identifying it. Most related to the current study, Currarini et al. (2009), show both homophily and degree disparity between types can be explained by an interplay of preferences and sorting, an insight used as a basis for my model; Graham (2014) shows homophily could be an artifact of heterogeneity and develops a method to account for it in dense graphs, which motivates the use of Directed Homophily to deal with heterogeneity. Another related work by Leung (2014) shows that inference about link formation is possible using a single network observation, based on the presence of effectively isolated parts in the network. Leung also uses a public version of the Medicare referrals data for illustration of his method, and shows that referrals between primary care doctors exhibit homophily on distance and institutional affiliation, and partly reflect reciprocity. He does not consider gender. Few papers study the impact of homophily. One is Golub and Jackson (2012), which shows that clustering due to homophily slows down information dissemination and learning.

The current paper adds to that literature the idea that homophily leads to propagation of inequality in networks, and documents it for a large professional network.

Central to this paper are interactions between physicians through referrals. Referrals are ubiquitous in medicine, and are commonly used to resolve diagnostic or therapeutic dilemmas and to manage conditions that fall outside a physician's scope of practice (Forrest et al., 2002). For patients, referrals provide information about available providers: Although in Medicare it is not obligatory to obtain a referral to see a specialist, a third of all physician encounters are referred. Referrals are most central to the practice of primary care physicians, but are also a routine part of specialized care (Barnett et al., 2012a). Furthermore, referral use is increasing: Barnett et al. (2012b) show the fraction of ambulatory visits resulting in referrals has almost doubled between 1999–2009. Economists have mostly studied referrals in the context of their impact on care value and cost, in light of conflicting incentives of patients, doctors, and insurers. For example, Ho and Pakes (2014) show that capitated insurance plans influence physicians referral decisions towards sending patients to cheaper, more distant, hospitals for childbirth. In contrast, the current focus is on the impact on referrals on physician labor demand. In its focus it is close to Johnson (2011) who showed the career path of cardiac surgeons in Medicare is affected by learning about their quality by doctors. Doctors are well aware of the importance of referral relationships for clinical and financial outcomes. There is plenty of on-line advice for physicians regarding how to establish and develop such relationships[6]. As per the reasons for choosing specific physicians to refer to, surveys show physicians consider the main factor after clinical appropriateness to be between-physician communication (Barnett et al., 2012a). Communication is also an oft mentioned process-measure for referral quality (Choudhry et al., 2014; Gandhi et al., 2000; Song et al., 2014; Mehrotra et al., 2011). The importance of communication for referrals hints as to why physicians may pick others of similar gender, age, or other attributes, since communication could be facilitated by such similarity.

In data and methods, this paper is part of a fast growing literature that uses large administrative data to address questions that previously seemed beyond reach. Extensive data support identification not just in revealing variables of interest, directly or through linkages across databases, it also opens new avenues for identification by allowing researchers to solve nested prediction or classification problems. For example, Currie et al. (2015); Currie and MacLeod (2013) demonstrate how large administrative databases of medical records can be used to predict the appropriate course of medical treatment (essentially a classification

---

[6]For example, "Three Keys to Setting Up a Referral Network for Your Practice", `www.physicianspractice.com/blog/three-keys-setting-referral-network-your-practice`. Accessed August 2015.

problem), a benchmark against which diagnostic skills can then be evaluated. In a similar spirit, I use extensive data on Medicare patients to find how medical specialties are related by referrals, and use this information to construct demand shifters for identifying the impact of homophily on specialty choice. And I use the names of nearly all U.S. physician to classify how informative first-names are about gender, to test for prejudice. Before analyzing the pay gap and homophily in physician referrals, I begin by describing the data.

2 Data

Since payments and referrals are jointly observed, confidential administrative data on Medicare physician claims are useful for studying homophily in physician referrals and its impact on pay disparities. With more than half a million doctors over 306 local markets, data are fairly representative of the United States. This section describes the data, and the way event-level claims are linked and aggregated to a panel of payments and referral networks.

Data Sources   The main data source for this study is the Carrier database, a panel of all Medicare physician-billed services for a random sample of 20% Medicare beneficiaries. Medicare is the federal health insurance program for people who are 65 or older, certain younger people with disabilities, and people with End-Stage Renal Disease. The data contain the complete record of claims for physician-billed services for the period 2008–2012. It is run by a government agency, the Centers for Medicare and Medicaid (CMS) and has standardized payments and billing systems. Data studies are from traditional ("Fee-For-Service") Medicare, which covers two-thirds of all Medicare patients (the other third is covered by private carriers).

   Referrals are not limited or driven by institutional constraints. A useful feature of traditional Medicare for the purpose of this study is that beneficiaries could see any provider that accepts them (90% do, the choice is of the physician, not the insurer), with no formal requirement to obtain a referral in order to see a specialist (this is in contrast to many managed care settings in which referrals are a means for preauthorization of secondary care). That is, beneficiaries are free to see physicians without referrals, and referrals reflect patients' voluntary choices to follow doctors' advice or suggestions when choosing specialists.

   I use the confidential version of the data, which contains both payment and referral information for each claim[7]. Only physician providers are included, based on CMS specialty

---

[7]For a detailed description of these data see "Carrier RIF Research Data Assistance Center (ResDAC)", `http://www.resdac.org/cms-data/files/carrier-rif`. Accessed May 2015. To protect the privacy of patients, no statistics are reported for demographic cells based on fewer than 11 individual patients. Thanks to the large sample size, such cells are rarely encountered.

code. For each encounter of a patient with a physician, the data contain the following: the date of service and its location, the type of service, patient gender, the physician specialty, and payments made to the physician by all payors; data also record the referring provider, if there was one. A small number of services are excluded, such as lab tests, which are often ordered by physicians directly, in which case the ordering physician is reported instead of the referring physician[8]. These data are combined with beneficiary residential zip-code and sex from the Master Beneficiary Summary File, the CMS database keeping track of all Medicare beneficiaries. Data are further combined with additional data on physician gender and other characteristics from Physician Compare, a public CMS database[9]. The included characteristics are: sex, specialty, medical school attended, hospital affiliation, and year of graduation (used to calculate experience).

The sample is fairly representative of U.S. physicians, with more than 90% of U.S. physicians providing Medicare services, although specialties related to elderly patients are over-represented. By volume, Medicare billed physician services are a quarter of the market for physician services in the United States, which has an annual volume of half a trillion dollars[10], about 3% of the U.S. GDP. The data I use are from traditional "Fee-For-Service" Medicare, which is responsible for two-thirds of all Medicare beneficiaries, with 35 million covered lives. Even though claims for 20% of all patients are observed, selection of physicians into the sample based on their workload is negligible: even for those with minimal workload, the probability of being sampled is very close to 1. The average physician has hundreds of Medicare patients every year. Seeing 30 patients is enough to be sampled with probability 0.999. For the same reason, the probability of missing links between physicians drops sharply as long as they see more than just a few patients. Because of the sampling rate, sampled pay and volume levels are approximately a fifth of their actual value.

Physician Payments To study the pay gap, I aggregate payments from claims to obtain annual pay for each physician, and link to physician characteristics from Physician Compare (see above). Mean pay and characteristics used for the analysis of the pay gap are shown in Table 1 and Figure 3. The average pay for men is considerably higher, but women are

---

[8]Claims are reported using CMS Health Insurance Claim Form 1500, which contains a fields (17, 17a) for the name and identifier of the referring or ordering provider. For details see CMS Claims Processing Manual (Rev. 3103, 11-03-14) Chapter 26, 10.4, Item 17. Services are excluded with BETOS codes for Tests, Durable Medical Equipment, Imaging, Other, and Unclassified Services. For detail description of these codes see https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MedicareFeeforSvcPartsAB/downloads/BETOSDescCodes.pdf. Accessed May 2015. About a third of the remaining claims record a referring physician provider.

[9]Physician Compare Database, https://data.medicare.gov/data/physician-compare Accessed May 2015.

[10]2012 National Health Expenditure Accounts (NHEA).

Table 1: Descriptive Statistics: Physicians

	All	Men	Women
A. Physician Demographics and Average Work Volume			
Gender	0.723	1	0
Experience (years)	22.4	24.2	17.9
Patients	62.2	69.2	43.8
Claims	151	170	103
Pay ($)	21,222	24,399	12,924
Incoming Referrals, as a Fraction of Total:			
Patients	0.365	0.388	0.305
Claims	0.331	0.353	0.275
Pay	0.331	0.353	0.275
Obs. (All Physicians)	530,357	383,525	146,832
B. Avg. Outgoing Referral Volume			
Patients	27.5	29.9	20.9
Claims	52.0	56.5	39.4
Pay ($)	8,785	9,663	6,362
Fraction Male Patients	0.430	0.463	0.339
Specialists (out-degree)	16.2	17.1	13.7
Referrals to Male, as Fraction of Outgoing:			
Patients	0.825	0.839	0.786
Claims	0.827	0.841	0.788
Pay	0.834	0.848	0.795
Obs. (Doctors)	383,173	281,238	101,935
C. Avg. Incoming Referral Volume			
Patients	30.5	33.9	20.1
Claims	57.6	64.4	36.8
Pay ($)	9,746	11,081	5,631
Fraction Male Patients	0.412	0.433	0.348
Doctors (in-degree)	18.0	19.9	12.3
Referrals from Male, as a Fraction of Incoming:			
Patients	0.774	0.793	0.717
Claims	0.775	0.793	0.718
Pay	0.777	0.795	0.719
Obs. (Specialist)	345,390	260,795	84,595

Notes: Physician demographics and average work volume are for all sampled physicians (Part A). Referred work volume (Parts B, and C) are for Doctors and Specialists, namely physicians with at least one outgoing referral (Part B) or incoming referral (Part C) and complete demographic characteristics. The terms "doctor" and "specialist" reflect the position in the dyad, not physician specialty. Volume measures are obtained from a random 20% sample of Medicare patients, so actual Medicare work volumes are approximately 5 times larger. Experience is years since medical school graduation.

less experienced, take more leaves, and practice different specialties. The number of leaves each physician had is approximated by calculating, for each physician, the share of inactive quarters—with no claims—in the sampled history. Quarters are used, and not shorter time intervals, to limit confounding of low workload with periods of inactivity due to sampling errors. The active history includes all sampled years, excluding the year of graduation and the current year, to avoid confounding changes in exact time of graduation with career interruptions. Combined together, these data allow to both document the pay gap and decompose it. The following section analyzes the gender pay gap.

Referral Networks For the study of homophily I construct the network of physician referrals from referral information recorded on claims. If one physician referred patients to another during the year a link is recorded, with the link weight depending on the volume of the relationship, measured as either of the following: the number of patients, the number of claims, and total dollar value of services referred during the year. Representing the referral relationships in Medicare over the period 2008–2012, this panel of directed and weighted networks is used for the study of homophily.

Augmenting network data with physician characteristics from Physician Compare allows me to calculate multiple dyad-specific attributes. For each dyad (pair of physicians) the following are included: practice locations (5-digit zipcode), whether physicians went to the same medical school or whether they share any hospital affiliation. To account for the impact of patient's gender preferences, the gender mix of patients referred is used. The use of physician and dyad attributes both supports the identification of homophily, and helps compare its magnitude against other determinants of referrals.

Men both send and receive more referrals on average, in part due to practicing different specialties. On average, conditional on making any referrals doctors refer to 16 specialists; and conditional on receiving any referrals specialists receive referrals from 18 doctors. But men send referrals to 5 more specialists and receive referrals from 6 more doctors. These gender differences are clearly related to occupational differences: as seen in Figure 1, the average number of working referral relationships each physician has (both incoming and outgoing) varies a lot by medical specialty:Primary care specialties, such as family medicine, mostly send referrals, whereas other specialties, such as neurology, mostly receive referrals. Surgical specialties (where men are the vast majority) both send and receive many referrals. Hence gender differences in participation in the different specialties translates to differences in both incoming and outgoing referral volume. It is therefore particularly important to control for underlying heterogeneity, specifically by specialty, when studying homophily in referrals.

Market Definition   Referrals are a very local phenomenon, mostly targeted at nearby specialists. To study the implications of homophily on the pay gap, I therefore relate physician participation and pay at the local market level. I define local markets based on Hospital Referral Regions (HRR) from the 2012 Dartmouth Atlas of Healthcare[11]. There are in total 306 HRR, corresponding roughly to a metropolitan area. Each zip-code maps to one and only one HRR. HRR are the smallest geographical areas that are effectively isolated networks: Less than 10% of referrals cross their boundaries.

To study the impact of homophily on pay, I construct, for each market (HRR), a monthly panel for 2008–2012 summarizing the fraction of primary-care claims handled by male doctors, and the fraction of services incurred by male patients (a control variable). In addition, total monthly payments are computed for each physician.

Cohort Data   To study the impact of homophily on female entry to medical specialties I reconstruct a retrospective cohort panel, by calculating the number of currently-active physicians who have entered each specialty in each of the years between 1965–2005. Cohort information is obtained from medical school graduation years, recorded in Physician Compare. For each CMS specialty and each year, physician counts are calculated, for those that have graduated before and exactly at that year, and the fraction male among them. Year-specialty cells with fewer than 500 specialists are omitted.

3   Documenting and Decomposing the Pay Gap for Medicare Physicians

Before discussing homophily in referrals, I document the current Medicare physician pay gap, and decompose it in light of known explanations. I find large differences in pay between male and female Medicare physicians, consistent with previous findings of large gender pay gaps both in medicine and in general. However, existing explanations account for only half of this gap.

In 2012, the average female physician in the sample received a total of $70,579 from Medicare, compared with the average male physician, who received $128,872: that is, 48% less (66.8 log points). Such a large pay gap is on par with previous estimates of substantial pay gaps among doctors, and more generally among highly skilled professionals. And while some of these earlier works used much smaller or less detailed data, they still make a useful benchmark. Using Current Population Surveys (CPS) from 1987 to 2010, Seabury et al.

---

[11]"Dartmouth Atlas of Healthcare", `http://www.dartmouthatlas.org/tools/downloads.aspx?tab=39`. Accessed May 2015

Figure 1: Average Number of Referral Relationships by Medical Specialty
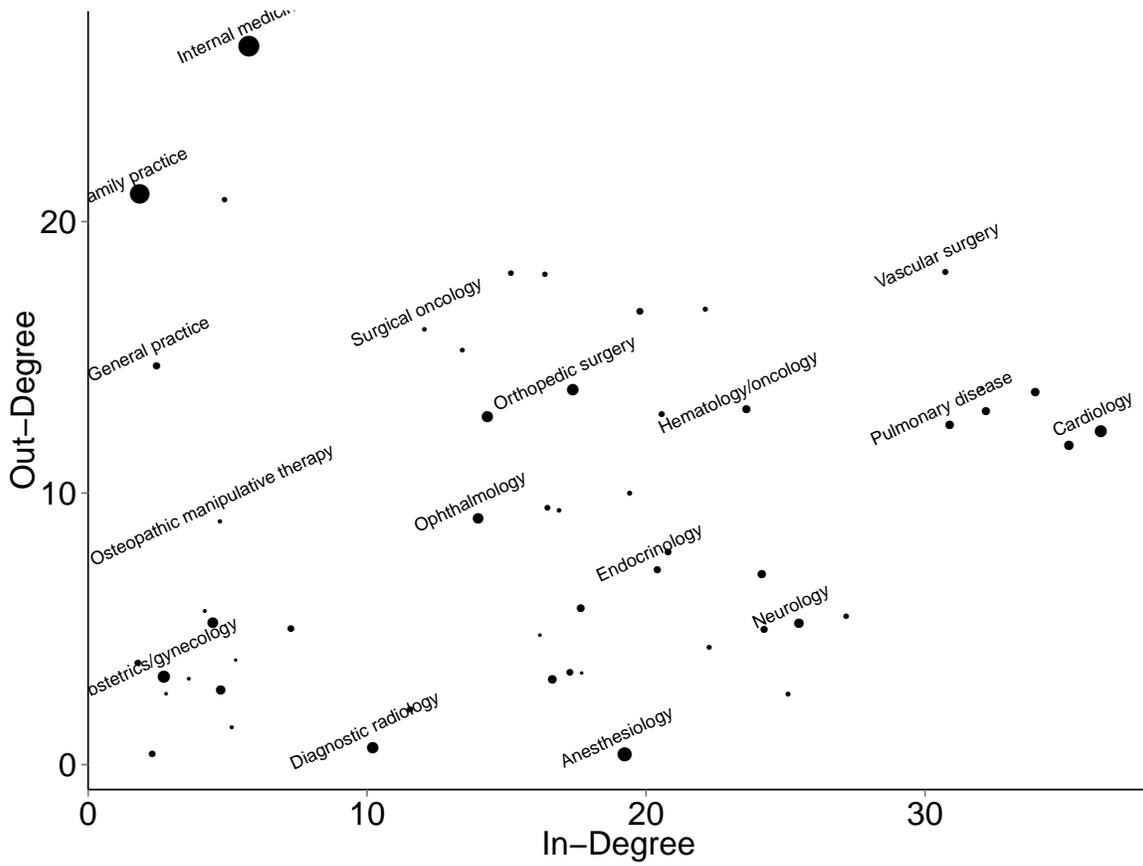
Notes: Degree-heterogeneity is to be expected since doctors in different specialties play different roles in routing patients: some mostly diagnose and refer out, others mostly receive referrals and treat. The figure shows degree distribution by specialty for 2012 referrals: Out-degree is the average number of physicians to whom a physician referred patients during the year. In-degree is the average number of physicians from whom a physician received referrals. Physicians with neither incoming nor outgoing referrals during the year were excluded. Point diameter is proportional to the square root of the number of practitioners in a specialty. Common specialties are labeled. See Table A3 for the data.

(2013) estimate a persistent median gross gap ranging between 16% and 25% (18–30 log points) among U.S. physicians. Using Physician Surveys administered between 1998 and 2005, Weeks et al. (2009) find women earn about a third less than men. These estimates of the gender pay gap in Medicare are also on par with pay gaps in other high-skilled occupations: Bertrand et al. (2010), using data from MBA graduates working in the financial and corporate sectors, found a gross gap of almost 60 log points 10 to 16 years after graduation. Clearly, such gross measurements represent the sum of multiple differences between male and female doctors, making it hard to understand root causes. Observing detailed physician characteristics allows for such decomposition.

To decompose the gross pay gap into its different components and evaluate their rela-

Figure 2: Gender Pay Gap by Physician Experience

2012 Pay by Gender and Experience
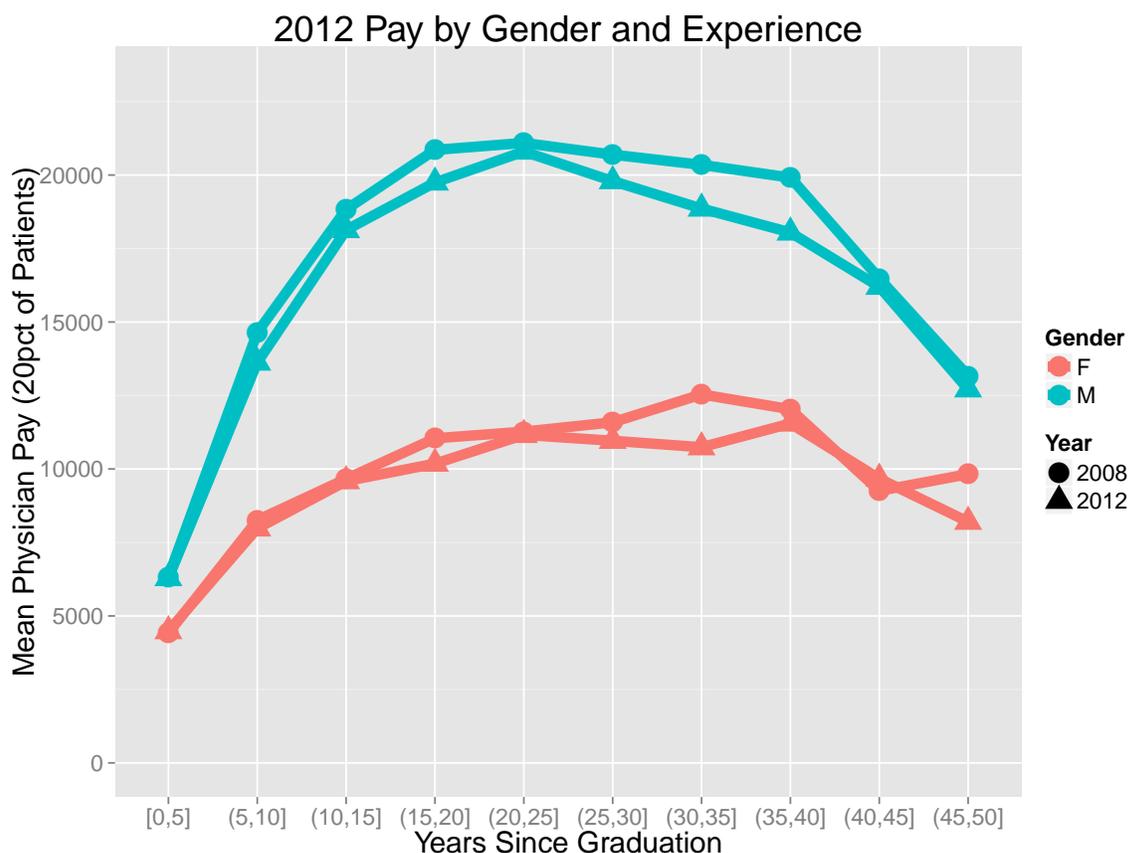
Notes: Physician Pay is total annual payments (by all payers) to physicians for Medicare services during the year (shape legend). Years Since Graduation refer to years since graduation from medical school at the time services were provided.

tive importance, I estimate a standard (log) annual pay equation with controls (Table 2). Previously studied causes of gender gaps: differences in specialization and more career interruptions for female, explain about a half of this gap. Differences in experience, location, and medical school attended explain little. Half of the gross gap remains unexplained. These findings are next discussed one by one.

About a third of the gender pay gap in Medicare is due to differences in specialization: women participate much more in lower paying specialties (Figure 3). While men are the majority of active physicians in almost all specialties, the fraction of women varies greatly. For example, women are 51% of active obstetrician-gynecologists but less than 6% of active orthopedic surgeons. Accounting for specialty in the pay equations reduces the coefficient on gender from 65.4 to 46.8 log points. That is, between-specialty gaps explain a third of the gap, while the remaining two thirds are difference in workload within specialty. These findings again resonate with previous works (e.g., Weeks et al., 2009).

Figure 3: Male Fraction of Physicians in Common Medical Specialties

Gender Composition of Common Physician Specialties
2012 Active Physicians
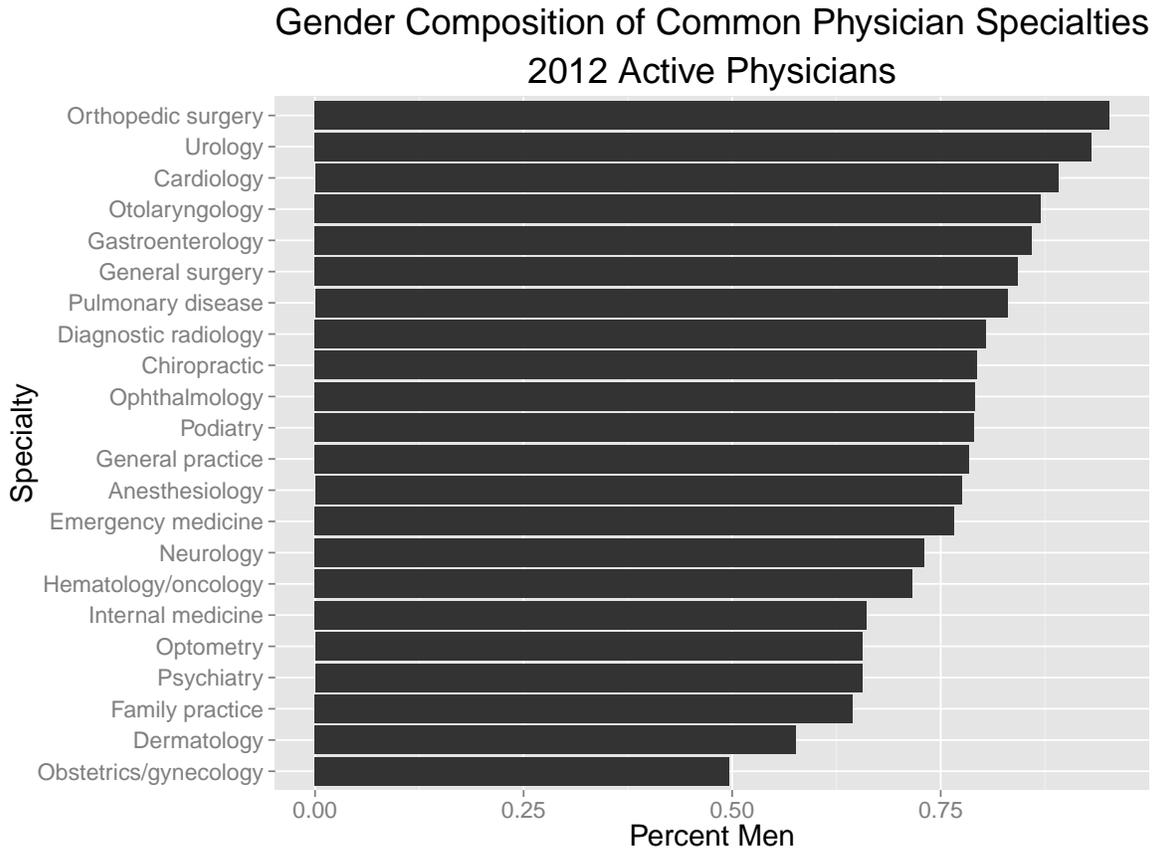
Notes: Percent of active physicians (with any claims) who are male, for the most common specialties by overall number of physicians. Columns are sorted so specialties with the greatest male shares are at the top.

Gender differences in career interruptions also explain part of the pay gap in Medicare. As previously shown by Bertrand et al. (2010) such differences also explain a large part of the gap among highly skilled professional in the financial and corporate sectors: More interruptions were reported by female MBA alumni on surveys, and those interruptions have had persistent, and sizable effect on their subsequent earnings. Career interruptions may be related to existing (if shifting) differences in family roles: women taking time off for having children and later taking care of them may explain both the initial interruption and the subsequent persistent change in work hours. Inactive quarters are indeed more common for women physicians in Medicare, and account for additional 10 log points of the Medicare pay gap (Column 4).

Experience differences, while substantial, explain little of the gap. Because women have been entering the profession in equal numbers only recently, female physicians are on average 7 years less experienced than male physicians. Experience is also a strong determinant of

Table 2: The Gender Pay Gap for Medicare Physicians

	(1)	(2)	(3)	(4)	(5)	(6)
	Dependent variable:					
	Log(Annual Pay)					
Male Physician	0.668***	0.654***	0.468***	0.361***	0.337***	0.340***
	(0.005)	(0.005)	(0.005)	(0.004)	(0.004)	(0.005)
Experience Quadratic	No	Yes	Yes	Yes	Yes	Yes
Specialty	No	No	Yes	Yes	Yes	Yes
No-Work Spells	No	No	No	Yes	Yes	Yes
City	No	No	No	No	Yes	Yes
Med. School	No	No	No	No	No	Yes
Constant	Yes	Yes	Yes	Yes	Yes	Yes
Observations	498,580	447,863	447,863	424,361	420,319	296,199
Adjusted R^2	0.033	0.052	0.290	0.407	0.441	0.471

Notes: Estimates from an OLS regression of annual pay on physician attributes. Experience is years since graduation. Specialty is a dummy for 54 CMS specialty code. No-work spells are previous quarters with no claims (see text for details). City is a dummy for one of 306 Dartmouth Hospital Referral Regions. Med. School is Physician Compare medical school ID. The number of observations vary due to incomplete data on some characteristics.

Medicare work volume: Experienced physicians work more, with the peak volume reached towards mid-career. In spite of these facts, experience differences explain very little of the gender gap in pay: Figure 2 shows the gross gaps within experience bins: In every experience level gaps in pay exist. Pay differences within cohorts remained stable between 2008 and 2012. Indeed, controls for quadratic experience hardly change the estimated gap (Column 2 of Table 2).

Neither the medical school physicians attended, nor where they are located explain much of the gap (Columns 5 and 6). Location differences could in principle contribute to the gap due to the geographic adjustment of the Medicare fee schedule, but they in fact do not contribute much.

About half of the gender pay gap in Medicare remains unexplained even after known explanations are account for: differences in specialty, career interruptions, experience, location, and education. The question therefore remains: why do female Medicare physicians work fewer hours than their male counterparts? Understanding the causes for this large difference in workload is important as beyond its direct effect on pay, lower workload by women could feed back to their specialization choices, and impact their wage. For example, Chen and

Chevalier (2012) show that lower expected workload alters the net-present-value of alternative career choices, possibly making it rational for women to invest less in specialization.

Next, a new explanation for the gap is discussed: gender homophily in referrals. The paper proceeds by modeling and estimating homophily in physician referral networks (Section 4), and then by studying its effects for the pay gap (Section 5) and finally, for patient outcomes (Section **??**).

4 Homophily in Referral Networks

In this section I show referrals exhibit gender homophily (i.e., doctors refer relatively more to their own gender). I define homophily and model it in directed networks to generate testable differences between its two potential mechanisms: preferences and sorting. I then estimate homophily using data on Medicare referrals, showing it is mostly driven by gender-biased preferences, not sorting.

4.1 Measuring Homophily in Directed Networks

Before examining evidence for gender homophily in Medicare referrals, I first propose a new measure of homophily for directed networks. This measure, *Directed Homophily*, is based is different from existing ones in that it compares referrals between the genders, and not to population baseline fractions, and thus accommodates heterogeneity in the propensity to refer or receive referrals better than previous ones. This section discusses this measure.

Consider the network of physician referrals in a given market[12], where a link exists between "doctor" j and "specialist[13]" k if j referred any patients to k. There are two genders: male and female ($g \in \{m, f\}, G \in \{M, F\}$, lowercase indexes are used for doctors and uppercase for specialists). Let n_{mF} be the average number of links a male doctor sends to female specialists (likewise define n_{gG}). The average fraction of referrals male doctors send to female specialists is:

$$r_{F|m} := \frac{n_{mF}}{n_{mF} + n_{mM}} \tag{1}$$

Likewise define $r_{G|g}$. We are now ready to define Directed Homophily.

---

[12]For concreteness I focus on gender homophily in referrals, but both this homophily measure and the following model are more broadly applicable to directed networks in general.

[13]Throughout this paper the terms "doctor" and "specialist" are used to denote the role of a physicians as a referral origin or target, regardless of their actual specialties, similar to how "ego" and "alter" often used in the sociology literature. Thus the same physician can be a "doctor" with respect to one link and a "specialist" with respect to another.

Table 3: Directed Homophily (DH)

To (Specialist)

	Female (F)	Male (M)
From (Doctor) Female (f)	20%	80%
Male (m)	15%	85%

$$DH = 85\% - 80\% = 20\% - 15\% = 5pp$$

Definition 1 (Directed Homophily). *Directed Homophily* is the difference between the fraction of outgoing referrals of male and female doctors to male specialists (or equivalently, to female specialists):

$$DH := r_{M|m} - r_{M|f} = r_{F|f} - r_{F|m}$$

Directed Homophily exists ($DH > 0$) if male doctors refer to male specialists more than their female counterparts[14]. Table 3 illustrates this definition using Medicare data. In Medicare, male doctors refer 85% of their patients to male specialists, compared to female doctors who refer 80% of their patients to male specialists, so $DH = 5pp$ (figures are rounded to the nearest integer). Instead of comparing outgoing referrals, one could define homophily based on the difference in incoming referral rates. It is easy to verify such a measure has always the same sign as Directed Homophily.

Directed Homophily captures a tendency to link within gender beyond what is expected from random sorting. Particularly, Directed Homophily is not driven by baseline imbalance in the gender distribution of doctors or specialists: if most specialists are men then both male and female doctors are expected to refer more to male specialists, but not differentially so.

Referrals are asymmetric: they have a direction, and Directed Homophily reflects that. Ignoring directions as if relationships were symmetric (e.g. Kossinets and Watts, 2006, 2009) can be problematic because in directed networks the gender distribution of outgoing and incoming links do not generally overlap. For example, in Figure A2(a), women send the majority but receive the minority of links; and only link directions reveal which of otherwise equivalent networks exhibits homophily.

Directed Homophily admits the use of weighted links. Weights reveal whether same-

---

[14]In general, referrals could also be biased toward the other gender as well, if $DH < 0$. In this case the network exhibits *Directed Heterophily*. Being the difference of referral rates denoted in percentage terms Directed Homophily is denoted in percentage points (or, in case referral rates are considered as fractions, as a scalar in the range $[-1, 1]$).

gender referrals are not only more likely, but also more voluminous. To adapt directed homophily to weighted links just redefine n_{gG} using weighted degrees, as follows: Let n_{jk} be the weight of the link from j to k (e.g. number of patients referred). The weighted out-degree of j is $d(j) = \sum_k n_{jk}$. The weighted out-degree to females is $d^F(j) = \sum_k \mathbb{1}_{g_k=F} n_{jk}$. Now n_{mF} is the average of $\dfrac{d^F}{d}$ over all male j, and so on for n_{gG}. The rest of the definition is verbatim.

To see what makes Directed Homophily a useful measure, contrast it with the commonly used measure of homophily: Inbreeding Homophily (see McPherson et al., 2001; Currarini et al., 2009; Bramoullé et al., 2012). Instead of comparing referral rates across genders, Inbreeding Homophily uses population shares as the baseline:

Definition 2 (Inbreeding Homophily). Male doctors exhibit *Inbreeding Homophily* if

$$r_{M|m} > M$$

Where M is the fraction of males in the specialist population. Likewise, female doctors exhibit inbreeding homophily if $r_{F|f} > F$. Where $F = 1 - M$ is the fraction of female specialists.

Note that Inbreeding Homophily by both genders immediately implies Directed Homophily, while the reverse is not true, e.g. if $r_{M|m} > r_{M|f} > M$.

Because Directed Homophily is not using population gender shares as a benchmark, it is more robust than Inbreeding Homophily to potential heterogeneity in the propensity to send or receive referrals. For example, suppose male specialists are more likely to receive referrals because they are more experienced. Then there are more referrals to men than their fraction in the specialists population (as in Figure 4b), but unlike Inbreeding Homophily, Directed Homophily would not confound such differences with homophily. Note that the cost of this robustness to heterogeneity is that Directed Homophily makes no distinction as to whether men or women are more homophilous: rates are compared against each other, and not to an external baseline. Thus, in case heterogeneity can be ruled out or fully accounted for, Inbreeding Homophily is more informative.

Before moving to model homophily, I show preliminary evidence for its presence in Medicare referrals.

4.2 Preliminary Evidence for Homophily in Medicare Referrals

Physicians are more likely to refer and receive referrals within their gender. Table 4 tabulates all sampled 2012 patient referrals by physician gender. Medicare referrals exhibit Directed Homophily: of all referrals by male doctors, only 15.23% were to female specialists (84.77% to

Figure 4: Directed Homophily and Heterogeneity

(a) Heterogeneity in Referring   (b) Heterogeneity in Receiving   (c) Homophily
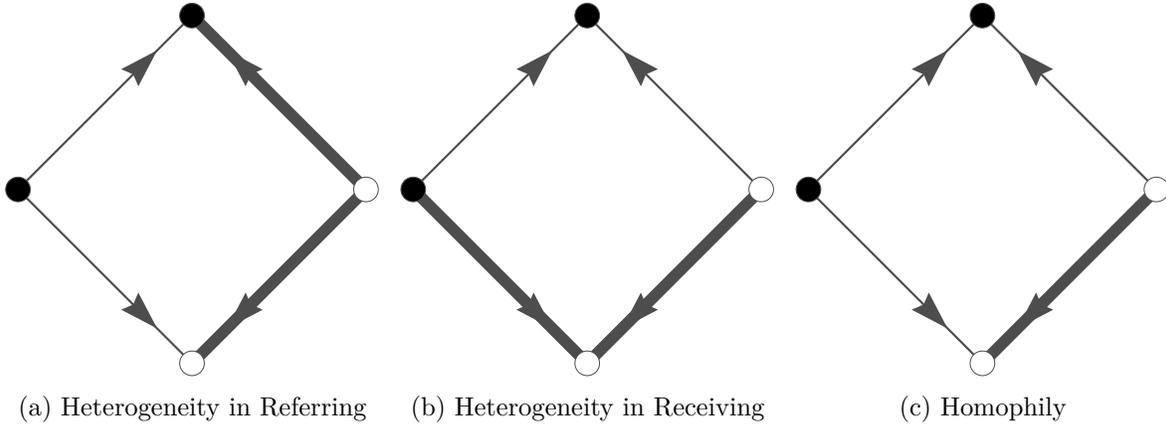
Directed Homophily (DH) is robust to systematic differences in the tendency to refer or receive referrals, as it compares referrals across genders, not to population fractions. For example, even though male (in white) are half of the population, in (a) they send the majority of referrals (line thickness denotes volume) and in (b) they receive the majority of referrals. Yet in both cases $DH = 0$, since the *fraction* of referrals from male to male is the same as from female to male. Only in case (c) $DH > 0$, since men refer not only refer more overall, but also refer dis-proportionally more to other men.

Table 4: Medicare Referrals by Gender

		A. Referrals		B. Percent of Outgoing				C. Percent of Incoming		
		To			To				To	
		F	M		F	M	Total		F	M
From	f	420,976	1,712,510	f	19.73	80.27	100	f	24.74	19.36
	m	1,280,691	7,130,872	m	15.23	84.77	100	m	75.26	80.64
								Total	100	100

Notes: Referral counts and percentages, by gender of referring and receiving physician. Since services are sometimes billed on several separate claims, multiple referrals of the same patient from a doctor to a specialist are counted as one. Source: 20% sample of Medicare physician claims for 2012.

male specialists), whereas of all referrals by female doctors 19.73% were to female specialists (80.27% to male specialists), so $DH = 19.73\% - 15.23\% = 4.5\%$ (Panel B of Table 4). Put differently, women refer 30% more often to other women than do men[15]. Were Directed Homophily zero, the rows of Panel B of Table 4 would have been equal to each other. Panel C of Table 4 shows the difference in terms of specialists' incoming work volume.

In itself, homophily (Table 4) does not imply physicians have gender-biased preferences, since it could reflect physician sorting by gender (e.g., into markets). It is therefore useful to consider homophily within local markets (Figure 5). Even within markets, male systematically refer more to male, suggesting gender-biased preferences might play a role. Also clear from this figure is that Directed Homophily is non-linear in the fraction of male specialists: it is always zero if there are only male specialists (or only female ones). A characterization of it thus requires accounting for the variation in the gender composition of different choice sets faced by different doctors. To clarify this dependency, and to study homophily's consequences for the pay gap, I develop a model of homophily in referrals.

4.3 A Model of Homophily in Referrals

I model referrals to characterize homophily in directed networks, its mechanisms, and their consequences for the pay gap. Doctors choose specialists to refer to from local opportunity pools. Both gender-biased preferences (i.e., a preference for working with same-gender others), and physician sorting on gender may cause homophily, but these mechanisms have different implications. In particular, only homophily due to preferences increases with the availability of choice and with market size. And only gender-biased preferences make the demand for specialists decrease if fewer doctors share their gender (and even more if fewer fellow specialists do). These predictions later illuminate my strategy to identify gender-biased referrals and their impact on female pay.

Mechanisms: Preferences versus Sorting

To study the causes and consequences of homophily, consider a model where doctors $j \in J$ choose specialists to refer patients to, from an opportunity pool $k \in K_j$. Denote the gender of doctors and specialists by $g_j \in \{f, m\}$, and $g_k \in \{F, M\}$. Doctors maximize a gender-sensitive utility function:

$$\max_{k \in K_j} U_j(k) = \beta \mathbb{1}_{g_j = g_k} + \delta X_{jk} + \varepsilon_{jk} \tag{2}$$

---

[15]That is, $\dfrac{19.73\%}{15.23\%} - 1$

Figure 5: Referrals to Male by Doctor Gender, Different Markets
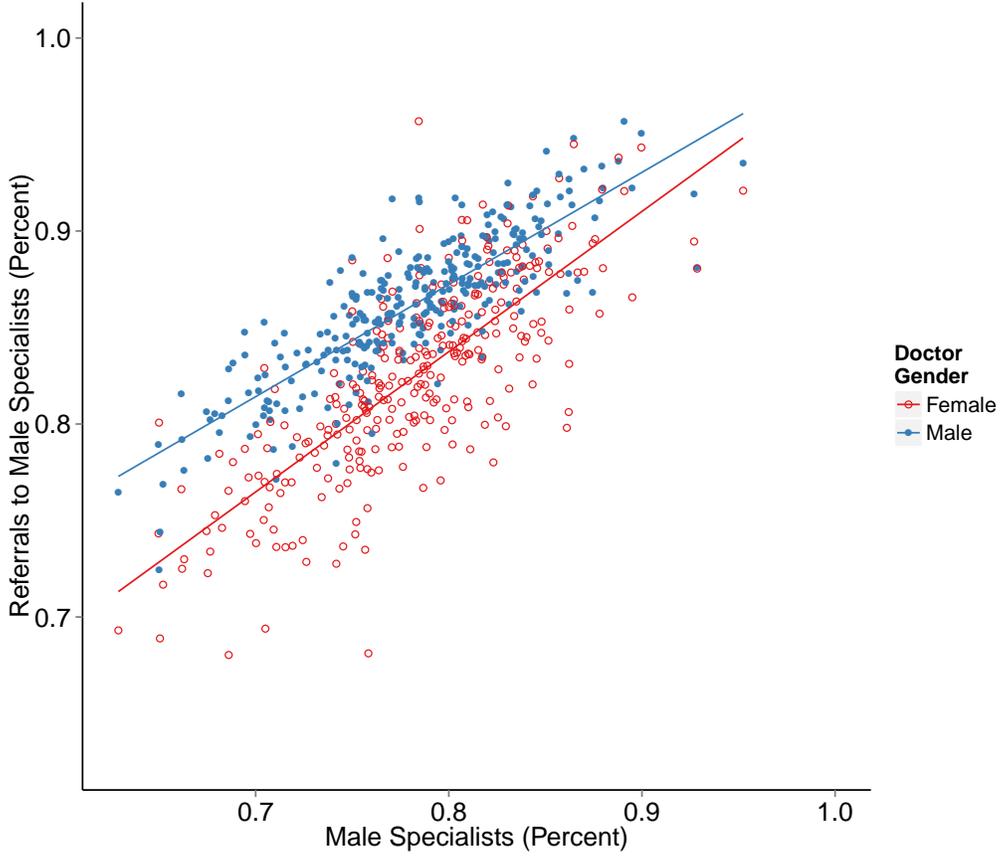
Notes: For each local physician market (Dartmouth Hospital Referral Region), average fractions of referrals from male and from female doctors to male specialists are plotted over the fraction of male specialists in the market. Each of these 306 local U.S. markets is thus represented by two vertically-aligned data points. On average, men refer more to men than women do, even after accounting for the variation between markets in the availability of male specialists.

Where $\mathbb{1}_{g_j=g_k}$ indicates both physicians are of same gender, $(g_j, g_k) \in \{(f, F), (m, M)\}$. The choice of specialists depends on individual and pairwise attributes (X_{jk}; e.g., experience and distance), but may also depend on gender: $\beta > 0$ represents *gender-biased preferences*. For now, I abstract from the multiplicity of medical specialties (explicitly introduced later, in Section 5.2). If ε_{jk} is independently and distributed Gumbel-extreme-value, equation (2) yields the conditional logit probability for a referral from j to k, given gender and other characteristics:

$$p_{jk} := Pr(Y_{jk} = 1|g, X) = \frac{e^{\eta_{jk}}}{\sum_{k' \in K_j} e^{\eta_{jk'}}} \tag{3}$$

Where $Y_{jk} = 1$ if j refers to k and $Y_{jk} = 0$ otherwise. $\eta_{jk} := \alpha_j + \beta\mathbb{1}_{g_j=g_k} + \delta X_{jk}$. That is, link formation is determined by pairwise characteristics. This excludes more strategic setups where links are formed in response to or in anticipation of other links.

21

Homophily due to Gender-Biased Preferences Biased preferences cause homophily. To see how, consider first the case there is one market with one common pool of specialists $K_j = K$, for all doctors $j \in J$, and no heterogeneity: $\delta = 0$. The following claim shows that gender-biased preferences lead to homophily.

Claim 1 (Preference-Based Homophily)**.** Within a market, there is Directed Homophily iff preferences are gender-biased.

To see why Claim 1 is true, note that the conditional probabilities of referrals to a male specialist are (see appendix for derivation of this and other results):

$$p_{M|m} = \frac{M}{M + \omega(1 - M)} \geq M \geq \frac{\omega M}{\omega M + (1 - M)} = p_{M|f} \tag{4}$$

where $p_{G|g} := Pr(g_k = G | g_j = g)$ denote probabilities conditional on doctors' gender, $\omega := e^{-\beta} \in (0, 1]$, and $M = \frac{1}{|K|} \sum_k \mathbb{1}_{g_k = M}$ is the fraction of male specialists (with slight abuse of notation). Equation (4) shows biased preferences result in Directed Homophily ($p_{M|m} > p_{M|f}$): doctors of each gender slightly discount the other (by a factor ω). Conversely, if preferences are unbiased ($\beta = 0$) referral rates to men are common to doctors of both genders:

$$p_{M|m} = M = p_{M|f} \tag{5}$$

This is the baseline: Directed Homophily is zero. Clearly, if the most specialists are men then men refer more to men than to women: $p_{M|m} > p_{F|m}$, which is not to be confused with $p_{M|m} > p_{M|f}$.

An implication of (4) is that homophily is higher when the fraction of male (or female) specialists is less extreme (as illustrated in Figure 6a). Conversely, when most specialists are of one gender there is less room for choice and thus homophily is weaker. This implication is confirmed by the evidence in Figure 5 for higher homophily in markets with more diverse pools (lower M).

Considering heterogeneity, suppose there exists a correlation between gender and specialist characteristics that determine the volume of incoming-referrals (e.g. men have greater capacity to receive referrals, or women are better specialists). In this case, (4) becomes:

$$p_{M|m} = \frac{M}{M + \omega\eta(1 - M)} \geq \frac{\omega M}{\omega M + \eta(1 - M)} = p_{M|f} \tag{6}$$

Regardless of gender-biased preferences, if $\eta < 1$ male specialists attract a disproportionally high fraction of referrals from both genders (Figure 6b). Conversely, if $\eta > 1$, female specialists attract more referrals, so whether $p_{M|m}$ and $p_{M|f}$ are each greater or smaller than

Figure 6: Preference-Based Homophily

(a) Homogeneous Case ($\eta = 1$) (b) Heterogeneous Case ($\eta = 0.3$)
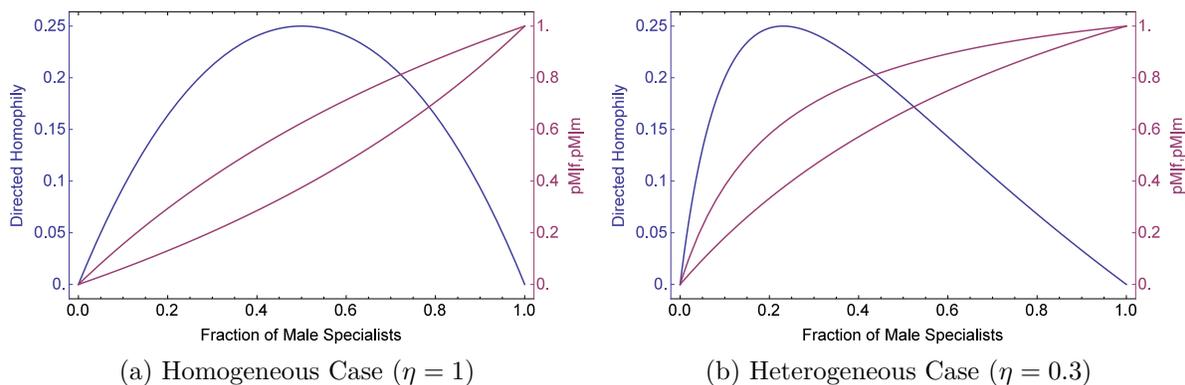
Notes: Probability of referrals to male specialists by male and female doctors, and the difference: Directed Homophily for different fractions of male specialists, M, with gender-biased preferences ($\omega = 0.6$). Case (a) shows gender-biased preferences in a homogeneous specialist population. Male specialists receive more referrals than their fraction in the population from males, and less than this fraction from females. Case (b) combines gender-biased preferences with heterogeneity: male specialists receive more referrals than their fraction in the population from both male and female doctors, but more from male than from female doctors.

M depends on η. But, as (6) shows, whether or not $p_{M|m} = p_{M|f}$ only depends on whether preferences are gender-biased. Therefore with heterogeneity ($\eta \neq 1$) Directed Homophily is better measure of homophily than Inbreeding Homophily because it is not using M as the benchmark, but rather compares referrals of both genders against each other. For simplicity, for the rest of this section again assume homogeneity.

Homophily due to Sorting   Apart from preferences, sorting also causes homophily, as it makes women and men more exposed to their own gender. To see how suppose instead of a single market there is a set C of separate markets. Markets may vary in size $\mu^c = \dfrac{J^c}{J}$ ($\sum_c \mu^c = 1$), each market has its own sets of doctors J^c and specialists K^c, and its distinct fractions of male doctors m^c and male specialists M^c (for $c \in C$). Referrals only occur within markets. That is, $K_j = K$ for all $j \in J^c$. The definition of a market is abstract, and can stand not only for geographically-defined markets, but also for other segmentations (e.g., institutional affiliations). The conditional probabilities of referrals to men now vary by market and are: $p^c_{M|m}$ and $p^c_{M|f}$. Define *Sorting* to be a positive correlation between the genders of doctors and specialists in a market[16] $\text{Cov}(m^c, M^c) > 0$ (equivalently, $\text{Cov}(f^c, F^c) > 0$). Homophily then arises at the aggregate, when all markets are pooled together:

---

[16] This definition extends to the more general case where K_j is specific to each doctor as: $\text{Cov}(m^j, M^{K_j}) > 0$, where $m^j = \mathbb{1}_{g_j = m}$ and M^{K_j} is the fraction of male in K_j.

Claim 2 (Sorting-Based Homophily)**.** With sorting, referrals exhibits homophily when pooled together across all markets:

$$p_{M|m} > M > p_{M|f}$$

even when $\beta = 0$.

Intuitively, if fractions of male doctors and specialists are correlated then referrals coming from male doctors are more likely to occur in markets with more male specialists. Homophily then appears at the aggregate level, even when preferences are unbiased and there is no homophily within each market.

The following proposition shows observed homophily is the combination of these two causes: preferences and sorting[17].

Proposition 1 (Homophily Decomposition)**.** *Homophily observed across all markets decomposes to preferences and sorting as follows:*

$$\overbrace{p_{M|m} - M}^{\text{Overall Homophily}} = \frac{1}{m}\Big(\overbrace{\mathrm{E}[m^c(p_{M|m}^c - M^c)]}^{\text{Biased Preferences}} + \overbrace{\mathrm{Cov}[m^c, M^c]}^{\text{Sorting}}\Big)$$

That is, the overall Inbreeding Homophily, observed when all markets are pooled together, is the sum of two terms: (a) the average market-specific preference-based homophily, weighted by market size μ^c and share of doctors, $\frac{m^c}{m}$, and (b) sorting into markets. Proposition 1 shows sorting and preferences are exhaustive: combined together, they fully account for the overall homophily observed. This decomposition holds for any distribution of ε with finite second moment.

A corollary of Proposition 1 is that when market boundaries are observed, preferences and sorting are separately identified: homophily observed within each market is due to preferences, while the rest is due to sorting. If $\mathrm{Cov}[m^c, M^c] < 0$ sorting dampens homophily, rather than augments it. In this case even if preferences are biased overall homophily could be zero.

Market Size Effects: A Market for Preference-Based Homophily   Another way to distinguish between preferences and sorting is to consider market size. Intuitively, since preference-based homophily depends on availability, it is lower in smaller markets, where choice is more limited, and there are fewer opportunities for gender preferences to be revealed. The next proposition shows this.

---

[17]The proposition is stated here for Inbreeding Homophily for clarity. Its equivalent for Directed Homophily is in the appendix.

Proposition 2 (Size Effects). *If specialists enter markets independently and homophily is strictly concave, preference-based homophily is greater in large markets. Namely, for every market size N there exists $N' > N$ such that $DH(N') > DH(N)$. Where $N = |K^c|$ and $DH(N) = E[p_{M|m} - p_{M|f}|N]$.*

Intuitively, smaller markets are more likely to have extreme specialist gender-mix, which restricts choices and prevents preferences from being revealed. Since this relationship to market size is unique to preferences and is not implied by sorting, Proposition 2 provides a test for the presence of gender-biased preferences.

Homophily's Consequences for Gender-Disparities in Specialist Demand

Homophily causes disparity in demand between the genders: When preferences are gender-biased, specialist receive fewer referrals the fewer doctors share their gender. The gender of fellow specialists matters too, in a more nuanced way: whether same-gender specialists substitute or complement each other depends on the gender distribution of doctors.

Proposition 3 (Demand for Specialists by Gender). *Average specialist demand depends on gender as follows:*

1. *With gender-neutral preference ($\beta = 0$) specialist demand is invariant to gender.*

2. *With gender-biased preference ($\beta > 0$) average demand for specialists is higher the more doctors share their gender.*

3. *Same-gender specialists are substitutes when most doctors share their gender, and complements when most doctors are of the opposite gender.*

The proof of Proposition 3 is by noting demand for male specialist—the average number of referrals received (denoted by D)—is a weighted-average of doctors' respective probability of referring to male:

$$D = mp_{M|m} + (1-m)p_{M|f} \tag{7}$$

where m is the fraction of male doctors in the market (superscript c is omitted for clarity as all magnitudes are within markets), and where for tractability assume $|J| = |K|$ (see appendix for general results). Substituting (4) into (7) and differentiating by m and M

25

yields:

$$\frac{\partial D}{\partial m} = \overbrace{p_{M|m} - p_{M|f}}^{\text{Directed Homophily}} \tag{8}$$

$$\frac{\partial D}{\partial M} = (1-m) \overbrace{\frac{w(1-w)}{(1-M(1-w))^2}}^{\text{Complements }(+)} + m \overbrace{\frac{-(1-w)}{(M+w(1-M))^2}}^{\text{Substitutes }(-)} \tag{9}$$

Intuitively, specialists are better-off when most doctors upstream share their gender. But the magnitude of this effect is mediated by the number of downstream specialists who are of the same gender. And whether downstream specialists substitute or complement each other depends on the gender distribution upstream: with biased-preferences, a specialist of the same gender as of most doctors upstream is most popular if most other specialists downstream are of the opposite gender. But the converse is true for a specialist of the opposite gender than most upstream doctors.

To further develop this intuition, consider first the special case $m = M$: identical fractions of male doctors and specialists. With gender-biased preferences, the relationships between a population gender and the percent of referrals it receives is S-shaped[18] (Figure 7). With biased preferences the majority gender receives dis-proportionally more referrals.

Figure 7: Overall Demand as a Function of Gender ($m = M$)
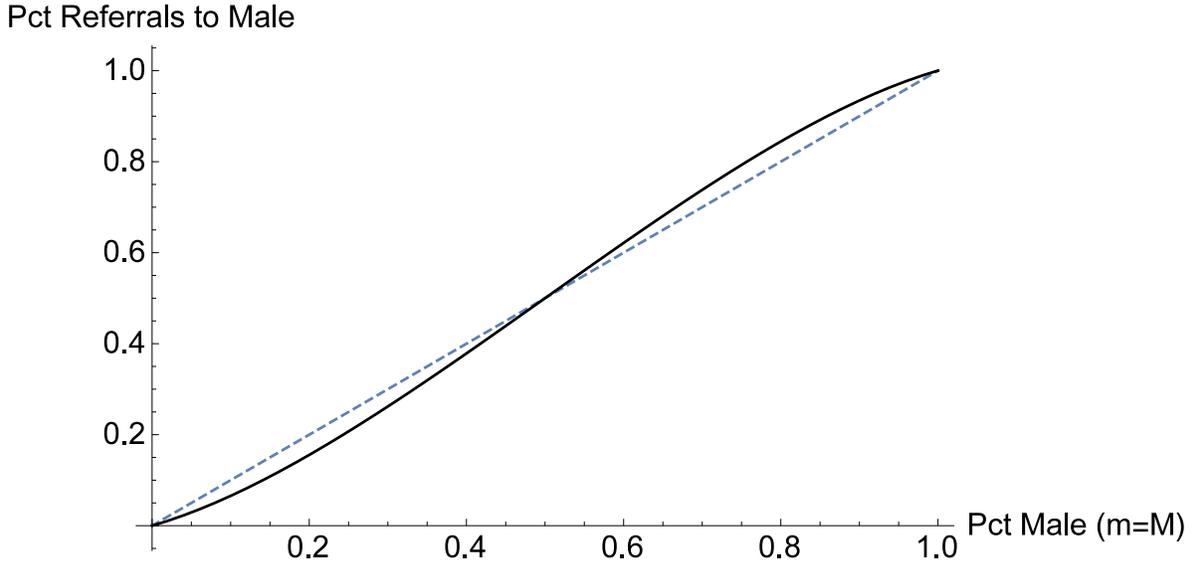
Notes: The figure plots the overall share of referrals going to male, as a function of the male population fractions, in the special case where $m = M$, for gender biased preferences ($\beta > 0$). The majority gender receives more than its share; the minority gender receives less than its share.

---

[18]This case is similar to previous results for undirected homophily of (Currarini et al., 2009, Section 4.5). There too the basic mechanism underlying the S-shaped curve is preferences, albeit the setting is different.

The relationship in the general case where $m \neq M$ is a bit more nuanced, as demand depends on the gender of physicians both upstream and downstream. This dependency is illustrated in Figure 8, which depicts the average demand for a female specialist, as a function of the both the fraction of female doctors (upstream), and the fraction of female specialists (downstream). With gender-biased preferences, a female specialist faces higher demand the more doctors are female (8). The effect of other female specialists depends on which gender is the majority upstream: if most doctors are male, female specialists are complement, whereas if most doctors are female they are substitutes (9). As female ares in fact both the minority of doctors and the minority of specialists in most markets (darker area of the surface), they suffer a lower demand due to both these effects: fewer doctors favor them, and it is easier for male doctors to choose male specialists over them.

Figure 8: Average Demand as a Function of Gender
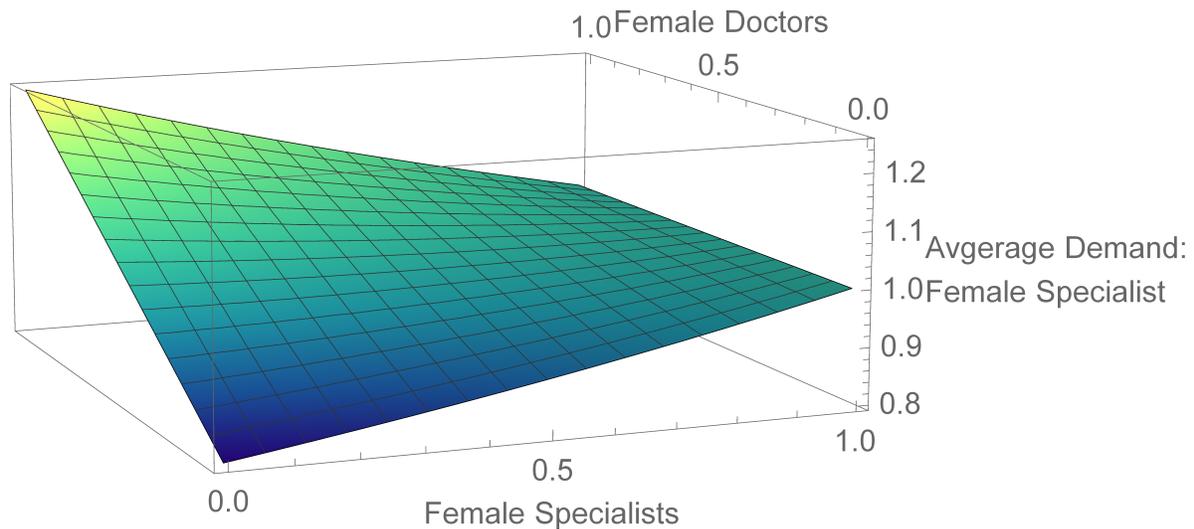
Notes: The figure plots the average female specialist demand as a function of the fractions of female doctors and female specialists, with gender-biased preferences ($\beta > 0$).

Since specialists of the majority gender face higher demand, sorting—so far taken as exogenous—could also be endogenous. An implication of Proposition 3 is that it is better for female specialists to enter markets where a greater share of referrals is handled by female doctors. For an extreme example, if men and women were to sort into fully segregated markets, there would be no room for gender preferences to impact demand. For similar reasons, even partial sorting is beneficial. In Section 5.2 I developed and empirically test one aspect of this idea: that biased preferences may contribute to occupational segregation.

To conclude, this model of referrals shows homophily is a composition of gender-biased preferences and sorting. When driven by gender-biased preferences, homophily is stronger in larger markets and it results in fewer referrals to the minority gender. Induced demand

disparities create incentives for specialists to segregate (i.e., join market segments where more referrals are handled by their own gender). I now turn to estimating homophily.

4.4 Empirical Strategy: Estimating Homophily and Gender-Biased Preferences

The above model shows homophily is a combination of biased preferences and sorting. In this section overall homophily and separately its part due to biased preferences are estimated. Identification involves two concerns. First, heterogeneity: if men systematically make and receive more referrals it could appear as if they tend to refer more among themselves. Second, what seems like preferences can be sorting: if specialists cluster by gender, male doctors are more exposed to male specialists (where clustering may be on other dimensions of the attribute space rather than just geographic location). To cope with heterogeneity within-doctor variation is used for identification. Three different approaches are taken to address the possibility of sorting: controlling for rich individual and pairwise characteristics, comparing physicians who face similar opportunity pools, and directly testing theoretical predictions that are unique to preference-driven homophily. Since referral networks are large, estimation also involves a technical challenge: it is difficult to compute statistics depending on the immense number of all physician pairs. With therefore use estimators that only rely on realized links (of which there are many fewer because networks are sparse). Alternatively, choice-based sampling is used, where existing links are compared to sampled pairs with no links.

Reduced Form Estimation of Homophily and Preference Bias   Directed Homophily can be estimated by regressing the fraction of patients each doctor j referred to male specialists[19], r_{jM}, on the doctor gender g_j and other characteristics X_j:

$$r_{jM} = \alpha_1 + \beta_1 g_j + \delta_1 X_j + \varepsilon_j \tag{10}$$

For doctors with any referrals. The ordinary least squares estimate of β_1 measures average Directed Homophily: how much more men refer to men, on average across markets. As shown above, this average masks heterogeneity due to differences in availability across markets: Directed Homophily is expected to be higher in markets where specialist pools are more gender-balanced. Specification (10) is therefore estimated separately for each market.

---

[19]That is, $r_{jM} = \dfrac{\sum_{k:g_k=M} n_{jk}}{\sum_k n_{jk}}$, where $n_{jk} \geq 0$ is the volume of referrals from j to k. The unweighted specification uses (with slight abuse of notation): $n_{jk} = \mathbb{1}_{\{n_{jk}>0\}}$.

To further estimate how much of homophily is due to sorting I use a variant of (10):

$$r_{jM} = \beta_2 g_j + \delta_2 X_j + \gamma_{h(j)} + \varepsilon_{jh} \tag{11}$$

Where now $\gamma_{h(j)}$ is a fixed-effect for the hospital with which doctor j is affiliated (for robustness, hospital interacted with medical specialty are also considered). As shown in Proposition 1, if doctors affiliated with the same hospital face approximately the same opportunity pool of specialists, K, then β_2 estimates average preference-based homophily.

These reduced-form specifications directly measure Directed Homophily, and have several advantages: First, they do not rely on assumptions regarding the opportunity pool faced by each doctor, as they depend only on realized—not potential—links. Second, they incorporate link weights, capturing homophily through same-gender links being more voluminous. Third, they accommodate heterogeneity, by normalizing referral rates within each doctor, and accounts for sorting on one dimension, using fixed effects. Yet these estimates do not account for sorting on multiple non-nested dimensions (e.g., hospital and medical school). And they are less portable, since they estimate average homophily, which implicitly depends on the distribution of opportunity pools. These issues are better addressed by structural estimates of the discrete-choice network formation model, discussed next.

Structural Estimation of Homophily and Preference Bias   I estimate the conditional logit model in (3) for the probability of referrals from doctor j to specialist k, given gender g and other characteristics X. The identifying variation comes from differences within each doctor's choice set, thus doctor level attributes are differenced-out, as is clear from comparing the log of the ratio of probabilities:

$$log \frac{p_{jk}}{p_{jk'}} = \beta(\mathbb{1}_{g_j=g_k} - \mathbb{1}_{g_j=g_{k'}}) + \delta(X_{jk} - X_{jk'}) \tag{12}$$

The data consist of an observation for each dyad (j, k), with associated physician and dyad (pairwise) characteristics, and a binary outcome standing for whether the dyad is linked. To account for differences between opportunity pools, X_{jk} includes specialist gender. The main parameter of interest is β, which estimates the gender-bias in preferences. Directed Homophily is a function of β and specialist availability (See Equation 4 above).

Identification: Homophily and Preference Bias   The primary concern is to identify the part of homophily due to gender-biased preferences separately from homophily due to sorting and individual heterogeneity in the propensity to refer or receive referrals. To address it, I use reduced form estimates shown above to be robust to heterogeneity, and use

hospital fixed effect to restrict comparison to doctors who likely fact similar opportunity pools. Structural estimates accommodate heterogeneity by using within-doctor variation, and mitigate potential sorting by including controls for multiple factors that are expected to impact the likelihood of referrals between pairs of physicians, including: location (distance), specialty, experience, patient gender, shared medical school, and shared hospital affiliation. The residual threat is from factors correlated with the gender of both physicians and that determine referrals, or put differently, by sorting on unobserved attributes. Note that for an omitted factor to confound the estimates, it must not only be related to referrals, but also correlated with the genders of both doctors and specialists. For example, if doctors mostly refer within-hospitals, omitting hospital-affiliation is only a problem if hospitals are also gender-segregated. Furthermore, characteristics unrelated to referral appropriateness that might be shaping preferences are not confounders, but rather underlying mechanisms (e.g., if men refer to men because they golf together, golf-club affiliation explains homophily, but does not explain it away). The identification assumption is therefore that no clinically-related factors correlated with both the probability of a referral and with the gender of both physicians are omitted.

I further validate the findings of gender-bias in preferences by directly testing two predictions of the model, which are derived in Section 4.3 above and shown to be unique to biased-preferences: First, an increase in homophily with market size (Proposition 2). Second, a correlation of doctors' gender distribution with specialist demand (Proposition 3).

Estimation: Homophily and Preference Bias Since the opportunity pools of specialists are very large (the number of possible links is square the number of physicians), considering all possible dyads is computationally difficult. I therefore estimate the model using choice-based sampling (also known as case-control sampling). That is, instead of considering all possible dyads, each *case*: linked dyad (j, k), is matched with two *controls*: unlinked dyads (j, k') and (j, k'') with k', k'' sampled at random from K_j, defined as all k within the same referral region and of the same specialty of k, the specialist to which j actually referred. Controls for the multiple other attributes observed are included. Weighted estimates are consistent under this sampling scheme (Manski and Lerman, 1977). Sampling is not required for the estimation of (10) and (11), as they rely solely on realized links.

4.5   Results: Homophily and Gender-Biased Preferences

Table 5 shows reduced-form estimates for Directed Homophily obtained from individual-physician data. On average, Medicare male doctors refer 4.3% more to male, compared with female doctors of similar specialty and experience (Column 3). One concern is that

Table 5: Reduced Form Estimates of Directed Homophily

| | Percent of Referrals to Male Specialists | | | | | |
| | OLS | | | | FE | |
	(1)	(2)	(3)	(4)	(5)	(6)
Male Doctor	0.053***	0.043***	0.040***	0.040***	0.029***	0.030***
	(62.7)	(49.1)	(44.8)	(44.0)	(30.5)	(32.6)
Male Patients (pct.)			0.029***	0.028***	0.031***	0.043***
			(16.5)	(14.7)	(16.1)	(23.4)
Cons.	0.79***	0.81***	0.80***	0.81***	0.82***	0.78***
	(1027.6)	(263.8)	(254.3)	(256.9)	(249.4)	(589.1)
Specialty (Doctor)	No	Yes	Yes	Yes	Yes	No
Experience (Doctor)	No	Yes	Yes	Yes	Yes	Yes
Obs. (Doctors)	385104	384985	384985	347534	347534	347534
Groups (Hospital/Spcl)					4819	66563
Rank	2	56	57	57	57	4
Mean Dep. Var.	0.82	0.82	0.82	0.83	0.83	0.83
R^2	0.012	0.038	0.039	0.041		
R^2 Within					0.034	0.0079

Derived from sample of 20pct of patients. (t-statistics in parentheses.)

Notes: Estimates of (10) and (11) are for the sample of doctors with any referrals. Percent of referral to male specialists is the fraction of each doctor's referrals-relationships who are with male specialists. Percent male patients is each doctor's fraction of referred patients who are male. For comparability with the fixed-effect estimates in Columns 5 and 6, Column 4 repeats the same OLS specification from Column 3 for the subsample of doctors affiliated with a hospital. For further details regarding of the sample and the rest of the variables, see Section 2.   * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; t statistics in parentheses.

patient preferences could lead to apparent physician homophily, if patients choose doctors of their own gender and request to see such specialists[20]. I address it by controlling for the gender-mix of patients (Column 4). Doctors with more male patients are more likely to refer to male specialists, suggesting patients too exhibit some homophily. But rather than confound physician homophily, patient homophily coexists with it. About a quarter of overall homophily can be attributed to systematic differences in the opportunity pool of specialists, as seen by the reduction in the magnitude of the estimate by 25% when fixed-effect are included for the doctors' hospital affiliation interacted with their medical specialty (Column 6). That is, male doctors affiliated with the same hospital, and of the same medical specialty (Column 7), refer 3% more to male specialists. As such doctors likely face similar opportunity pools of specialists, this suggests most homophily is driven not by sorting, but by preference for same-gender specialists.

Older doctors (with above-median experience) exhibit greater directed homophily than younger ones (Table A5). Older doctors could be more homophilous because they have more biased preferences. But this age-pattern of gender homophily could be the consequences of a longer accumulation process, in light of later findings that same-gender links are more persistent.

Homophily estimates are virtually unchanged when links are weighted. Appendix Table A6 shows estimation results for different measures of referral volume: number of patients, number of claims, or overall dollar value of services.

That homophily is mostly due to preference is further supported by the fact it is stronger in larger markets (Figure 9). The greater variance in smaller markets is due to the greater variability in the gender composition of smaller pools, and is expected regardless of the cause of homophily. But the increase in homophily with size is a marker of homophily being driven by gender biased preferences (Proposition 2).

Table 6 shows estimates of the link formation model (2). All else equal, doctors are 10% more likely to refer to specialists of the same gender. Estimates represent odds ratios, but due to sparsity they approximately equal the increase in link probability with an increase in the attributes[21]. Distance (proximity) and hospital affiliation are the strongest determinants of referrals, with referrals far more likely between providers sharing an affiliation and within the same zip-code. Modest sorting on location and hospital affiliated is confirmed by the slight decrease in same-gender estimates when they are included as controls.

The estimated gender-bias is comparable with the reduced-form homophily estimates

---

[20]For example, Reyes (2006) shows female patients are more likely to visit female obstetrician-gynecologists.

[21]Estimates close to zero approximately equal the percentage increase in probability. That is, using the notation of (2), around zero $\beta \approx \dfrac{p_{jk}|g_j = g_k}{p_{jk'}|g_j \neq g_{k'}} - 1$.

Table 6: Conditional Logit Estimates: Link Probability

			Link(j,k)		
	(1)	(2)	(3)	(4)	(5)
Link					
Same Gender	0.11***	0.10***	0.10***	0.096***	0.11***
	(56.5)	(51.3)	(51.2)	(46.7)	(37.4)
Male Specialist	-0.029***	-0.024***	-0.024***	-0.00034	-0.016***
	(-14.9)	(-12.3)	(-12.4)	(-0.16)	(-5.22)
Similar Experience		0.0091***	0.0091***	0.0088***	0.0090***
		(111.1)	(110.3)	(102.9)	(73.1)
Same Hospital			1.16***	1.02***	0.95***
			(261.1)	(224.7)	(158.5)
Same Zipcode				2.28***	2.22***
				(487.4)	(373.6)
Same School					0.27***
					(72.8)
Obs. (Dyads)	14,222,742	14,217,381	14,217,381	13,991,067	6,607,612
Clusters (Doctors)	375,032	374,908	374,908	366,968	242,352
Pseudo R Sqr.	0.000	0.001	0.013	0.105	0.106

Notes: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; t statistics in parentheses. Results of conditional logit estimates of (3) for 2012. Data consists of all linked dyads and a matched sample of unlinked dyads, by location and specialty (see text for details). The dependent binary variable is 1 if there was a link between the doctor j and the specialist k during the year. Same gender is a dummy for the specialist and doctors being of the same gender. Male specialists is a dummy for the specialist being male. Similar Experience is the negative of the absolute difference in physicians' year of graduation. Schooling information is only partly available.

Figure 9: Homophily and Market Size
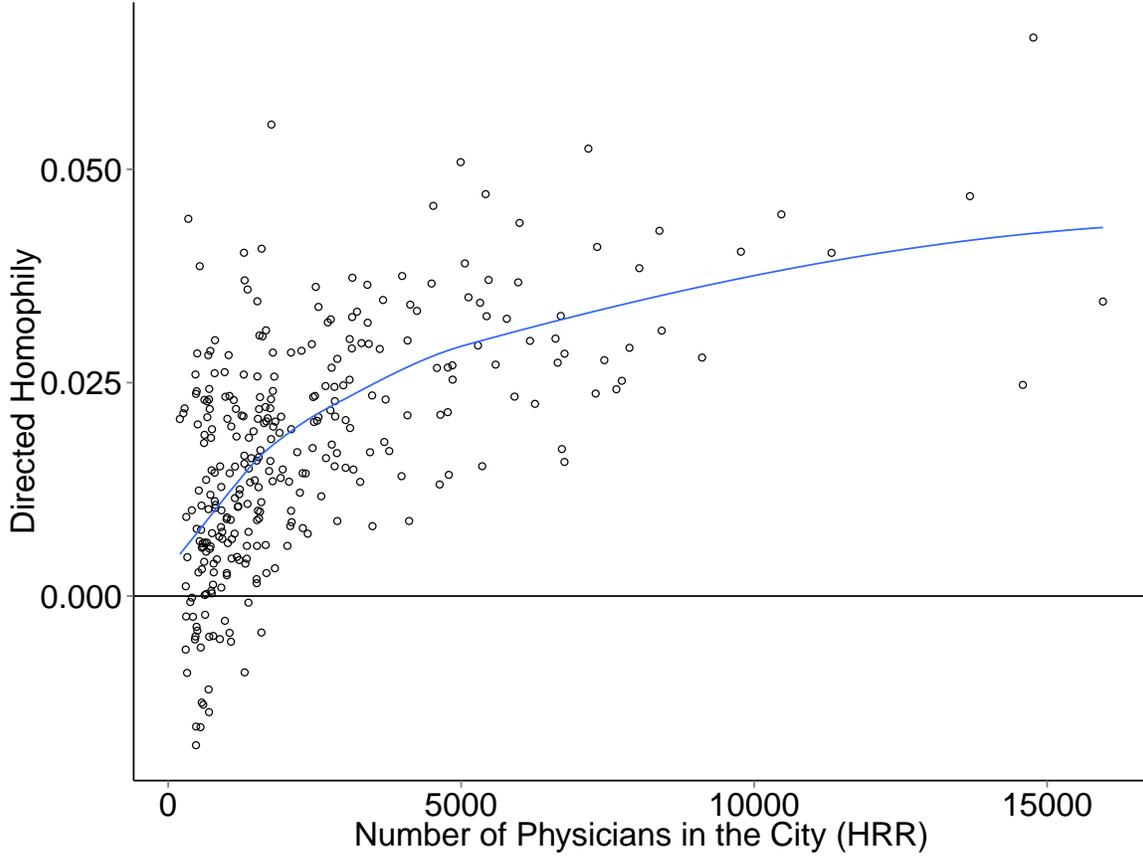
Notes: Homophily estimates of (10), estimated separately for each local physician market (Dartmouth Hospital Referral Region), over the overall number of physician in the market (men an women). Lined are fitted linear and quadratic fitted to the data shown.

from Table 5, as seen by substituting $\hat{\beta} = 0.1$ in (4): facing an opportunity pool with 80% male specialists (roughly the U.S. average), the estimated gender bias of 10% implies directed homophily of 3.2%, net of sorting (the implied directed homophily at the maximally-diverse pool, with the same number of men and women, is 5.0%).

In addition to gender homophily, referrals also show homophily on other dimensions: doctors refer disproportionally to specialists of similar experience, and to specialists that previously went to the same medical school. A doctor and a specialist one year closer in age (approximated by graduation year) have 0.9% greater probability of referrals between them. The other dimension of affinity: having went to the same medical school is also a strong determinant of referrals, with doctors being 30% more likely to refer to same-school graduates. Since medical-school data is partial, estimates with and without inclusion of same-school dummies are presented; they do not differ much.

Measuring the estimated effect of gender against the effects of other attributes implies a

social-distance between the genders: there is a comparable effect on the likelihood of referrals for being of different gender and for having a ten years age difference. And the effect on referrals of being of the same gender is about a third of that of having graduated from the same medical school.

In Appendix C I further study the dynamics of homophily over time, and find same-gender physicians are more likely to maintain referral relationships over time. I find same-gender links are between 1.5–4.5% relatively more likely to persist (i.e., stay active the year after). This suggests same-gender referrals are more common as a consequence of a dynamic process in which same-gender relationships are more likely to survive over time.

In sum, both reduced-form estimates and structural estimates point to a significant gender homophily, mostly due to gender-bias in referral choices. That is, men still refer to men more than their female counterparts facing a similar set of specialists. That homophily is not only due to sorting, on either observed or unobserved dimensions, is bolstered by the findings that homophily is stronger in larger markets, as predicted from the model only for gender-biased preferences. Results imply that increasing the diversity of the opportunity pools would increase homophily, rather than decrease it gender-biased individual preferences are more manifest in diverse pools, which permit choice. Another implication is that homophily should divert demand away from female (the gender minority), and generate a gap in pay. I next turn to test this implication directly.

5 Homophily's Consequences for the Pay Gap

The previous section has shown homophily in referrals is ubiquitous and is largely due to gender-biased preferences. In this section I directly test the model's prediction that gender-biased referrals should divert work away from women, the minority, towards men, the majority. Such an effect could work on both the intensive and the extensive margins: female specialist may be working less, or they may be discouraged from choosing specialties relying on referrals from male. I find evidence for both.

5.1 Homophily's Impact on The Gender Workload Gap

The model above predicts women should receive fewer referrals because they are the minority of doctors and specialists (Proposition 3). Because female specialist pay is predicted to be negatively correlated with to the fraction of male doctors in their market, inter-temporal variations in this fraction identify the link between homophily and pay, and reveal its strength. I find that a higher monthly fraction of claims handled by primary care doctors (who make

most referrals) is associated with greater pay for male and lower pay for female non-primary-care specialists. The effect is large: were half of all primary-care referrals handled by women (as opposed to the current 30%), demand shifted to female specialists would reduce the average gender-pay gap for specialists by between 10–16%.

Strategy

To directly estimate the effect homophily has on the pay gap, I use the implied correlation between the share of work coming from male and specialist demand. I focus on primary-care physicians, who handle the majority of outgoing referrals. Using a monthly panel of physician payments I estimate:

$$log(Pay_{k,t}) = (\beta_M \mathbb{1}_{g_k=M} + \beta_F \mathbb{1}_{g_k=F})m_{c(k,t),t} + \gamma_t + \alpha_k + \varepsilon_{k,t} \tag{13}$$

For all non-primary care physicians k, and months t. The dependent variable $Pay_{k,t}$ is the specialist total monthly Medicare payments; The variable $m_{c(k,t),t}$ is the percent of outgoing primary-care referrals sent by male at specialist k's market at month t; Specialist and time fixed effects, α_k and γ_t, are included. Of interest is the difference $\beta_M - \beta_F$: the differential impact a higher fraction of male doctors has on male and female specialists' pay, tested against the null of unbiased-referrals, where this difference is zero.

The inclusion of specialist fixed effects means the identifying variation is within individual physicians, over time. This specification therefore allows for systematic differences between male and female specialists, differences we know exist (e.g., due to maternity-related no-work spells), and does not confound them with homophily. It also allows for workload to be correlated across specialties. Indeed, it is likely that when primary care doctors see more patients, so do specialists, for example because of seasonality. The identifying assumption is that no omitted factors simultaneously boost the workload of male primary-care physicians and decrease the workload of female physicians in non-primary care specialties. To rule out the possibility that patient homophily may confound the results, controls are included for the monthly fraction of services incurred by male patient[22].

The empirical model (13) is estimated using a monthly panel of individual-physician pay for the period 2008–2012. The data are described in Section 2 above.

---

[22]By including a term $(\delta_M \mathbb{1}_{g_k=M} + \delta_F \mathbb{1}_{g_k=F})\mu_{c(k,t),t}$ where μ is the is the percent of services incurred by male patients at k's market at t. Here too the effect is allowed to differ by specialist gender.

Results: Homophily's Impact on The Gender Workload Gap

When more referrals are handled by male primary-care physicians, demand for male-specialists increases, while demand for female-specialists decreases. Specifically, each 1.0% monthly increase in the fraction of referrals handled by male is associated with 0.47% increase in male workload and a 0.27% decrease in female workload. Results hardly change when controls for patient gender are included, suggesting the effect is not due to homophily on behalf of patients. These results are all identified from within-specialists variation in workload, so they are not an artifact of systematic differences in between male and female specialists labor supply.

Table 7: Male Fraction of Primary Care and Specialist Workload

	(1) Log monthly pay	(2) Log monthly pay
Female Specialist X Pct male PCP (city)	-0.26***	-0.27***
	(0.054)	(0.054)
Male Specialist X Pct male PCP (city)	0.49***	0.47***
	(0.029)	(0.029)
Month Dummies	Yes	Yes
M,F x Pct Male patients (city)	No	Yes
Obs. (Phys x Month)	18087629	18087629
Clusters (Phys)	418939	418939
R Sqr.	0.0323	0.0322

Notes: Fixed-effect estimates of (13) with and without controls for patient gender. For each specialists (non-primary-care physician), monthly pay is the the total monthly pay for Medicare services billed. Specialist gender is interacted with the fraction of claims handled by male primary-care physicians in the same market during the month. In Column (2) it is also interacted, separately, with the percent of services incurred by male patients in the market (as controls). Standard errors are clustered by specialist.

This relationship supports my previous results, showing homophily is indeed due mostly to gender-bias in referrals. As shown above (Proposition 3), demand for specialists varies with their gender and upstream gender fractions only when homophily is caused by gender-biased preferences, not sorting.

The magnitude of the effect of referrals on gender pay disparities is fairly large: considering the counterfactual scenario where female handle exactly half of outgoing primary-care referrals, instead of their current share (about 35%). In such case, the pay gap would decrease by an estimated $(.50 - .35) \times (0.47 + 0.27) = 11\%$, for specialties other than primary care. These estimates do not account for indirect benefits from referred patients, such as returning patients (in specialties where patients are monitored repeatedly), or from additional

patient obtained through word of mouth. However, neither they hold constant that overall volume of patients, and therefore should be taken as suggestive, rather than conclusive.

A more appealing method to quantify the counterfactual impact on the gender pay gap is to use the model estimated in Section 4 (Figure 10). We calculate the counterfactual demand for the average upstream and downstream gender fraction in the United States. Overall, biased-preferences of the level estimated from Medicare referrals result in 5% lower demand for female, relative to male. This is in addition to any other factors that may contribute to the gap, such as differences in labor supply. Note the asymmetric effect of preference-bias on demand by gender: it increases the demand for men while decreasing the demand for women. This is because men are the majority upstream.

Figure 10: Counterfactual Workload Gap for Different Levels of Preference-Bias
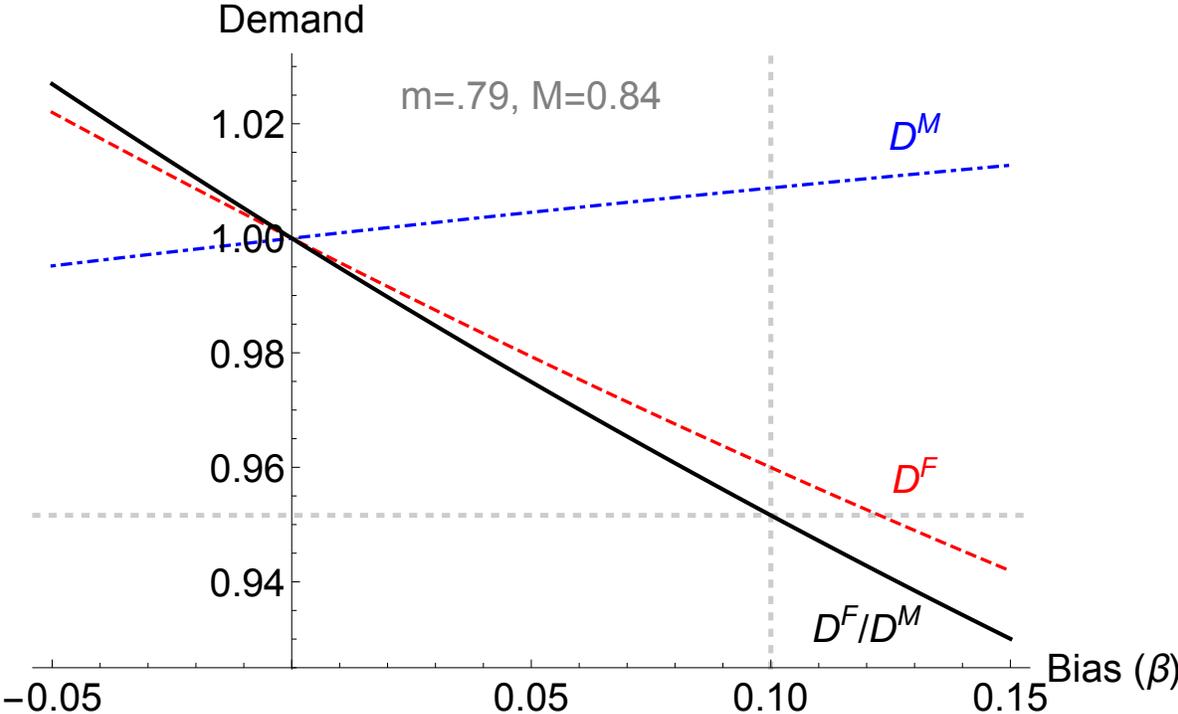
Notes: Average demand for female and male specialists are predicted using the model, given current upstream and downstream male fractions (m, and M). The thick line gives the fraction of the gap contributed solely due to homophily. Eliminating the bias would reduce the gap by about 5% (in popular terms, restoring 5 "cents-per-dollar" to women). See appendix for the same calculation with different values of m and M.

Even by conservative estimates, each year female specialists forego to their male colleagues thousands of dollars worth of work, due to a combination of biased preferences and most current referral being handled by men.

These estimates of the short-run, intensive impact of biased-referrals on the gender pay gap could be further augmented by long-run, extensive effects on specialty choice, discussed

next.

## 5.2   Homophily's Impact on The Gender Specialization Gap

Women choose lower paying specialties. As shown in Section 3, close to 20% of the unconditional pay gap are due to differences in specialty choice. Although recent decades are marked by great equalization in entry into medicine (from 10% female graduates in the 1970s to a half today), still many fewer women enter high-paying, surgical and medical specialties (Figure 11). Differences in specialty choice may exist because women, balancing career and family, put a premium on flexibility or prefer shorter training (for example, surgical specialties require more years of residency, fellowships, and later permit less flexibility in schedule). The great changes in entry of the last decades are due, according to such explanations, to shifting family roles, attitudes and preferences of women.

Yet occupational differences may be further driven by homophily: when a substantial fraction of work in a specialty relies on referrals from men, which advantage men, women may be less inclined to join. In this section I present cross-sectional and longitudinal evidence showing such "Boys' Club" effect is empirically plausible. Currently, there are more women in specialties where more referrals come from women. And analysis of cohort data for 1965–2005 shows female entry to specialties was higher when a greater fraction of their referrals came from female. To identify the impact of homophily on entry, I first use data on patient flows to estimate how medical specialties are interconnected by referrals, then use these estimates to approximate the share of referrals made by female to each medical specialty at each past period.

Strategy: Identifying Homophily's Impact on Medical Specialty Choice

My empirical strategy to identify the link between homophily and specialization relies on testing whether over time, women enter specialties where a larger share of referrals comes from women. The main concern is to separately identify the effect of homophily in referrals from unobserved specialty characteristics that may attract women. To address it, I use variation between cohorts in the fraction of referrals each specialty received from females at the time of specialty choice, variation induced by differential female entry to *related* specialties over time. For example, cardiac surgeons often receive referrals from cardiologists, but rarely if ever from dermatologists. So because of homophily female entry to cardiac surgery should have responded to the fraction of female cardiologists, but not to the fraction of dermatologists. To extend this idea to all medical specialties, I leverage the fact clinical relations between specialties, which are essentially a technology parameter unrelated to gender, and

Figure 11: Female Participation Rates by Medical Specialty, 1965–2005
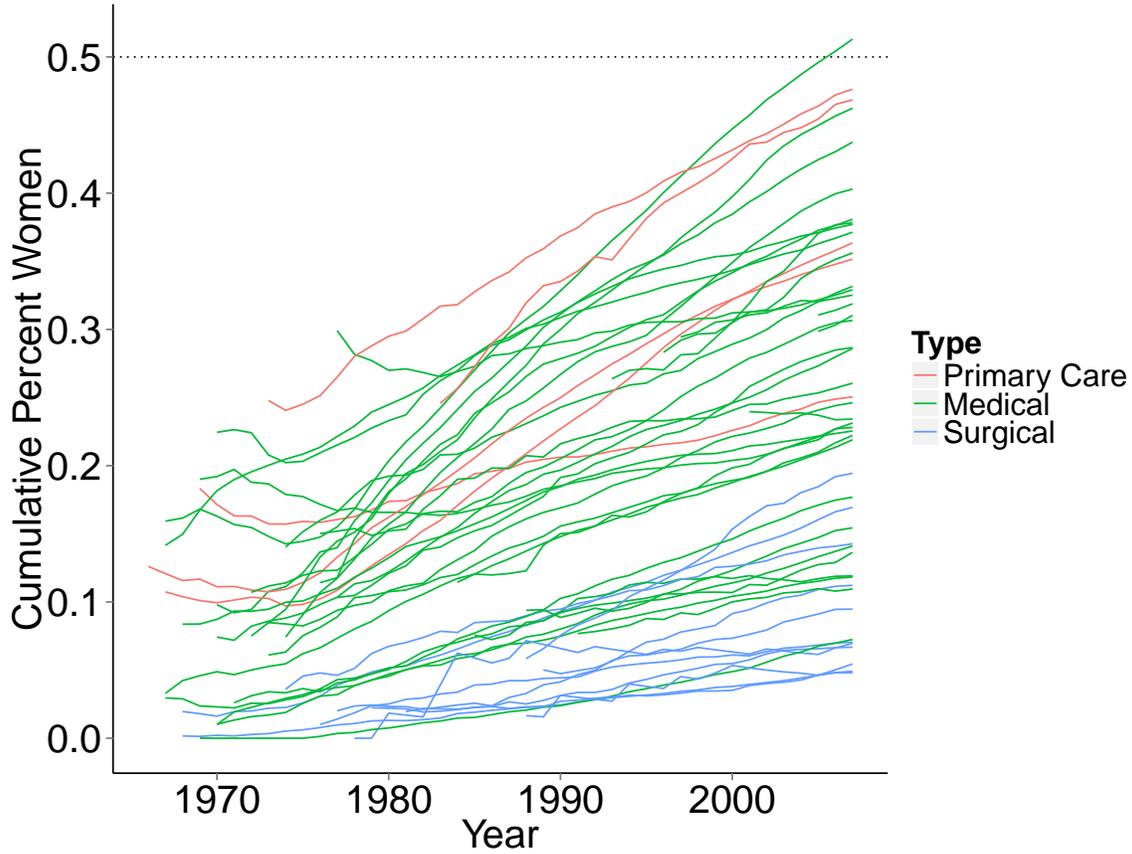
Notes: Percent of active female physicians by medical specialty, for different specialty categories. Source: Cohort data reconstructed from current data on graduation years. Specialty-year cells with fewer than 500 physicians were omitted.

are revealed from extensive data on patient flows.

To provide theoretical basis for the following test, I being by formalizing the interconnectedness of specialties used for identification, by extending the model from Section 4.3 to multiple specialties.

Extension of the Model to Multiple Specialties Let (with the usual abuse of notation) $S = (1, \ldots, S)$ be the set of medical specialties, and suppose doctors choose specialists as in (2), separately for each specialty. The relative volume of referrals between specialties depends on medical practice and can be summarized in a transition matrix:

$$\mathbf{R} = \left[r_{s|s'} \right]_{S \times S} \tag{14}$$

Where $r_{s|s'}$ is the average fraction of specialty s referrals coming from specialty s', so $\sum_{s' \in S} r_{s|s'} = 1$. The fraction of referrals made by men to each specialty depends on exactly which other specialties refer to it. To see how, for each specialty $s \in S$ denote by M_s be the fraction of men practicing it, and by m_s the fraction of referrals to it made by men. In vector notation: $\mathbf{M} = (M_1, \ldots, M_S)'$, and $\mathbf{m} = (m_1, \ldots, m_S)'$. Then:

$$\mathbf{m} = \mathbf{R}'\mathbf{M} \tag{15}$$

Namely, for each specialty, the fraction of referrals coming from male is the weighted average of the fractions of men in clinically related specialties[23]. For example, if referrals occur only within-specialty, $\mathbf{R}$ is the identity matrix and $\mathbf{m} = \mathbf{M}$. The matrix $\mathbf{R}$ is generally asymmetric.

The demand for specialists of each medical specialty now depends on the fractions of men in all related specialties, in proportion to how related they are. Formally, the demand for male specialists in specialty s varies with the fraction of males in another specialty s' as follows:

$$\frac{\partial D_s}{\partial M_{s'}} = \frac{\partial D_s}{\partial m_s}\frac{\partial m_s}{\partial M_{s'}} = (p_{M|m} - p_{M|f})r_{s|s'} \tag{16}$$

for all pairs of specialties $s \neq s' \in S$. This is an extension of (8) to multiple specialties.

Identification and Estimation: Homophily and Medical Specialty Choice by Gender Variation in the fraction of female in related specialties is used to identify the impact of homophily on entry, based on the relationships given in (16): Observing the changes in $\mathbf{M}$ over time weighted by $\mathbf{R}$ gives a set of demand shifters that vary between specialties and over time. I proceeds in two step. First, estimate $\mathbf{R}$ from current observed referrals flows across all specialties. Second, using a panel of gender fractions in all specialties M_{st} and rates of entry, $\Delta M_{st} = M_{s,t} - M_{s,t-1}$, I estimate how has entry responded to contemporaneous gender mixes of doctors in related specialties[24]:

$$\Delta M_{st} = \beta_4 \hat{m}_{st} + \eta_s + \gamma_t + \varepsilon_{st} \tag{17}$$

Where the scalar $\hat{m}_{st} = (\hat{\mathbf{m}}_t)_s = (\hat{\mathbf{R}}'\mathbf{M}_t)_s$ is the estimated fraction of referrals to specialty s made by male at period t (to avoid endogeneity same-specialty referrals are omitted, by zeroing the diagonal of $\hat{\mathbf{R}}$). Fixed effects η_s and γ_t for specialty and year capture potential specialty-specific attributes that impact referrals and a flexible time-trend in entry. The

---

[23]That is, $m_s = \sum_{s' \in S} r_{s|s'} M_{s'}$

[24]To keep previous notation, the following is articulated in terms of male, rather than female fractions, but by symmetry the same holds verbatim for female

parameter β_4 captures the response of male entry to male participation in related upstream specialties (likewise for female), where relatedness is estimates in the first stage. The identification assumptions are thus (a) the clinical-relatedness of specialties has not changed (absent historical data, I estimate it from current data), and (b) there are no unobserved time-varying factors beyond referrals that cause male entry to be correlated with the stock of male in clinically-related specialties (that is, beyond time trends in entry that impacts all specialties together and time-invariant differences in entry rates to the different specialties).

Since data only cover 2008–2012, the model (17) is estimated using retrospective cohort data covering a longer period of time: 1965–2005. This period saw substantial variation in entry, and is more suitable for studying long-term effects. Clearly, a limitation of retrospective cohort data is that exits are not observed, so results should be interpreted in light of the possibility that they reflect gender differences in exit rates. The panel is truncated at 2005, to permit medical school graduates several years to appear on claims.

Results: Homophily and Medical Specialty Choice by Gender The first-stage estimates, $\hat{\mathbf{R}}$ are visualized in Figure A3. Clearly, primary care is the main source of referrals, and there are some within-specialty referrals as well. There is, however, considerable variation across specialties, which is used for identification in the second stage.

Table 8: Female Entry and Specialty Fraction of Referrals from Female

	Percent Female Entrants			
	(1)	(2)	(3)	(4)
Percent Referrals from Female	1.255***	1.167***	1.372***	1.198***
	(27.17)	(15.23)	(35.00)	(7.65)
Year Fixed-Effects	No	Yes	No	Yes
Specialty Fixed-Effect	No	No	Yes	Yes
Obs. (Specialty-Year)	1501	1501	1501	1501
R-sqr.	0.296	0.306	0.842	0.851
Clusters	time	time	time	time
N. Cluster	41	41	41	41

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Notes: Results of fixed-effect estimates of (17) using cohort data. Observation represent specialty-year pairs (s, t), for specialty-years with at least 500 physicians. Percent Referrals from Female is the fraction of active female in all other specialties $s' \in S \setminus \{s\}$ at year t, weighted by the estimated dependency of s on each of s' for referrals from the first stage, depicted in appendix Figure A3. Standard errors were clustered by year.

I find that over 1965–2005 more women entered specialties with a greater fraction of referrals from women, beyond specialty-specific entry rates and a common flexible time-trend

(Table 8). Bearing the limitation of reconstructed cohort panel in mind, these results still suggest homophily had an impact on female specialization: it affects the pay gap also through this extensive margin. Results mean some of the occupational segregation in medicine is due not to supply side differences, but rather to demand effects of gender-biased referrals. Results also suggest further female entry to upstream specialties like primary care would facilitate entry of female to more lucrative specialties by increasing the demand to their services. Quantifying how much of the gap is explained by entry differences requires more assumptions

6 Extensions

This section reports two additional analyses that further explore the causes and consequences of homophily. The first uses data on physician names to test the hypothesis that homophily is driven by outright discrimination. The second, uses data on patients to check whether homophily has impact on their care. I find no evidence of prejudice based on first names, and no evidence for impact on patient outcomes. Combined with earlier findings on the presence of preference-bias and adverse effects on female workload and specialty choice, these results propose a more nuanced picture of gender homophily: while it may not be ill-disposed (although no-prejudice is not confirmed, just not rejected), and not outrageously harmful, it is still detrimental for female physicians.

6.1 Testing for Prejudice Using Specialist Names

"What's in a name?"

In this section I further leverage data on physician first names to test for prejudice. I perform a quasi-audit study, comparing whether specialists with gender-ambiguous names (e.g. Alex, or Robin) are treated differently than specialists with generized names (e.g. David, or Jennifer). If prejudiced doctors choose referrals based on a specialist names alone, without even knowing them, specialists whose names are uninformative of gender show experience milder homophily. I find no evidence for that: doctors choices discriminate even between ambiguously named specialists, suggesting they know the specialists they work with, at least enough to know their gender not just based on their name.

Identifying prejudice is challenging because it involved elicitation of beliefs, and is often reserved for audit studies. The problem is separating the impact beliefs have on choices from the impact of other privately observed information. Audit studies elicit beliefs by manipulating signals while holding constant the information available to subjects. For example,

Bertrand and Mullainathan (2004) elicit racial discrimination by sending fictitious resumes with names sounding either very African American or very White. Since names convey little information other than on individual race, a differential response indicates the presence of discriminatory beliefs regarding racial groups[25].

With observational data one cannot control for private information, so I use a different design: instead of considering responses to very genderized names, I consider responses to gender-ambiguous names. By definition, such names are uninformative of the gender of the specialists, so doctors who discriminate based on names alone could not discriminate against such individuals. Therefore, responses to ambiguous names identify whether choices respond to perceived and not actual gender of individuals. This is tested against the null that referrals to ambiguously names individuals are still homophilous, which means doctor prefer same-gender specialists based on more than just their name.

This strategy is implemented in two steps. First, I classify the names of all specialists by their *name masculinity* $\gamma \in [0,1]$, defined as the fraction of name-bearers who are male[26]. Higher name masculinity corresponds to a higher probability that the name-holder is male. Names with masculinity $\gamma = 0$ are only given to female (e.g., Jennifer), conversely, names with masculinity $\gamma = 1$ are only given to men (e.g., David). Names with index between 0 and 1 are sometimes given to both genders (e.g., Alex or Robin). The most ambiguous names have $\gamma = 0.5$: they are as likely to be male as they are to be female (neglecting the prior).

In the second stage, I then test whether Directed Homophily is lower for specialist with ambiguous names, against the null that homophily is unrelated to name-masculinity. The point is that knowing the names of specialists with ambiguous names do not reveal their gender, so discrimination against them suggests doctors base referral decisions on more than just names[27]. The identifying variation is in the informativeness of name as a signal for

---

[25]Note that whether discriminatory beliefs exist, and whether such beliefs are rationalized by group differences in productivity are separate questions.

[26]Formally, denote the gender of specialist k by $g_k \in \{M, F\}$ and the first name of an individual k by $m_k \in M = \{m_1, m_2, \ldots, m_M\}$ I assign an index $\gamma : M \mapsto [0,1]$ measuring *name masculinity*:

$$\gamma(m) = \frac{|\{k \in K : m_k = m, g_k = M\}|}{|\{k \in K : m_k = m\}|}$$

[27]More formally, let Γ denote a level of name-masculinity (an element of a partition of $[0,1]$ to bins, e.g.,$\{[0,.1), [.1,.2), \ldots, [.8,.9), [.9,1]\}$). Define $r_{g|M}^{\Gamma}$ to be the average fraction of referrals to male specialists received from doctors of gender g (for $g = m, f$), calculated over the set of specialists $k \in K$ with name-masculinity $\gamma(m_k) \in \Gamma$. I then test for discriminatory beliefs by considering the relationship between ambiguity and homophily. That is, if homophily is driven by discriminatory beliefs then one should expect (H_1) $DH(\Gamma) = r_{m|M}^{\Gamma} - r_{f|M}^{\Gamma}$ to be lower for ambiguous names ($\Gamma \ni \frac{1}{2}$ or close to that), than it is for unambiguous names (Γ close to 0 or 1). (I use the version of DH defined by comparing incoming referrals

Figure 12: Homophily and First-Name Masculinity
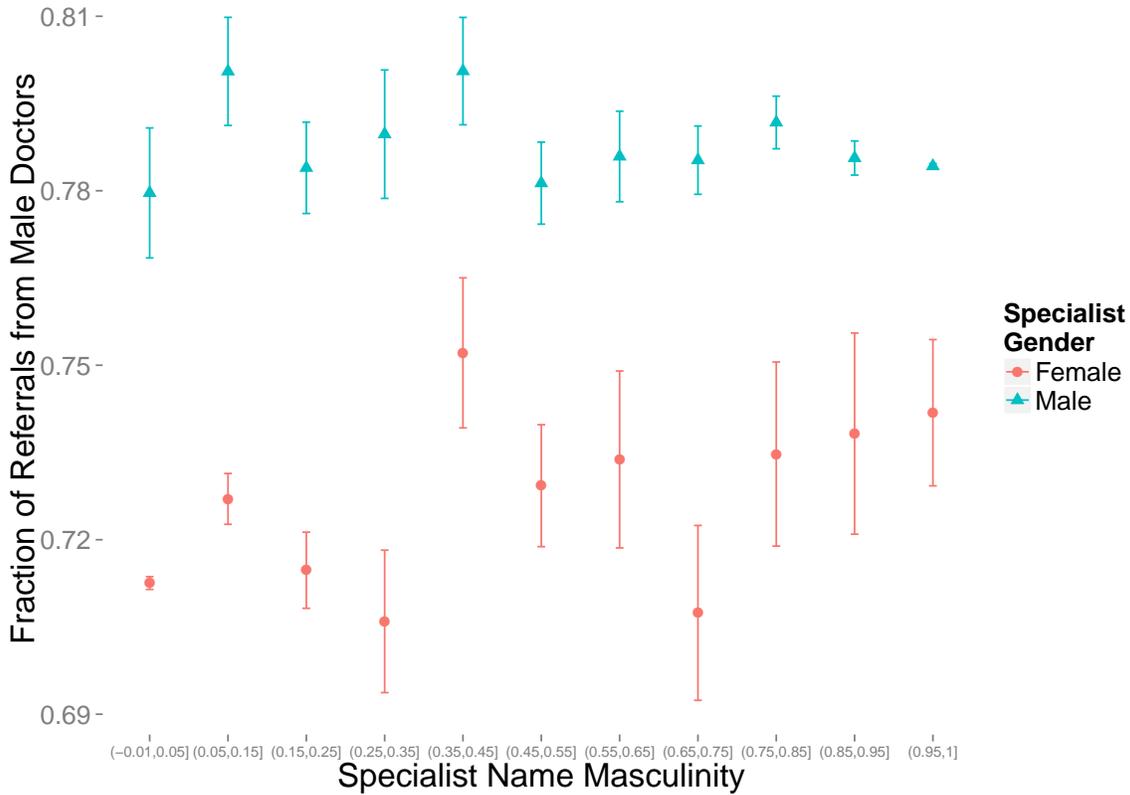
Notes: The figure shows the fraction of referrals specialists of different name-masculinity received from male doctors. Name-masculinity is defined as the share of name-holders who are male. The two extreme bins contain all unambiguous names (about 93% of specialists): left-most bin contain all specialists with feminine names (e.g. Jennifer), the right-most bin contains all specialists with masculine names (e.g., David). The middle bins contain more ambiguous names (e.g., Alex or Robin). That male doctors treat male and female with ambiguous names differently suggest that they do not base their referral decisions on name alone.

gender, across different levels of name masculinity.

Figure 12 shows the results of the first-name analysis, plotting rates of referrals from males and females to male specialists with different levels of name-masculinity[28]. Male and female referrals to ambiguously-named specialists (mid-range masculinity) still shows significant homophily: men refer more and women refer less even to ambiguously named specialists. Thus the null hypothesis is not rejected: referrals seems not to be prejudiced, at

---

across specialists, not outgoing referrals across doctors. The two measures always have the same sign.) The null (H_0) that there is no difference in DH between more and less ambiguous names.

[28]The standard errors for the estimates of referral rates to men are decreasing in the specialist name masculinity because of sample size: by construction, there are many more men with higher name-masculinity than there are men with lower name-masculinity. Conversely for women.

least not based on names alone.

6.2 Homophily and Patient Outcomes

I find no significant impact of homophily on patient mortality or cost.

The above analysis has shown that doctors are more inclined to refer to specialists of their own gender, which hurts female physicians, the minority gender. Next, its effects on patients are studied. Gender-homophily in referrals could harm patients, if patient care is compromised by preferring inappropriate specialists over appropriate ones of the opposite gender. On the contrary, homophily could have no effect on outcomes if gender differences in specialist quality are insubstantial, or if gender-preferences only break ties between otherwise-equally appropriate specialties. Homophily could even have a positive effect, for example if physicians of the same gender communicate better.

To test whether homophily has consequences for patients, I compare outcomes of patients who have been referred and treated by same-gender and mix-gender providers, using the following specification:

$$Y_{ijk}^{t+1} = \beta_5 \mathbb{1}_{g_j = g_k} + \delta_5 X_i^{t-1} + \alpha_j + \alpha_k + \varepsilon_{ijk} \tag{18}$$

where each observation is a triple (i, j, k) of patient i referred to specialist k by doctor j at period t, for a single base year t. The dependent variable Y is subsequent patient outcomes: period $t+1$ overall cost of Medicare services, a proxy for sickness severity. The parameter of interest β_5 captures the impact of same-gender dyads.

To control for doctor and specialist differences, fixed effects for both are included. Accounting for such differences is important because otherwise physicians treating more complex patients, who are of higher risk, would appear to have worse outcomes, a typical selection problem. To control for patient heterogeneity, X includes detailed patient characteristics, including demographics, utilization, and cost for $t-1$. Cost and utilization for t are excluded, to avoid endogeneity: they may be outcomes of the physicians encounters of interest. As the same patient could encounter multiple dyads, standard errors are clustered by patient.

Patients referred and treated by providers of the same gender fare similarly as those who saw providers of opposite gender, as seen in Tables A7 and A8. These table shows the estimates of (18) for two different outcomes: patient log annualized cost and mortality. Patient referred and treated by physicians of the same gender have slightly lower subsequent cost and mortality rates, but these differences all but vanish when physician fixed-effects are included, suggesting male physicians treat sicker patients. There seems to be no impact of the provider gender mix on patient.

46

Evidence is far from conclusive, but it suggests homophily has on average no effect on patient outcomes: gender biased referrals do not appear to compromise patients. Further research is required to evaluate the impact of homophily, as zero averages effect may conceal heterogeneity in effects. Overall, widespread homophily impacts physicians but not patients.

7 Conclusion

This paper have examined a demand-side explanation for the physician gender pay gap: gender-homophily in referrals (physicians refer to their same gender). Existing supply-side explanations explain only part of the gender pay gap in Medicare, which is a combination of women choosing lower-paying specialties and working fewer hours. A model of referral choices extends existing theory of homophily to directed networks, and shows that it implies a relationship between gender of physicians and demand for their services. The model also yields testable differences between two mechanisms causing homophily: preferences and sorting, and clarifies how homophily contributes to the pay gap.

The model raises two conceptual points. First, referral networks (more broadly, directed networks) provide an opportunity to measure homophily in a way robust to underlying heterogeneity, by comparing the gender-distribution of referrals not to baseline gender fractions but between genders. Second, in directed networks, both upstream and downstream gender distributions matter for how homophily impacts downstream demand. Being the minority of both, female specialists face a lower demand for their services because they receive fewer referrals than men.

Estimates of the model and its predictions using Medicare data on physician referrals and reimbursement, spanning all specialties and local U.S. markets for 2008–2012, show physician referrals exhibit significant gender homophily, partly due to sorting, bust mostly due to gender-biased preferences. All else equal, physicians are less likely to refer to the other gender, and more likely to refer to their own gender. As long as men are the majority of referring physicians, homophily in referrals diverts demand away from female. On the intensive margin homophily is currently causing female to forgo to their male colleagues thousands of dollars worth of work each year. Moreover, evidence exists for effects on the extensive margin, namely on entry: women stay away from specialties to which more referrals come from men, both historically and currently. Evidence for this "Boys' Club" effect is admittedly weaker, as it is based on retrospective cohort data. But if it indeed exists, it will further contribute to pay disparity. On a brighter note, results suggest that current entry trends will have favorable impact on female physicians in the future, particularly the equalization of female participation in primary care which is a main source of referrals for

many other specialties.

Even among the highly-educated, preference biases in favor of same-gender, and more broadly in favor of similar others are ubiquitous. Such biases, while slight and innocuous from an individual perspective, translate to systematic disadvantage to minority groups, particularly when resources are made available through connections. Data on such connections and the social interactions facilitating their formation can be elicited from transaction-level data of the kind that is increasingly becoming available, and its analysis could therefore shed light on the propagation of inequality, on dimensions beyond gender and in domains beyond medicine.

References

Antoninis, Manos (2006), "The wage effects from the use of personal contacts as hiring channels." *Journal of Economic Behavior & Organization*, 59, 133–146.

Barnett, Michael L, Nancy L Keating, Nicholas A Christakis, A James OMalley, and Bruce E Landon (2012a), "Reasons for choice of referral physician among primary care and specialist physicians." *Journal of general internal medicine*, 27, 506–512.

Barnett, Michael L, Bruce E Landon, A James O'Malley, Nancy L Keating, and Nicholas A Christakis (2011), "Mapping physician networks with self-reported and administrative data." *Health services research*, 46, 1592–1609.

Barnett, Michael L, Zirui Song, and Bruce E Landon (2012b), "Trends in physician referrals in the united states, 1999-2009." *Archives of internal medicine*, 172, 163–170.

Bertrand, Marianne, Claudia Goldin, and Lawrence F Katz (2010), "Dynamics of the gender gap for young professionals in the financial and corporate sectors." *American Economic Journal: Applied Economics*, 228–255.

Bertrand, Marianne and Sendhil Mullainathan (2004), "Are emily and greg more employable than lakisha and jamal? a field experiment on labor market discrimination." *The American Economic Review*, 94, 991–1013.

Bramoullé, Yann, Sergio Currarini, Matthew O Jackson, Paolo Pin, and Brian W Rogers (2012), "Homophily and long-run integration in social networks." *Journal of Economic Theory*, 147, 1754–1786.

Burks, Stephen V, Bo Cowgill, Mitchell Hoffman, and Michael Gene Housman (2014), "The value of hiring through employee referrals." *Qaurterly Journal of Economics*.

Chen, M Keith and Judith A Chevalier (2012), "Are women overinvesting in education? evidence from the medical profession." *Journal of Human Capital*, 6, 124–149.

Choudhry, Niteesh K, Joshua M Liao, and Allan S Detsky (2014), "Selecting a specialist: Adding evidence to the clinical practice of making referrals." *JAMA*, 312, 1861–1862.

Croson, Rachel and Uri Gneezy (2009), "Gender differences in preferences." *Journal of Economic literature*, 448–474.

Currarini, Sergio, Matthew O Jackson, and Paolo Pin (2009), "An economic model of friendship: Homophily, minorities, and segregation." *Econometrica*, 77, 1003–1045.

Currie, Janet and W Bentley MacLeod (2013), "Diagnosis and unnecessary procedure use: Evidence from c-section." Technical report, National Bureau of Economic Research.

Currie, Janet, W Bentley MacLeod, and Jessica Van Parys (2015), "Physician practice style and patient health outcomes: The case of heart attacks." Technical report, National Bureau of Economic Research.

Dustmann, Christian, Albrecht Glitz, and Uta Schönberg (2011), "Referral-based job search networks."

Echenique, Federico and Roland G Fryer (2007), "A measure of segregation based on social interactions." *The Quarterly Journal of Economics*, 441–485.

Esteves-Sorenson, Constança, Jason Snyder, et al. (2012), "The gender earnings gap for physicians and its increase over time." *Economics Letters*, 116, 37–41.

Forrest, Christopher B, Paul A Nutting, Barbara Starfield, and Sarah Von Schrader (2002), "Family physicians' referral decisions." *J Fam Pract*, 51, 215–222.

Gandhi, Tejal K, Dean F Sittig, Michael Franklin, Andrew J Sussman, David G Fairchild, and David W Bates (2000), "Communication breakdown in the outpatient referral process." *Journal of General Internal Medicine*, 15, 626–631.

Goldin, Claudia (2014), "A grand gender convergence: Its last chapter." *The American Economic Review*, 104, 1091–1119.

Goldin, Claudia and Cecilia Rouse (2000), "Orchestrating impartiality: The impact of" blind" auditions on female musicians." *The American Economic Review*, 90, 715–741.

Golub, Benjamin and Matthew O Jackson (2012), "How homophily affects the speed of learning and best response dynamics."

Graham, Bryan S (2014), "An empirical model of network formation: detecting homophily when agents are heterogenous."

Hackl, Franz, Michael Hummer, and Gerald Pruckner (2013), "Old boys' network in general practitioner's referral behavior."

Ho, Kate and Ariel Pakes (2014), "Physician payment reform and hospital referrals." *The American Economic Review*, 104, 200–205.

Johnson, Erin M (2011), "Ability, learning and the career path of cardiac specialists." *Cambridge, Massachusetts Institute of Technology.*

King, Gary and Langche Zeng (2001), "Logistic regression in rare events data." *Political analysis*, 9, 137–163.

Kossinets, Gueorgi and Duncan J Watts (2006), "Empirical analysis of an evolving social network." *Science*, 311, 88–90.

Kossinets, Gueorgi and Duncan J Watts (2009), "Origins of homophily in an evolving social network." *American Journal of Sociology*, 115, 405–450.

Kugler, Adriana D (2003), "Employee referrals and efficiency wages." *Labour economics*, 10, 531–556.

Leung, Michael (2014), "A random-field approach to inference in large models of network formation." *Available at SSRN.*

Lo Sasso, Anthony T, Michael R Richards, Chiu-Fang Chou, and Susan E Gerber (2011), "The $16,819 pay gap for newly trained physicians: the unexplained trend of men earning more than women." *Health Affairs*, 30, 193–201.

Manski, Charles F and Steven R Lerman (1977), "The estimation of choice probabilities from choice based samples." *Econometrica*, 1977–1988.

Massey, Douglas S and Nancy A Denton (1988), "The dimensions of residential segregation." *Social forces*, 67, 281–315.

McGee, Andrew, Peter McGee, and Jessica Pan (2015), "Performance pay, competitiveness, and the gender wage gap: Evidence from the united states." *Economics Letters*, 128, 35–38.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook (2001), "Birds of a feather: Homophily in social networks." *Annual review of sociology*, 415–444.

Mehrotra, Ateev, Christopher B Forrest, and Caroline Y Lin (2011), "Dropping the baton: specialty referrals in the united states." *Milbank Quarterly*, 89, 39–68.

Reyes, Jessica Wolpaw (2006), "Do female physicians capture their scarcity value? the case of ob/gyns." Technical report, National Bureau of Economic Research.

Seabury, Seth A, Amitabh Chandra, and Anupam B Jena (2013), "Trends in the earnings of male and female health care professionals in the united states, 1987 to 2010." *JAMA internal medicine*, 173, 1748–1750.

Song, Zirui, Thomas D Sequist, and Michael L Barnett (2014), "Patient referrals: a linchpin for increasing the value of care." *JAMA*, 312, 597–598.

Weeks, William B, Tanner A Wallace, and Amy E Wallace (2009), "How do race and sex affect the earnings of primary care physicians?" *Health Affairs*, 28, 557–566.

Appendices

A   Proofs

Proof. (Claim 1) Pick any j such that $g_j = m$. Summing up probabilities of referrals to all available specialists gives:

$$
\begin{aligned}
p_{M|m} &= \sum_{k:g_k=M} p_{jk} \\
&= \frac{\sum_{k:g_k=M} e^{\beta \mathbb{1}_{g_j=g_k}}}{\sum_k e^{\beta \mathbb{1}_{g_j=g_k}}} \\
&= \frac{\sum_{k:g_k=M} e^{\beta}}{\sum_{k:g_k=M} e^{\beta} + \sum_{k:g_k\neq M} e^{0}} \\
&= \frac{Me^{\beta}}{Me^{\beta} + 1 - M}
\end{aligned}
$$

The probability $p_{M|f}$, and (6) are similarly derived. $\qquad\square$

Proof. (Claim **??**) The overall conditional probability is a weighted average of market-specific conditional probabilities (weights are proportional to both market size and the relative share

of male doctors in each market). Using Bayes rule:

$$p_{M|m} = \sum_{c \in C} p_{c|m} p_{M|m,c} = \sum_{c \in C} \mu^c \frac{m^c}{m} p_{M|m}^c$$

$$\geq \sum_{c \in C} \mu^c \frac{m^c}{m} M^c = \frac{1}{m} E[m^c M^c]$$

$$> \frac{1}{m} E[m^c] E[M^c] = M$$

The first inequality is due to preferences: $p_{M|m}^c \geq M^c$ (equality being the case $\omega = 1$), and the second is due to segregation. By symmetry, the same proof works for female. $\square$

Alternatively, for the more general definition, segregation* ($\mathrm{Cov}[m_j, M^j] > 0$), the proof follows immediately from Claim 1: with unbiased preferences $p_{M|m} = E[M^j | g_j = m] > M$, by segregation*. QED. Note that segregation* is indeed more general, as by covariance decomposition, $\mathrm{Cov}[m_j, M^j] = \mathrm{Cov}[m^c, M^c]$ under separate markets with common $K_j = K^c$ in each.

Proof. (Proposition 1)

$$p_{M|m} - M = \sum_{c \in C} \mu^c (\frac{m^c}{m} p_{M|m}^c - \frac{m^c}{m} M^c + \frac{m^c}{m} M^c - M^c)$$

$$= \sum_{c \in C} \mu^c (\frac{m^c}{m} (p_{M|m}^c - M^c) + M^c (\frac{m^c}{m} - 1))$$

$$= E[\frac{m^c}{m} (p_{M|m}^c - M^c)] + \mathrm{Cov}[\frac{m^c}{m}, M^c]$$

$\square$

Proposition 4 (Directed Homophily Decomposition). *The overall Directed Homophily decomposes as follows:*

$$E[\frac{m^c}{m} p_{M|m}^c - \frac{1 - m^c}{1 - m} p_{M|f}^c] + \frac{1}{m(1 - m)} \mathrm{Cov}[m^c, M^c] \tag{19}$$

Proof. (Proposition 4) Applying the proof of Proposition 1 to female (by symmetry) and substituting $p_{M|f} = 1 - p_{F|f}$ yields :

$$M - p_{M|f} = E[\frac{1 - m^c}{1 - m} (M^c - p_{M|f}^c)] + \mathrm{Cov}[\frac{m^c}{1 - m}, M^c]$$

Hence

$$p_{M|m} - p_{M|f} = \mathrm{E}[\frac{m^c}{m}(p^c_{M|m} - M^c) + \frac{1-m^c}{1-m}(M^c - p^c_{M|f})]$$
$$+ \frac{1}{m(1-m)}\mathrm{Cov}[m^c, M^c]$$

rearranging yields the result. □

Proof. (Proposition 2) Let $h(M) = p_{M|m} - p_{M|f}$, so $DH(N) = \mathrm{E}[h(M)|N]$. Recall that: $h(M)$, given in Claim 1, is concave in M (strictly for $\beta > 0$) and satisfies $h(0) = h(1) = 0$. Since entry is independent, the fraction of male specialists in a given market c, denoted M^c, is a random variable distributed $F(N) := Binomial(N^c, M)/N^c$, with mean M. By the Law of Large Numbers, for every $\varepsilon > 0$ there exists N' such that for $N'' > N'$ the probability of $|F(N'') - M| > \varepsilon$ is arbitrarily small, so $h(F(N''))$ is arbitrarily close to $h(M)$. the result then follows from the concavity of h by Jensen's inequality. □

Proof. (Proposition 3) Pick any male specialist k. The demand k faces in market c is obtained by aggregating over all doctors in that market (as all variables are market-specific I suppress the superscript c):

$$D_M = \sum_{j \in J} p_{jk} = \sum_{j \in J} \frac{e^{\beta s(j,k)}}{\sum_{k' \in K} e^{\beta s(j,k')}}$$
$$= \sum_{j \in J, g_j=1} \frac{e^{\beta s(j,k)}}{\sum_{k' \in K} e^{\beta s(j,k')}} + \sum_{j \in J, g_j=0} \frac{e^{\beta s(j,k)}}{\sum_{k' \in K} e^{\beta s(j,k')}}$$
$$= \frac{1}{N}(\sum_{j \in J, g_j=1} \frac{1}{M + \omega(1-M)} + \sum_{j \in J, g_j=0} \frac{\omega}{\omega M + (1-M)})$$
$$= \frac{n}{N}(\frac{m}{M + \omega(1-M)} + \frac{\omega(1-m)}{\omega M + (1-M)})$$

Where $n = |J|$ and $N = |K|$. When $\omega = 1$ then $D_M = \frac{n}{N}$ which is independent of both M and m. Suppose $\omega < 1$. To see 2 is true rewrite:

$$D_M = \frac{n}{NM}\left(mp_{M|m} + (1-m)p_{M|f}\right)$$
$$= \frac{n}{NM}\left(p_{M|f} + m(p_{M|m} - p_{M|f})\right)$$

and note that $\partial D_M / \partial m > 0$ since $p_{M|m} - p_{M|f} > 0$ for every $\beta > 0$. To see 3 is true take

the derivative of D_M with respect to M:

$$\frac{\partial D_M}{\partial M} = \frac{n(1-w)}{N}\left(\underbrace{\frac{(1-m)w}{(1-M(1-w))^2}}_{\text{Complements}} - \underbrace{\frac{m}{(M+w(1-M))^2}}_{\text{Substitutes}}\right)$$

The denominators of the terms labeled "Complements" and "Substitutes" are both positive. Therefore, for m near enough zero, "Complements" dominates and the derivative $\partial D_M/\partial M$ is positive, whereas for m near enough one "Substitutes" dominates and the derivative is negative. For intermediate values of m, the sign of the derivative may depend on M. □

B   Counterfactuals

In section 5.1 I have estimated the impact of homophily on the pay gap directly, as well as provided counterfactual estimates of the impact of estimated bias on demand. In this section, I provide model counterfactual for how much would have the gap been reduced had there been no bias in preferences, for different upstream and downstream gender fraction.

Table A1: Homophily Pay-Gap Counterfactuals (Female Cent per Male Dollar)

Upstream (m)	Fraction Males Downstream (M)						
	0.4	0.5	0.6	0.7	**0.8**	0.9	1
0.4	1.019	1.020	1.021	1.022	**1.023**	1.024	1.025
0.5	0.999	1	1.001	1.002	**1.003**	1.004	1.005
0.6	0.979	0.980	0.981	0.982	**0.983**	0.984	0.985
0.7	**0.960**	**0.961**	**0.962**	**0.962**	0.963	**0.964**	**0.965**
0.8	0.941	0.942	0.942	0.943	**0.944**	0.944	0.945
0.9	0.923	0.923	0.923	0.924	**0.924**	0.925	0.925
1	0.905	0.905	0.905	0.905	**0.905**	0.905	0.905

Notes: Using estimated bias $\hat{\beta} = 0.1$, the table shows calculated earnings gaps: $\frac{D^F}{D^M}(m, M; \beta)$, due to homophily related workload differences, for different gender distributions upstream and downstream. The formula is given below. The row and column in bold are the current U.S. averages across all specialties. At current rates the gender-bias in referrals alone is contributing 1-.963=3.7% to the physician earnings gap.

Table A1 shows estimation of the impact on the pay gap contributed by the estimated preference-bias $\hat{\beta} = 0.1$ (i.e., doctors are 10% more likely to work with same-gender others), buy calculating the demand disparity with different populations upstream and downstream:

$$\frac{D^F}{D^M}(m, M; \beta) = \frac{M}{1-M}\frac{mp_{F|m} + (1-m)p_{F|f}}{mp_{M|m} + (1-m)p_{M|f}} \tag{20}$$

Where $p_{G|g} = p_{G|g}(G; \beta)$ for $g \in \{m, f\}$ and $G \in \{M, F\}$, are given in (4) above. Note that with balanced gender ($M = m = 0.5$) there is no gap, even when preferences are biased: in this case homophily only affects the composition of demand, not its level.

C   Homophily Dynamics

The above analysis relied on a cross-section data. Here longitudinal data on the evolution of the network of referrals over several years is used here to estimate the dynamics in referral relationships. I find same-gender links persist longer in time.

For the study of link persistence, I estimate the following specification:

$$
p^{persist}_{jk,t+1|jk,t} := Pr(Y_{jk,t+1} = 1 | Y_{jk,t} = 1, g, X) = \frac{e^{\eta_{jkt}}}{1 + e^{\eta_{jk't}}} \tag{21}
$$

using data on all dyads (j, k) such that $Y_{jk,t} = 1$, where $Y_{jk,t} = 1$ if j referred to k at period t and $Y_{jk,t} = 0$ otherwise, and $\eta_{jkt} := \alpha_j + \beta \mathbb{1}_{g_j = g_k} + \delta X_{jkt}$. That is, (21) estimates the probability of links (referral relationships) existing at t would still exist at $t + 1$. Each dyad is included only once: for the first year it is observed. Since this specification is restricted to existing links, no sampling is necessary: all observed dyads are used.

Results: Link Persistence and Homophily Dynamics

Existing link are relatively more likely to persists between same-gender providers. Table A2 shows different estimates of link persistence, obtained from the sample of all initially connected dyads (physicians with referral relationships at the base year, defined as the first year they were observed in the data). Both logit and linear estimates with two-way fixed effects (for doctors and for specialists) show that same-gender links are more likely than cross-gender links to carry on to the following year (Columns 1–2). Columns (3) and (4) estimate separately for male and female doctors the probability of links persisting, again using physician fixed-effects to account for individual heterogeneity in the persistence of relationships. Consistent with the findings above, that male are much more likely to receive referrals, both male and female doctors' relationships with male specialists are more persistent, but persistence is significantly higher for male doctors than it is for female doctors ($p < 0.001$). That is, same-gender relationships persist relatively longer in time.

Table A2: Estimates: Link Persistence

	Link Persists Next Year			
	(1)	(2)	(3)	(4)
	Logit	FE	FE	FE
Same Gender	0.044***	0.014***		
	(16.2)	(24.0)		
Male Doctor	0.069***			
	(16.3)			
Male Specialist	0.16***		0.029***	0.0062***
	(57.4)		(50.4)	(5.89)
Similar Experience	0.0046***	0.0011***	0.0016***	0.00085***
	(38.1)	(39.5)	(55.3)	(15.8)
Same Hospital	0.12***	0.027***	0.030***	0.027***
	(28.5)	(29.5)	(31.6)	(14.3)
Same Zipcode	0.16***	0.097***	0.092***	0.076***
	(55.1)	(145.1)	(129.9)	(56.3)
Same School	0.088***	0.013***	0.015***	0.014***
	(26.9)	(17.1)	(20.0)	(9.09)
Constant	-0.81***			
	(-193.7)			
Specialty (Specialist)	No	No	Yes	Yes
Obs. (j,k)	7255778	7204471	5734596	1496658
Rank	8	5	58	58
R^2		0.20	0.10	0.11
N. Cluster	280750	255507	191647	64579
FE1 (Doctors)		255507	191647	64579
FE2 (Specialists)		237363		

Notes: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; t statistics in parentheses. Results of link persistence estimates. Column (1) shows estimates (3) for 2008–2012. Data consists of an observation for each linked dyad (j, k), for the first year it was observed in the data. The dependent binary variable is 1 if the link between the doctor j and the specialist k continued during the subsequent year. Same gender is a dummy for the specialist and doctors being of the same gender. Male specialists/doctor is a dummy for the specialist/doctor being male. Similar Experience is negative the absolute difference in physicians' year of graduation. Column (2) shows linear estimates with two-way fixed effect (for doctor and for specialist) using the same data. Columns (3) and (4) show linear estimates with one fixed-effects (for doctor), separately for female (3) and male (4) doctors. Sample size is restricted by the availability of medical school data. Results excluding school affiliation are very similar. All standard errors are clustered by doctor.

D   Additional Tables and Figures

Figure A1: Homophily Through Link Weights

(a) Homophily via Weights   (b) Unweighted
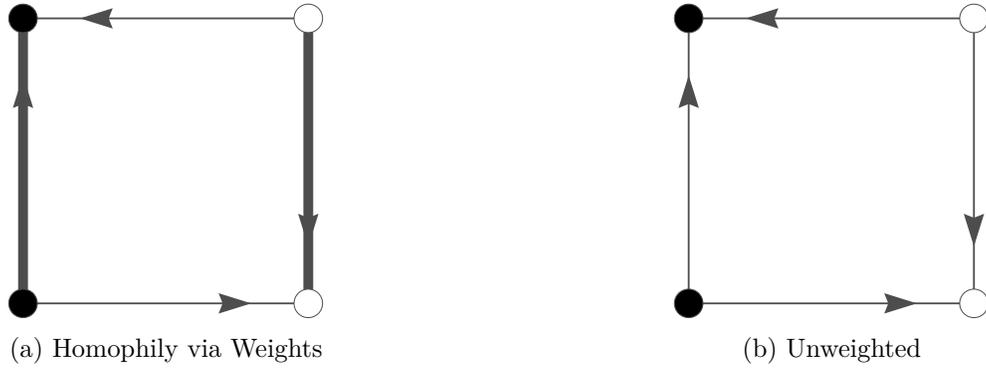
Ignoring referral volume (link weights) may yields wrong measures of homophily: network (a) exhibits homophily: more patients are referred within-gender (higher weights between same-color nodes); this homophily is concealed if link weights are ignored (b).

Figure A2: Homophily Through Link Direction

(a) Homophily   (b) No Homophily   (c) Undirected
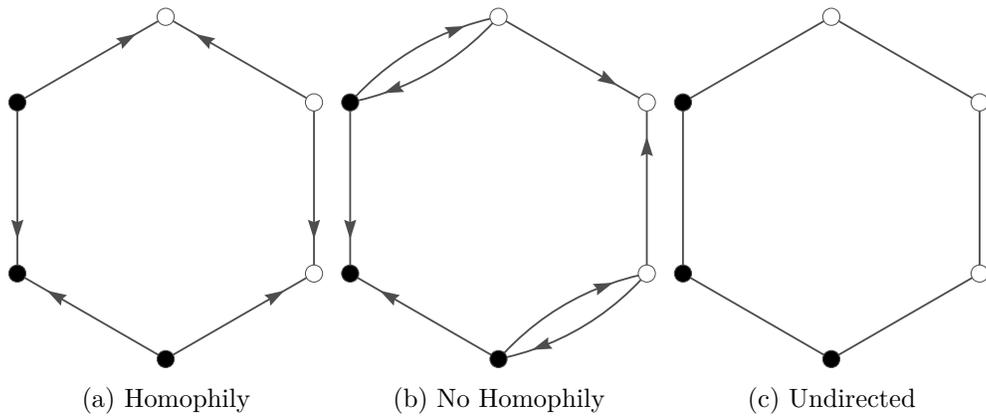
Looking at directed networks as if they are undirected yields wrong measures of homophily: the network (a) exhibits homophily while (b) does not, a difference concealed in their undirected counterpart (c). Also concealed is the asymmetry in gender shares: e.g., in (a) most most senders are black, but most receivers are white; that is, most doctors are of one gender but most specialists are of the other.

Table A3: 2012 Average Degree by Specialty

	Specialty	Indegree	Outdegree	Physicians
1	Internal medicine	5.8	26.4	86,220
2	Family practice	1.9	21.0	74,638
3	Anesthesiology	19.2	0.4	33,434
4	Obstetrics/gynecology	2.7	3.2	22,871
5	Cardiology	36.3	12.3	21,714
6	Orthopedic surgery	17.4	13.8	19,411
7	Diagnostic radiology	10.2	0.6	18,768
8	General surgery	14.3	12.8	18,011
9	Emergency medicine	4.5	5.2	16,065
10	Ophthalmology	14.0	9.1	15,702
11	Neurology	25.5	5.2	11,469
12	Gastroenterology	35.2	11.7	11,178
13	Psychiatry	4.8	2.7	10,861
14	Dermatology	16.6	3.1	8,624
15	Pulmonary disease	30.9	12.5	8,272
16	Urology	33.9	13.7	8,234
17	Otolaryngology	24.2	7.0	7,666
18	Nephrology	32.2	13.0	7,105
19	Hematology/oncology	23.6	13.1	7,019
20	Physical medicine and rehabilitation	17.7	5.7	6,224
21	General practice	2.5	14.7	4,853
22	Endocrinology	20.4	7.2	4,534
23	Infectious disease	24.2	5.0	4,492
24	Neurosurgery	19.8	16.7	4,010
25	Radiation oncology	17.3	3.4	3,933
26	Rheumatology	20.8	7.8	3,765
27	Plastic and reconstructive surgery	7.3	5.0	3,759
28	Pathology	2.3	0.4	3,627
29	Allergy/immunology	11.5	2.0	2,768
30	Pediatric medicine	1.8	3.8	2,695
31	Medical oncology	20.6	12.9	2,507
32	Vascular surgery	30.7	18.1	2,486
33	Critical care	16.5	9.5	2,046
34	Thoracic surgery	15.2	18.1	1,886
35	Interventional Pain Management	27.2	5.5	1,655
36	Geriatric medicine	4.9	20.8	1,597
37	Cardiac surgery	16.4	18.0	1,526
38	Colorectal surgery	22.1	16.8	1,161
39	Pain Management	22.3	4.3	1,055
40	Hand surgery	19.4	10.0	1,047
41	Interventional radiology	25.1	2.6	938
42	Gynecologist/oncologist	13.4	15.3	834
43	Surgical oncology	12.1	16.0	684
44	Hematology	16.9	9.4	667
45	Osteopathic manipulative therapy	4.7	9.0	463
46	Nuclear medicine	5.2	1.4	289
47	Preventive medicine	4.2	5.7	217
48	Maxillofacial surgery	3.6	3.2	164
49	Oral surgery	2.8	2.6	108
50	Addiction medicine	1.7	3.7	77
51	Peripheral vascular disease	32.0	13.9	62
52	Neuropsychiatry	16.2	4.8	61
53	Podiatry	17.7	3.4	40

Notes: A link represents referral relationships with another physician from any specialty.

Table A4: Percent Men and Percent of Referrals from Men, by Medical Specialty

		Dependent variable:	
		Percent Men	
	(1)	(2)	(3)
Pct Ref From Male	1.260***	1.210***	1.220***
	(0.160)	(0.149)	(0.150)
Training Duration		0.066***	0.064***
		(0.013)	(0.013)
Pct Work in Weekend			−0.137
			(0.151)
Constant	−0.237*	−0.486***	−0.474***
	(0.123)	(0.124)	(0.125)
Observations	172	172	172
Adjusted R²	0.264	0.361	0.361

Note:	*p<0.1; **p<0.05; ***p<0.01

Notes: Percent Men is the fraction of male specialists. Percent Referrals from Male is the fraction of referrals from male doctors in all other specialties. Training Duration is residency and fellowship duration in years. Percent Work in Weekend is the fraction of claims for the specialty that record services incurred in weekends.

Table A5: Homophily Estimates for Different Age Groups

| | Percent of Referrals to Male Specialists | | |
	Young	Old	All
Male Doctor	0.038***	0.044***	0.040***
	(0.0011)	(0.0015)	(0.00090)
Male Patients (pct)	0.028***	0.031***	0.029***
	(0.0024)	(0.0026)	(0.0018)
Constant	0.79***	0.81***	0.80***
	(0.0078)	(0.0040)	(0.0032)
Specialty (Doctor)	Yes	Yes	Yes
Experience (Doctor)	Yes	Yes	Yes
Obs. (Doctors)	200670	184315	384985
Rank	57	57	57
Mean Dep. Var.	0.82	0.83	0.82
R^2	0.035	0.041	0.039

Notes: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; standard errors in parentheses. OLS estimates of (10) are shown for three subgroups: young doctors (below median experience of 24 years, Column 1); old doctors (above median experience, Column 2); and all doctors together (Column 3). Despite the similar opportunity pools they face, older doctors exhibit stronger average directed homophily than younger ones.

Table A6: Homophily Estimates with Weighted Links

| | Percent ... to Male Specialists: | | | |
| | (1) | (2) | (3) | (4) |
	Links	Patients	Claims	Dollars
Male Doctor	0.038***	0.040***	0.040***	0.040***
	(43.2)	(44.8)	(42.7)	(41.4)
Percent Male Patients	0.029***	0.029***	0.029***	0.029***
	(16.6)	(16.5)	(16.1)	(15.4)
Cons.	0.80***	0.80***	0.80***	0.81***
	(262.2)	(254.3)	(243.8)	(243.9)
Specialty (Doctor)	Yes	Yes	Yes	Yes
Experience (Doctor)	Yes	Yes	Yes	Yes
Obs. (Doctors)	384985	384985	384985	383054
R^2	0.0384	0.0394	0.0360	0.0368

Derived from sample of 20pct of patients. (t-statistics in parentheses.)

Notes: OLS estimates of (10) using different definitions of link weights: The first column show results for unweighted links. Columns 2–4 shows results for different weights: number of patients, number of claims, and Dollar value of services.

Table A7: Physician Dyad Gender Mix and Subsequent Patient Spending

	(1) 2012 Cost	(2) 2012 Cost
Same Phys Gender	-0.0165***	-0.00554**
	(-8.28)	(-2.77)
Male Patient	0.0421***	0.0316***
	(22.89)	(16.94)
Male Doctor	0.0354***	
	(17.14)	
Male Specialist	0.0149***	
	(7.19)	
2010 Cost	0.126***	0.127***
	(159.97)	(164.23)
2010 Drugs	0.0345***	0.0333***
	(138.58)	(126.94)
Patient Demographics, Chronic Cond., Utilization	Yes	Yes
Experience	Yes	No
Specialty	Yes	No
Obs. (Patient, Doctor, Specialists)	7120083	7424095
Rank	169	58
R^2	0.240	0.349
Clusters	dyad	dyad
No. Cluster	4444359	4581862
FE1 (Doctors)		302997
FE2 (Specialists)		276488

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table A8: Physician Dyad Gender Mix and Patient Mortality

	(1) Mortality (1 Year)	(2) Mortality (1 Year)
Same Phys Gender	-0.00303***	-0.00169***
	(-8.43)	(-4.52)
Male Patient	0.0270***	0.0245***
	(85.46)	(76.33)
Male Doctor	0.00759***	
	(20.24)	
Male Specialist	0.00459***	
	(12.17)	
2010 Cost	-0.000100	0.000253**
	(-1.12)	(2.78)
2010 Drugs	-0.000642***	-0.000577***
	(-12.11)	(-10.63)
Patient Demographics, Chronic Cond., Utilization	Yes	Yes
Experience	Yes	No
Specialty	Yes	No
Obs. (Patient, Doctor, Specialists)	7120083	7424095
Rank	169	58
R^2	0.104	0.236
Clusters	dyad	dyad
No. Cluster	4444359	4581862
FE1 (Doctors)		302997
FE2 (Specialists)		276488

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Figure A3: Estimates of **R**
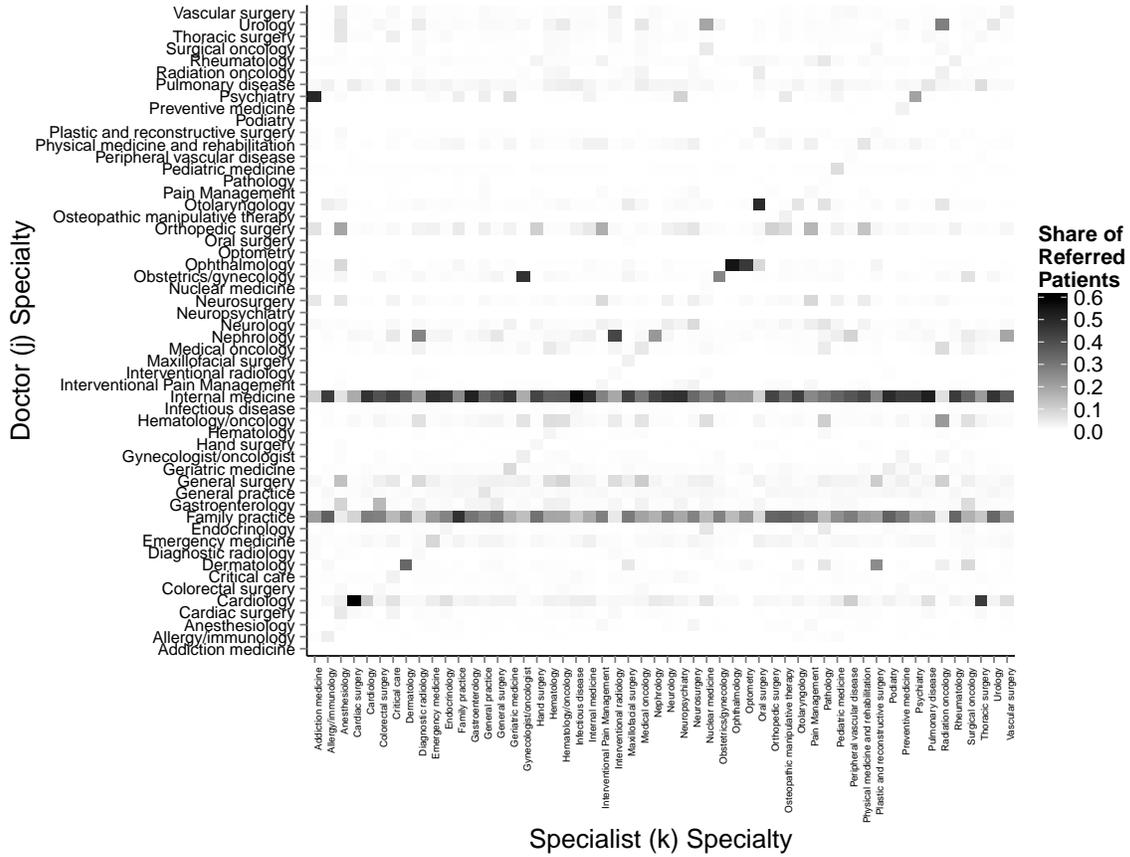
Notes: Estimated transition probabilities between different specialties (numbers denote CMS specialty codes). Cell (j, k) represents the fraction of referrals to specialty k coming from specialty j, with darker colors represent higher fraction. Columns each sum to one. The dark horizontal lines are the two main primary care specialties: family medicine (8) and internal medicine (11).