Team Relationships and Performance: Evidence from Healthcare Referral Networks

Leila Agha
Dartmouth College and NBER

Keith Marzilli Ericson
Boston University's Questrom School of Business and NBER

Kimberley H. Geissler
University of Massachusetts at Amherst School of Public Health and Health Sciences

James B. Rebitzer
Boston University's Questrom School of Business and NBER

January 28, 2021

Abstract

We examine the teams that emerge when a primary care physician (PCP) refers patients to specialists. When PCPs concentrate their specialist referrals—for instance, by sending their cardiology patients to fewer distinct cardiologists—repeat interactions between PCPs and specialists are encouraged. Repeated interactions provide more opportunities and incentives to develop productive team relationships. Using data from the Massachusetts All Payer Claims Database, we construct a new measure of PCP team referral concentration and document that it varies widely across PCPs, even among PCPs in the same organization. Chronically ill patients treated by PCPs with a one standard deviation higher team referral concentration have 4% lower health care utilization on average, with no discernible reduction in quality. We corroborate this finding using a national sample of Medicare claims and show that it holds under various identification strategies that account for observed and unobserved patient and physician characteristics. The results suggest that repeated PCP-specialist interactions improve team performance.

**Acknowledgements*: We acknowledge funding from the Boston University Questrom School of Business, the National Institute of Health Care Management, the National Institute of Health grant PO1 AG005842, and Agency for Healthcare Research and Quality grant 1R03HS025515-O1A1. We are especially grateful to Benjamin Lubin for his thoughtful contributions to the project. Buqu Gao provided excellent research assistance. We thank Michael Barnett, Jay Bhattacharya, Jon Skinner, Doug Staiger, and the Anthem Payment Innovation team for helpful comments, along with seminar participants at Boston University, Dartmouth College, Emory University, Massachusetts Institute of Technology, RAND, Rensselaer Polytechnic Institute, University of Pennsylvania, University of Toronto, and Yale School of Public Health, and conference attendees at the ASHEcon, Caribbean Health Economics Symposium, iHEA, Midwest Health Economics Conference, ASSA Annual Meeting, and the NBER . We are also grateful to Jean Roth and Mohan Ramanujan for assistance obtaining and managing the Medicare claims data.

1. Introduction

Teams are pervasive in economic organizations, making team performance an important determinant of productivity. The efficiency of teams is especially important in healthcare because a wide array of specialist and primary care clinicians must work together to treat patients with complex health problems. Compounding the clinical complexity is the reality that contracts between referring parties are either absent or quite incomplete, and the law greatly restricts the use of incentives in referrals.[1]

In this paper, we investigate whether the structure of referrals between primary care physicians (PCP) and specialists affects team performance, measured primarily by the cost of healthcare. Our analysis is motivated by the idea that repeated interactions between PCPs and specialists help overcome barriers to high performance. Specifically, when PCPs concentrate their patient referrals within a narrower group of specialists (e.g., refer their cardiology patients to a smaller set of cardiologists), this encourages repeat interactions. Repeat interactions improve team-specific knowledge, coordination, and trust, in part by facilitating investments in working well together. We propose a novel empirical measure—*team referral concentration*—to capture these effects, and we use this measure to study the impact of PCP-specialist team structure on patient costs.

Empirically, we find that patients of PCPs with more-concentrated specialist referrals have lower total healthcare utilization. For commercially insured, chronically ill patients in Massachusetts, those treated by PCPs with below-median team referral concentration have 6% higher utilization on average compared to those treated by PCPs with above-median team referral concentration. This relationship holds after controlling for detailed patient, physician, and insurer characteristics. We replicate our results in Medicare data and find similar effects in that population.

[1] The "Stark laws" effectively prohibit physicians from being financially compensated for referrals. As a result, a referring physician cannot write an incentive contract. Kolber (2006) provides more detail. Theoretical work by Garicano and Santos (2004) highlights the importance of contracting for efficient referrals when diagnosis is costly; this finding suggests that Stark law restrictions may exacerbate inefficient referral patterns.

Our empirical strategies address two major potential confounding factors. First, doctors who have lower team referral concentration may also have more costly practice styles along other dimensions; for example, they may be "cowboys" who have a taste for high-cost interventions (see Cutler et al. 2019). However, we find that the effect of team referral concentration on spending persists even using within-PCP, across-specialty variation in team referral concentration. The effect is similarly present when comparing patients who consult the same specialist but whose referring PCP has different levels of team referral concentration. Two-way fixed effects models that include both specialist and PCP fixed effects exploit both sources of variation and show similar results.

The second identification threat is that patients choosing lower team referral concentration PCPs may be in worse health (along unobservable dimensions) and demand more costly care. Our analysis controls for a rich set of health conditions used in risk adjustment models. For Medicare beneficiaries, we also employ an additional identification strategy to study patients who switch doctors as the result of a move across geographic regions. In these analyses, we use patient fixed effects to control for stable differences in patient demand for care, and we still find that switching to a PCP with higher team referral concentration is associated with reductions in care utilization.

Our study differs from much of the prior research on care coordination in that we measure relationships within physician teams who may interact repeatedly over the course of treating multiple patients, rather than focusing on how an individual patient's care is spread across physicians. Team-based measures are quite different from individual-based measures of care coordination. To see this, imagine a PCP who refers each of her patients with diabetes to a cardiologist and an endocrinologist. The PCP could refer each of her patients with diabetes to a single cardiologist and a single endocrinologist. Alternatively, she could refer each of these patients to a different cardiologists and endocrinologists. In both cases, the distribution of an individual patient's care across physicians is the same: each patient sees his PCP, cardiologist, and endocrinologist. But in the former case, the physician team works together more

frequently and has higher team referral concentration.

The consistency of our findings across both a commercially insured, working-age population and a Medicare population is noteworthy. It suggests that insurance characteristics are not likely to be driving the relationship between team referral concentration and healthcare spending. The persistence of this pattern also reduces concerns that unique features of healthcare delivery in one state drive our findings.

While our main empirical result focuses on the connection between team referral concentration and costs, we also examine healthcare quality. A priori, team referral concentration has an ambiguous effect on quality. Higher team referral concentration enables repeated interactions between PCPs and specialists, which may increase quality through improved coordination. On the other hand, considering a smaller set of specialists for potential referrals may reduce quality, if the match between patient and specialist varies idiosyncratically and having some familiarity with a wider set of specialists enables more-tailored referral choices. Care quality is multi-dimensional and difficult to assess across a large population of diverse patients, but we investigate several utilization-based measures of care quality. We find no evidence that greater team referral concentration reduces quality as measured by inpatient admissions, emergency department visits, or distance traveled to specialists.

The paper proceeds as follows. Section 2 discusses prior work on team organization and productivity as well as the specific setting of physician teamwork. Section 3 discusses the mechanisms through which repeated interactions may influence team performance. Section 4 introduces our empirical measure of team referral concentration. Section 5 describes the Massachusetts All Payer Claims Data (APCD). Section 6 presents results on how team referral concentration influences care utilization. Section 7 examines the relationship between team referral concentration and healthcare quality. Section 8 extends our work to the Medicare sample and provides an alternative approach to identification. Section 9 concludes.

2. Background

A. Team Organization and Productivity

Our study contributes to a growing body of empirical literature on the economics of team organization and productivity.[2] In work that examines software development teams, Faraj and Sproull (2000) highlight the role of coordination of expertise on team performance. In the healthcare context, Gittell et al. (2008, 2020) find that nursing home teams and outpatient surgical teams with greater relational coordination also have higher levels of staff engagement, job satisfaction, and care quality. Prior studies report a link between team familiarity (i.e., repeated interactions between team members) and team performance for inventor teams (Jaravel et al. 2018), software development teams (Huckman et al. 2009), surgical teams (Reagans et al. 2005), and physician-nurse teams in emergency medicine (Kim et al. 2019). Huckman and Pisano's (2006) study of cardiac surgeons is especially relevant. They find that a surgeon's mortality outcomes at a hospital improve as they perform more surgeries at a particular hospital, but operations performed elsewhere do not affect that hospital's results. This result suggests that repeated interaction with the team at a hospital or learning about some other component of the hospital's working environment improves performance.

In addition to contributing to the general literature on team performance, our paper also relates to a large body of literature arguing that better care coordination may reduce healthcare costs and improve quality.[3] This research inspired important policy initiatives that aim to improve care coordination (e.g.

[2] See Bloom and Van Reenan (2011) for a detailed discussion of how team management practices are related to productivity. Other research has studied team formation. Hamilton, Nickerson, and Owan (2003) examine the impact of team versus individual work for productivity in a garment plant and examine who chooses to join a team. Experimental economists have examined the formation of teams in the lab (e.g. Weber 2006; Feri, Irlenbusch, and Sutter 2010; Grosse, Putterman, and Rockenback 2011). There is also a literature in psychology on the performance of teams, reviewed in Kozlowski and Ilgen (2006).

[3] See Agha, Ericson, and Zhao (2020), Agha et al. (2019), Hussey et. al. (2014), Romano et. al (2015), Berwick and Hackbarth (2012), and Milstein and Gilbertson (2009). In economics, care coordination is often referred to by the obverse term, "care fragmentation." For a discussion and review of the literature on fragmentation, see Cebul et. al. (2008), Frandsen and Rebitzer (2014), and Rebitzer and Votruba (2011).

Accountable Care Organizations and Patient Centered Medical Homes) and investments in health information technology and electronic health records.

B. Clinical Teamwork

Managing effective teamwork across care providers is often described as a crucial function of primary care (Bodenheimer 1999; Press 2014), and professional advice to specialist physicians recommends fostering relationships with referring physicians to build a successful specialty practice (Arulrajah 2017; Dyrda 2012). Stille (2005) identifies a successful primary care model as one in which referring occurs within a "tight web of consultants in which physicians know one another well and can share work effectively."

Referrals are often driven by personal connections: in a survey covering over 2000 referrals, family physicians recommended a specific specialist to the patient in 86% of referrals, and personal knowledge of the specialist was the most commonly cited reason for selecting a specific specialist (Forrest et al. 2002). A subsequent survey found that over 60% of PCP referrals cite ease of communication with the specialist as a reason for referral choice (Barnett et al. 2012b). These descriptions of physician teams align with economic theory suggesting that familiarity and repeated interactions are important to productive team relationships (cf. Crawford 1990).

Referrals between PCPs and specialists are a central aspect of clinical teamwork. Mehrotra et al. (2011) report breakdowns and inefficiencies in all aspects of the specialty referral process. Many referrals do not contain sufficient data for medical decision making, and imperfect information flows degrade care coordination. For instance, PCPs often do not know whether a patient actually visited the referred specialist or are not informed of the recommended care plan. Coordination is further complicated by ambiguity about the specialist's role. In some cases, a specialist acts as a consultant to assist the PCP, while in other cases, a specialist becomes a co-manager of care with the PCP.

Poor communication and ambiguous lines of responsibility between PCPs and specialists can lead to a

wide variety of undesirable and costly outcomes. These include missed or delayed diagnoses, referrals to inappropriate specialists, costly cascades of low-value testing and follow-up care, and duplicative testing and redundant patient visits.

Team-specific investments and learning from previous interactions can enhance collaboration between specialists and referring PCPs. For example, institutionalizing regular PCP-specialist conversations about respective roles can clarify lines of responsibility. Similarly, establishing channels for pre-referral check-ins between PCPs and specialists can enhance information flows and detect unnecessary or misdirected referrals (Mehrotra et al. 2011). Even in the context of episodic specialty care such as orthopedic surgery or gynecological cancer, the prior literature finds a strong correlation between the providers' self-reported relational coordination and patient's surgical outcomes (Gittell et al 2000), describes the importance of provider-to-provider communication (Zuchowski et al 2017), and shows that managing referrals deliberately can improve healthcare delivery (Harrington et al 2005). Andreatta and Marzano (2012) write that interdisciplinary approaches to obstetric and gynecologic care are common when managing complex cases and further identify "human factors associated with team performance as a prime improvement area for clinical patient care."

There is also some evidence that concentrating referrals among a small set of specialists is associated with improved team performance. Simon et al. (2017) compared the attributes of six "high value" primary care practice sites with four "average" sites. High-value sites were those whose average patient outcomes placed them in the top quintile of both the cost and quality metrics derived from a large sample of commercial health insurance plan enrollees. Compared to the average practice, the high-performing primary care sites relied on a narrowly selected list of trusted specialists, and they also devoted more attention to ensuring that patients completed referrals to these specialists. Although this study involved a small sample and the interpretation is not causal, the pattern suggests that the structure of referrals may matter for clinical teams.

3. Teams and Repeated Interactions

Conventionally, teams exhibit clearly defined membership, task interdependencies, and shared responsibility for outcomes (see e.g. Sundstrom, De Meuse, and Futrell 1990). Teams formed by PCP/specialist referrals look a bit different. They are similar to conventional teams in that patient care tasks are interdependent and the team shares responsibility for patient outcomes, but they differ in that the teams we study are often ad hoc groupings brought together for a specific type of patient or set of issues (see Andreatta 2010). Team boundaries and roles in our setting can be more variable than in archetypical teams, and our teams share some features with temporary organizations (Bechky 2006), project teams (Edmondson and Nembhard 2009), and fluid teams (Huckman and Staats 2011; Valentine et al. 2018). PCP-specialist relationships can span formal organizations, occur within a formal organization, or both. These characteristics of PCP/specialist referral teams can be found elsewhere—for example, in settings that rely heavily on freelancers and independent contractors such as construction, university collaboration, or development of open source software. Because our teams are most similar to fluid or temporary teams, care must be taken in generalizing our results to more traditional teams.

Coordinating complex tasks and sharing information within teams is difficult (Marschak and Radner 1972; Becker and Murphy 1992; Dessein and Santos 2006), but repeated interactions can facilitate more-productive teams by increasing the knowledge and trust required for successful team operations.

The economic theory of knowledge acquisition—the theory of human capital—makes a fundamental distinction between general and specific knowledge (Lazear, 2009). General knowledge improves productivity in any setting, and so the benefits of knowledge acquisition do not depend on the length or intensity of interactions with a specific team or team member. Specific knowledge, however, raises productivity within a particular relationship. The term *team-specific capital* is commonly used to refer to skills, experiences, trust, and knowledge that are useful for the productivity of a specific relationship (Jaravel, Petkova and Bell 2018; Chillemi and Gui 1997).

In our setting, when a PCP and a specialist share a patient, they can work together more or less effectively, depending on their shared team-specific capital. Team-specific capital may develop in two ways. First, it may be a costless and incidental by-product of repeated interaction; for instance, when they work together, teams can learn by doing (Nagypál 2007). Second, team-specific capital can result from effort and investment, for instance by developing effective patterns of communication or figuring out how to get electronic medical records systems to interoperate efficiently. The returns to investment in team-specific capital increase with the amount of production that takes place between the PCP and specialist. Teams with more repeated interaction will have more team-specific capital both because it develops organically as members interact and because members invest more in the relationship.

The problem of trust is intertwined with the problem of building team-specific capital. Trust within teams is also facilitated by repeated interactions. Consider, for example, Baker, Gibbons, and Murphy's (2002) model of relational contracts within groups. In their analysis, failure to act in a trustworthy manner puts at risk valuable future interactions between team members. The greater the expectation of future interactions, the more powerful the incentives to sustain trust. Increased trust and communication in relationships is connected to the concept of relational coordination (Gittell et al. 2000, 2020). Effective teams rely upon informal understandings, codes of conduct, quid pro quos, and task allocations to guide the behavior of individual members (Gibbons and Henderson 2012). Acquiring shared knowledge about these informal understandings is an essential part of building trusting relational contracts.

Finally, note that in forming teams there may be a tradeoff between team-specific capital and specialization (Marschak and Radner 1972; Epstein et al. 2010). In our context, when PCPs concentrate their patient referrals to a narrower group of physicians within a specialty (e.g. refer their cardiology patients to a smaller set of cardiologists), this encourages repeat interactions and raises team-specific capital. These enhanced collaborative relationships, however, come at a potential cost: worse patient-specialist matches due to a more limited set of referred specialists. The precise source of specialist match

value can vary but it plausibly includes things like the specialist's ability or experience with the patient's specific disease, the patient's travel time, or appointment schedules. In forming a team, the PCP balances this potential loss of match value against the gains that are enabled by greater team-specific capital with a smaller set of specialists.

4. Empirical Implementation of Team Referral Concentration

We measure team referral concentration, which identifies teams with more opportunities for repeated interactions and stronger incentives to build relationship-specific capital and trust. Note that we cannot directly measure team-specific capital, but studying team referral concentration is an attractive alternative. As described in the previous section, we predict that teams with more repeated interactions will develop greater team-specific capital, thereby improving the productivity of the team relationship.

To build intuition for our team referral concentration measure, consider Figure 1. The lefthand panel of the figure depicts four patients, each of whom sees one PCP and two specialists. The differences across patients in their chosen specialists are due to the referral practices of their PCPs. PCP A, who treats patients 1 and 2, refers each of them to a different set of specialists, while PCP B refers patients 3 and 4 to the same set of specialists. These referral patterns give rise to very different levels of repeat interactions for the two PCPs. As depicted on the right side of the figure, PCP A interacts with each of four specialists only once, while PCP B interacts with each of two specialists twice. PCP B's referrals are more concentrated within a smaller set of specialists than PCP A's.

If repeat interactions improve team-specific capital, it follows that the patients of PCP B will have superior coordination of care than the patients of PCP A—even though each individual patient sees the same number of physicians. In our empirical specifications, we compare the healthcare utilization of otherwise similar patients whose PCPs have different levels of team referral concentration.

Our empirical measure, *team referral concentration*, calculates a PCP's Herfindahl–Hirschman Index (HHI) of shared patients within each specialty. Note that we conceptualize team relationships as a

concentration measure rather than simply measuring the number of shared patients, because we want to avoid conflating differences in a PCP's patient panel size or the PCP's tendency to make outside referrals with differences in how they interact as a team.

The HHI functional form has been validated for related measures of patient care continuity across physicians by a large body of prior work (see Pollack et al. 2016 for a review). We extend this framework to capture the team aspects of physician collaboration. This functional form choice was also informed by a stylized model of team formation developed in a related working paper (Agha et al. 2018). The model finds that investments in team-specific capital are inversely proportional to the number of specialists to whom the PCP refers in a specialty. Within this framework, team-specific capital is the mechanism by which more-concentrated teams are more productive. The HHI function we apply here is a generalization of the relationship derived in this model.

To measure team referral concentration, we begin by identifying patient-sharing relationships. A PCP has a *shared patient* with a specialist physician if that patient visits both the PCP and the specialist during the same year. While not all shared patients are the result of deliberate referrals, survey evidence from Barnett et al. (2011) finds that physicians who share more patients are more likely to report a professional relationship, corroborating our assumption that patient-sharing is a useful indicator of relationship strength. Further, some of the benefits of team referral concentration may arise organically through repeated interactions (as described in Section 3), even in the absence of deliberate referrals.

For each PCP *d* and specialist *s* in specialty *j* (e.g. endocrinology), we calculate the number of shared patients m_{ds}. To translate these shared patients into "market shares," we calculate the PCP's total number of patient-specialist links for that specialty, or $M_{dj} = \sum_{s \in j} m_{ds}$. For each PCP-specialist pair *ds*, the specialist's share of total PCP referrals in that specialty is defined as: $share_{ds} = \frac{m_{ds}}{M_{dj}}$. Our PCP-level measure of team referral concentration within each specialty *j* is the HHI: $ReferralCon_{dj} = \sum_{s \in j}(share_{ds})^2$. When PCPs refer equally to each specialist, this measure reduces to 1/N, where N is the

number of referred specialists.

For our primary analysis, we average the specialty-specific team referral concentration measure across various specialties that a PCP could refer to, weighting each specialty equally. The resulting measure of team referral concentration, $ReferralCon_a$, describes the *network* of connections among physicians, where connections are defined by patients shared between PCPs and specialists. In auxiliary analyses, we consider the PCP's specialty-specific team referral concentrations, exploiting variation across specialties.

Like other network measures, the closer the sample of patients in the data is to the underlying population, the more accurate our measure of team referral concentration will be. Because the sample of patients included in the Massachusetts APCD includes a large proportion of the underlying population, our results in that sample are less vulnerable to measurement error. Measurement error is a much bigger concern in our Medicare results where we rely on a 20% sample of Medicare beneficiaries. We discuss the complications this creates in Section 8.

Figure 1 demonstrates that our team-based referral measure is conceptually different from measuring the spread of an individual's patient care across distinct physicians. However, these two features of care delivery are unlikely to be independent of each other. More-concentrated physician teams may be able to avoid unnecessary or low-value referrals, possibly reducing the number of physicians each patient sees. Because concentrating patient visits in a smaller number of physicians can be either a mechanism for enhanced team coordination or a potential confounder, we report results with and without controls for patient-specific physician concentration.

Following prior practice in the empirical literature, we measure the spread of an individual patient's care across physicians, commonly referred to as a measure of patient care continuity (Pollack et al. 2016). We construct the patient care continuity HHI by first defining n_{ip} as the number of visits during the year by patient *i* to each physician p (who may be a PCP or a specialist). The patient care continuity HHI is then $\sum_{p \in physicians} \left(\frac{n_{ip}}{N_i}\right)^2$, where $N_i = \sum_{p \in physicians} n_{ip}$ is the total number of visits by patient *i* to all

physicians p.

5. Data on the Commercially Insured (Massachusetts APCD)

Our primary data come from the 2012 Massachusetts All Payer Claims Database (APCD), version 2.1; in Section 8, we replicate and extend our findings in a national sample of Medicare beneficiaries. As Ericson and Starc (2015) describe in further detail, the Massachusetts APCD provides insurance claims records for commercial insurance, Medicaid, and Medicare Advantage. An advantage of using the APCD to construct our network-based measure of referral concentration is that it allows us to observe shared-patient relationships across many different payers. Trogdon et al. (2019) find that patient-sharing networks derived from a single payer can differ across payers, underscoring the benefit of using more comprehensive all-payer claims for this type of analysis.

We create two extracts from the APCD: a broad sample that allows us to characterize PCP referral patterns and an analysis sample on which we run our regressions relating team referral concentration to total spending. We describe both extracts in detail below.

A. Data on PCP Team Referral Concentration

We use a *broad sample* to construct our PCP-level measure of team referral concentration. In the broad sample, we limit claims to evaluation and management visits for patients aged 21 and older with primary health insurance information available in the APCD. This includes patients enrolled in commercial health insurance, self-insured employer plans, Medicaid Fee-for-Service, Medicaid managed care, and Medicare Advantage whose claims are processed by the 12 largest payers.

Using the National Provider Identifier (NPI) associated with each claim, we link claims to physician specialty and demographic information in the National Plan and Provider Enumeration System data. Guided by the most common sources of outpatient specialty visits identified in the 2011 National Ambulatory Medical Care Survey, we define PCP team referral concentration using nine common

specialties: cardiology, dermatology, endocrinology, obstetrics and gynecology (OB/GYN), orthopedics, ophthalmology, otolaryngology, urology, and general surgery.[4] Details on how we identify primary care physicians and each of these nine specialty types are in Appendix C. Physicians outside of these categories did not enter the calculation of team referral concentration but were included in other visit and cost measures.

We construct team referral concentration based on physician links. Each link represents a PCP-specialist pair who share at least one patient, with the strength of the link determined by the number of patients in common. While many patient-sharing ties are created by direct referrals from PCPs to specialists (cf. Barnett et al. 2011), others may arise when a patient visits both physicians without a direct referral. Because patient-sharing may require coordination and teamwork between physicians even when a direct referral was not made, team relationships may still influence care outcomes in these cases.

We calculate team referral concentration using shared patient links with the method described in Section 4 above. Referral concentration for each PCP is first calculated separately by specialty for each of the nine specialties and then averaged equally across specialties to define a single PCP-level measure of team referral concentration. If we do not observe any referrals by the PCP to a particular specialty, we impute a team referral concentration of 1 (fully concentrated) for that specialty. Subsequent analyses also exploit within-PCP variation in referral concentration by specialty.

This measure of team referral concentration is based on patient-sharing in the sample year 2012, across all patients observed in the APCD. Although it captures Medicare Advantage enrollees, this sample will not capture patient-sharing of patients enrolled in Fee-for-Service Medicare. The 2012 National Ambulatory Medical Care Survey reports that Medicare is the payer for 25% of office visits (NAMCS 2012).

[4] These specialties cover eight of the ten specialties with the highest volume of outpatient visits, as reported in the National Ambulatory Medical Care Survey (2011). From the top ten list, we exclude psychiatry because it relies on specialized billing codes for office visits, and we exclude oncology due to widespread subspecialization in treating different cancer types. We add endocrinology as a common specialty used by patients with chronic disease (particularly diabetes).

Our measurement error simulation results discussed in Section 8A suggest that changing our sample coverage in this range would have little discernible impact on our findings.

B. Data on Patient Outcomes

Our *analysis sample* is limited to chronically ill patients residing in Massachusetts, aged 21–64, who were continuously enrolled with the same commercial insurer or self-insured employer for all of 2012. See Appendix Figure A1 for details of the sample construction. We focus on chronically ill patients because we expect team productivity will matter most for patients who often require specialized care. The restriction to continuously enrolled patients helps remove noise or confounds associated with insurance churn and facilitates calculation of annual spending and utilization.

We use an adapted definition of chronic illness from Frandsen et al. (2015) to identify a patient as having a chronic illness if they have at least one inpatient hospitalization or two outpatient claims indicating one of the following conditions: coronary artery disease, cerebrovascular disease, peripheral arterial disease, mesenteric vascular disease, other ischemic vascular disease or conduction disorders, heart failure, migraine or cluster headache, hypertension, hyperlipidemia, diabetes mellitus, asthma, chronic obstructive pulmonary disease, hypercoagulability disorders, osteoarthritis, and/or rheumatoid arthritis (see Appendix B).[5]

Patients are attributed to a PCP based on the "plurality primary care physician algorithm" adapted from Pham et al. (2009). We assign each enrollee to the doctor trained in Internal Medicine or Family Practice that the enrollee visited the most for evaluation and management; ties are broken by assignment to the PCP with the highest total billed claims. Attributions to physicians with primary care specialization training are more stable within-patient across years, suggesting they better identify the provider taking responsibility for a patient's longitudinal care (Higuera and Carlin 2017). We drop patients from the

[5] Our results are robust to the number of claims with a given diagnosis code required for inclusion of the sample. As seen in Appendix Table A2, our primary results in Table 3 are very similar in magnitude when we require four outpatient claims (rather than two outpatient claims or one inpatient stay as required in the main results).

sample who cannot be assigned to a plurality PCP with this algorithm.

We capture health status by hierarchical condition category (HCC) risk scores and 162 binary condition categories calculated by applying the Massachusetts "Market-wide Risk Adjustment" algorithm to individual insurance claims (Kautter et al. 2014). HCCs are defined using a diagnosis-based algorithm that assigns individuals binary indicators for each condition category if they have claims that indicate a given condition (e.g., diabetes without complications).

We calculate total spending at the enrollee level for all inpatient and outpatient claims in 2012. Spending outcomes are based on the insurer-allowed amount, which consists of the insurer-paid amount and any patient cost-sharing. Higher annual patient spending can correspond to more procedures being performed or to the same number of procedures being performed but by a higher-priced provider or in a higher-price setting (e.g., hospital versus physician's office).

To distinguish the contribution of price and quantity changes in our aggregate spending outcome, we also create a measure of utilization using standardized prices. Standardized prices are defined as the mean price per CPT code, procedure modifier, and quantity of procedure units.[6] These standardized prices are constant for each service across insurers and providers. After applying standardized prices to each claim, we aggregate these amounts to the patient level to create a measure of annual care utilization for 2012.

C. Descriptive Statistics

Team referral concentration has a mean value of 0.14 and varies widely across PCPs, with a median value of 0.126 and a standard deviation of 0.07; its distribution is displayed in Appendix Figure A2. Table 1 provides descriptive statistics for patients and PCPs, split by whether the PCP has above-median or below-median team referral concentration. A complete correlation table is reported in Appendix Table A1. PCPs with above-median concentration have an average referral concentration of 0.19, compared to

[6] Before calculating standardized prices, we winsorized the payment data, rounding all non-zero payments in the bottom 1% up to the 1st percentile prices and all payments in the top 1% down to the 99th percentile price.

0.09 among below-median physicians. The differences between these are equivalent to a PCP increasing the number of specialists in each specialty category they refer to from 5.3 to 11.1, if they referred with equal frequency to each specialist within a specialty.

Before restricting to the chronically ill analysis sample, Table 1, Panel A compares PCP referral patterns and panel size using the broad sample of patients that underlie the calculation of team referral concentration. Relative to below-median concentration PCPs, above-median PCPs, on average, have 39% more patients in common with each consulted specialist, even though their total number of attributed patients is smaller. PCPs with high team referral concentration are more likely to have capitated or HMO contracts; under these contracts, PCPs are likely to internalize the patients' care costs to a greater extent and thus may have stronger incentives to concentrate referrals and develop team-specific capital.

Now limiting to the analysis sample, Table 1, Panel B compares patient characteristics among patients with high and low team referral concentration PCPs. The distribution of disease categories and demographic composition of patients are quite similar across patients seeing PCPs with high versus low team referral concentration. Patients treated by PCPs with high team referral concentration have a slightly higher average patient care continuity HHI. A more-concentrated patient care continuity HHI may in fact be one of the channels through which team referral concentration has an effect: a better coordinated PCP-specialist team could reduce the number of unique physicians a patient sees by preventing redundant or low-value referrals.

Table 2 examines differences in team referral concentration by organizational affiliation and shows there is variation across physician contracting networks. The largest, highest priced hospital system in Massachusetts, Partners Community Health Care (see Seltz et al. 2016; this system is now also known as Mass General Brigham), has an average team referral concentration of 0.13, near the average team referral concentration in our analysis sample (0.14). Team referral concentration also varies substantially across different PCPs *within* a physician contracting network.

Figure 2 displays a binned scatterplot of team referral concentration and the average of log utilization. While higher levels of team referral concentration are associated with lower average spending throughout the distribution, the negative relationship is strongest for lower levels of team referral concentration distribution. This would be consistent with very uncoordinated care being more expensive. Alternatively, PCPs with the sickest (i.e., most costly) patients may refer to many different specialists within a specialty (e.g., two cardiologists with different subspecialty expertise) due to clinical need. While we do not see large differences in patient characteristics for PCPs above and below the median, we nevertheless develop an identification strategy in the next section to address this and other potential confounders.

6. Team Referral Concentration and Spending for the Commercially Insured in Massachusetts

A. Empirical Approach and Identification

We now investigate the relationship between team referral concentration and spending. We pursue three identification strategies, beginning with a simple controlled regression. Baseline regressions take the following form:

$$\log y_i = \alpha ReferralCon_{-i} + \beta X_i + \gamma Z_i + \varepsilon_i \,,$$

where y_i is the annual healthcare spending of patient *i.* X_i is a set of patient characteristics. Z_i is a set of the assigned PCP's characteristics. $ReferralCon_{-i}$ denotes the team referral concentration of patient *i*'s PCP. In the regression analyses, we use a jackknifed calculation of the PCP's referral concentration, $ReferralCon_{-i}$, that omits the contribution of the current patient *i* to the doctor's team referral concentration. The jackknifing procedure overcomes an important endogeneity threat: that a *patient's* own preferences or health status necessitates specific referrals, thus reducing the *physician's* team referral concentration and driving up the patient's own spending.

All regressions include a rich vector of patient and insurer controls including patient sex, a piecewise

linear control for age (i.e., spline with five knots), fixed effects for each of the 162 hierarchical condition categories, and patient 5-digit ZIP code fixed effects. Insurer controls include a fixed effect for each payer and for each of the 13 types of insurance plans defined by the APCD (i.e. Health Maintenance Organization [HMO], Preferred Provider Organization [PPO], Exclusive Provider Organization [EPO], etc.). Given the inclusion of these rich controls, the baseline specification is identified by variation in PCP referral concentration among patients having similar health status, demographics, insurers, and insurance types and residing in the same ZIP code. We then augment this baseline specification with a series of additional controls for patient and physician characteristics.

There are two main threats to identification in these baseline controlled regression specifications. First, PCPs with varying referral concentration may also differ in their practice style along other dimensions. If, for example, physician taste for more-intensive care correlates with low team referral concentration, this could bias our estimates. To account for this possibility, we run additional specification checks that directly control for PCP and specialist fixed effects, exploiting differences in PCP referral concentrations across different specialties. We describe this approach in more detail and report results in Section 6D below.

A second threat to causal interpretation in the baseline specifications is the possibility that patients seeing physicians with low team referral concentration are in worse health. While we include detailed controls for patient demographics and hierarchical condition categories in our baseline analysis, there could nevertheless be unobserved differences in health. To assess this possibility, we will analyze the experience of Medicare beneficiaries who change PCPs due to a move, controlling for patient fixed effects. We describe the mover specifications in more detail and report results in Section 8C.

Together, these strategies aim to identify the impact of PCP referral concentration on the costs of care, accounting for other differences in PCP practice style and the possibility of endogenous sorting of patients to PCPs.

B. Main results

Baseline results are in Table 3. Columns (1)–(3) run regressions in which the dependent variable is care utilization measured at standardized prices, while columns (4)–(6) use total spending as the dependent variable, combining both price and utilization effects. Columns (1) and (4) report results with the baseline controls for patient and insurance characteristics described above. The findings confirm the strong relationship between within-team referral concentration and utilization, as depicted in Figure 2.

The estimated magnitude of team referral concentration's effect on utilization and spending is economically significant. To interpret magnitudes, it is helpful to remember that the above- and below-median average measures of team referral concentration differ by 0.1. Thus, the coefficients in columns (1) and (4) imply that patients seen by PCPs with above-median team referral concentration have 5.6% (= −0.563*0.1) lower medical care utilization and 10.5% (= −1.05*0.1) lower total spending on average, when compared to similar patients seen by PCPs with below-median team referral concentration. Alternatively, a one standard deviation (0.07) increase in team referral concentration leads to 3.9% lower utilization and 7.4% lower spending.

These results indicate that patients of PCPs with higher team referral concentration use fewer services and see lower-priced providers. Our findings control for both insurer (e.g., Anthem, United, etc.) and plan type (e.g. HMO, PPO, etc.), so pricing variation due to differences in insurance plan breadth and quality are unlikely to be the primary driver of this result. Results reported in Appendix Table A3 decompose spending into billings for care delivered by the patient's plurality PCP and billings for all other care. We find that PCPs with higher team referral concentration do not themselves bill for less care; instead, patients of PCPs with higher team referral concentration incur lower spending across all their other providers.

PCPs who put greater importance on containing the total costs of care will be inclined to both concentrate their referrals (and reap the cost savings facilitated by team-specific capital) and refer to

lower-priced providers. In Appendix Table A1, we show that PCPs with more patients enrolled in HMOs, with more visits paid by capitation contracts, and more likely to be participating in a Blue Cross Blue Shield (BCBS) Alternative Quality Contract (a precursor to accountable care organizations) all have higher average team referral concentration. In Section 6D, we introduce physician fixed effects and physician network fixed effects to allay concerns that differences across doctors or firms may contribute to the utilization results reported in Table 3.

As a robustness check, we augment the baseline regression in Table 3 with a control for patient care continuity HHI*;* results are in columns (2) and (5). The relationship between team referral concentration and healthcare use attenuates somewhat in these specifications. Compared to similar patients and holding fixed patient care continuity HHI, patients seen by PCPs with average above-median team referral concentration have 3.8% lower utilization and 8.7% lower total spending than patients seen by PCPs with average below-median team referral concentration.

If patient care continuity HHI captures unobserved heterogeneity in patient demand for care, the specification with these controls would be preferred. Alternatively, patient care continuity HHI may be a consequence of team referral concentration. For example, a PCP's repeat interactions with one specialist may improve clarity and agreement on which patients do not require referrals. In this case, controlling for patient care continuity HHI would result in excessively conservative estimates of the effects of team referral concentration. Given the evidence presented in Section 8C that accounting for unobserved variation in patients' demand for care with patient fixed effects does not attenuate our results, we favor the interpretation that controlling for patient care continuity HHI is likely to underestimate the true effect.

In columns (3) and (6) we add new controls for other dimensions of PCP heterogeneity: five-knot splines of the average HCC risk score of a PCP's patients; the number of patients the PCP treats; an indicator for whether the PCP's training is in Internal Medicine or Family Medicine; and fixed effects for the 5-digit PCP ZIP Code. We find similar results after including these variables, providing reassuring

evidence that variation in referral concentration is not reflecting major differences in the size of the physician's patient panel, training, or case mix. The PCP ZIP code fixed effects are particularly interesting because the resulting estimates account for local geography, allaying the concern that team referral concentration may reflect geographic differences in access to specialists.

Finally, in Appendix Table A4, we analyze each specialty-specific measure of team referral concentration in turn, limiting the sample to patients who saw at least one specialist of the corresponding type. The estimated effect of team referral concentration is uniformly negative and statistically significant for every specialty we study. These patterns suggest that referral concentration plays an important role in team performance across a wide array of clinical contexts. The largest effects of referral concentration occur in orthopedics and surgery. While future work should investigate the specialty-specific channels by which referral concentration impacts team performance, we can speculate that the effect of referral concentration may be larger when there is more scope for treatment choice (e.g., in cases in which there is uncertainty over the clinical value of procedures) and when total spending is higher. It is notable that orthopedic and surgical procedures are relatively high-cost events and have been the subject of Medicare policy reforms, such as bundled payments (Clement et al. 2017). These results suggest that particular attention to surgical referral decisions may be valuable to physicians and administrators.

C. Team Referral Concentration Within Insurance and Contracting Networks

Team-specific capital and care utilization may be jointly mediated by organizational boundaries. As described in Table 2, we find substantial variation in team referral concentration both within and across PCP contracting networks. To assess whether our main results are driven by comparisons across physician contracting networks with different levels of PCP-specialist integration, reputations, or practice styles, we run an analysis including fixed effects for the PCP contracting network. For the PCPs who are not reported to be part of a contracting network, we include a separate fixed effect.

Results of this analysis with PCP contracting network fixed effects are reported in Appendix Table A5.

Even after accounting for variation in utilization related to PCPs' organizational affiliations, we find a negative and statistically significant relationship between team referral concentration and care utilization. The magnitude of the coefficient, −0.40, is slightly smaller than our baseline estimate of −0.56 in Table 3. This result suggests that while physician contracting networks may play a role in shaping team referral concentration, there is still a large scope for differences in team referral concentration within each PCP organization, with more-concentrated physicians incurring lower care utilization.

In a further test reported in Appendix Table A6, we restrict the sample to only PCPs that are part of Partners HealthCare. Partners HealthCare (now called Mass General Brigham) is the largest contracting network in Massachusetts; it includes the academic medical centers of Massachusetts General Hospital and Brigham and Women's Hospital. Each of the 925 PCPs in this network is vertically integrated with many specialists of every variety, and thus has wide scope to select a level of referral concentration within its integrated network. Even within Partners HealthCare, we find that PCPs with greater referral concentration have lower levels of care utilization; the regression coefficient of −0.68 is even larger than the corresponding estimate in Table 3. These findings underscore that even within a single organization, we see diverse team structures. While firm-level investments may ease coordination frictions, our empirical findings suggest they do not fully substitute for repeated team interactions.

Finally, we consider whether the effect of team referral concentration is driven by physicians' participation in different insurance networks. Our baseline analyses in Table 3 control for features of the patient's insurance plan: the insurer and insurance type. Appendix Table A6 replicates Table 3, but is limited to a sample of patients who have the most generous large insurance contract in the data (Blue Cross Blue Shield PPO, see Ericson and Starc 2015 for more information). In this limited sample, we estimate a slightly larger relationship between team referral concentration and utilization than reported in Table 3, column (1). This evidence, in tandem with our analysis of Medicare enrollees reported in Section 8 (all of whom share the same insurer), implies that insurance network restrictions are not driving

our findings.

D. Within-Physician Variation in Team Referral Concentration

A limitation of our approach so far is that we cannot distinguish the effects of team referral concentration from other, unobserved, dimensions of physician practice style. To address this concern, we perform additional analyses that include PCP fixed effects, specialist fixed effects, or both. This analysis will implicitly account for the role of physician contracting networks discussed in the prior section, since physician fixed effects will absorb any effect of physician contracting network.

For these analyses, we restrict the sample to patients who saw at least one specialist from one of our nine specialty categories. (Unlike the base analysis sample, this subsample excludes patients who saw no specialists.) Instead of using a PCP's average team referral concentration across all nine specialties, patients are assigned the (jackknifed) team referral concentration of their PCP averaged only across the specific specialties that the patient consulted. Throughout all the analyses reported in this section, we include an additional set of controls: a set of indicator variables for the specific combination of specialties consulted by the patient (e.g., cardiology only; cardiology and endocrinology; etc.). These controls capture baseline differences in patient utilization that may depend on the specialist types required.

Table 4, column (1) estimates the relationship between team referral concentration and utilization in this subsample, echoing the control variables used in Table 3, column (1). We find a 0.1 increase in team referral concentration is associated with 3.6% lower spending. This effect is statistically and economically significant, but smaller than the estimate in Table 3. The key difference is that this sample conditions on the patient consulting at least one specialist. Because the decision to consult a specialist may be an outcome of team referral concentration, with more concentrated PCPs avoiding low-value specialists, we interpret results from this subsample as conservative.

Next, we consider a specification that introduces PCP fixed effects. This analysis exploits differences in team referral concentration within PCP across each of the nine specialties. For example, if a PCP is highly

concentrated in her cardiology referrals but not in her endocrinology referrals, we predict that her cardiology patients will have lower relative utilization. This approach to identifying the effects of team-specific capital is conservative, for reasons beyond the subsample conditioning described in the prior paragraph. These estimates may be attenuated to the extent that within-PCP variance is driven by measurement error. Further, the inclusion of PCP fixed effects will lead us to underestimate the role of team referral concentration if there are spillovers in team performance across specialties.

Results with PCP fixed effects are reported in Table 4, column (2). The findings remain negative and statistically significant, but the magnitude is about one-half the size of the effect reported in column (1). The persistence of a negative, statistically and economically significant effect after including PCP fixed effects suggests that unobserved differences in PCP quality or practice style (uniform across conditions the PCP treats) do not explain the estimated effect of team referral concentration. This specification also removes concern about possible confounding variation in PCP practice settings such as practice size, ownership structure, IT implementation, or integration with health plans.

In Table 4, column (3), we consider the role of specialist practice style. Another potential concern with the results reported so far is that physicians with differing team referral concentration refer to specialists of differing quality. Perhaps PCPs with less-concentrated referrals have higher costs because they are referring to specialists with more-intensive practice styles. To address this, we estimate specifications with fixed effects for the identity of the patient's plurality specialist in each specialty consulted. These regressions effectively compare patients who share the same specialist, but who are referred by different PCPs with different levels of team referral concentration for that specialty.

We continue to find that patients of PCPs with higher team referral concentration have significantly lower levels of utilization, even when they are referred to identical specialists. The magnitude suggests that moving from a below- to an above-median team referral concentration PCP is associated with a 2.4% reduction in utilization, significant at the 1% level. This effect is similar in magnitude (slightly larger) than

the estimates that included PCP fixed effects and is about two-thirds the size of the effect estimated in this sample without any physician fixed effects (column (1)). These findings provide further support for the notion that team relationships between PCPs and specialists promote lower-cost care.

Finally, we report results that simultaneously control for both PCP fixed effects and specialist fixed effects. These two-way fixed effect specifications combine the two sources of variation highlighted above: (1) within-PCP variation in specialty-specific team referral concentration, and (2) within-specialist variation in team referral concentration from different referring PCPs. Results of the two-way fixed effect model with both PCP and specialist fixed effects are reported in Table 4, column (4). We continue to predict that higher team referral concentration leads to lower spending, and the result remains statistically significant at the 1% level. The magnitude of this relationship is smaller: a 0.1 increase in team referral concentration leads to a 1% reduction in care utilization. This estimate is conservative for the three reasons explained above. First, this subsample conditions on seeing a specialist, which could be a direct result of team referral concentration. Second, the effects of team referral concentration may spill over across specialties a PCP refers to. Third, a larger share of the residual variation in referral concentration conditional on PCP and specialist fixed effects may reflect measurement error in team relationships. Nevertheless, it is reassuring that team referral concentration still predicts lower spending after the addition of these rich controls.

7. Utilization-Based Quality Metrics

Our results have focused on the effect of team referral concentration on the cost of health care. Our discussion of the tradeoff between repeated interactions and specialization in Section 3 suggests that increased team-specific capital will lower costs, but that the effect on care quality is ambiguous. While quality could increase due to improved coordination, quality might also decrease if the smaller set of familiar specialists leads to a lower quality match between the specialist and the patient's idiosyncratic needs. Although quality of care is multi-dimensional and difficult to measure, Table 5 examines the

relationship between team referral concentration and a set of utilization-based quality indicators for the Massachusetts APCD Sample.

We find suggestive evidence that quality of care increases with team referral concentration. Higher team referral concentration has a negative effect on emergency department visits, with a 0.1 unit increase in team referral concentration leading to a 4% (or 0.9 percentage point) drop in the probability of an emergency department visit, from a mean of 22% (see Table 5, column (1)). Higher team referral concentration is also associated with fewer inpatient visits and ambulatory care–sensitive hospitalizations, although the results are not statistically significant in the most controlled specification. We find no significant change in duplicate imaging.

Turning to patient-specialist match, we do not have measures of clinical match quality, but we can examine the distance patients travel to see specialists. We measure the straight-line distance between the patient ZIP code and the ZIP code of the patient's plurality specialist. In our preferred specification, patients of PCPs with 0.1 higher team referral concentration travel 1.6 miles fewer to see their plurality specialist; these results condition on fixed effects for the patient's ZIP code of residence. Greater team referral concentration does not come at the expense of patient travel time.

8. Team Referral Concentration in Medicare

Our results so far have focused on the Massachusetts All Payer Claims Data. These data offer remarkable breadth for measuring referral networks precisely at the physician level, but they are limited in three important ways. First, the APCD is limited to a single state, whose healthcare institutions may not be nationally representative. Second, our extract of the APCD data is essentially a cross-sectional data set; we do not have an extended panel of claims. As a result, we cannot use the Massachusetts data to estimate a model with patient fixed effects. Finally, the APCD covers patients enrolled in many different insurance networks; while our analysis controls for insurer and plan type, there is still a possibility of confounding from local variation in network coverage within insurers. We address these limitations by

analyzing a national sample of Medicare beneficiaries. Our Medicare sample covers seven years, and all patients have the same, broad insurance network of participating providers.

A. Measurement Error in Medicare

The Medicare sample covers 20% of Medicare Fee-for-Service enrollees, allowing us to observe only a small fraction of each doctor's total patient panel. This limitation creates non-classical measurement error for network measures, including our team referral concentration variable. In this section, we use simulations to explore the likely influence of measurement error on our results.

From the full Massachusetts APCD, we draw repeated subsamples that range from # = {10%, 20%, …, 100%} of the full sample of patients. Within each subsample, we construct team referral concentrations and then re-estimate our baseline regression specification in the subsample. We then calculate the multiplier $\lambda_{\#} = \alpha_{100}/\alpha_{\#}$ that tells us how to scale estimated coefficients on team referral concentration for each #% subsample to match the coefficient in the full sample.

In Appendix Figure A3, we plot how the estimated multiplier $\lambda_{\#}$ varies with the sample size. As expected, bias from measurement error falls as the sample size increases. The multipliers for specification 1 are moderate: if we observed 20% of the APCD sample, we would want to multiply our estimate by about 3 to scale up to the estimated effect of team referral concentration in the full sample. Adding further controls to the regression greatly exacerbates the measurement error problem. With 20% of the APCD sample, specification 2 has a multiplier of about 13, with a 95% CI from 7 to 26; we find similar results for specification 3. On the basis of these simulations, we conclude that estimating specification 1 in Medicare claims provides insight on the generalizability of our Massachusetts results but that specifications 2 and 3 are less informative.

This simulation exercise provides a range of multipliers suggesting the direction and approximate magnitude of measurement error bias. As it is not obvious which scaling factor calculated in the Massachusetts APCD corresponds to the 20% Medicare sample, we report unadjusted regression

coefficients, and then discuss how imposing a scaling factor would influence the interpretation.

B. Replicating Cross-Sectional Specification in Medicare

The Medicare analysis relies on a 20% sample of Medicare Fee-for-Service enrollees, with data from the Medicare Carrier, Inpatient, and Outpatient claims files, and demographic information from the Master Beneficiary Summary File. The broad sample used to calculate team referral concentration uses Carrier claims for all patients in the raw file, limited to evaluation and management visits. For regression analysis, we construct an analytic sample of continuously enrolled patients aged 66–99. We do not restrict to patients with a chronic condition, since previous research has found that over 80% of elderly Medicare enrollees have a chronic illness (Finkelstein et al. 2016). Following the same attribution rule as in the APCD, we assign patients to their plurality PCP and drop unassigned patients. As in the APCD analysis, we exclude PCPs who do not refer to at least five of our nine included specialties.

We begin by replicating our cross-sectional analysis in the Medicare sample, using data from 2012. Cross-sectional regression results are reported in Table 6, Panel A; Appendix Table A7 reports summary statistics. Because Medicare prices are administratively set and primarily adjusted only for geographic location, we report only utilization outcomes and not spending. All models include ZIP code fixed effects that will capture variation driven by geographic price adjustment. We find that a 0.1 increase in measured team referral concentration is associated with a 1.4% decline in utilization. The analogous coefficient estimated in Massachusetts data (Table 3, column (1)) was about four times as large, similar to the magnitude of attenuation predicted in our measurement error simulations. We find statistically significant negative coefficients on team referral concentration in specifications 2 and 3, as we add controls for patient care continuity and PCP characteristics. The magnitudes are smaller in these controlled specifications, which is expected given that measurement error is likely to cause severe attenuation.

C. Medicare Movers Analysis

We have documented a negative relationship between team referral concentration and costs in two

very different patient populations: chronically ill, commercially insured working age patients in the Massachusetts APCD, and elderly patients in Medicare. In both cases, our results rely on cross-sectional regressions, and given the limitations of claims data, there is still room for selection on unobservable patient characteristics to bias these findings. To account more richly for patient heterogeneity, in this section we identify the effect of team referral concentration from the experience of Medicare patients who change their PCP as a result of a move. This approach allows for the inclusion of a patient fixed effect to control for differences in patient demand for care that are stable over time, building on work by Finkelstein, Gentzkow, and Williams (2016), and Agha, Frandsen, and Rebitzer (2019).

For the movers analysis, we expand the Medicare sample to include data from 2006–2012. We calculate PCP team referral concentration on an annual basis using the full 20% Medicare sample (not restricted to physicians who treat movers). The analysis sample restricts to enrollees who move to a new hospital referral region over this period.[7]

We use a difference-in-differences strategy to examine how utilization changes when a moving patient switches to a PCP with a different level of team referral concentration. We estimate the following difference-in-differences equation:

$$\log y_{it} = \alpha \Delta ReferralCon_{-i} Post_{it} + \beta_i + \gamma Z_{it} + \rho_{R_{it}} + \varepsilon_{it}.$$

The variable $\Delta ReferralCon_{-i}$ measures the change in the jackknifed referral concentration of the patient's post-move PCP compared to the patient's pre-move PCP.[8] We interact $\Delta ReferralCon_{-i}$ with $Post_{it}$, an indicator variable for the post-move period. The regression also includes fixed effects $\rho_{R_{it}}$ for the event year relative to the move (denoted $R_{i,t}$), allowing movers' annual demand for care to depend

[7] The sample restricts to patients with exactly one move over this period and requires that at least 75% of a patient's claims are in the hospital referral region that corresponds to their listed address ZIP code in each year (excluding the year of the move).
[8] Patients are assigned to their plurality PCPs on an annual basis, allowing for patients to switch PCPs across years, even in the absence of a move. Pre-move team referral concentration is calculated as the average level of PCP team referral concentration over the year(s) prior to the move. Similarly, post-move team referral concentration is calculated as the average level of PCP team referral concentration over the year(s) after the move. (Note the year of the move is excluded from both calculations.) The change in PCP team referral concentration is the difference of post- and pre-move average team referral concentrations.

on the timing of their move. (For example, year $R_{i,t}$= −1 corresponds to the year before the move, year 0 for the year of the move, etc.) Characteristics of the patient's plurality PCP and calendar year fixed effects are included in the control vector Z_{it}. Patient fixed effects, β_i, account for fixed differences in patient health over time; we do not include time-varying controls for patient comorbidities because evidence suggests that there are regional differences in comorbidity coding (Song et al. 2010, Finkelstein et al. 2017), which could be endogenously related to changes in team referral concentration. The role of measurement error in this specification is described in Appendix D.

Table 6, Panel B reports results from the difference-in-differences mover analysis. The baseline specification estimates an increase in team referral concentration of 0.1 is associated with a 2.2% decrease in care utilization, significant at the 1% level. The estimated effect sizes are larger in the difference-in-differences results than the cross-sectional comparison. This pattern suggests that bias from endogenous sorting of patients to doctors is not driving our main result and may even lead us to understate the impact of team referral concentration. Further results reported in Appendix Table A8 find that movers who switch to PCPs with higher team referral concentration also experience a decline in emergency department visits, inpatient visits, and ambulatory care–sensitive hospitalizations.

Figure 3 allows us to graphically assess the possibility of both pre- and post-trends in the difference-in-difference analysis. Instead of considering a single post dummy, this specification interacts the change in referral concentration with a full vector of event-time indicators for years before and after the move. The figure displays stable pre-trends in the years leading up to the move; patients who eventually switch to more-concentrated PCPs are not a trend of declining spending relative to other patients prior to their move. The year of the move, denoted as "Year 0" in the graph, is partially treated, with the patient moving at some point during this year. The change in spending is persistent for several years following the move.

We complement this difference-in-differences analysis with an instrumental variables (IV) approach. The IV analysis eliminates one possible source of endogeneity bias: specifically, the choice of destination

PCP might be endogenously related to *changes* in the patient's health status that coincide with their move. Similar to the approach taken by Abaluck et al. (2020), Agha, Ericson, and Zhao (2020), and Laird and Nielson (2016), the IV exploits a mean reversion property: patients who are initially treated by PCPs with high team referral concentration tend to switch to PCPs with more typical (lower) team referral concentrations. Thus, the patient's PCP team referral concentration in the pre-move period is used as an IV for the *change* in the patient's PCP team referral concentration, $\Delta ReferralCon_{\_i}$. The exclusion restriction requires parallel trends among patients with different initial exposure to team referral concentration. Note that the IV approach may still suffer bias from measurement error. More details on the IV approach can be found in Appendix E.

The first-stage estimates reported in Appendix Table A9 confirm the predicted mean reversion pattern: patients with higher initial PCP team referral concentration tend to experience a decline in team referral concentration after they move. Table 6, Panel C reports the second stage of the IV estimation; the results remain statistically significant, with magnitudes that are in between the cross-sectional estimates and the difference-in-differences estimates. The IV estimates are slightly smaller than the difference-in-differences estimates, which is consistent with a small amount of bias in the difference-in-differences approach (resulting from changes in the patient's health status affecting the choice of destination PCP), although the differences between the two estimates are small and not statistically distinguishable.

9. Conclusion

Teams are pervasive in economic organizations, but researchers know very little about how the structure of teams influences economic performance. Team production is especially challenging in healthcare because PCPs and specialists must work together closely to address the often complex and changing needs of patients. We study team production in healthcare by examining the patient-sharing patterns between PCPs and specialists. Concentrating referrals among a smaller team of specialists facilitates the development of team-specific capital. Empirically, we find that patients of PCPs who

concentrate their referrals among a smaller set of specialists have lower health care utilization. This reduction in utilization is observed in both commercially insured and Medicare populations, is statistically and economically significant, and holds under various identification strategies that account for unobserved patient and physician characteristics.

More specifically, for chronically ill adults in Massachusetts, those treated by PCPs with below-median team referral concentration have 6% higher utilization compared to those treated by above-median PCPs, after controlling for detailed patient and insurance characteristics. We find evidence suggesting that this reduction in utilization is driven in part by less frequent visits to the emergency department and fewer hospitalizations. Smaller effects are found in specifications using PCP and specialist fixed effects; these specifications are conservative since they are likely to absorb some physician-level investments in team-specific capital into the fixed effects and exacerbate any attenuation from measurement error.

A negative relationship between team concentration and utilization persists in a national 20% sample of Medicare beneficiaries, despite the noise introduced by calculating the network-based team concentration measure with only a fraction of each PCP's patient panel. In the Medicare sample, we also study patients who switch PCPs as the result of a move. Using a difference-in-difference analysis with patient fixed effects, we find that an increase in team referral concentration is associated with lower utilization.

Our analysis has several limitations that may inspire future research. While theory and our empirics suggest that concentrated referrals facilitate both organic learning and relationship-specific investments that improve team performance, we do not directly observe the learning or investments directly. More work on the channels through which concentrated team referrals lead to reduced utilization would be valuable, as we cannot distinguish whether this improved performance results from familiarity and trust or greater investments in coordinating mechanisms—such as shared protocols and information systems—that support these relationships (Gittell 2002; Faraj and Xiao 2006). Indeed, both channels could

contribute. Another issue is whether the gains from concentrated team referrals are different within or between organizations. If the gains from concentrated team referrals are greatest within organizations, these must be weighed against the higher prices that may be charged by these integrated practices (Baker et al. 2014). Finally, our measures of quality of care are limited due to the challenges of measuring care quality in insurance claims data.

Policy or management interventions that shape referral networks are a promising avenue for future work, and our measure of team referral concentration can be used to evaluate these interventions. Our results suggest that encouraging concentrated referral networks may be cost-reducing. PCPs participating in the BCBS Alternative Quality Contract, a precursor to the Accountable Care Organization payment model, had higher team referral concentration on average, as did PCPs who had a higher share of HMO patients and a higher share of patients with capitation payment contracts. These relationships suggest that alternative payment models may encourage greater team referral concentration and stronger PCP-specialist relationships.

Finally, our approach to analyzing the effect of team structure on team performance may also be useful in non-healthcare settings. Teams can be organized in ways that encourage or reduce repeat interactions between team members with specialized knowledge. Design choices that facilitate repeat interactions enhance team-specific capital at the cost of reducing gains from specialization and matching. Understanding the determinants of this tradeoff and its consequences for economic performance in other environments is an important area of study.

10. References

Abaluck, J, Bravo, MMC, Hull, P, & Starc, A (2020). Mortality effects and choice across private health insurance plans. Working paper 27578, National Bureau of Economic Research.

Agency for Healthcare Research and Quality. 2016. “Prevention Quality Indicators Technical Specifications - Version 6.0.” Online: https://www.qualityindicators.ahrq.gov/modules/pqi_techspec.aspx.

Agha L, Ericson KM, Geissler KH, Rebitzer JB (2018) Team formation and performance: Evidence from healthcare referral networks. Working paper 24338, National Bureau of Economic Research.

Agha L, Frandsen B, Rebitzer JB (2019) Fragmented division of labor and healthcare costs: Evidence from moves across regions. *Journal of Public Economics* 169:144–159.

Agha, L, Ericson KM, & Zhao, X (2020). The impact of organizational boundaries on healthcare coordination and utilization. Working paper 28179, National Bureau of Economic Research.

Alchian AA, Demsetz H 1972. Production, information costs, and economic organization. *American Economic Review* 62(5): 772–795.

Andreatta P, Marzano, D (2012) Healthcare management strategies: interdisciplinary team factors. *Current Opinion in Obstetrics and Gynecology.* 24(6): 445–452.

Andreatta PB (2010) A typology for health care teams. *Health Care Management Review*. 35(4):345–354.

Arulrajah N (2017) Build Your Dermatology Practice with Referral Marketing. *Practical Dermatology*. (December), https://practicaldermatology.com/articles/2017-dec/build-your-dermatology-practice-with-referral-marketing.

Atalay E, Hortaçsu A, Li MJ, Syverson, C (2017) How wide is the firm border? Working paper 23777, National Bureau of Economic Research, Cambridge.

Baker G, Gibbons R, Murphy KJ (2002) Relational contracts and the theory of the firm *The Quarterly Journal of Economics*. 117(1):39–84.

Baker LC, Bundorf MK, Kessler DP (2014) Vertical integration: Hospital ownership of physician practices is associated with higher prices and spending. *Health Affairs.* 33(5): 756–763.

Baker LC, Bundorf MK, Kessler DP (2016) The effect of hospital/physician integration on hospital choice. *Journal of Labor Economics.* 50: 1–8.

Barnett ML, Landon BE, O'Malley AJ, Keating NL, Christakis NA (2011) Mapping physician networks with self-reported and administrative data. *Health Services Research*. 46(5): 1592–1609.

Barnett ML, Keating NL, Christakis NA, O'Malley AJ, Landon BE (2012b) Reasons for choice of referral physician among primary care and specialist physicians. *Journal of General Internal Medicine.* 27(5):506–512.

Becker GS, Murphy KM (1992) The division of labor, coordination costs, and knowledge. *Quarterly Journal of Economics*. 107(4):1137–1160.

Bechky BA (2006) Gaffers, gofers, and grips: Role-based coordination in temporary organizations. *Organization Science.* 17(1):3–21.

Berwick DM, Hackbarth AD (2012) Eliminating waste in US health care. *JAMA.* 307(14):151-4-1516.

Bloom N, Van Reenen J (2011) Human resource management and productivity. *Handbook of Labor Economics*, edited by David Card and Orley Ashenfelter, 4:1697–1767. Elsevier.

Bodenheimer T, Lo B, Casalino L (1999) Primary care physicians should be coordinators, not gatekeepers. *Journal of the American Medical Association.* 281(21):2045–2049.

Brot-Goldberg Z, de Vaan M (2018) Intermediation and vertical integration in the market for surgeons. Working paper.

Clement RC, Bhat SB, Clement ME, Krieg JC (2017). Medicare reimbursement and orthopedic surgery: past, present, and future. *Curr Rev Musculoskelet Med* 10, 224–232. https://doi.org/10.1007/s12178-017-9406-7.

Cebul RD, Rebitzer JB, Taylor LJ, Votruba ME (2008) Organizational fragmentation and care quality in the U.S. healthcare system. *Journal of Economic Perspectives*. 22(4):93–114.

Chillemi O, Gui B (1997) Team human capital and worker mobility. *Journal of Labor Economics.* 15(4):567–585.

Coase. R (1937) The nature of the firm, *Economica* 4(16): 386–405.

Crawford, VP (1990) Relationship-specific investment. *Quarterly Journal of Economics.* 105(2): 561–574.

Cutler D, Skinner J, Stern AD, Wennberg D (2019) Physician beliefs and patient preferences: A new look

at regional variation in health care spending. *American Economic Journal: Economic Policy*. 11(1):192–221.

Dessein W, Santos T (2006) Adaptive organizations. *Journal of Political Economy*. 114(5):956–995.

Dyrda L (2012) 25 key concepts for young orthopedic & spine surgeons to build a successful practice. *Becker's Spine Review*. (May 4), https://www.beckersspine.com/spine/item/11832-25-key-concepts-for-young-orthopedic-spine-surgeons-to-build-a-successful-practice.html.

Edmondson AC., & Nembhard IM (2009). Product development and learning in project teams: The challenges are the benefits. *Journal of Product Innovation Management*, 26(2), 123-138.

Epstein AJ, Ketcham J, Nicholson S (2010) Specialization and matching in professional services firms. *RAND Journal of Economics*. 41(4):811–834.

Ericson K, Starc A (2015) Measuring consumer valuation of limited provider networks. *American Economic Review*. 105(5): 115–119.

Faraj S, Sproull L (2000) Coordinating expertise in software development teams. *Management Science.* 46(12):1554–1568.

Faraj S, Xiao Y (2006) Coordination in fast-response organizations. *Management Science* 52(8):1155–69.

Feri F, Irlenbusch B, Sutter M (2010) Efficiency gains from team-based coordination—large-scale experimental evidence. *American Economic Review*. 100(4):1892–1912.

Finkelstein A, Gentzkow M, Williams H (2016) Sources of geographic variation in health care: Evidence from patient migration. *Quarterly Journal of Economics.* 131(4):1681–1726.

Finkelstein A, Gentzkow M, Hull P, Williams H (2017) Adjusting risk adjustment – accounting for variation in diagnostic intensity. *New England Journal of Medicine.* 376(6):608–610.

Forrest CB, Nutting PA, Starfield B, Von Schrader S (2002) Family physicians' referral decisions. *J Fam Pract.* 51(3):215–222.

Frandsen BR, Joynt KE, Rebitzer JB, Jha AK (2015) Care fragmentation, quality, and costs among

chronically ill patients. *American Journal of Managed Care*. 21(5):355–326.

Frandsen B, Rebitzer JB (2014) Structuring incentives within accountable care organizations. *Journal of Law, Economics, and Organization.* 31(S1): i77–i103.

Garicano L, Santos T (2004) Referrals. *American Economic Review*. 94(3):499–525.

Geissler K, Lubin B, Ericson KM (2020) The role of organizational affiliations in physician patient-sharing relationships. *Medical Care Research and Review*. 77(2):165–175.

Gibbons R, Henderson R (2012) Relational contracts and organizational capabilities. *Organization Science.* 23(5): 1350–1364.

Gittell, JH (2002) Coordinating mechanisms in care provider groups: Relational coordination as a mediator and input uncertainty as a moderator of performance effects. *Management Science* 48 (11):1408–1426. https://doi.org/10.1287/mnsc.48.11.1408.268.

Gittell JH, Fairfield KM, Bierbaum B, Head W, Jackson R, Kelly M, Laskin R, Lipson S, Siliski J, Thornhill T, Zuckerman J (2000) Impact of relational coordination on quality of care, postoperative pain and functioning, and length of stay: A nine-hospital study of surgical patients. *Medical Care.* 38(8):807–819.

Gittell JH, Logan C, Cronenwett J, Foster TC, Freeman R, Godfrey M, Vidal DC (2020) Impact of relational coordination on staff and patient outcomes in outpatient surgical clinics. *Health Care Management Review.* 45(1):112–120.

Gittell JH, Weinberg D, Pfefferle S, Bishop C (2008) Impact of relational coordination on job satisfaction and quality outcomes: a study of nursing homes. *Human Resource Management Journal.* 18(2):154–170.

Grosse S, Putterman L, Rockenbach B (2011) Monitoring in teams: Using laboratory experiments to study a theory of the firm. *Journal of the European Economic Association.* 9(4):785–816.

Hamilton BH, Nickerson JA, Owan H (2003) Team incentives and worker heterogeneity: An empirical

analysis of the impact of teams on productivity and participation. *Journal of Political Economy.* 111(3):465–497.

Harrington JT, Barash HL, Day S, Lease J (2005) Redesigning the care of fragility fracture patients to improve osteoporosis management: A health care improvement project. *Arthritis & Rheumatism*. 53:198–204.

Higuera, L., & Carlin, C. (2017). A comparison of retrospective attribution rules. *Am J Manag Care*, *23*(6), e180–e185.

Huckman RS, Pisano GP (2006) The firm specificity of individual performance: Evidence from cardiac surgery. *Management Science.* 52(4):473-488.

Huckman RS, Staats BR, Upton DM (2009) Team familiarity, role experience, and performance: Evidence from Indian software services. *Management Science.* 55(1):85–100.

Huckman, R. S., & Staats, B. R. (2011). Fluid tasks and fluid teams: The impact of diversity in experience and team familiarity on team performance. *Manufacturing & Service Operations Management,* 13(3), 310-328.

Hussey PS, Schneider EC, Rudin RS, Fox S, Lai J, Pollack CE (2014) Continuity and the costs of care for chronic disease. *JAMA Internal Medicine*. 174(5):742–748.

Jaravel X, Petkova N, Bell A (2018) Team-specific capital and innovation. *The American Economic Review.* 108(4-5):1034–1073.

Jensen, MC, Meckling, WH., 1976. Theory of the firm: Managerial behavior, agency costs and ownership structure. Journal of Financial Economics, 3(4):305–360.

Kautter J, Pope GC, Ingber M, Freeman S, Patterson L, Cohen M, Kennan P (2014) The HHS-HCC risk adjustment model for individual and small group markets under the Affordable Care Act. *Medicare & Medicaid Research Review*. 4(3):E1–E46.

Kim S, Song H, Valentine MA (2019) Staffing temporary teams with an optimal lookback window for

team familiarity and partner variety: Implications for emergency departments. SSRN Working paper https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3176306.

Kolber MJ (2006) Stark Regulation: A historical and current review of the self-referral laws. *HEC Forum: An Interdisciplinary Journal on Hospitals' Ethical and Legal Issues.* 18(1):61–84.

Kozlowski SWJ, Ilgen DR (2006) Enhancing the effectiveness of work groups and teams*. Psychological Science in the Public Interest*. 7(3):77–124.

Laird J, Nielsen T (2016) The effects of physician prescribing behaviors on prescription drug use and labor supply: Evidence from movers in Denmark. Working paper, Harvard University, Cambridge.

Lammers EJ, Adler-Milstein J, Kocher K (2014) Does health information exchange reduce redundant imaging? Evidence from emergency departments. *Medical Care*. 52(3):227–234.

Lazear EP (2009) Firm-specific human capital: A skill-weights approach. *Journal of Political Economy*. 117(5):914–940.

Marschak J, Radner R (1972) *Economic Theory of Teams*. Yale University Press. New Haven, CT.

Massachusetts Health Quality Partners (2016) MHQP 2016 Massachusetts Provider Database (MPD). Online: http://www.mhqp.org/products_and_tools/?content_item_id=226.

McWilliams JM (2016) Cost containment and the tale of care coordination. *New England Journal of Medicine.* 375(23):2218–2220.

Mehrotra A, Forrest CB, Lin CY (2011) Dropping the baton: Specialty referrals in the United States. *The Milbank Quarterly*. 89(1):39–68.

Milstein A, Gibertson E (2009) American medical home runs. *Health Affairs*. 28(5):1317–1326.

Nagypál É (2007) Learning by doing vs. learning about match quality: Can we tell them apart? *The Review of Economic Studies*. *74*(2):537–566.

National Ambulatory Medical Care Survey. 2014. State and National Summary Tables. Available from: http://www.cdc.gov/nchs/ahcd/ahcd_products.htm.

National Ambulatory Medical Care Survey. 2011. State and National Summary Tables. Available from: http://www.cdc.gov/nchs/ahcd/ahcd_products.htm.

Pham HH, O'Malley AS, Bach PB, Saiontz-Martinez C, Schrag D (2009) Primary care physicians' links to other physicians through Medicare patients: The scope of care coordination. *Annals of Internal Medicine*. 17;150(4):236–42.

Pollack CE, Hussey PS, Rudin RS, Fox SD, Lai J, Schneider EC (2016) Measuring care continuity: A comparison of claims-based methods. *Medical Care*. 54(4):e30–e34.

Press MJ (2014) Instant replay—A quarterback's view of care coordination. *New England Journal of Medicine.* 371: 489–491.

Reagans R, Argote L, Brooks D (2005) Individual experience and experience working together: Predicting learning rates from knowing who knows what and knowing how to work together. *Management Science.* 51(6):851–1013.

Rebitzer JB, Votruba ME (2011) Organizational economics and physician practices. Working paper 17535, National Bureau of Economic Research.

Romano MJ, Segal JB, Pollack CE (2015) The association between continuity of care and the overuse of medical procedures. *JAMA Internal Medicine*. 175(7):1148–1154.

Seltz D, Auerbach D, Mills K, Wrobel M, Pervin A (2016) Addressing price variation in Massachusetts. *Health Affairs Blog*, http://healthaffairs.org/blog/2016/05/12/addressing-price-variation-in-massachusetts/

Simon M, Choudhry NK, Frankfort J, Margolius D, Murphy J, Paita L, Wang T, Milstein A (2017) Exploring attributes of high-value primary care. *Annals of Family Medicine*. 15(6):529–534.

Song Y, Skinner J, Bynum J, Sutherland J, Wennberg JE, Fisher ES (2010) Regional Variations in Diagnostic Practices. *New England Journal of Medicine.* 363:45-53.

Stille CJ, Jerant A, Bell D, Meltzer D, Elmore JG (2005) Coordinating care across diseases, settings, and

clinicians: A key role for the generalist in practice. *Annals of Internal Medicine*. 142:700–708

Sundstrom E, De Meuse KP, Futrell D (1990) Work teams: Applications and effectiveness. *American Psychologist*. 45(2):120–133.

Trogdon JG, Weir WH, Shai S, Mucha PJ, Kuo TM, Meyer AM, Stitzenberg KB (2019) Comparing shared patient networks across payers*. Journal of General Internal Medicine.* 34(10):2014–2020.

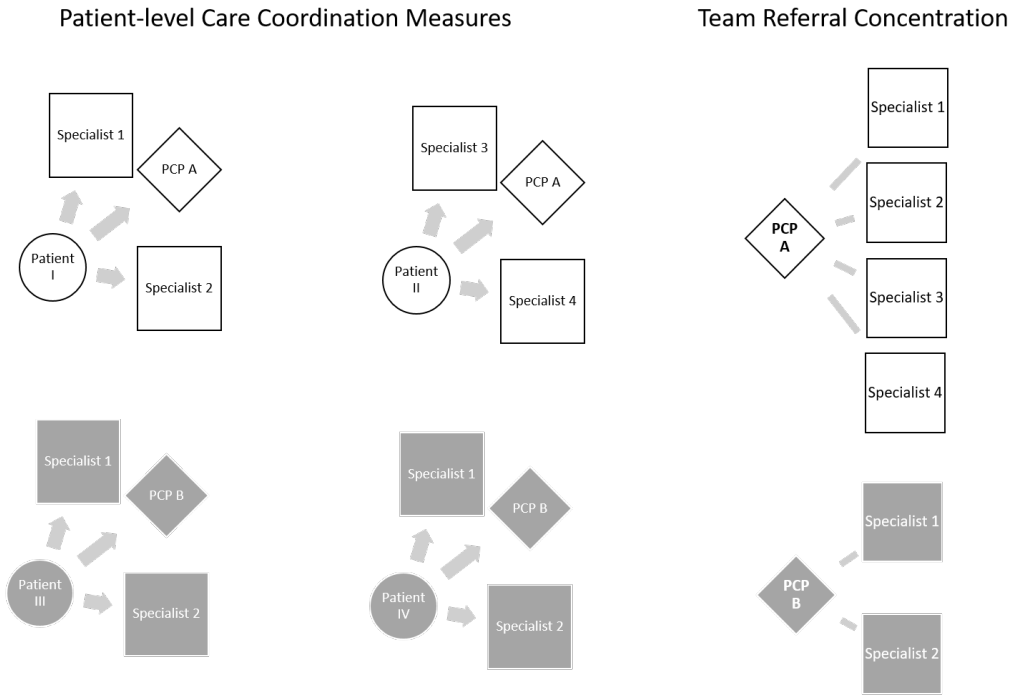
Valentine MA, Tan TF, Staats BR, Edmondson AC (2018) Fluid Teams and Knowledge Retrieval: Scaling Service Operations. *Manufacturing & Service Operations Management* 21(2):346–60.

Weber RA (2006) Managing growth to achieve efficient coordination in large groups. *American Economic Review*. 96(1): 114–126.

Zuchowski JL, Chrystal JG, Hamilton AB, Patton EW, Zephyrin LC, Yano EM, Cordasco KM (2017) Coordinating care across health care systems for veterans with gynecologic malignancies: A qualitative analysis. *Medical Care.* 55(Suppl 7 Suppl 1):S53–S60.

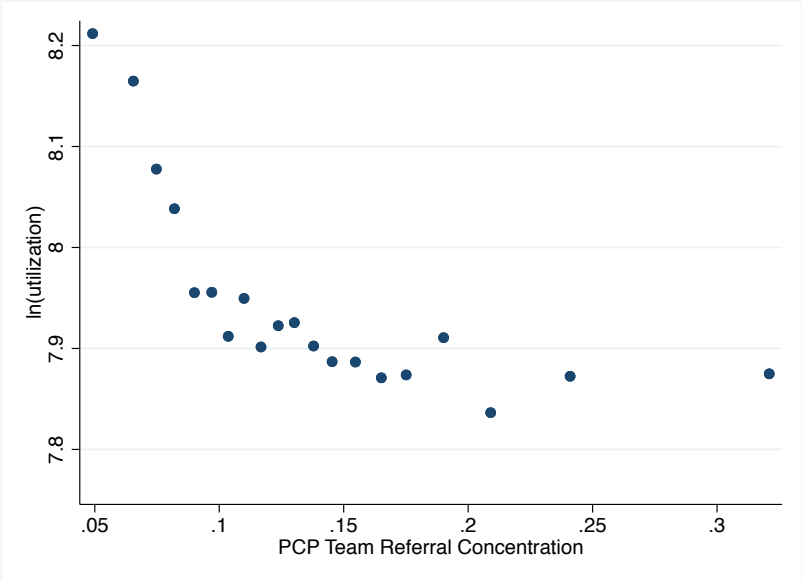
Figures and Tables

Figure 1. Patient-level Measures of Care Coordination versus Team Referral Concentration



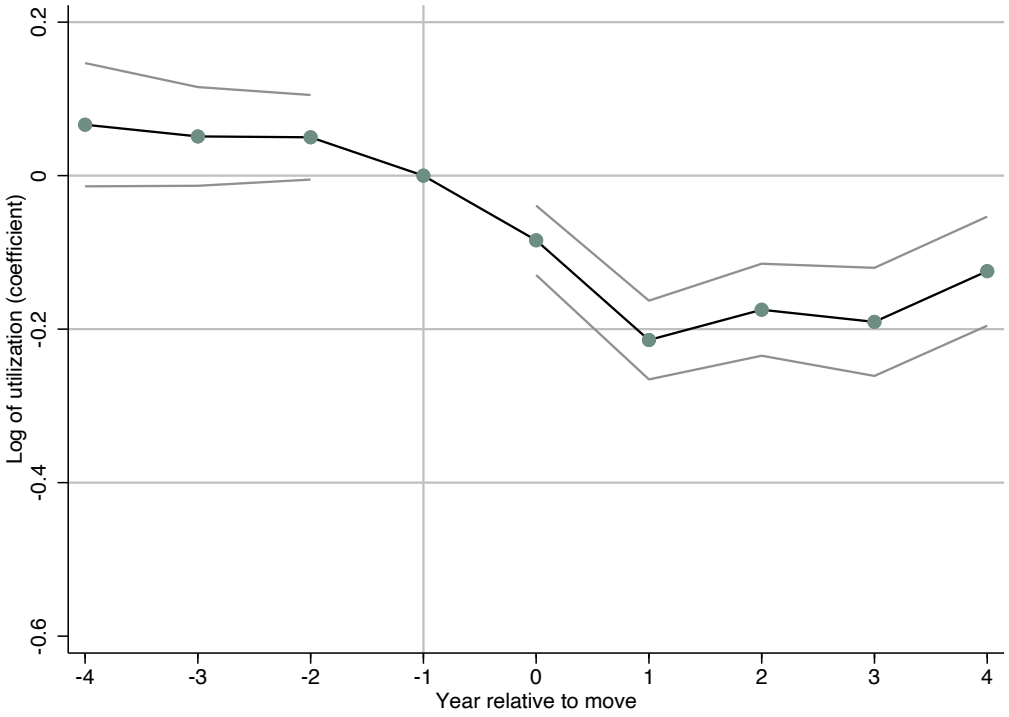
Notes. This is a stylized diagram contrasting the referral patterns of PCP A, who has low team referral concentration, and PCP B, who has high team referral concentration.

Figure 2: Binned Scatterplot of PCP Team Referral Concentration and Patient Utilization



Notes. Sample: MA APCD Analysis Sample. Patients are grouped into twenty equally sized groups based on their PCP's Team Referral Concentration (x-axis). Average ln(utilization) for each group is plotted on the y-axis.

Figure 3. Medicare Mover Analysis: Team Referral Concentration and Total Utilization



Notes. Regression results estimated in Medicare 20% sample, limited to patients who moved across regions during 2006–2012. The sample includes up to five years before and after each patient's move, including the year of the move. The unit of observation is a patient-year; N = 118,970. PCP Team Referral Concentration is jackknifed. Standard errors are clustered at the patient level. This graph plots coefficients (and 95% confidence intervals) on the interaction between the change in PCP team referral concentration at the time of the patient's move and indicator variables for the year relative to move. Year 0 is the year of the move; year −1 is the excluded category (normalized to zero). The regression includes individual patient fixed effects, year fixed effects, and a series of indicators for event year relative to the move.

Table 1. Descriptive Statistics of the Massachusetts APCD Sample

	Overall Mean	Overall Standard Deviation	Below-Median Team Referral Concentration	Above-Median Team Referral Concentration
Panel A: PCP characteristics (broad sample)				
PCP's team referral concentration	0.14	(0.07)	0.09	0.19
Number of shared patients per referral	3.50	(1.51)	2.93	4.08
Number of attributed patients	910	(518)	935	884
Pr(Internal Medicine)	0.73		0.76	0.71
PCP is male	0.61		0.61	0.61
Fraction capitation-paid PCP visits	0.06	(0.16)	0.03	0.08
Fraction HMO Patients	0.60	(0.11)	0.59	0.61
Number of insurer networks	15.2	(3.66)	15.08	15.33
Panel B: Patient characteristics (analysis sample)				
Mean spending ($)	7730	(20505)	8629	6828
Median spending ($)	2867		3142	2606
Male	0.49		0.49	0.50
Age	49.6	(10.0)	49.4	49.8
Pr(Any inpatient admission)	0.08		0.08	0.07
Patient care continuity HHI	0.45	(0.27)	0.43	0.47
Pr(Diabetes)	0.13		0.12	0.14
Pr(Heart Condition)	0.14		0.14	0.13
Pr(Bipolar and Major Depressive)	0.07		0.07	0.07
Pr(Asthma)	0.10		0.10	0.10
Number of patients	284,604		142,538	142,066

Notes. Data from Massachusetts All Payer Claims Data. Physician measures in Panel A are calculated using the Broad Sample of the PCP's entire patient panel, including patients who are not chronically ill. Reported means in this panel are weighted according to the number of patients each physician treats who are retained in the chronically ill sample. Patient characteristics in Panel B report results from the chronically ill Analysis Sample. Comorbidities are defined by HCC codes. Diabetes: 15–20. Heart Condition: 79–88, 92–93. Bipolar and Major Depressive: 55. Asthma: 110. The columns represent mean values for patients whose PCP has levels of team referral concentration that are respectively below or above the median. PCP characteristics are weighted by number of assigned patients. The probability of any inpatient admissions excludes admissions for pregnancy-related admissions. Missing observations for fraction of capitation-paid PCP visits, fraction of HMO patients, and number of insurer networks (n = 17).

Table 2. Average Team Referral Concentration by Physician Contracting Network

Physician Contracting Network	Average Team Referral Concentration	Std. Dev. Of Team Referral Concentration	N PCPs
Fallon Clinic	0.19	0.06	95
Atrius Health	0.16	0.06	254
Southcoast Physicians Network	0.16	0.07	54
Caritas Christi Network Service	0.15	0.07	204
Baycare Health Partners	0.13	0.06	136
Beth Israel Deaconess P.O.	0.13	0.12	309
New England Quality Care Alliance	0.13	0.07	229
Partners Community Health Care	0.13	0.09	925
UMass Memorial Health Care	0.13	0.08	234
Lahey Clinic	0.11	0.07	91
No Physician Contracting Network	0.21	0.16	2154

Notes. Unit of observation is a PCP. Physician Contracting Network is obtained by linking NPIs to the 2010 Massachusetts Provider Database maintained by Massachusetts Health Quality Partners, which defines Physician Contracting Network as an "organization of medical groups and/or practice sites with an integrated approach to quality improvement that enters into contracts with payers on behalf of its provider members" (Massachusetts Health Quality Partners 2016).

Table 3. Relationship Between PCP Team Referral Concentration and Healthcare Utilization

	Dependent variable: ln(utilization)			*Dependent variable: ln(spending)*		
	(1)	(2)	(3)	(4)	(5)	(6)
PCP Team Referral Concentration	-0.563***	-0.384***	-0.379***	-1.050***	-0.869***	-0.670***
	(0.047)	(0.046)	(0.059)	(0.060)	(0.060)	(0.075)
PCP controls	No	No	Yes	No	No	Yes
Patient Care Continuity HHI	No	Yes	Yes	No	Yes	Yes
Patient and Insurance controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Sample: MA APCD 2012 Analysis Sample. Unit of observation is a patient, N = 284,604. Standard errors are clustered at the PCP level. PCP Team Referral Concentration is jackknifed. Insurance controls are a fixed effect for each payer and for each of 13 types of insurance plans (i.e., HMO, PPO, EPO, indemnity, etc.). Patient controls are patient ZIP code fixed effects, sex, age (included as a five-knot spline), and comorbidity controls (fixed effects for each of the 162 hierarchical condition categories (HCCs)). PCP controls are the average HCC risk score of the PCP's commercial patients (as a five-knot spline), total number of patients (as five-knot spline), indicator for PCP sex, an indicator for PCP's specialty (either family medicine or internal medicine), and PCP modal ZIP code fixed effects. *p < 0.1; **p < 0.05; ***p < 0.01

Table 4. Robustness of Relationship Between Referral Concentration and Healthcare Utilization to PCP and Specialist Fixed Effects

	Dependent variable: ln(utilization)			
	(1)	(2)	(3)	(4)
PCP Team Referral Concentration in relevant specialties	-0.359*** (0.028)	-0.161*** (0.030)	-0.238*** (0.030)	-0.097*** (0.0348)
PCP fixed effect	No	Yes	No	Yes
Specialist fixed effect	No	No	Yes	Yes
Patient and insurance controls	Yes	Yes	Yes	Yes

Notes Sample: MA APCD 2012 Analysis Sample, limited to patients who saw at least one specialist in one of the nine included specialties. Unit of observation is a patient, N = 167,792. PCP Team Referral Concentration is jackknifed. Standard errors are clustered at the PCP level. All specifications include fixed effects for the combination of specialties consulted by the patient. Specialty-specific team referral concentration measures, averaging across only specialties that the index patient has consulted. Patient and insurer controls are the same as in Table 3, column (1). Column (1) reproduces Table 3, column (1) on the restricted sample with the specialty-specific measure of referral concentration and new controls for the mix of specialists seen. Columns (2)–(4) add additional fixed effects as noted. *p < 0.1; **p < 0.05; ***p < 0.01

Table 5. Relationship Between PCP Team Referral Concentration and Utilization-Based Quality Outcomes

	(1)	(2)	(3)	Mean of dependent variable
Dependent variable:				
Any emergency department visit	-0.085***	-0.037**	-0.132***	0.222
	(0.017)	(0.018)	(0.023)	
Any inpatient visit (excluding pregnancy)	-0.068***	-0.062***	-0.023*	0.076
	(0.010)	(0.010)	(0.013)	
Any ambulatory care sensitive hospitalization	-0.007**	-0.006**	-0.0005	0.008
	(0.003)	(0.003)	(0.005)	
Any duplicate imaging	-0.009	-0.006	0.003	0.043
	(0.007)	(0.007)	(0.009)	
Miles from patient ZIP to plurality specialist ZIP	-15.9***	-15.6***	-5.1***	11.3
	(1.04)	(1.04)	(0.97)	
Patient and insurance controls	Yes	Yes	Yes	
Patient care continuity HHI	No	Yes	Yes	
PCP controls	No	No	Yes	

Notes. Sample: MA APCD 2012 Analysis Sample. Sample size is 284,604 for all outcomes except for the "miles from patient ZIP to plurality specialist ZIP." Sample size for this distance outcome, limited to those who see at least one specialist in the included specialties, is 168,277. PCP Team Referral Concentration is jackknifed. Standard errors are clustered at the PCP level. Any hospitalization for ambulatory care–sensitive condition (ACSC) uses the indicator for an inpatient hospitalization and then uses software provided by the Agency for Healthcare Research and Quality (2016). Any duplicate imaging indicates whether a patient had the same imaging modality on the same body part within 30 days. This measure is adapted from Lammers et al. (2014), and includes head/neck CT, head/neck MRI, chest CT, chest MRI, chest x-ray, spine CT, spine MRI, pelvis CT, pelvis MR, lower extremity CT, lower extremity MRI, upper extremity CT, upper extremity MRI, abdominal CT, abdominal MRI, cardiac MR, abdominal ultrasound, pelvis ultrasound, and vein ultrasound. Miles from patient ZIP to plurality specialist ZIP uses straight line distances between 5-digit ZIP Codes and is set to missing if greater than 500 miles, and to zero for patients who see their plurality specialist in their home ZIP code. Columns (1)–(3) mirror specifications from Table 3, columns (1)–(3). *p < 0.1; **p < 0.05; ***p < 0.01

Table 6. Medicare Analysis: Team Referral Concentration and Total Utilization

	A. Cross-sectional analysis (2012) *Dependent variable: Ln(utilization)*		
	(1)	(2)	(3)
PCP Team Referral Concentration	-0.143***	-0.051***	-0.059***
	(0.009)	(0.004)	(0.004)
Number of observations	3,425,743	3,425,743	3,425,743
	B. Difference-in-differences *Mover analysis (2006-2012)* *Dependent variable: Ln(utilization)*		
	(4)	(5)	(6)
(Δ Team Referral Concentration)*Post	-0.217***	-0.095***	-0.089***
	(0.022)	(0.017)	(0.017)
Number of individual patients	22,121	22,121	22,121
Number of observations (patient X year)	99,334	99,334	99,334
	C. Instrumental variables (2nd stage) *Mover analysis (2006-2012)* *Dependent variable: Ln(utilization)*		
	(7)	(8)	(9)
(Δ Team Referral Concentration)*Post	-0.195***	-0.065***	-0.059***
	(0.024)	(0.019)	(0.019)
Number of individual patients	22,121	22,121	22,121
Number of observations (patient X year)	99,334	99,334	99,334
Patient controls	Yes	Yes	Yes
Patient care continuity HHI	No	Yes	Yes
PCP controls	No	No	Yes

Note. Data: patients in Medicare 20% sample. Panel A reports the cross-sectional analysis using 2012 Medicare data. Panels B and C analyze patients who switch PCPs at the time of a move. Panel B uses a difference-in-differences specification. Panel C uses an instrumental variables strategy; the endogenous variable is the patient's change in PCP team referral concentration and the instrumental variable is the patient's pre-move PCP's level of Team Referral Concentration. The sample in Panels B and C restrict to patients who move across regions during 2006–2012 and covers up to five years before and after each patient's move, dropping the year during which the move took place. PCP Team Referral Concentration is jackknifed in all specifications. Standard errors are clustered at the PCP level in Panel A and at the patient level in Panel B. Specifications 1–3 all control for patient sex and race, a five-knot spline in patient age, and fixed effects for each of 27 comorbidities coded in the Chronic Condition Warehouse. In lieu of these patient-level covariates, Specifications 4–9 include patient fixed effects. PCP controls in specification 3, 6, and 9 include PCP specialty (family medicine or internal medicine), gender, and a five-knot spline in number of patients. Specifications likely suffer from substantial attenuation bias due to measurement error. First-stage regressions for the IV are reported in Appendix Table A9. *p < 0.1; **p < 0.05; ***p < 0.01

Online Appendix

For "Team Relationships and Performance: Evidence from Healthcare Referral Networks"

Leila Agha, Keith Marzilli Ericson,
Kimberley H. Geissler, James B. Rebitzer

Appendix A: Additional Empirical Results

Figure A1. Sample Construction

This flow chart describes sample construction in the Massachusetts APCD. Team referral concentration is calculated in the broad sample. A series of sample restrictions yields the final analytic sample. For step 4, we calculated the modal ZIP code for each physician as the location where they practiced most days. In order to exclude physicians who may be treating many out-of-state patients who are not in our data set, we exclude all patients whose attributed PCP's modal ZIP code is not in Massachusetts.

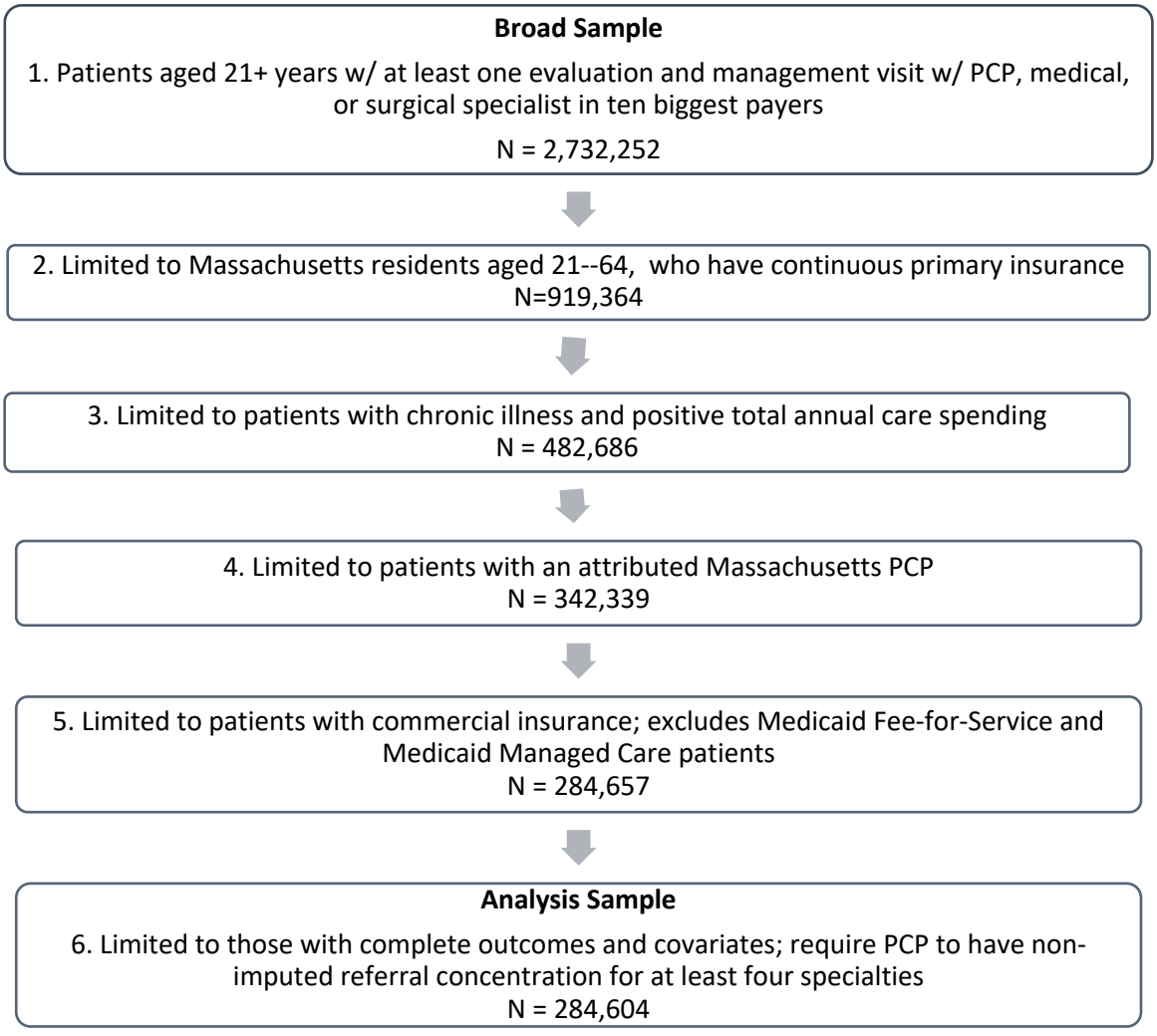
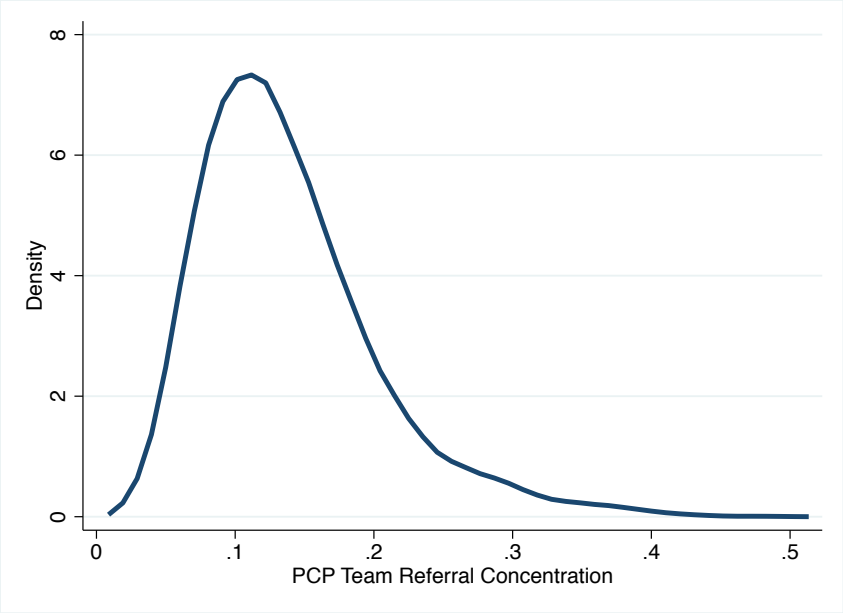
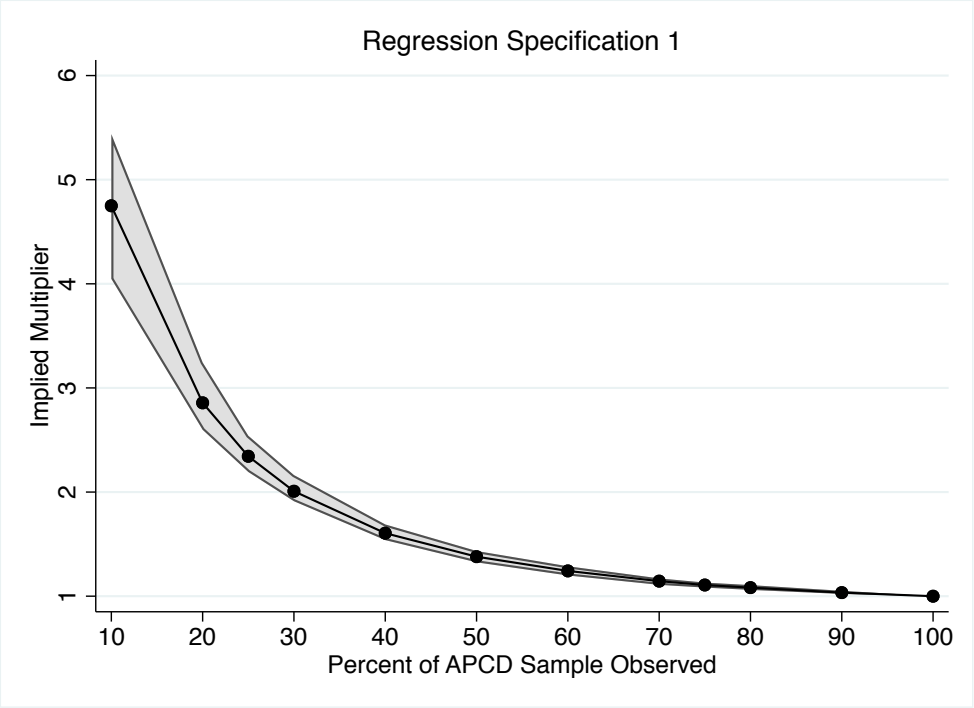


Figure A2. Distribution of PCP Team Referral Concentration



Note. Sample: MA APCD 2012 Analysis Sample. One observation per patient.

Figure A3. Measurement Error Multiplier Simulations Using Subsamples of APCD Data



Notes. Sample: MA APCD 2012 Analysis Sample. Due to computational limitations, the TRC is not jackknifed. This figure plots the average multiplier required to re-scale the regression coefficient estimated in a randomly drawn subsample of the APCD to the result calculated using the 100% sample. Shaded area shows 5th and 95th percentiles bootstrapped from 50 random samples per percent subsample. This specification testing used the same regression equation as column (1) of Table 3.

Table A1. Correlations Between Variables

	PCP's team referral concentration	Average patient HCC score	Patient care continuity HHI	Number of Insurer Networks	Internal Medicine	Number of patients	PCP is Male	Fraction HMO Patients	Fraction capitation paid PCP visits	Fraction patients with BCBS-HMO	Number of shared patients per referral	Patient Age	Patient is Male	Spending	Any inpatient admission	Any ED visits	Pr(Diabetes)	Pr(Heart Condition)	Pr(Bipolar and Major Depressive)	Pr(Asthma)
PCP's team referral concentration	1																			
Average patient HCC score	-0.10	1																		
Patient care continuity HHI	0.07	-0.11	1																	
Number of insurer networks	-0.003	-0.18	0.10	1																
Internal Medicine	-0.09	0.10	-0.05	-0.12	1															
Number of attributed patients	-0.06	-0.14	0.02	0.33	-0.08	1														
PCP is male	0.01	0.02	0.10	0.11	-0.001	0.17	1													
Fraction HMO Patients	0.13	-0.13	0.03	0.13	-0.19	0.17	-0.08	1												
Fraction capitation paid PCP visits	0.19	-0.06	-0.03	-0.11	0.13	0.15	-0.07	0.13	1											
Fraction patients with BCBS-HMO	0.07	-0.22	0.06	-0.02	-0.14	0.09	0.05	0.37	-0.10	1										
Number of shared patients per referral	0.43	-0.08	0.007	0.17	0.04	0.53	0.13	0.14	0.11	0.09	1									
Patient age	0.02	0.02	0.003	0.007	0.04	-0.02	0.06	-0.06	-0.03	-0.002	0.04	1								
Patient is male	0.01	0.01	0.20	0.03	0.01	0.03	0.36	-0.02	-0.005	0.009	0.02	0.01	1							
Spending	-0.05	0.32	-0.23	-0.07	0.03	-0.04	-0.01	-0.04	-0.03	-0.06	-0.04	0.04	-0.03	1						
Any inpatient admission	-0.03	0.25	-0.22	-0.06	0.02	-0.03	0.0003	-0.04	-0.02	-0.07	-0.02	0.02	-0.01	0.51	1					
Any ED visits	-0.01	0.15	-0.36	-0.03	0.0002	-0.03	-0.005	-0.02	-0.04	-0.07	-0.02	-0.08	-0.02	0.24	0.33	1				
Pr(Diabetes)	0.02	0.04	-0.004	0.01	0.0003	-0.009	0.02	0.004	-0.0001	-0.03	0.01	0.12	0.06	0.05	0.06	0.03	1			
Pr(Heart condition)	-0.03	0.14	-0.18	-0.04	0.03	-0.03	0.04	-0.05	-0.03	-0.04	-0.006	0.11	0.07	0.22	0.23	0.18	0.05	1		
Pr(Bipolar and major depressive)	-0.01	0.03	-0.1	-0.02	0.002	-0.01	-0.04	-0.0006	0.01	-0.02	-0.01	-0.04	-0.08	0.09	0.09	0.10	-0.0009	0.02	1	
Pr(Asthma)	-0.005	-0.001	-0.08	-0.003	-0.007	0.006	-0.04	0.02	0.01	0.002	-0.003	-0.16	-0.10	0.005	0.01	0.06	-0.05	-0.03	0.03	1

Notes. Correlations reported from patient-level data in the MA APCD 2012 Analysis Sample. We use the percent of patient panel in a Blue Cross-Blue Shield HMO plan as a proxy for participation in BCBS Alternative Quality Contract (AQC), a precursor to Affordable Care Organizations. Starting in 2009, BCBS Massachusetts implemented the AQC, which pays providers based on a risk-adjusted global budget. By 2012, the year of our analysis, 85% of eligible physicians in the BCBS network were participating. PCPs were eligible if they were part of a group collectively caring for at least 5,000 BCBS HMO or POS members (Chernew et al. 2011; Song et al. 2014).

Sources:

Chernew ME, Mechanic RE, Landon BE, Safran DG (2011) Private-payer innovation in Massachusetts: The "alternative quality contract.". *Health Affairs* 30(1):51–61.
Song Z, Rose S, Safran DG, Landon BE, Day MP, Chernew ME (2014) Changes in health care spending and quality 4 years into global payment. *New England Journal of Medicine* 371(18):1704–1714.

Table A2. Robustness of Table 3 results to definition of chronically ill sample

	Dependent variable: ln(utilization)			*Dependent variable: ln(spending)*		
	(1)	(2)	(3)	(4)	(5)	(6)
PCP Team Referral Concentration	-0.517***	-0.366***	-0.370***	-1.002***	-0.850***	-0.667***
	(0.0483)	(0.0478)	(0.0607)	(0.0619)	(0.0611)	(0.0765)
PCP controls	No	No	Yes	No	No	Yes
Patient Care Continuity HHI	No	Yes	Yes	No	Yes	Yes
Patient and Insurance controls	Yes	Yes	Yes	Yes	Yes	Yes

Notes. Sample: MA APCD 2012 Analysis Sample, further limited to those with four outpatient claims including a diagnosis code used to define chronically ill patients. Unit of observation is a patient, N = 236,067. Standard errors are clustered at the PCP level. PCP team referral concentration is jackknifed. Insurance controls are a fixed effect for each payer and for each of 13 types of insurance plans (i.e., HMO, PPO, EPO, indemnity, etc.). Patient controls are patient ZIP code fixed effects, sex, age (included as a five-knot spline) and comorbidity controls (fixed effects for each of the 162 hierarchical condition categories (HCCs)). PCP controls are the average HCC risk score of the PCP's commercial patients (as a five-knot spline), total number of patients (as five-knot spline), indicator for PCP sex, an indicator for PCP's specialty (either family medicine or internal medicine), and PCP modal ZIP code fixed effects. *p < 0.1; **p < 0.05; ***p < 0.01

Table A3. Decomposing PCP Spending and All Other Spending

	Dependent variable: Ln(Spending by plurality PCP only) (1)	*Dependent variable: Ln(Spending by all providers except plurality PCP)* (2)
PCP Team Referral Concentration	0.127	-1.237***
	(0.104)	(0.075)
Patient and insurer controls	Yes	Yes
N patients	284,294	280,997

Notes. Sample: MA APCD 2012 Analysis Sample. PCP's team referral concentration is jackknifed. Standard errors are clustered at the PCP level. The column (1) dependent variable is logged total spending assigned to the plurality PCP at insurer-allowed amounts; column (2) dependent variable is logged total spending for all other claims at insurer-allowed amounts. The controls in each regression mirror the specifications in Table 3, column (1). *p < 0.1; **p < 0.05; ***p < 0.01

Table A4. Relationship Between Utilization and Specialty-Specific Team Referral Concentration

	Dependent Variable: ln(utilization)								
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Sample:	Orthopedics	Dermatology	OB/GYN	Cardiology	Otolaryngology	Surgery	Urology	Endocrinology	Ophthalmology
Specialty-specific PCP Team Referral Concentration	-0.700***	-0.339***	-0.395***	-0.307***	-0.226***	-0.423***	-0.224***	-0.259***	-0.218***
	(0.090)	(0.036)	(0.082)	(0.072)	(0.061)	(0.082)	(0.058)	(0.0387)	(0.107)
Patient and insurance controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Number of Observations	51,814	49,987	48,653	31,436	19,769	17,984	17,208	16,046	11,742

Notes Sample: MA APCD 2012 Analysis Sample; sample in each column is restricted to patients who saw at least one specialist in the given specialty. PCP's team referral concentration is jackknifed and is specialty-specific. Standard errors are clustered at the PCP level. All specifications include fixed effects for the set of specialties consulted by the patient. Patient and insurer controls are the same as in Table 3, column (1). *p < 0.1; **p < 0.05; ***p < 0.01

Table A5. PCP Team Referral Concentration and Utilization, With Physician Contracting Network Fixed Effects

	Dependent variable: Ln(utilization)	
	(1)	(2)
Sample:	PCP in Any Contracting Network	PCP in Partners Healthcare
PCP Team Referral Concentration	-0.403***	-0.681***
	(0.048)	(0.126)
Patient and insurance controls	Yes	Yes
Physician contracting network FEs	Yes	N/A
N patients	284,604	61,793

Notes. Column (1) sample is MA APCD 2012 Analysis Sample. Column (2) sample is MA APCD 2012 Analysis Sample, limited to patients with PCPs affiliated with Partners Healthcare. PCP's team referral concentration is jackknifed. Standard errors are clustered at the PCP level. Patient and insurer controls are the same as in Table 3, column (1). In column (1), a fixed effect for the physician contracting network is also included, including an indicator of not being in a physician contracting network. *p < 0.1; **p < 0.05; ***p < 0.01

Table A6. PCP Team Referral Concentration and Utilization, Limited to Blue Cross-Blue Shield PPO Enrollees

	Dependent variable: Ln(utilization)
Sample:	Only patients enrolled in BCBS PPO
PCP Team Referral Concentration	-0.610***
	(0.070)
Patient controls	Yes
N patients	74,711

Notes. Sample: MA APCD 2012 Analysis Sample, limited to patients enrolled in a Blue Cross-Blue Shield PPO plan. PCP's team referral concentration is jackknifed. Standard errors are clustered at the PCP level. Patient controls are the same as in Table 3, column (1); insurer controls are not included due to the sample limitations. *p < 0.1; **p < 0.05; ***p < 0.01

Table A7. Medicare Sample Descriptive Statistics

	Below-Median PCP Team Referral Concentration (2012)	Above-Median PCP Team Referral Concentration (2012)	Sample of moving beneficiaries (2006-2012)
Patient characteristics:			
Age	77.0	76.7	79.7
Male	0.41	0.40	0.34
Mean spending ($)	12,236	10,221	10,834
Median spending ($)	4158	3657	4387
Pr(Any inpatient admission)	0.26	0.22	0.29
Patient Care Continuity HHI	0.31	0.41	0.38
Pr(Diabetes)	0.37	0.29	0.26
Pr(Heart Condition)	0.44	0.41	0.50
Pr(Depression)	0.16	0.15	0.18
Pr(Asthma)	0.05	0.05	0.05
PCP characteristics:			
PCP Team Referral Conc.	0.31	0.69	0.50
Pr(Internal Medicine)	0.63	0.49	0.59
PCP is male	0.79	0.74	0.76
Number of observations	1,712,875	1,712,868	99,334

Notes. Data: patients in Medicare 20% sample. Columns (1) and (2) report the cross-sectional sample using 2012 Medicare data. Unit of observation is a patient. Column (3) sample restricts to patients who moved across regions during 2006–2012, and covers up to five years before and after each patient's move, dropping the year during which the move took place. Unit of observation is a patient-year. Similar to our findings in Massachusetts, we find that a patient's age, sex, and disease burden are similar among patients seeing PCPs with above- and below-median team referral concentration. Because of measurement error, there is both more concentration and more variation in the measured PCP team referral concentration.

Table A8. Utilization-Based Quality Outcomes in Medicare Mover Sample

	(1)	(2)	(3)	
	Difference-in-differences results			Dep. var. mean
Dependent variable:				
Any emergency department visit	-0.039***	-0.006	-0.013*	0.401
	(0.007)	(0.007)	(0.007)	
Any inpatient visit	-0.054***	-0.025***	-0.031***	0.294
	(0.007)	(0.007)	(0.018)	
Any hospitalization for ambulatory care sensitive condition	-0.024***	-0.016***	-0.020***	0.089
	(0.005)	(0.005)	(0.005)	
Patient fixed effects	Yes	Yes	Yes	
Patient care continuity HHI	No	Yes	Yes	
PCP controls	No	No	Yes	

Notes. Data: patients in Medicare 20% sample, limited to patients who move across regions during 2006–2012 and covers up to five years before and after each patient's move, dropping the year during which the move took place. Unit of observation is a patient-year. Sample size is 99,334. PCP team referral concentration is jackknifed in all specifications. Standard errors are clustered at the patient level. This table reports the coefficient on the change in team referral concentration multiplied by a post-move indicator from a difference-in-difference specification analyzing patients who switch PCPs at the time of a move. Calculation of hospitalization for ambulatory care–sensitive conditions uses software provided by the Agency for Healthcare Research and Quality (2016). All specifications include patient fixed effects. PCP controls in specification 3 include PCP specialty (family medicine or internal medicine), gender, and a five-knot spline in number of patients. Specifications likely suffer from substantial attenuation bias due to measurement error; simulations suggest measurement error is particularly acute after additional controls are incorporated in specifications 2 and 3. *p < 0.1; **p < 0.05; ***p < 0.01

Table A9. Medicare Analysis: First Stage of Instrumental Variables Approach

	A. Instrumental variables (1st stage) Mover Analysis (2006–2012) Dependent variable: (Δ Team Referral Concentration)*Post		
	(1)	(2)	(3)
(Pre-period Team Referral Concentration)*Post	-0.968***	-0.967***	-0.950***
	(0.006)	(0.006)	(0.007)
F-test of excluded instrument	25,035	24,497	17,724
Patient fixed effect	Yes	Yes	Yes
Patient care continuity HHI	No	Yes	Yes
PCP controls	No	No	Yes
Number of individual patients	22,121	22,121	22,121
Number of observations (patient X year)	99,334	99,334	99,334

Notes. Sample same as Table 6, Panel C. Standard errors are clustered at the patient level. Specifications 1–3 match the specifications in Table 6, Panel C. *p < 0.1; **p < 0.05; ***p < 0.01

Table A10. Complete Regression Results for Table 3, Replacing Splines with Simple Linear Controls

	Dependent variable: ln(utilization)			*Dependent variable: ln(spending)*		
	(1)	(2)	(3)	(4)	(5)	(6)
PCP Team Referral	-0.564***	-0.385***	-0.530***	-1.050***	-0.870***	-0.814***
Concentration	(0.0466)	(0.0462)	(0.0584)	(0.0603)	(0.0598)	(0.0743)
Patient Care Continuity		-1.590***	-1.607***		-1.604***	-1.620***
HHI		(0.00864)	(0.00822)		(0.0093)	(0.00878)
Patient age	-0.000745***	0.00175***	0.00180***	4.95E-05	0.00256***	0.00254***
	(0.000196)	(0.000174)	(0.000172)	(0.000212)	(0.00019)	(0.000183)
Patient is male	-0.275***	-0.145***	-0.138***	-0.267***	-0.136***	-0.124***
	(0.00406)	(0.00369)	(0.00359)	(0.00445)	(0.00415)	(0.00375)
PCP Internal medicine			-0.0200***			-0.0128
			(0.0075)			(0.00819)
PCP total patients			-5.89e-05***			-5.47e-05***
			(1.52E-05)			(1.59E-05)
PCP avg. HCC risk score			-0.0179***			-0.0200***
			(0.00339)			(0.0036)
PCP is male			-0.0158***			-0.0213***
			(0.00516)			(0.00602)
HCC indicator variables	Yes	Yes	Yes	Yes	Yes	Yes
Insurance Type FE	Yes	Yes	Yes	Yes	Yes	Yes
Payer FE	Yes	Yes	Yes	Yes	Yes	Yes
Patient ZIP FE	Yes	Yes	Yes	Yes	Yes	Yes
Physician ZIP FE	No	No	Yes	No	No	Yes
Observations	284,604	284,604	284,604	284,604	284,604	284,604
R-squared	0.511	0.609	0.612	0.513	0.604	0.61

Notes. Sample: MA APCD 2012 Analysis Sample. PCP's team referral concentration is jackknife. Standard errors are clustered at the PCP level. Specifications are nearly identical to those reported in Table 3, but omit splines and use linear forms for these control variables instead. *p < 0.1; **p < 0.05; ***p < 0.01

Appendix B. Chronic Illness Definitions for Massachusetts Population

We use an adapted definition of chronic illness from Frandsen et al. (2015). A patient is identified as having a chronic illness if they received an ICD-9 diagnostic code for one inpatient hospitalization or at least two outpatient claims in one of the following categories:

- Coronary artery disease: 410.xx–414.xx
- Cerebrovascular disease: 433.xx–438.xx, 441.xx–442.xx
- Peripheral arterial disease: 443.xx–445.xx
- Mesenteric vascular disease: 557.xx
- Other ischemic vascular disease or conduction disorders: 391.xx, 394.xx–398.xx, 440.xx, 426.xx–427.xx
- Heart failure: 402.01, 402.11, 402.91, 401.01, 404.03, 404.11, 404.13, 404.91, 404.93, 428.xx
- Migraine and cluster headache: 346.xx, 339.xx
- Hypertension: 401.xx–405.xx
- Hyperlipidemia: 272.xx
- Diabetes mellitus: 249.xx–250.xx, 362.0x
- Asthma: 493.xx
- Chronic obstructive pulmonary disease: 491.xx–492.xx, 494.xx, 496.xx, 416.xx
- Hypercoagulability disorders: 415.xx, 451.xx–454.xx
- Osteoarthritis: 715.xx, 717.xx, 721.xx, 726.xx
- Rheumatoid arthritis: 714.xx, 720.x

Appendix C. Physician Classification

We classify physicians by specialty using specialty taxonomy codes reported in the National Plan and Provider Enumeration System (NPPES). Our extract of the NPPES data was downloaded in February 2014. We used the physician's primary specialty if available. If multiple specialties were listed, but none were indicated as primary, we used the most specialized as their classification. The specific taxonomy code definitions are listed below.

Classification	Taxonomy Code	Physician Specialty
Primary Care Physician	207Q00000X	Family Medicine
Primary Care Physician	207QA0000X	Family Medicine–Adolescent Medicine
Primary Care Physician	207QA0505X	Family Medicine–Adult Medicine
Primary Care Physician	207QG0300X	Family Medicine–Geriatric Medicine
Primary Care Physician	207R00000X	Internal Medicine
Primary Care Physician	207RA0000X	Internal Medicine–Adolescent Medicine
Primary Care Physician	207RG0300X	Internal Medicine–Geriatric Medicine
Primary Care Physician	208000000X	Pediatrics
Primary Care Physician	2080A0000X	Pediatrics–Adolescent Medicine
Primary Care Physician	208D00000X	General Practice
Cardiology	207RC0000X	Internal Medicine–Cardiovascular Disease
Cardiology	207RC0001X	Internal Medicine–Clinical Cardiac Electrophysiology
Cardiology	207RI0011X	Internal Medicine–Interventional Cardiology
Cardiology	207UN0901X	Nuclear Medicine–Nuclear Cardiology
Cardiology	2080P0202X	Pediatrics–Pediatric Cardiology
Cardiology	2086S0129X	Surgery–Vascular Surgery
Cardiology	208G00000X	Thoracic Surgery (Cardiothoracic Vascular Surgery)
Dermatology	207N00000X	Dermatology
Dermatology	207ND0101X	Dermatology–MOHS–Micrographic Surgery
Dermatology	207ND0900X	Dermatology–Dermatopathology
Dermatology	207NI0002X	Dermatology-Clinical & Laboratory Dermatological Immunology
Dermatology	207NP0225X	Dermatology–Pediatric Dermatology
Dermatology	207NS0135X	Dermatology–Procedural Dermatology
Endocrinology	207RE0101X	Internal Medicine–Endocrinology, Diabetes & Metabolism
Endocrinology	2080P0205X	Pediatrics–Pediatric Endocrinology
OB/GYN	207V00000X	Obstetrics & Gynecology
OB/GYN	207VB0002X	Obstetrics & Gynecology-Bariatric Medicine

Classification	Taxonomy Code	Physician Specialty
OB/GYN	207VC0200X	Obstetrics & Gynecology-Critical Care Medicine
OB/GYN	207VE0102X	Obstetrics & Gynecology-Reproductive Endocrinology
OB/GYN	207VF0040X	Obstetrics & Gynecology–Female Pelvic Medicine and Reconstructive Surgery
OB/GYN	207VG0400X	Obstetrics & Gynecology–Gynecology
OB/GYN	207VH0002X	Obstetrics & Gynecology–Hospice and Palliative Medicine
OB/GYN	207VM0101X	Obstetrics & Gynecology–Maternal & Fetal Medicine
OB/GYN	207VX0000X	Obstetrics & Gynecology–Obstetrics
OB/GYN	207VX0201X	Obstetrics & Gynecology–Gynecologic Oncology
OB/GYN	2088F0040X	Urology–Female Pelvic Medicine and Reconstructive Surgery
Ophthalmology	207W00000X	Ophthalmology
Orthopedic	204C00000X	Neuromusculoskeletal Medicine, Sports Medicine
Orthopedic	204D00000X	Neuromusculoskeletal Medicine & OMM
Orthopedic	207QS0010X	Family Medicine–Sports Medicine
Orthopedic	207RS0010X	Internal Medicine–Sports Medicine
Orthopedic	207X00000X	Orthopaedic Surgery
Orthopedic	207XP3100X	Orthopaedic Surgery-Pediatric Orthopaedic Surgery
Orthopedic	207XS0106X	Orthopaedic Surgery-Hand Surgery
Orthopedic	207XS0114X	Orthopaedic Surgery-Adult Reconstructive Orthopaedic Surgery
Orthopedic	207XS0117X	Orthopaedic Surgery-Orthopaedic Surgery of the Spine
Orthopedic	207XX0004X	Orthopaedic Surgery–Foot and Ankle Surgery
Orthopedic	207XX0005X	Orthopaedic Surgery–Sports Medicine
Orthopedic	207XX0801X	Orthopaedic Surgery–Orthopaedic Trauma
Orthopedic	2080S0010X	Pediatrics–Sports Medicine
Orthopedic	2083S0010X	Preventive Medicine–Sports Medicine
Otolaryngology	207Y00000X	Otolaryngology
Otolaryngology	207YP0228X	Otolaryngology–Pediatric Otolaryngology
Otolaryngology	207YS0012X	Otolaryngology–Sleep Medicine
Otolaryngology	207YS0123X	Otolaryngology–Facial Plastic Surgery
Otolaryngology	207YX0007X	Otolaryngology–Plastic Surgery within the Head & Neck
Otolaryngology	207YX0602X	Otolaryngology–Otolaryngic Allergy
Otolaryngology	207YX0901X	Otolaryngology–Otology & Neurotology
Otolaryngology	207YX0905X	Otolaryngology–Otolaryngology/Facial Plastic Surgery
Surgery	208600000X	Surgery
Urology	208800000X	Urology
Urology	2088P0231X	Urology–Pediatric Urology

Appendix D. Measurement Error in Medicare Difference-in-Differences Regressions

In this section, we derive the expected impact of measurement error on the difference-in-differences results in the Medicare sample.

For simplicity, we consider a first-differenced specification where we keep one observation per patient who moves across regions. The dependent variable is the patient's change in care utilization (denoted $\Delta \log y_i$) and the independent variable is the change in patient's PCP team referral concentration (denoted ΔTRC_i) after the move. (For brevity, we notate team referral concentration as TRC in this appendix rather than ReferralCon used in the text.) The regression takes the following form:

$$\Delta \log y_i = \alpha \Delta TRC_i + \beta + \varepsilon_i \, .$$

In the absence of any measurement error, we would have a coefficient on the change in referral concentration that takes the following form:

$$\hat{\alpha} = \frac{Cov(\Delta \log y_i \, , \Delta TRC_i)}{Var(\Delta TRC_i)} .$$

We do not observe ΔTRC_i directly in the Medicare data because we only have a 20% sample of Medicare patients for each doctor. As a result, team referral concentration is measured with error. We denote these noisy signals $\widetilde{\Delta TRC}$ and suppress subscripting notation below for simplicity.

Specifically, define:

$$\widetilde{\Delta TRC} \; = \; \Delta TRC + \Delta \mu \, .$$

We will consider two cases. First, we will assume a case with classical measurement error, so that $\Delta \mu$ is independently distributed and therefore is not correlated with the change in team referral concentration nor with care utilization. Then we will consider the more realistic case that $\Delta \mu$ is not independently distributed.

In the classical measurement error case, the independence assumption implies that $\Delta \mu$ is uncorrelated with ΔTRC and $\Delta \log y$. When we estimate the difference-in-differences specification, we will find the following coefficient:

$$\hat{\alpha}^{classical \; m.e.} = \frac{Cov(\Delta \log y \, , \Delta TRC)}{Var(\Delta TRC) + Var(\Delta \mu)} .$$

This coefficient suffers from attenuation bias, as in the classical derivations; this is seen in the addition of the term to the denominator.

Now consider the more complicated, but also more realistic, possibility that the error in the team referral concentration measure is related to the level of team referral concentration. It is easy to see why the

independence assumption may be violated in our setting if you consider the behavior of measurement error near the bounds of the referral concentration measure. A doctor who is perfectly concentrated and only refers to one specialist of each type will have no error in his team referral concentration measure when measured using a 20% sample. As long as we observe one referred patient, we would be able to perfectly calculate his TRC as equal to 1. By contrast, consider a doctor who is not at all concentrated in his referrals. Within each specialty, he refers each of his patients to a different specialist. The more patients we observe, the closer his TRC comes to 0, but in any finite subsample of his patient panel, we will overestimate his TRC. Extending this intuition, we expect measurement error in team referral concentration to be negatively correlated with the true referral concentration.

The difference-in-differences regression coefficient now becomes:

$$\hat{\alpha}^{non-independent\ m.e.} = \frac{Cov(\Delta \log y\,, \Delta TRC) + Cov(\Delta \log y\,, \Delta \mu)}{Var(\Delta TRC) + Var(\Delta \mu) + 2Cov(\Delta \mu, \Delta TRC)}.$$

Unlike the classical measurement error case, the sign and size of the bias is no longer obvious and will depend on the particular relationships in our setting. We expect that $Cov(\Delta \log y\,, \Delta TRC) < 0$, given the predictions of our model and the results in the Massachusetts data, which have minimal measurement error. By contrast, we expect that $Cov(\Delta \log y\,, \Delta \mu) > 0$, given the intuition about measurement error and its relationship to team referral concentration described in the previous paragraph. As long as $|Cov(\Delta \log y\,, \Delta TRC)| > |\,Cov(\Delta \log y\,, \Delta \mu)|$, the changes in the numerator will tend to attenuate the measured coefficient.

In the denominator, the variance terms are positive. We expect that $Cov(\Delta \mu, \Delta TRC) < 0$. This implies that the net effect of measurement error on the denominator depends on the relative size of the $Var(\Delta \mu)$ and $Cov(\Delta \mu, \Delta TRC)$ terms. If $|Var(\Delta \mu)| > |2\ Cov(\Delta \mu, \Delta TRC)|$, then the denominator will be inflated, and there will be attenuation bias. On the other hand, if $|Var(\Delta \mu)| < |2\ Cov(\Delta \mu, \Delta TRC)|$, then the denominator will be smaller relative to the case without measurement error. In this case, the coefficient could be inflated (or even, in the extreme, wrong-signed).

In sum, the net effect of measurement error on the coefficient is theoretically ambiguous in the difference-in-differences setting. The coefficient could be either inflated or attenuated depending on the strength of the correlation of measurement error with the other terms. Note that the cross-sectional regressions would have a very similar formulation for bias from measurement error; eliminating the Δ terms from the formulas above would yield the OLS coefficients. The simulations of measurement error we run in the Massachusetts APCD suggest that the attenuating terms dominate, at least in the cross-sectional specification (cf. Section 8A).

To compare attenuation bias in the difference-in-differences specifications to attenuation bias in the cross-sectional Medicare results, we must also draw one more distinction. When we estimate the mover

results, we average the patient's PCP team referral concentration over the year(s) of the pre-move period to form the patient's PCP's pre-move team referral concentration (and similarly for the post period). This change will tend to reduce the noise in our signal of PCP team referral concentration, reducing relative to the single year from the static OLS model. This change would lead us to predict a smaller role for bias from measurement error in the mover specifications.

Appendix E. Instrumental Variables Analysis

This section provides more detail on the instrumental variables results reported in Section 8. The key endogenous variable of interest is the change in PCP team referral concentration experienced by a patient who switches PCPs due to a move. In the destination region, the patient may endogenously select a new provider on the basis of changes in the patient's health status. To surmount this potential source of endogeneity bias in the difference-in-differences approach, the IV analysis uses the patient's pre-move PCP's team referral concentration as an IV for their experienced change in team referral concentration.

The first-stage IV equation proceeds as follows:

$$\Delta ReferralCon_{-i}Post_{it} = \tilde{\alpha} PreMoveReferralCon_{-i}Post_{it} + \tilde{\beta}_i + \tilde{\gamma} Z_{it} + \tilde{\rho}_{R_{it}} + \widetilde{\varepsilon_{it}} .$$

And the reduced form is given as:

$$\log y_{it} = \alpha PreMoveReferralCon_{-i}Post_{it} + \beta_i + \gamma Z_{it} + \rho_{R_{it}} + \varepsilon_{it} .$$

The exclusion restriction requires parallel trends among patients with different initial exposures to team referral concentration. This relaxes the assumption required for interpreting the difference-in-differences approach, which requires parallel trends among patients who experience different *changes* in team referral concentration. A more detailed discussion of identifying assumptions can be found in Agha, Ericson, and Zhao (2020).

The first stage of this IV procedure is reported in Appendix Table A9. The second stage is reported in Table 6. Results are discussed and interpreted in Section 8.

Measurement error (as described in Appendix D above) will also bias the IV estimation. We have non-classical measurement error, due to the correlation in the error of the instrument, $PreMoveReferralCon_{-i}Post_{it}$, and the error in the endogenous variable $\Delta ReferralCon_{-i}Post_{it}$.