

NBER WORKING PAPER SERIES

TEAM FORMATION AND PERFORMANCE:
EVIDENCE FROM HEALTHCARE REFERRAL NETWORKS

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

Working Paper 24338
<http://www.nber.org/papers/w24338>

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
February 2018, Revised June 2019

We acknowledge funding from the Boston University Questrom School of Business, the National Institute of Health Care Management, and the National Institute of Health grant PO1 AG19783. 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Leila Agha, Keith Marzilli Ericson, Kimberley H. Geissler, and James B. Rebitzer. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

Team Formation and Performance: Evidence from Healthcare Referral Networks
Leila Agha, Keith Marzilli Ericson, Kimberley H. Geissler, and James B. Rebitzer
NBER Working Paper No. 24338
February 2018, Revised June 2019
JEL No. D85,I10,I11,L2,M5

ABSTRACT

How does team specific capital affect productivity? We examine the teams that primary care physicians (PCPs) assemble when referring patients to specialists. Our theoretical model finds that team specific capital is greater when PCPs concentrate their referrals within a smaller set of specialists. Empirically, we find patients of PCPs with concentrated referrals have lower healthcare costs, with no discernable reduction in quality. This effect exists for commercially insured and Medicare populations; is statistically and economically significant; and holds under identification strategies that account for unobserved patient and physician characteristics.

Leila Agha
Department of Economics
Dartmouth College
6106 Rockefeller Hall
Hanover, NH 03755
and NBER
leila.gha@dartmouth.edu

Kimberley H. Geissler
University of Massachusetts at Amherst
School of Public Health and Health Sciences
715 N Pleasant St
325 Arnold Hall
Amherst, MA 01003
kgeissler@umass.edu

Keith Marzilli Ericson
Boston University
Questrom School of Business
595 Commonwealth Avenue
Boston, MA 02215
and NBER
kericson@bu.edu

James B. Rebitzer
Professor of Management, Economics, Public Policy
Markets, Public Policy and Law Department
Boston University School of Management
595 Commonwealth Ave.
Boston, MA 02215
and NBER
rebitzer@bu.edu

1. Introduction

Teams are pervasive in economic organizations, so the performance of teams matters a great deal for organizational efficiency. Potential obstacles to optimal team performance include free-riding (Holmstrom 1982) and incomplete contracting (Williamson 1985; Hart 2017), but even in the absence of these problems, coordinating complex tasks and sharing information within teams is difficult (Marschak and Radner 1972, Becker and Murphy 1992, Dessein and Santos 2006).

The efficiency of teams is important and challenging in healthcare because a wide array of specialist and primary care clinicians must work together to treat patients who often have complex health problems. Managing this complexity is made more difficult by the fact that these clinical teams often involve referrals across different organizations; contracts between referring parties are either absent or quite incomplete; and the use of incentives in referrals is greatly restricted as a matter of law.¹ Prior research in non-health settings suggests that team-specific capital improves team performance (Mailath and Postlewaite 1990; Chillemi and Gui 1997; Baghai, Silva, and Ye 2018; Jaravel et al. 2018), in part because repeated interactions among team members may encourage investments in team relationships (Crawford 1990).² In this paper, we investigate whether repeated interactions between primary care providers and specialists similarly affect team performance, particularly the costs of healthcare delivery.

Primary care physicians (PCPs) regularly refer patients to specialists, but researchers know little about how the structure of these referrals influences the cost or quality of care that patients receive. Our analysis rests on the assumption that repeated interactions between PCPs and specialists facilitate investments in team relationships.³ Specifically, when PCPs concentrate their patient referrals within a narrower group of providers within a specialty (e.g. refer their cardiology patients to a smaller set of cardiologists), this encourages repeat interactions and so incentivizes greater investments in team-specific capital. While this enhanced team-specific capital may foster improvements in team performance, the benefits of referring to a smaller set of specialists come at some cost: a potentially less favorable patient-specialist match. We formalize this tradeoff between team-specific capital and patient-specialist match in a model of team formation. The model shows that the concentration of specialist referrals within a PCP's team is a meaningful proxy for investment in team-specific capital. In our empirical work, we use

¹ The "Stark laws" effectively prohibits 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.

² See also Topel (1991) and Jacobson, LaLonde and Sullivan (1993) for examples of empirical evidence on the role of specific capital in employment relationships.

³ Barnett et al. (2012b) survey PCPs about the reasons they choose a referral, and find that patient experience and ease of communication are important reasons.

this result to develop a novel measure of PCP *team referral concentration*. We apply this measure to study the effects 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. This association exists for both commercially insured and Medicare populations; is statistically and economically significant; and holds under various identification strategies. For commercially insured, chronically-ill patients in Massachusetts, those treated by PCPs with the average below-median team referral concentration have 4% lower utilization, compared to those treated by PCPs with the average 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 similarly sized effects in that population, after accounting for measurement error issues (described in more detail below).

There are two major threats to identification in our analysis. 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. 2018). To address this concern, we demonstrate that the effect of team referral concentration on spending persists even using within-PCP, cross-specialty variation in team referral concentration, or when comparing patients who consult the same specialist but are referred from PCPs with different levels of team referral concentration.

The second endogeneity threat is that patients who choose lower referral concentration PCPs may be in worse health (along unobservable dimensions) and demand more costly care. For Medicare beneficiaries, we use 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 and an instrumental variables strategy based on regression to the mean to remove potential endogeneity in choice of PCP. In the Medicare population, we also find that an increase in the team referral concentration of a patient’s PCP is associated with lowered utilization.

Our study contributes to a growing empirical literature on the economics of team organization and productivity.⁴ In work that examines the role of coordination of expertise within software development teams, Faraj and Sproull (2000) highlight the role of coordination on team performance. Prior studies have demonstrated 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.

⁴See 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, and Grosse, Putterman, and Rockenback 2011). There is also a literature in psychology on the performance of teams, reviewed in Kozlowski and Ilgen (2006).

2009) and surgical teams (Reagans et al. 2005). An advantage of our research setting is that we have micro data on team formation, tasks and performance across a large number of organizations, encompassing thousands of PCP-specialist teams.

In addition to contributing to the general economics 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.⁵ This prior research has motivated important policy initiatives that aim to improve care coordination, including policies designed to alter organizational form and incentives (e.g. Accountable Care Organizations and Patient Centered Medical Homes) and public subsidies to health information technology such as electronic health records.

Much of the preceding care coordination literature has focused on the relationship an individual patient has with their set of providers rather than on relationships within teams of providers.⁶ To see the significance of this distinction, imagine a PCP who refers each of her patients with diabetes to two specialists: a cardiologist and an endocrinologist. The PCP can choose to make each of her patient's referrals to the same cardiologist and the same endocrinologist, or could refer to a different cardiologist and different endocrinologist for each patient. In both cases, the distribution of an individual patient's care across providers is the same: each patient sees her PCP, cardiologist, and endocrinologist. But in the former case, the provider team works together more frequently, thus facilitating investment in team-specific capital.

While team-specific capital investment may be facilitated by vertically integrated relationships unifying primary care and specialty care within the same organization, as studied by Brot-Goldberg and de Vaan (2018), our analysis abstracts from firm boundaries.⁷ We find empirically large variation in team structure even within organizations, and further show that team structure remains an important predictor of care utilization after accounting for the physician's contracting organization.

⁵ For example, see Agha et al. (2017); Hussey et. al. (2014); Romano et. al (2015); 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). For a contrary view on the role of care coordination in lowering costs, see McWilliams (2016). The problem of building organizations and institutions that coordinate the activities of specialized providers has implications that extend beyond healthcare costs. See, for example, Meltzer (2001) and Meltzer and Chung (2010) on hospitalists.

⁶ There is also another literature that maps out patient-sharing networks among physicians (see e.g. Barnett et al. 2012a). This literature uses characteristics of the social network inferred from claims data that are distinct from our measure of referral concentration. Jackson, Rogers, and Zenou (2017) review the underlying concepts of network structure (e.g. clustering, centrality).

⁷ Geissler et al. (2018) using all payer data from Massachusetts find that practice sites and medical groups, but not physician contracting networks, are important determinants of referrals.

Our empirical analysis relies primarily on the Massachusetts All Payer Claims Database (APCD),⁸ a database of nearly all patients in the state with commercial insurance, Medicaid, or Medicare Advantage. The APCD is uniquely suited for our study design because it allows us to observe a larger fraction of a given provider's links (via shared patients) to other providers than has been available in other data sets from Medicare or commercial insurers. When estimating team referral concentration, we show that omitting a large fraction of a PCP's patient referrals—as is routine in more commonly used Medicare or commercial insurer data sets—introduces measurement error and attenuation bias. Accounting for this bias, we report results in a national sample of Medicare fee-for-service enrollees that are of similar magnitude to the APCD results from Massachusetts.

The consistency of our findings across both a commercially insured, working age population and a Medicare population is noteworthy. First, it suggests that insurance characteristics are not likely to be driving the observed relationship between team referral concentration and healthcare spending. Secondly, this pattern 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. As we note in our theoretical model, team referral concentration has an ambiguous effect on quality: higher team referral concentration enables greater relationship-specific investments (increasing quality), but the smaller set of specialists involved in care may reduce patient-specialist match quality. Using quality measures observable in claims data, we find no evidence that greater team referral concentration reduces quality as measured by inpatient admissions, emergency department visits, or distance traveled to specialist. In fact, we observe a statistically significant reduction in emergency department visits associated with greater team referral concentration in our working age, Massachusetts APCD sample. Care quality is, however, notoriously multi-dimensional and hard to assess across a large population of diverse patients. In light of these limitations, our interpretation is that team referral concentration reduces costs but does not negatively impact our quality measures.

The paper proceeds as follows. Section 2 describes the institutional context of teamwork among physicians. Section 3 lays out a theory of team formation and provides a theoretical motivation for our measure of team referral concentration. Section 4 introduces our empirical measure and describes the empirical approach. Section 5 describes the data. Section 6 presents results from the Massachusetts APCD. Section 7 extends our work to the Medicare sample and provides an alternative approach to

⁸ See Ericson and Starc (2015) for a more detailed description of the Massachusetts APCD.

identification. Section 8 examines the relationship between team referral concentration and utilization-based measures of healthcare quality. Section 9 concludes.

2. Institutional Context of Clinical Teamwork

Managing effective teamwork across care providers is often described as a crucial function of primary care (Bodenheimer 1999), and coordination failures are thought to be a key determinant of wasteful spending and medical errors (Berwick and Hackbarth 2012). Stille (2005) identifies a successful primary care model as one where referring occurs within a “tight web of consultants in which physicians know one another well and can share work effectively”. More recently, Press (2014) described relationships between physicians as key to delivering high quality care to complex patients. These descriptions of physician teams align with economic theory suggesting that investment in team-specific capital is important to productive team relationships (cf. Crawford 1990).

Referrals between PCPs and specialists are a central aspect of clinical teamwork. Mehrotra et al. (2011) reviewed the literature on specialist referrals and reports breakdowns and inefficiencies in all aspects of the specialty referral process. One of the core issues identified in their review involves information handoffs. At the most basic level, many referrals do not contain sufficient data for medical decision making. Poor information flows also degrade care coordination. PCPs often do not know whether a patient actually visited the referred specialist and, if so, what the recommended plan of care was.

Coordination is further complicated by ambiguity about the role that the PCP expects the specialist to take on. As Mehrotra et. al (2011) describe it, sometimes a specialist acts as a consultant to assist PCPs in diagnosis or in referrals for procedures. In other cases, a specialist becomes a co-manager of care along with the PCP—as can happen when treating a chronic condition like congestive heart failure. Other times, as in the case of a patient with end stage renal disease, the specialist becomes the principal caregiver. Each of these roles requires a different level of specialist involvement and misunderstandings about which role the specialist is to take on are commonplace. Expected division of responsibility between specialists and PCPs is not solely determined by the shared patient’s clinical condition. To be successful, a specialist needs to develop an understanding of what this specific PCP expects of them regarding the care of a particular shared patient.

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 (which can spark a costly cascade of testing, follow-up testing, and even hospitalizations); as well as duplicative testing and patient visits. These problems can be ameliorated by team-specific investments that 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 also detect unnecessary or misdirected referrals and so reduce repeat specialist visits (Mehrotra et. al. 2011 p. 56).

The medical literature describes benefits of team coordination among providers that go beyond management of chronic conditions. Even in the context of episodic specialty care such as orthopedic surgery or gynecological cancer, the prior literature has found a strong correlation between the providers' self-reported relational coordination and patient's surgical outcomes (Gittell et al 2000), described the importance of provider-to-provider communication (Zuchowski et al 2017), and shown that care teams are important for outcomes (Harrington et al 2005; Pollack et al 2012). Andreatta and Marzano (2012) write that interdisciplinary approaches to obstetric and gynecologic care are common when managing complex cases, and further identifies "human factors associated with team performance as a prime improvement area for clinical patient care."

To the extent that investments in team specific capital create positive spillovers from one shared patient to another, one would expect effective PCPs to strategically concentrate their referrals in a relatively small group of specialists. There is some evidence in support of this. In a detailed qualitative study, M. 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 very small sample and the results are not causal, the pattern is what we would expect if investments in team specific capital had important spillover effects.

3. A Model of Team Formation

In forming teams there is a tradeoff between investment in team-specific capital and specialization. In our context, when PCPs concentrate their patient referrals within a narrower group of providers within a specialty, e.g. refer their cardiology patients to a smaller set of cardiologists, this encourages repeat interactions and so incentivizes greater investments in 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.⁹

⁹ Existing evidence suggests that matching patients to appropriately subspecialized doctors within a clinical specialty may improve clinical outcomes (Epstein et al. 2010). However, to our knowledge, there is no evidence on whether PCPs effectively direct referrals to realize these gains.

We formalize this tradeoff between investment in team-specific capital and benefits of specialization in the following model of team formation. The model shows that team referral concentration is a useful indicator of team-specific capital. It also suggests factors that may cause team referral concentration to vary across physicians independent of the clinical characteristics of the PCP's patients. Both of these results will prove useful in interpreting the subsequent empirical work.

A. Model Set-Up

In our model, PCPs all have the same number of chronically ill patients, normalized to measure one. Each patient is referred to a specialist and there is only one specialty.¹⁰ The PCP chooses how many specialists N to work with and how much effort r_s she puts into developing team-specific capital with each specialist s . Then, the PCP selects a specialist for each patient referral.

The PCP's utility from a patient's referral depends on the quality and cost of care the patient receives when referred to a specialist. Thus, a PCP's utility for referring patient i to specialist s is given by

$$U_{is} = Q_s - \theta c_s + \varepsilon_{is},$$

where Q_s is a measure of the quality of care provided by the specialist s and c_s is the (overall) healthcare costs the patient will incur if they are referred to the specialist. Note that c_s includes all healthcare costs generated by both PCP and specialist actions.¹¹ The relative weight the PCP places on costs versus quality is given by $\theta > 0$. Finally, ε_{is} is the idiosyncratic match value between the patient i and specialist s . We assume that ε_{is} is distributed i.i.d. Type 1 Extreme Value.

Given these assumptions, the PCP's total expected utility from all her referrals is:

$$E[\max_{s \in N}(U_{is})] = \ln \left[\sum_{s=1}^N e^{V_{is}} \right]$$

where V_{is} is the deterministic component of a referral's utility ($V_{is} = Q_s - \theta c_s$). This is the standard log-sum formula for welfare with differentiated products (Small and Rosen 1981).¹² We see that without any effort costs of working with a specialist, the benefits of matching each patient to the specialist that is best for them creates an incentive to work with as many specialists as possible.

¹⁰ Another interesting dimension that we discuss in the empirical section is the PCP's decision of whether to refer to the specialist at all. To highlight the coordination/specialization tradeoff in team formation, our model abstracts away from that decision and assumes a fixed number of referrals.

¹¹ PCPs may in fact care differentially about costs generated by their own care versus costs generated by the specialist, but we abstract away from this issue.

¹² The assumptions underlying the log-sum formula can be restrictive. For our purposes, however, it offers a parsimonious expression of the welfare gain from increasing the number of specialists.

We now turn to the PCP's effort costs, which include a set-up cost to work with each specialist and a relationship-specific investment that improves team productivity by raising care quality and reducing spending. We assume that PCPs must pay a fixed set-up cost of effort $\varphi > 0$ the first time they refer a patient to a particular specialist. For example, this could be due to the search cost of identifying the specialist or the startup cost of establishing a communication channel with the specialist. Once paid, the PCP does not need to pay the setup cost for subsequent patients referred to that specialist.

Next, the PCP can put effort into building team-specific capital with a specialist. We assume that heterogeneity in specialist characteristics is limited to the idiosyncratic match value, ε_{is} . That is, all specialists have the same average healthcare cost c and quality Q ex ante, before PCPs make investments in team-specific capital. We make the following assumption for how team-specific capital affects healthcare cost and quality:

Assumption 1: "Team-specific capital". If a PCP invests effort r_s in developing team-specific capital with a specialist s , this benefits all of that PCP's patients who see that specialist via lower healthcare costs c and higher quality Q as follows:

$$Q_s = Q + \omega r_s$$

$$c_s = c - r_s$$

where $\omega > 0$ scales the relative effect team-specific capital has on quality relative to cost.

We also assume that the cost of effort is quadratic. Now, we can write the PCP's objective function as the sum of her utility of patient referrals and disutility of effort:

$$\max_{N, \{r_s\}_{s \in N}} \ln \left[\sum_{s=1}^N e^{V_{is}} \right] - \varphi N - \sum_{s=1}^N \left[\frac{1}{2} r_s^2 \right]$$

We also make a technical assumption to guarantee that the optimal choice of team capital investment is symmetric across specialists: $\omega + \theta < \sqrt{2}$.

B. Model Results: Team Referral Concentration and Investments in Team-Specific Capital

Solving the model outlined above, we have the following results (proofs in Appendix).

Result 1: The PCP's optimal team structure has a number of specialists N^* . The PCP invests the same amount of effort r_s^* in developing team-specific capital with each specialist and refers to each specialist with probability $\frac{1}{N^*}$.

Next, we investigate how variation in the number of specialists a PCP works with translates into cost and quality.

Result 2: Investment in team-specific capital r_s^* is inversely proportional to N^* , the number of specialists the PCP refers to. A PCP with a higher fixed cost φ of working with an additional specialist will work with fewer specialists, invest more in team-specific capital with each specialist, and have lower expected healthcare costs for their patients.

To see this, from the first order condition for the choice of effort, we find that

$$r_s^* = \frac{1}{N^*}(\omega + \theta),$$

where $\frac{1}{N^*}$ is the probability the PCP refers to that specialist. Higher fixed costs φ of adding additional specialists to the referral pool lowers the optimal number of specialists and increases r_s^* . Note also that, holding fixed the number of specialists in the team, a higher value of θ (PCP disutility from healthcare costs) raises investment in team-specific capital. Investment also increases when ω is higher—that is, when investment has a stronger effect on quality.

C. Discussion

The model analyzes how the tradeoff between specialization and team-specific capital influences team formation. PCPs who work with a smaller team of specialists have more concentrated specialist referrals, increasing team-specific capital, but their patients experience a worse idiosyncratic match with their specialists.

The model does not specify the precise source of specialist match value 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 loss of quality against the gains that are enabled by greater team-specific capital with a smaller set of specialists.

The model is also agnostic about the source of gains from team-specific capital but the medical literature suggests this likely includes improved coordination of care driven by better personal relationships between the PCP and the specialist; improved communication; or investment in coordination enhancing technology such as interoperable electronic medical records and e-referral systems.

Our model identifies factors that would cause N^* (the concentration of referrals among specialists) to vary independent of the clinical characteristics of a PCP's patient panel. Specifically it highlights the importance of the fixed cost of adding another specialist to the team, φ , as a source of variation across providers. PCPs who have worked with specialists in the past have already paid the start-up cost of

establishing a relationship. This would effectively lower φ . Variation in PCP knowledge about specialist options in the area could also lead to variation in φ , as could variation in the size and scope of local multi-specialty physician practices. The parameter φ could also respond to regional variation in the demand and supply of specialists. If, for example, a PCP is operating in an area where demand for specialists greatly exceeds supply, this may increase the cost to a PCP of establishing a relationship with a new specialist.¹³

Given our assumption that improved team-specific capital reduces costs, then N^* will also vary with θ , the weight PCPs place on reducing total care costs. This assumption is reasonable given prior research that stresses cost savings from eliminating duplicative diagnostic studies; avoiding cascades of testing and low-value therapeutic interventions; and reducing polypharmacy (the use of a large number of medications).¹⁴

Although we do not explicitly model the PCPs decision to refer a patient to any specialist at all, our model suggests that cross PCP variation in N^* likely correlates with the number of specialists seen by an *individual* patient. PCPs who put greater weight on reducing total care costs (i.e. PCPs with higher θ) will also tend to be to be judicious in their use of specialists because specialists tend to deliver expensive care. In addition, as the clinical literature suggests, it is also likely that enhanced coordination between PCPs and specialists will lead to fewer, and more appropriate, specialist referrals (Bodenheimer 1999). The correlation between team referral concentration and the number of specialists an individual patient sees is an important empirical issue that we take up in Section 4 below.

Finally, it is worth noting that in some cases the PCP may not chose team structure optimally. This can happen, for example, when organizational rules or externally imposed narrow networks limit the set of specialists to whom a PCP can practically refer. In this case, however, the logic of our model suggests that optimal investment in team-specific capital will still be inversely proportional to team size.

4. Empirical Implementation of Team Referral Concentration

In this section, we introduce an empirical measure of team-based referrals that generalizes from the preceding model. To build intuition for our team referral concentration measure, consider Figure 1. The left hand side 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

¹³ Another effect of excess demand for specialists is that congestion or patient wait times for a visit may increase. In our model, congestion simply reduces the quality benefit of making a specialist referral. We thank Jay Bhattacharya for sharing this insight regarding specialist congestion.

¹⁴ Bodenheimer (1999) outlines the cost saving mechanisms highlighted above.

repeat interactions for the two PCPs. As depicted on the right hand 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 human capital, it follows that the patients of PCP B will have superior coordination of care than the patients of PCP A – even though the care of each individual patient involves the same number of providers. In our empirical specifications we compare otherwise similar patients whose PCPs have different levels of team referral concentration and measure the relationship to healthcare costs and utilization.

Our theory motivates a measure of team referral concentration, $1/N$, using the special case where PCPs refer equally to each specialist in their team. We generalize this measure by calculating 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. A PCP has a *shared patient* with a specialist physician if that patient visits both the PCP and the specialist during our sample window (calendar year 2012). Survey evidence from Barnett et al. (2011) finds that physicians who share more patients in common are more likely to report a professional relationship, corroborating our assumption that patient sharing is a useful indicator of relationship strength.

For each PCP d , for each specialist s in each specialty j (e.g. endocrinology), we calculate the number of shared patients m_{ds} . Then, to translate these shared patients into “market shares”, we calculate the PCP’s total number of patient-specialist links for that specialty,¹⁵ or $M_{dj} = \sum_{s \in j} m_{ds}$. For each PCP-specialist pair ds , we then calculate that specialist’s share of total PCP referrals in that specialty: $share_{ds} = \frac{m_{ds}}{M_{dj}}$.

Our PCP-level measure of team referral concentration within each specialty j is then the HHI: $ReferralCon_{dj} = \sum_{s \in j} (share_{ds})^2$. When PCPs refer equally to each specialist, this measure reduces to $1/N$, the same measure of team referral concentration that emerged in our model.

Our theoretical model also considered only a single specialty. To accommodate the reality of referrals across many specialties, our empirical measure of team referral concentration averages across the various specialties weighting each specialty equally.

¹⁵ Note that this is not the number of patients seeing a specialist, but the number of patient-specialist links. We define it this way because a PCP sharing a patient with a specialist is a referral. Sharing the patient with two specialists then counts as two referrals. The market share we use is the market share of “referrals” within a given specialty.

The resulting measure of team referral concentration, $ReferralCon_d$, describes the *network* of connections among providers where connections are defined by patients shared between PCPs and specialists. Like other network measures, the closer the sample of patients in the data is to the underlying population, the more accurate will be our measure of team referral concentration. Because the sample of patients included in the Massachusetts APCD includes a large fraction 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 7.

Figure 1 demonstrates that our team-based referral measure is conceptually distinct from measures of the number of distinct providers an individual patient sees. As we have already noted in the theory section, however, these two features of care delivery are unlikely to be independent of each other. Concentrating an individual patient's visits among a smaller number of providers likely increases the continuity of care a patient receives and is a plausible channel through which teams can achieve more cost-effective team coordination. Alternatively, reducing individual patient visits to distinct providers may reduce costs entirely independently of any differences in team-specific capital. Because concentrating patient visits in a smaller number of providers can be either a mechanism for enhanced team coordination or a potential confounder, we report results with and without controls for patient care continuity.

Following prior practice in the empirical literature, we measure patient care continuity using an HHI that summarizes the concentration of a patient's visits across different providers.¹⁶ Defining n_{ip} as the number of visits during the year by patient i to each provider p (who may be a PCP or a specialist), we construct the patient care continuity HHI as $\sum_{p \in \text{physicians}} \left(\frac{n_{ip}}{N_i}\right)^2$, where $N_i = \sum_{p \in \text{physicians}} n_{ip}$ is the total number of visits by patient i to all physicians p .

5. Data on the Commercially Insured (Mass. APCD)

Our primary data comes from the 2012 Massachusetts All Payer Claims Database (APCD), version 2.1; in Section 7, we replicate and extend our findings in a national sample of Medicare beneficiaries. We create two extracts from the APCD:¹⁷ a broad sample that allows us to characterize PCP's referral patterns, and an analysis sample on which we run our regressions relating team referral concentration to total spending.

¹⁶ Prior empirical work has used either patient level Herfindahl-Hirschman indices of patient visits across providers or the closely related Bice-Boxerman continuity of care index. See Pollack et al. (2016) for a review.

¹⁷ In all cases, we limit to 12 large payers with complete claims versioning information.

A. Data on PCP team referral concentration within teams

We use a *broad sample* to construct our PCP-level measure of team referral concentration. In the broad sample, we limited claims to evaluation and management visits for patients aged 21 and older with primary health insurance available in the APCD. This includes patients enrolled in commercial health insurance, self-insured employers, 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 insurance claim, we link claims to physician specialty and demographic information in the National Plan and Provider Enumeration System data.¹⁸ We analyzed five common specialties: cardiology, dermatology, endocrinology, obstetrics and gynecology (OB/GYN), and orthopedics.¹⁹ We categorized each physician as a PCP or as one of these five specialist types; if the physician did not fall into one of these categories, they were excluded from the construction of team referral concentration. Physicians outside of these specialty categories were included in the construction of other visit and cost measures.²⁰

Team referral concentration is constructed 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 providers without a directed referral. Because patient sharing may require coordination and teamwork between providers 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 five specialties, and then averaged equally across specialties to define a single PCP-level measure of team referral concentration. Subsequent analyses also exploit within-PCP variation in referral concentration by specialty.

¹⁸ Our extract of the National Plan and Provider Enumeration System 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.

¹⁹ We chose these five specialties using a nationally representative sample (National Ambulatory Medical Care Survey); orthopedics, OB/GYN, cardiology, and dermatology visits are the most common medical specialist visits (excluding psychiatry and ophthalmology), and endocrinology was included as a common specialty used by individuals with chronic conditions. Taxonomy codes for included specialties within each category are in Appendix C.

²⁰ We conduct a sensitivity analysis to eliminate PCPs who treat very few patients in the data, who may appear to have extreme values (0 or 1) of the referral concentration measure due to sharing a small number of patients with other providers. When we limit the analysis to PCPs who treat at least 50 patients in the MA APCD, we find results that are very similar in direction, magnitude, and significance.

B. Data on patient outcomes

Our *analysis sample* limits to chronically ill patients residing in Massachusetts, aged 21-64, who are continuously enrolled with the same commercial insurer or self-insured employer for all of 2012.²¹ We focus on chronically ill patients because we expect team productivity to matter most for patients with complex conditions that often require the care of specialists. The restriction to continuous enrollment helps remove noise or confounds associated with insurance churn, and facilitates calculation of annual spending and utilization.

Following Frandsen et al. (2015), we defined chronically ill enrollees as those with at least one claim with an ICD-9 diagnosis code indicating one or more 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 and cluster headache, hypertension, hyperlipidemia, diabetes mellitus, asthma, chronic obstructive pulmonary disease, hypercoagulability disorders, osteoarthritis, and/or rheumatoid arthritis.²²

We assigned each patient to a PCP based on the “plurality primary care physician algorithm” developed by Pham et al. (2007), which assigns each enrollee to the PCP with the highest number of evaluation and management visits, with ties broken by assignment to the PCP with the highest total billed claims. We drop patients from the sample who cannot be assigned to a plurality PCP with this algorithm.²³

We capture health status by hierarchical condition category (HCC) risk scores and binary condition categories calculated using the Massachusetts “Market-wide Risk Adjustment” calculator and an individual’s claims during the year.²⁴ These HCC risk scores are calculated 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).²⁵

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

²¹ Medicaid Fee-for-Service, Medicaid Managed Care, and Medicare Advantage enrollees are included in the calculation of PCP team referral concentration, but they are not included in the analysis sample due to data limitations. Medicaid Fee-for-Service and Medicaid Managed Care patients (aged 21-64) are included in the number of patients treated by the PCP.

²² Specific ICD-9 codes for identifying these individuals are included in Appendix B.

²³ 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 physicians for whom this ZIP code was not in Massachusetts. Patients matching to these PCPs are in turn excluded from our analysis.

²⁴ In line with software instructions, we limited to inpatient hospital, outpatient facility, and professional claims (Kautter et al. 2014).

²⁵ These indicator variables, along with demographic characteristics, are used to assign individuals an HCC risk score given their plan’s metal level (based on actuarial value). We assumed all individuals to be in “gold” plans, indicating an actuarial value of 80%, without cost sharing reductions.

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 price provider or in a higher price setting (e.g., hospital vs. 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.²⁶ 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 0.131 and varies widely across PCPs, with a median value of 0.119; it has a standard deviation of 0.064; its distribution is displayed in Figure 2. Table 1 provides descriptive statistics of patients and PCPs, split by whether the PCP has above or below the median team referral concentration. PCPs with above median concentration have an average referral concentration of 0.18, compared to 0.08 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 5.6 to 12 specialists, if they referred with equal frequency to each specialist within a specialty. The distribution of disease categories, demographic composition of patients, and rates of hospitalization are quite similar across patients seeing PCPs with high versus low team referral concentration.

Figure 3 shows how various PCP characteristics are related to team referral concentration. PCPs with higher team referral concentration have fewer patients on average, and patients who are younger and with lower risk scores. Consistent with predictions of the model, high team referral concentration PCPs 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.²⁷ However, the magnitudes of these relationships are all relatively small: the largest documented effect is a one standard deviation increase in the average risk score of a PCP's patient panel which predicts only about a quarter of a standard deviation decline in team referral concentration.

Note also that Table 1 does not show large differences in patient care continuity HHI between PCPs above and below the median team referral concentration, with patients treated by high team referral

²⁶ 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.

²⁷ 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 ACOs. 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).

concentration PCPs also having a slightly higher 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 providers seen by a patient by preventing redundant or low-value referrals. Alternatively, PCPs with high team referral concentration may be more careful about referral decisions generally, and have both a higher threshold for whether they refer a patient to a specialist at all and thus have a more concentrated pattern of referrals.

Table 2 examines differences in team referral concentration by organizational affiliation, and shows that there is substantial variation across different physician contracting networks.²⁸ Team referral concentration also varies substantially across different PCPs *within* a physician contracting network. Also of note, the largest, highest price hospital system in Massachusetts (Partners Community Health Care, see e.g. Seltz et al. 2016) has an average team referral concentration of 0.11, near the average team referral concentration in our analysis sample.

Figure 4 displays a binned scatterplot of team referral concentration and the average of log utilization, along with a best-fit linear regression line. While higher levels of team referral concentration are associated with lower average spending throughout the distribution, the negative relationship is strongest in the lower part of the team referral concentration distribution. This would be consistent with very uncoordinated care being more expensive. An alternative interpretation is that PCPs with the sickest (i.e. most costly) patients refer to many different specialists within a specialty (e.g., two cardiologists with different sub-specialty expertise) due to clinical need. While the uncontrolled, cross-sectional comparison shown in Figure 4 could be contaminated by omitted variable bias, it is worth noting that Table 1 and Figure 3 do not show large differences in patient characteristics for PCPs above and below the median. In the next section, we use a regression approach to isolate variation in referral concentration from patient care continuity HHI and other possible confounders.

²⁸ For this analysis, we link the physician NPIs to the 2010 Massachusetts Provider Database (MPD) maintained by Massachusetts Health Quality Partners. The MPD has information on the organizations and physician contracting networks to which the PCP belongs. Physician Contracting Network is defined by the Massachusetts Provider Database 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.” See Massachusetts Health Quality Partners (2016).

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:

$$\log y_i = \alpha \text{ReferralCon}_{-i} + \beta X_i + \gamma Z_i + \varepsilon_i$$

where y_i is patient i 's spending²⁹, X_i is a set of patient characteristics and 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 severe health status necessitates more unusual 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, 5-knot splines for both age and HCC risk score, and patient ZIP code fixed effects to capture local heterogeneity in patient demand for care. Insurer controls include a fixed effect for each payer and an indicator 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.). We then augment this baseline specification with a series of additional controls for patient and physician characteristics. Given the inclusion of these rich controls, the baseline specification is identified by variation in PCP referral concentration among patients who reside in the same ZIP code, and have similar health status, demographics, insurer, and insurance type.

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 fixed effects, exploiting differences in PCP referral concentration across different specialties. For example, if a PCP is more concentrated in her cardiology referrals than her endocrinology

²⁹ We exclude any individuals with zero or negative net spending which can result from, for instance, reversed claims. All individuals in our sample must have had healthcare spending during the year in order to pass our chronic illness sample screen and to be assigned a PCP.

referrals, her cardiology patients should have relatively higher costs, compared to a peer physician with the inverse pattern. We describe this approach in more detail and report results in Section 6D below.

A second major threat to causal interpretation in the baseline specifications is the possibility that patients seeing low team referral concentration physicians are in worse health. While observed patient characteristics such as age and HCC risk score do not suggest obvious differences in health status across low and high team referral concentration PCPs, there could be differences in health that are not observed by the econometrician. To assess this possibility, we will analyze the experience of Medicare beneficiaries who change PCPs due to a move. The mover design allows us to control for patient fixed effects and exploit a plausibly exogenous change in patient referral concentration using an instrumental variable strategy. We describe the mover specifications in more detail and report results in Section 7C.

Together, these three 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. Column 7 parallels column 1, but is run on the 20% sample of Medicare beneficiaries. We discuss its interpretation in the next section, noting here that the Medicare magnitude is not directly comparable to the Massachusetts APCD results due to measurement error in the Medicare measure of team referral concentration.

Columns 1 and 4 report results with only the baseline controls for patient and insurer characteristics described above. The findings confirm the strong relationship between within team referral concentration and spending visually seen in Figure 4.

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 6.3% ($= -0.626 * 0.1$) lower medical care utilization and 11.6% ($= -1.16 * 0.1$) lower total spending on average, when compared to similar patients seen by PCPs with below median team referral concentration. Alternatively, a 1 standard deviation increase in team referral concentration leads to 4.0% lower utilization and 7.4% lower spending.

These results indicate that patients of PCPs with higher team referral concentration use fewer services and also see lower priced providers. While we are not able to identify the precise mechanisms causing the negative association between team referral concentration and provider pricing, it is helpful to consider possible explanations. 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. Instead, the pricing effect suggests that PCPs with higher team referral concentration tend to either have lower prices themselves or send patients to lower-priced specialists and hospitals.

Our model suggests that PCPs who put greater weight on containing the total costs of care will be inclined to both concentrate their referrals and refer to lower priced providers. However, there are other possible explanations for the correlation between high referral concentration and lower prices. For example, physicians participating in narrow insurance networks may concentrate their referrals and also have lower prices due to reduced bargaining power. For these reasons, we do not interpret this finding as evidence of a causal relationship between referral concentration and the price of care. Rather, we include spending as an auxiliary outcome because these findings are suggestive of other differences between more and less concentrated providers. In Sections 6C, 6D and 6E below we introduce provider fixed effects and physician network fixed effects to allay concerns that differences across doctors or firms are driving the utilization results reported in Table 3.

As a robustness check, we augment the regression with a control for patient care continuity HHI; results are in columns (2) and (5). The relationship between team referral concentration and spending 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 3.7% lower utilization and 9.0% lower total spending than patients seen by PCPs with average below median team referral concentration.

If variation in patient care continuity HHI simply captures exogenous and unobserved patient heterogeneity, the results in columns (2) and (5) ought to be preferred over those in columns (1) and (4). But enhanced care continuity (i.e. concentrating patient visits among a smaller set of providers) might be in part a result of team referral concentration. This might happen if a PCP's investments in improving coordination with one specialist reduces the need to refer the patient to a different type of specialist (within the same specialty) to collect additional information, or improves clarity and agreement on which patients do not require referrals. In this case, introducing a variable for patient care continuity HHI likely results in excessively conservative estimates of the relationship between team referral concentration and utilization.

Finally, in columns (3) and (6), we add new controls for other dimensions of PCP heterogeneity, including the average HCC risk score of a PCP's commercially insured working age patients, a 5-knot spline in the number of working-age patients the PCP treats, an indicator for whether the PCP's specialty is in Internal Medicine or Family Medicine, and fixed effects for the 5-digit PCP ZIP Code. The additional controls in this specification do not substantively change the magnitude of our results compared to the prior specification, 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 rely on variation in team referral concentration within a practice location, rather than between different practice locations.³⁰

C. Within-PCP 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 PCP practice style. To address this concern, we perform an additional analysis that includes PCP fixed effects and exploits differences in team referral concentration across specialties. The PCP will have stronger incentives to invest in team-specific capital for the specialists with whom the PCP shares patients more frequently. For example, if a PCP is highly concentrated in her cardiology referrals but not in her endocrinology referrals, then our model predicts the PCP will invest more in team-specific capital with cardiologists (relative to endocrinologists) given the higher volume of shared patients.

This approach to identifying the effects of team-specific capital is conservative, as it is identified only across specialties within PCP and so removes from consideration any efforts a PCP may have made in improving coordination with all specialists to whom she refers. The relationship between team referral concentration and spending may also be understated if team referral concentration in any one specialty effects a PCP's "bandwidth" for forming relationships with physicians in different specialties. Finally, estimates may be attenuated to the extent that within-PCP variance is driven by measurement error.

To estimate spending and utilization based on within PCP variation in team referral concentration, we restrict the sample to patients who saw at least one specialist in exactly one of our 5 specialty categories. (Unlike the base sample, this subsample excludes patients who saw no specialists or those who consulted more than one type of specialist.) Instead of using a PCP's average team referral concentration across all the specialties in our set, a patient is assigned the (jackknifed) team referral concentration of their PCP for the specialty in which they saw a specialist. The key independent variable of interest, team referral concentration, is now matched to the specific specialty the patient consulted.

³⁰ Results are similar when we do not include a fixed effect for PCP ZIP code.

Table 4 displays the regression results. Columns (1) and (4) include similar controls to those used in Table 3 column (1) to estimate the relationship between team referral concentration and utilization/spending in this specific subsample. The controls include patient characteristics (but not patient care continuity HHI) and insurer variables, but exclude PCP characteristics. Table 4 also introduces a new set of controls in all reported specifications: a set of indicator variables for the specialty consulted (e.g. cardiology, endocrinology, etc.). These controls capture baseline differences in patient utilization that may depend on the type of specialty the patient requires.

In Table 4, columns (2) and (5), we add PCP fixed effects, exploiting within-PCP variation in team referral concentration across specialties. The results remain negative and statistically significant, but the magnitude is about one-half (utilization) to one-third (spending) the size of the effect reported in columns (1) and (4).³¹ This attenuation is not surprising because, as discussed above, within PCP estimates remove important sources of variation in team referral concentration and a larger component of the remaining variation may be due to measurement error. The persistence of a negative, statistically and economically significant effect after controlling for PCP fixed effects is suggestive that our main findings in Table 3 are not entirely the result of some unobserved, fixed PCP characteristic that is correlated with team referral concentration. These results suggest that unobserved differences in PCP quality or practice style (uniform across conditions the PCP treats) do not drive the estimated effect of team referral concentration.

D. Within Specialist Variation in Team Referral Concentration

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 “better” and therefore more expensive specialists. Alternatively, perhaps PCPs with less concentrated referrals have higher costs because they are still learning to identify the low cost specialists and will continue to experiment until they identify the best possible concentrated set of specialists. We address these concerns by estimating our set of regressions using fixed effects for the specific specialist the patient sees.

These regressions echo the sample used in the PCP fixed effect analysis. In particular, we restrict the sample to patients referred to at least one specialist in exactly one specialty. Similar to our plurality rule for PCP assignment, we then assign each patient to their plurality specialist within the specialty consulted. The regressions then include a fixed effect for the identity of the patient’s plurality specialist; the independent variable of interest is the PCP’s team referral concentration for the relevant specialty. These

³¹ To interpret the magnitude of the PCP fixed effect specification, compare a specialty in which a PCP refers equally to 5 versus 10 specialists (a difference in team referral concentration of 0.10); this increase in team referral concentration is associated with 1.4% lower care utilization.

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.

Results of these regressions with specialist fixed effects are in Table 4, columns (3) and (6). We continue to find that PCPs with higher team referral concentration have significantly lower levels of utilization, even when their patients are referred to identical specialists. The magnitude suggests that moving from a below to above median team referral concentration PCP (a change of 0.1) is associated with a 1.9% 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 (compare to column 1). These findings provide further support for the notion that team relationships between PCPs and specialists promote lower cost care.

E. Within Firm Variation in Team Referral Concentration

Team-specific capital and intensity of care utilization may be jointly mediated by firm 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 results thus far are primarily 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's contracting network. For the PCPs who are not reported to be part of a contracting network, we include a separate fixed effect for not belonging to a network. Unlike the specifications with physician fixed effects which restrict to the potentially endogenous set of patients who receive specialty referrals, we estimate this coefficient on the full sample of chronically ill patients.

Results of this analysis are reported in Appendix Table A1. Even after accounting for variation in utilization that is related to network affiliations, we find a negative and statistically significant relationship between team referral concentration and care utilization. The magnitude of the coefficient, -0.132, is similar to (and not statistically distinguishable from) the effect size reported above in the physician fixed effect specifications, suggesting that the attenuation of these results relative to the less controlled specifications reported in Table 3 may be driven by their absorption of firm-level effects. To the extent that firms shape referrals and reduce frictions associated with team coordination, we might expect that contracting network fixed effects may reflect part of the effect of team specific capital on utilization outcomes. From this perspective, this specification is a conservative estimate relative to the less controlled specifications in Table 3.

In a further test, we restrict the sample to only PCPs that are part of the Partners HealthCare, the largest contracting network in Massachusetts which includes the prestigious academic medical centers of Massachusetts General Hospital and Brigham and Women's Hospital. Each of the 909 PCPs in this network

is vertically integrated with many specialists of every variety, and thus has tremendous scope to select a level of referral concentration within their integrated network. Even within Partners HealthCare, we find that PCPs with greater referral concentration have lower levels of care utilization. The coefficient is statistically distinguishable from zero and slightly larger than the magnitude reported for all organizations pooled. Taken together, these findings underscore that even within a single organization, we see a diverse set of team structures. While firm-level investments may ease coordination frictions, our empirical findings suggest they do not fully substitute for repeated team interactions.

7. Utilization-based quality metrics

Our results have focused on the effect of team referral concentration on the cost of health care. Our model predicts that increased team-specific capital will lower costs, but that the effect on care quality is ambiguous. While quality could increase due to increased coordination effort, quality might also decrease through a reduction in patient-specialist match quality. While quality of care is multi-dimensional and difficult to measure, we examine in Table 5 the relationship between team referral concentration and a set of utilization-based quality indicators for the Massachusetts APCD Sample.

We find no evidence that higher referral concentration PCPs deliver lower quality care, and some suggestive evidence that quality may increase with team referral concentration. Our most controlled specification finds that higher team referral concentration has a negative effect on an emergency department visits, with a 0.1 unit increase in team referral concentration leading to a 5% (or 1.2 percentage point) drop in the probability of an emergency department visit from a mean of 21.5%. While our point estimates also suggest higher team referral concentration is associated with fewer inpatient visits, the estimate is imprecise: the 95% confidence interval cannot reject a zero effect, but also includes a 7% (or 0.5 percentage point) reduction in hospitalizations for a 0.1 increase in team referral concentration, from a base hospitalization rate of 7%. Taken together, these results suggest that more concentrated teams may reduce utilization in part by reducing emergency department visits and hospitalizations. We have more acute power limitations when studying ambulatory care sensitive hospitalizations and duplicate imaging, and find no statistically significant effects on either outcome.

Turning to the patient-specialist match quality, we do not have measures of clinical match quality, but we can examine the distance patients travel to see specialists. To measure this, we use the straight-line distance between the patient ZIP code and the ZIP code of the patient's plurality specialist.³² We find patients of PCPs with higher team referral concentration travel shorter distances to see their plurality

³² Patients who see their plurality specialist in their home ZIP code travel a distance of zero. Distances between ZIP codes of greater than 500 miles are considered missing. The sample for this regression is those who see at least one specialist; this is different from the regression sample used in Table 4 as the distance measures in Table 5 are not limited to patients who only saw one type of specialty.

specialist; these results condition on fixed effects for the patient’s ZIP code of residence. While this statistically significant effect is small in magnitude—a 0.1 unit increase in leads to about 4/10ths of a mile shorter distance—it indicates that team referral concentration doesn’t come at the expense of patient travel time.³³

8. Team Referral Concentration in Medicare

Our results so far have focused on the Massachusetts All Payer Claims Data. This data offers remarkable breadth for measuring referral networks precisely at the physician level but is also limited in two important ways. First, the APCD is limited to a single state, whose healthcare institutions may not be nationally representative. Secondly, our extract of the APCD data is essentially a cross-sectional data set; even adding more recent years would make for a very short panel. As a result, we cannot use the Massachusetts data to estimate a model with patient fixed effects. We address both these deficiencies by analyzing team referral concentration in a national sample of Medicare beneficiaries.

A. Measurement Error in Medicare

A natural way to begin our analysis of Medicare beneficiaries would be to replicate the analysis we ran using the APCD data. Unfortunately, the Medicare data is only a 20% sample of Medicare fee for service enrollees over the age of 65. Thus we only observe a small fraction of the total patients each doctor sees and this creates acute measurement error problems for any network measures including our team referral concentration variable. In this section, we demonstrate the empirical impact of measurement error on the baseline specifications. In Section 7C, we discuss the theoretical derivation of the measurement error in our instrumental variable estimates.

Consider a PCP who has 5 patients, each referred to a different specialist within a single specialty. The PCP should have a team referral concentration of 0.2. However, if we only observe 1 out of the 5 patients in the data, we will measure a team referral concentration of 1 for that PCP. Similarly, a PCP who refers all her patients to the same specialist will also have a referral concentration of 1. As this example illustrates, measurement error in this setting differs from classical measurement error in two ways. First, the noise biases our estimate of team referral concentration upwards and second the size of the bias is correlated with the underlying true referral concentration.

To illustrate the impact of measurement error on our results, we run a series of simulations in the Massachusetts APCD. In these simulations, we draw subsamples of patients from the APCD and then

³³ We don’t calculate travel time to specialists for the Medicare movers because they are moving geographical regions, making the interpretation of a change in this variable more complex.

estimate team referral concentration from just these patient subsamples. For example, we compare results from the full 100% Massachusetts APCD sample to results from a randomly drawn 20% Massachusetts APCD sample.

We draw a series of subsamples of patients in the APCD, using percentage samples that range from $\# = \{10\%, 20\%, \dots, 100\%\}$ of the full sample of patients. For each percentage subsample, we repeat 50 random draws to account for sampling error. Within each subsample, we construct $ReferralCon_i^\#$ and estimate a regression of the form:³⁴

$$\log y_i = \alpha_\# ReferralCon_i^\# + \beta X_i + \gamma Z_i + \varepsilon_i$$

We then calculate the multiplier $\lambda_\# = \alpha_{100}/\alpha_\#$ that tells us how to scale estimated coefficients on team referral concentration for each $\#$ % subsample. The scaling factor $\lambda_\#$ will depend on the set of controls used.

In Figure 5, we plot the results of how the multiplier $\lambda_\#$ depends on the size of the subsample used to measure team referral concentration for regression specification from Table 3, Column 1. Figure 5 illustrates that the magnitude of the attenuation bias falls with the size of the patient sample. Further, the multipliers for specification 1 are modest: if we observed 20% of the APCD sample, we would want to multiply our estimate by about 2 to gauge impact of team referral concentration as measured from the 100% sample.

However, adding additional controls to the regression greatly exacerbates the measurement error problem. The control variables are correlated with the “signal” in our estimated referral concentration – most especially our control for patient care continuity HHI – and so the remaining variation in referral concentration has a proportionally larger “noise” component. Appendix Figure A1 shows that attenuation bias is much more severe for regression specification 2, which adds a control for patient care continuity HHI. For a 20% sample, specification 2 has a multiplier of about 8, with a 95% CI from 4 to 16. At a 10% sample, the problem is severe: the estimated mean multiplier is 48 with a 95% CI that includes -49 to 225, implying we may not even estimate the correct direction of the effect.

The very large distortions at the 10% APCD sample for specification 2 may be relevant for a 20% Medicare sample because even a 100% sample of Medicare fee-for-service beneficiaries likely represents a smaller share of the doctors’ patient panels than the APCD would include. We also find large attenuation bias in

³⁴ Precisely, we estimate the regression using the value of $ReferralCon_i^\#$ for each doctor constructed from the subsample of patients. We then run the regression using all the patients, but with the noisily measured $ReferralCon_i^\#$. Using the full sample of patients in the regression should not affect the expected multiplier, but it does reduce sampling variation in measuring the multiplier.

specification 3 (which adds PCP controls in addition to the patient care continuity HHI), though not as severe as specification 2.

On the basis of these simulations, we conclude that specifications 2 and 3 are uninformative in the Medicare data, but that specification 1 may shed useful light on the generalizability of our Massachusetts results. In recognition of these uncertainties, we report unadjusted coefficients, and then assess how imposing the APCD scaling factor to the national 20% Medicare sample would influence the interpretation of these results.

B. Replicating Results in Medicare

Summary statistics on the Medicare sample are reported in Appendix Table A2. Similar to the Massachusetts findings, summary statistics demonstrate that patient age, sex, and disease burden are similar across patients seeing PCPs with above and below median team referral concentration.

Table 3, column 7 shows the regression results for specification 1 in Medicare.³⁵ Note that we only estimate utilization equations, not spending. This is because prices are administratively set in Medicare and primarily adjusted only for geographic location. The geographic variation will be largely eliminated by the patient ZIP code fixed effects we include as controls.

The estimated effect is quite substantial. An increase in measured team referral concentration of 0.1 is associated with a 2.9% decline in utilization. This analogous coefficient estimated in the parallel specification on Massachusetts data (Table 3, column 1) was about twice as large, consistent with the magnitude of attenuation predicted in our measurement error simulations.

An advantage of the Medicare sample is it allows a more fine-grained analysis of charges. Specifically Medicare bills are easily decomposed into three categories: provider submitted claims from the Carrier files, Inpatient claims submitted by hospitals, and Outpatient claims for hospital-based outpatient care. Results from this decomposition exercise are reported in Appendix Table A4. Higher team referral concentration is associated with statistically significant reductions in all three types of billings in the least noisy specification 1 (without controls for patient care continuity HHI or PCP characteristics). Patients of PCPs with higher referral concentration are slightly less likely to have an inpatient stay.

³⁵ Appendix Table A3 reports regression results for specifications 2 and 3 as well, though our measurement error exercise suggests these specifications are not reliable. We find statistically significant negative coefficients on team referral concentration in both specifications 2 and 3, though the magnitude of the effect is small in specification 3.

C. Medicare movers and reversion to the mean identification strategy

We have documented a positive 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 relied on cross-sectional regressions and given the limitations of claims data there was still scope for selection on unobservable patient characteristics to bias the findings.

To address this possibility, 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 builds on work by Laird and Nielsen (2016); Agha, Frandsen, and Rebitzer (2017); and Finkelstein, Gentzkow, and Williams (2016). The technique allows the inclusion of a patient fixed effect to control for differences in patient demand for care that are stable over time. Moreover, we also use an instrument for the change in referral concentration to address the possibility that the choice of new PCP is endogenous to a change in health status.

For the movers analysis, we expand the Medicare sample to include data from 2007-2012.³⁶ We calculate PCP team referral concentration on an annual basis using the full 20% Medicare sample (not restricted to physicians who treat movers). We restrict the analysis sample to enrollees who move to a new hospital referral region over this period.³⁷

The first way to measure the response of utilization to change in team referral concentration is with a difference-in-difference strategy: we add a patient fixed effect, β_i , to our estimation equation, and include $\Delta ReferralCon_{i,d}$, which measures the change in the jackknifed referral concentration of the patient's post-move PCP compared to the patient's pre-move PCP.³⁸ We interact $\Delta ReferralCon_{i,d}$ with $Post_{it}$, an indicator variable for being in the post-move period. We include a vector of fixed effects $\rho_{R_{i,t}}$ for the event year relative to the move (denoted $R_{i,t}$), allowing movers' annual demand for care to depend 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., and we include indicator variables for each year in event time.) Further, we include

³⁶ We limit this movers analysis to the Medicare sample, since the Massachusetts APCD is not ideal for a number of reasons: there are not many regions in Massachusetts, our panel is short, and the APCD data do not let us accurately track a patient who changes insurers or employers.

³⁷ 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).

³⁸ Note that patients are assigned to PCPs using our plurality assignment rule 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 referral concentration over the year(s) prior to the move. Similarly, post-move team referral concentration is calculated as the average level of PCP 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.

characteristics of the patient's plurality PCP and year fixed effects in the control vector Z_{it} . We then get the following difference-in-differences equation:

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

Unlike the earlier specifications, we omit controls for patient comorbidities because the patient fixed effect should account for fixed differences in patient health over time. Patient comorbidities can change over time, and in principle could be tracked using this data. We do not include such controls in this regression 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.

Results from the difference-in-differences specification are in Panel A of Table 6. The role of measurement error in this specification is derived in Appendix E. The baseline specification estimates an increase in team referral concentration of 0.1 is associated with a 6.1% decrease in care utilization, significant at the 1% level. However, the difference-in-differences framework faces an identification threat: $\Delta ReferralCon_{-i}$ may be endogenous to real changes in patients' clinical conditions. If movers with deteriorating health selected new PCPs (after moving) who relied on more diverse specialists, our estimate of parameter α would be biased.

Analogous to the approach of Laird and Nielsen (2016), we apply an instrumental variables strategy that exploits mean reversion in PCP team referral concentration to identify exogenous variation in PCP team referral concentration. The patient's PCP team referral concentration in the pre-move period is used as an instrumental variable for the *change* in the patient's PCP team referral concentration, $\Delta ReferralCon_{-i}$.

The first stage equation proceeds as follows:

$$\Delta ReferralCon_{-i} Post_{it} = \tilde{\alpha} PreMoveReferralCon_{-i} Post_{it} + \tilde{\beta}_i + \tilde{\gamma} Z_{it} + \tilde{\rho}_{R_{i,t}} + \tilde{\varepsilon}_{it}$$

And the reduced form is given as:

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

These regressions measure the causal effect of team referral concentration on utilization under two key assumptions. First, the patient's initial PCP team referral concentration must be uncorrelated with future *changes* in the patient's demand for care after the move. For example, if poor underlying health status is correlated with the initial PCP's low level of referral concentration, and health status mean reverts, then

this identification approach will overstate the relationship between team referral concentration and spending. Conversely, if patients who initially sorted to PCPs with low team referral concentration are on deteriorating health trends relative to other movers, then this instrumental variables approach will understate the relationship between team referral concentration and spending. Second, as with the baseline set of regression results, the PCP's referral concentration must be independent of other dimensions of PCP practice style that might influence care utilization.

Note that the instrumental variable approach may still suffer bias from measurement error due to the correlation in the error of the instrument, $PreMoveReferralCon_{-i}Post_{it}$, and the error in the endogenous variable $\Delta ReferralCon_{-i}Post_{it}$.

Table 6 reports first stage (Panel B) and two-stage least squares (Panel C) results of these instrumental variable Medicare utilization regressions on the movers sample. The specification in column 1 includes baseline controls for patient fixed effects, year fixed effects, and a series of indicators for event year relative to the move. In column 2, we add in a time varying control for the patient care continuity HHI. Finally, in column 3, we add additional controls for the PCP's specialty, sex, and patient volume.

The first stage estimates confirm the predicted mean reversion pattern. Patients treated by PCPs with high team referral concentration prior to their move experience relative reductions in PCP team referral concentration after their move, compared to patients initially treated by PCPs with low team referral concentration.

The two stage least squares estimates of the impact of team referral concentration on utilization are consistently larger than the cross-sectional estimates in Medicare data. This pattern is consistent with a reduced degree of attenuation bias due to measurement error of PCP referral concentration using the instrumental variables strategy; it may also reflect a larger true effect size due to the movers being an older sample of patients with a higher prevalence of heart conditions than the 2012 Medicare 20% sample (Appendix Table A2). In Table 6, Panel C, column 1, an increase in team referral concentration of 0.1 is associated with a 5.8% reduction in utilization. Adding control variables for patient care continuity HHI and PCP characteristics reduces the estimated effect size only slightly, so that a 0.1 increase in team referral concentration is associated with a 4.7% reduction in utilization. These findings provide further evidence that unobserved patient characteristics associated with the PCP's team referral concentration are not driving our main results.³⁹

³⁹ We also examine the impact of team referral concentration on utilization-based quality metrics (as in Table 5) using this strategy, but we are underpowered to detect even large effects in this sample (see Appendix Table A5).

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 primary care physicians (PCPs) and specialists have to work closely together 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. Our theoretical model highlights how concentrating referrals among a smaller team of specialists facilitates beneficial investments in team-specific capital. These concentrated team referrals, however, come at the cost of potentially lower match value between patients and specialists. Empirically, we find that patients of PCPs who concentrate their referrals among a smaller set of specialists have lower healthcare utilization, but we find no evidence of a decline in quality along measured dimensions. 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 4% lower utilization compared to those treated by above median PCPs, after controlling for detailed patient and insurer 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 specialist fixed effects that rely on variation in team referral concentration across different referring PCPs. These specifications are, however, conservative since they are likely to absorb some organization-level and doctor-level investments in team-specific capital into the fixed effects.

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 small fraction of the PCP's total patient panel. In the Medicare sample, we also study those who switch doctors as the result of a move across regions. Using both a difference-in-difference analysis with patient fixed effects and a regression to the mean instrumental variables strategy, we find that an increase in team referral concentration is associated with lowered utilization.

Our analysis has a number of limitations that may inspire future research. While our theory and our empirics suggest that concentrated referrals facilitate relationship-specific investments that improve team performance, we do not directly observe these investments directly. More work on the channels through which concentrated team referrals lead to reduced utilization would therefore be valuable. Another issue is whether the gains from concentrated team referrals are different within or between

organizations.⁴⁰ 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.⁴¹ Another important limitation is that 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 is 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. Narrow insurance networks, for example, may promote team referral concentration by limiting the set of in-network specialists, although care would be required not to disrupt existing investments in PCP-specialist relationships. Certain cost containment incentives, such as those found in Accountable Care Organizations, patient centered medical homes, or physician payment adjustments by the Medicare Merit-based Incentive Payment System may potentially induce PCPs to form more concentrated referral networks.

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 enhance 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 the match value from drawing from a larger pool of specialized talent. Understanding the determinants of this tradeoff and its consequences for economic performance in other environments is a potentially important area of study.

⁴⁰ Baker et al. (2016) find that hospital ownership of physician practices leads physicians to increase patient admissions to the owning hospital, possibly increasing concentration of admissions. More generally, Atalay et al. (2017) show the importance of firm boundaries for structuring transactions.

⁴¹ See Baker et al. (2014) for estimates of the relationship between vertical integration and pricing.

10. References

- 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, Leila, Brigham Frandsen and James B. Rebitzer. 2017. "Causes and Consequences of Fragmented Care Delivery: Theory, Evidence, and Public Policy." *NBER Working Paper 23078*.
- Alonso, Ricardo, Wouter Dessein, and Niko Matouschek. 2008. "When Does Coordination Require Centralization?" *American Economic Review* 98 (1): 145–79. doi:10.1257/aer.98.1.145.
- Andreatta, Pamela and David Marzano. "Healthcare management strategies: interdisciplinary team factors." *Current Opinion in Obstetrics and Gynecology*, 24(6): 445-452.
- Atalay, E., Hortaçsu, A., Li, M.J. and Syverson, C., 2017. "How wide is the firm border?" National Bureau of Economic Research Working Paper 23777. doi:10.3386/w23777.
- Baghai, Ramin and Silva, Rui and Ye, Luofu, Teams and Bankruptcy. 2018. SSRN Working Paper. <http://dx.doi.org/10.2139/ssrn.3042304>
- Baker, Laurence C., M. Kate Bundorf, and Daniel P. Kessler. 2014. "Vertical Integration: Hospital Ownership of Physician Practices Is Associated With Higher Prices And Spending." *Health Affairs* 33(5): 756-763.
- Baker, Laurence C., M. Kate Bundorf, and Daniel P. Kessler. 2016. "The effect of hospital/physician integration on hospital choice." *Journal of Labor Economics* 50: 1-8.
- Barnett, Michael L., Bruce E. Landon, A. James O'Malley, Nancy L. Keating, Nicholas A. Christakis. 2011. "Mapping Physician Networks with Self-Reported and Administrative Data." *Health Services Research*, 46(5): 1592-1609. doi:[10.1111/j.1475-6773.2011.01262.x](https://doi.org/10.1111/j.1475-6773.2011.01262.x)
- Barnett, Michael L., Nicholas A. Christakis, A. James O'Malley, Jukka-Pekka Onnela, Nancy L. Keating, and Bruce E. Landon. 2012a. "Physician Patient-Sharing Networks and the Cost and Intensity of Care in US Hospitals." *Medical Care* 50 (2):152–60. <https://doi.org/10.1097/MLR.0b013e31822dcef7>.
- Barnett, Michael L., Nancy L. Keating, Nicholas A. Christakis, A. James O'Malley, and Bruce E. Landon. 2012a. "Reasons for Choice of Referral Physician Among Primary Care and Specialist Physicians." *Journal of General Internal Medicine* 27 (5):506–12. <https://doi.org/10.1007/s11606-011-1861-z>.
- Barton H. Hamilton, Jack A. Nickerson, and Hideo Owan. 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.
- Baumgardner, James R. 1988. "Physicians' Services and the Division of Labor Across Local Markets." *Journal of Political Economy* 96 (5): 948–82. doi:10.2307/1837242.
- Becker, Gary S. and Kevin M. Murphy. 1992. "The Division of Labor, Coordination Costs, and Knowledge" *Quarterly Journal of Economics*, 107(4): 1137-1160.
- Berwick, Donald M., and Andrew D. Hackbarth. 2012. "Eliminating Waste in US Health Care." *JAMA* 307(14):151-4-1516. doi:10.1001/jama.2012.362
- Bloom, Nicholas, and John Van Reenen. 2011. "Chapter 19 - Human Resource Management and Productivity." In *Handbook of Labor Economics*, edited by David Card and Orley Ashenfelter, 4:1697–1767. Elsevier. <http://www.sciencedirect.com/science/article/pii/S0169721811024178>.

- Bodenheimer, Thomas, Bernard Lo, Lawrence Casalino. 1999. "Primary Care Physicians Should Be Coordinators, Not Gatekeepers." *Journal of the American Medical Association*. 281(21): 2045–2049. doi:10.1001/jama.281.21.2045
- Brot-Goldberg, Zarek, and Mathijs de Vaan. 2018. "Intermediation and Vertical Integration in the Market for Surgeons." Working paper.
- Cebul, Randall D., James B. Rebitzer, Lowell J. Taylor, and Mark E. Votruba. 2008. "Organizational Fragmentation and Care Quality in the U.S. Healthcare System" *Journal of Economic Perspectives*, 22(4): 93-114.
- Chernew, Michael E., Robert E. Mechanic, Bruce E. Landon, and Dana Gelb Safran. 2011. "Private-payer innovation in Massachusetts: the 'alternative quality contract'." *Health Affairs*. 30(1): 51-61.
- Chillemi, Ottorino, and Benedetto Gui. "Team Human Capital and Worker Mobility." *Journal of Labor Economics* 15, no. 4 (1997): 567-85. doi:10.1086/209838.
- Crawford, Vincent P. 1990. "Relationship-Specific Investment." *Quarterly Journal of Economics*. 105(2): 561-574.
- Cutler, David, Jonathan Skinner, Ariel Dora Stern, and David Wennberg. 2018. "Physician Beliefs and Patient Preferences: A New Look at Regional Variation in Health Care Spending." NBER Working Paper 19320. <https://doi.org/10.3386/w19320>.
- Epstein, Andrew J. , Jonathan Ketcham, and Sean Nicholson. 2010. "Specialization and matching in professional services firms." *RAND Journal of Economics*, 41(4): 811-834.
- Ericson, Keith and Amanda Starc. 2015. " "Measuring Consumer Valuation of Limited Provider Networks." *American Economic Review*, 105(5): 115-19.
- Feri, Francesco, Bernd Irlenbusch, and Matthias Sutter. 2010. "Efficiency Gains from Team-Based Coordination—Large-Scale Experimental Evidence." *American Economic Review*, 100(4): 1892-1912.
- Finkelstein, Amy, Matthew Gentzkow, and Heidi Williams. 2016. "Sources of Geographic Variation in Health Care: Evidence From Patient Migration." *Quarterly Journal of Economics*, 131 (4): 1681–1726. doi:10.1093/qje/qjw023.
- Finkelstein, Amy, Matthew Gentzkow, Peter Hull, and Heidi Williams. 2017. "Adjusting Risk Adjustment – Accounting for Variation in Diagnostic Intensity" *New England Journal of Medicine*, 376 (6): 608-610.
- Frandsen, Brigham R., Karen E. Joynt, James B. Rebitzer, and Ashish K. Jha. 2015. "Care Fragmentation, Quality, and Costs among Chronically Ill Patients" *American Journal of Managed Care*, 21(5): 355-326.
- Frandsen, Brigham and James B. Rebitzer. 2014. "Structuring Incentives within Accountable Care Organizations" *Journal of Law, Economics, and Organization*, 31(S1): i77-i103.
- Garicano, Luis, and Thomas N. Hubbard. 2009. "Specialization, Firms, and Markets: The Division of Labor within and between Law Firms." *Journal of Law, Economics, and Organization*, 25 (2): 339–71. doi:10.1093/jleo/ewn003.
- Garicano, Luis, and Tano Santos. 2004. "Referrals." *American Economic Review*, 94 (3): 499-525.

- Geissler, Kimberley, Ben Lubin, and Keith M Ericson. (2018). "The role of organizational affiliations in physician patient-sharing relationships." *Medical Care Research and Review*. Published ahead of print, 20 April 2018; p. 1-18.
- Gittel, Jody Hoffer, Kathleen M. Fairfield, Benjamin Bierbaum, William Head, Robert Jackson, Michael Kelly, Richard Laskin, Stephen Lipson, John Siliski, Thomas Thornhill, and Joseph Zuckerman. "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, no. 8 (2000): 807-19.
- Grosse, Stefan, Louis Putterman, and Bettina Rockenbach. 2011. "Monitoring in Teams: Using Laboratory Experiments to Study a Theory of the Firm." *Journal of the European Economic Association* 9 (4):785–816. <https://doi.org/10.1111/j.1542-4774.2011.01026.x>.
- Harrington, J. T., Barash, H. L., Day, S. and Lease, J. (2005), Redesigning the care of fragility fracture patients to improve osteoporosis management: A health care improvement project. *Arthritis & Rheumatism*, 53: 198-204.
- Hart, Oliver. 2017. "Incomplete Contracts and Control." *American Economic Review*, 107(7): 1731-52.
- Holmström, Bengt. 1982. "Moral hazard in teams." *Bell Journal of Economics*, 13(2): 324-340.
- Huckman, Robert S., Bradley R. Staats, David M. Upton. 2009. "Team Familiarity, Role Experience, and Performance: Evidence from Indian Software Services." *Management Science* 55(1): 85-100.
- Hussey, Peter S., Eric C. Schneider, Robert S. Rudin, Steven Fox, Julie Lai, and Craig E. Pollack. 2014. "Continuity and the Costs of Care for Chronic Disease" *JAMA Internal Medicine*, 174(5): 742-748.
- Jackson, Matthew O., Brian W. Rogers, and Yves Zenou. 2017. "The Economic Consequences of Social-Network Structure." *Journal of Economic Literature* 55 (1):49–95. <https://doi.org/10.1257/jel.20150694>.
- Jacobson, L., LaLonde, R., & Sullivan, D. (1993). Earnings Losses of Displaced Workers. *American Economic Review*, 83(4), 685-709.
- Jaravel, Xavier, Neviana Petkova, and Alex Bell. 2018. "Team-Specific Capital and Innovation." *The American Economic Review* 108(4-5):1034-73.
- Kautter, John et al. 2014. "The HHS-HCC Risk Adjustment Model for Individual and Small Group Markets under the Affordable Care Act." *Medicare & Medicaid Research & Review*. DOI: <http://dx.doi.org/10.5600/mmrr.004.03.a03>.
- Kolber, Morey J. 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. <http://dx.doi.org/10.1111/j.1529-1006.2006.00030.x>.
- Laird, Jessica and Torben Nielsen. 2016. "The Effects of Physician Prescribing Behaviors on Prescription Drug Use and Labor Supply: Evidence from Movers in Denmark." Harvard University Working Paper.
- Lammers, Eric J., Julia Adler-Milstein, and Keith Kocher. 2014. "Does Health Information Exchange Reduce Redundant Imaging? Evidence From Emergency Departments." *Medical Care*, 52 (3): 227-234.

- Mailath, George J., Andrew Postlewaite. 1990. "Workers Versus Firms: Bargaining Over a Firm's Value." *The Review of Economic Studies*, 57 (3):369-380.
- Marschak, J and R. Radner. 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, J. Michael. 2016. "Cost Containment and the Tale of Care Coordination." *New England Journal of Medicine*, 375 (23): 2218–20. doi:10.1056/NEJMp1610821.
- McWilliams, J. Michael, Michael E. Chernew, and Bruce E. Landon. 2017. "Medicare ACO Program Savings Not Tied To Preventable Hospitalizations Or Concentrated Among High-Risk Patients." *Health Affairs*, 36 (12): 2085-2093. doi.org/10.1377/hlthaff.2017.0814.
- Mehrotra, Ateev, Christopher B. Forrest, and Caroline Y. Lin. 2011. "Dropping the Baton: Specialty Referrals in the United States" *The Milbank Quarterly*, 89(1): 39-68.
- Meltzer, David. 2001. "Hospitalists and the Doctor Patient Relationships" *Journal of Legal Studies*, 30: 589.
- Meltzer, David O. and Jeanette W. Chung. 2010. "Coordination, Switching Costs and the Division of Labor in General Medicine: An Economic Explanation for the Emergence of Hospitalists in the United States" NBER Working Paper #16040.
- Milstein, Arnold and Elizabeth Gibertson. 2009. "American Medical Home Runs" *Health Affairs*, 28(5): 1317-1326.
- National Ambulatory Medical Care Survey. 2014 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. PMID: 19221375.
- Pollack CE, Weissman G, Bekelman J, Liao K, Armstrong K. Physician social networks and variation in prostate cancer treatment in three cities. *Health Serv Res*. 2012;47(1 Pt 2):380–403.
- Pollack, Craig E., Peter S. Hussey, Robert S. Rudin, Steven D. Fox, Julie Lai, and Eric C. Schneider. 2016. "Measuring Care Continuity: A Comparison of Claims-Based Methods" *Medical Care*. 54(4):e30-4.
- Press, Matthew J. (2014) "Instant Replay—A Quarterback's View of Care Coordination." *New England Journal of Medicine*. 371: 489-491.
- Reagans, Ray, Linda Argote, and Daria Brooks. 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):869–81. <https://doi.org/10.1287/mnsc.1050.0366>.
- Rebitzer, James B. and Mark E. Votruba. 2011. "Organizational Economics and Physician Practices." NBER Working Paper #17535.
- Romano, M. J., J. B. Segal, and C. E. Pollack. 2015. "The Association between Continuity of Care and the Overuse of Medical Procedures." *JAMA Internal Medicine*. 175(7): 1148-54.
- Samer Faraj, Lee Sproull. 2000. "Coordinating Expertise in Software Development Teams." *Management Science* 46(12):1554-1568. <https://doi.org/10.1287/mnsc.46.12.1554.12072>

- Seltz, David, David Auerbach, Kate Mills, Marian Wrobel, and Aaron Pervin. (2016) "Addressing Price Variation in Massachusetts." *Health Affairs Blog*. Online: <http://healthaffairs.org/blog/2016/05/12/addressing-price-variation-in-massachusetts/>
- Simon, M., N. K. Choudhry, J. Frankfort, D. Margolius, J. Murphy, L. Paita, T. Wang, and A. Milstein. 2017. "Exploring Attributes of High-Value Primary Care" *Annals of family medicine*, 15(6): 529-534.
- Small, K. A. and Rosen, H. S. (1981) "Applied Welfare Economics with Discrete Choice Models." *Econometrica*. 49, pp. 105-130.
- Song, Yunjie, Jonathan Skinner, Julie Bynum, Jason Sutherland, John E. Wennberg, and Elliott S. Fisher. 2010. "Regional Variations in Diagnostic Practices". *New England Journal of Medicine*, 363:45-53. DOI: 10.1056/NEJMsa0910881
- Song, Z., et al. 2014. "Changes in health care spending and quality 4 years into global payment." *New England Journal of Medicine*, 371(18): 1704-1714.
- Stille, Christopher J., Anthony Jerant, Douglas Bell, David Meltzer, Joann G. Elmore. 2005. "Coordinating Care across Diseases, Settings, and Clinicians: A Key Role for the Generalist in Practice." *Annals of Internal Medicine*, 142:700–708. doi: 10.7326/0003-4819-142-8-200504190-00038
- Topel, Robert. 1991. "Specific Capital, Mobility, and Wages: Wages Rise with Job Seniority." *Journal of Political Economy* 99 (1): 145–76.
- Weber, Roberto A. 2006. "Managing Growth to Achieve Efficient Coordination in Large Groups." *American Economic Review*, 96(1): 114–26
- Williamson, O. 1985. *The Economic Institutions of Capitalism*, New York: Free Press.
- Zuchowski, Jessica L., Joya G. Chrystal, Alison B. Hamilton, Elizabeth W. Patton, Laurie C. Zephyrin, Elizabeth M. Yano, and Kristina M. Cordasco. "Coordinating care across health care systems for veterans with gynecologic malignancies: A qualitative analysis." *Medical Care* 55 (2017): S53-S60.

Figures and Tables

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

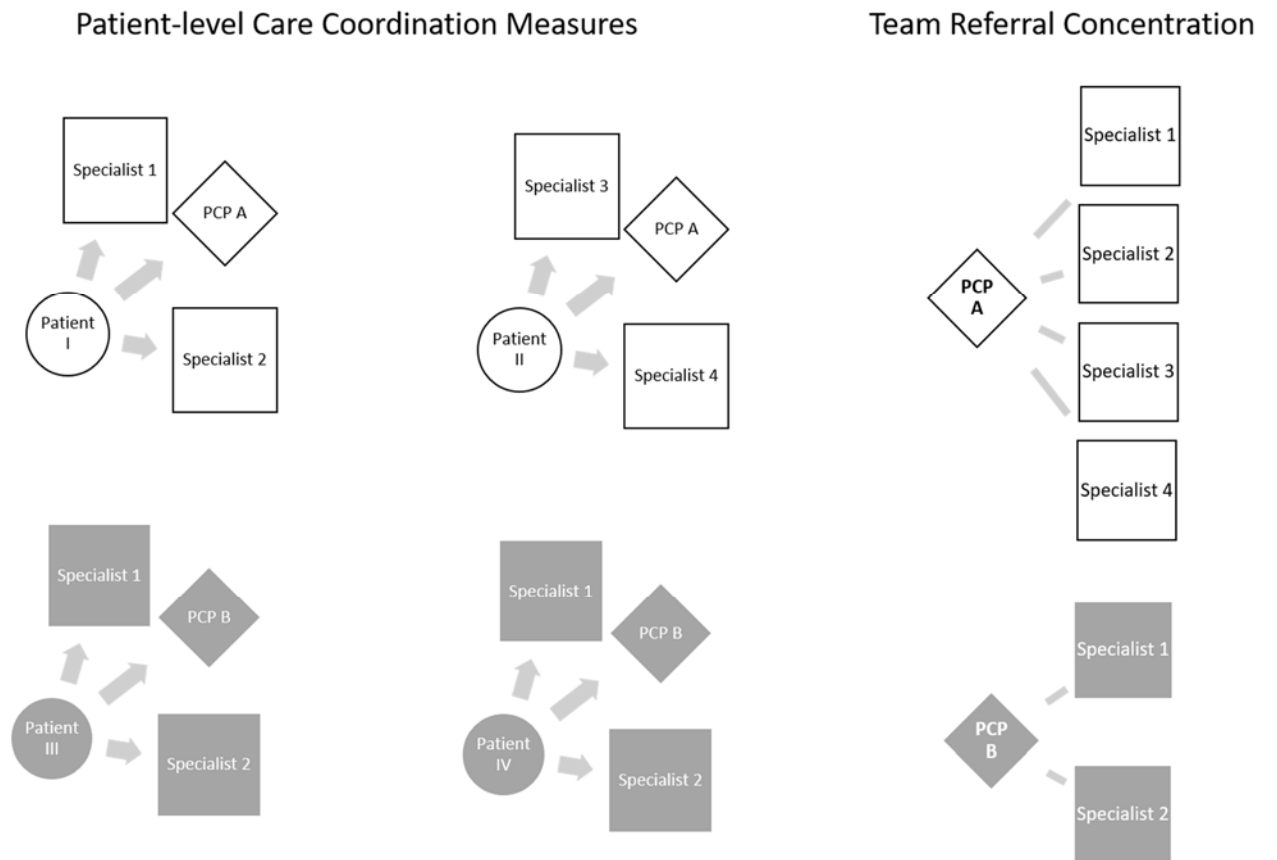
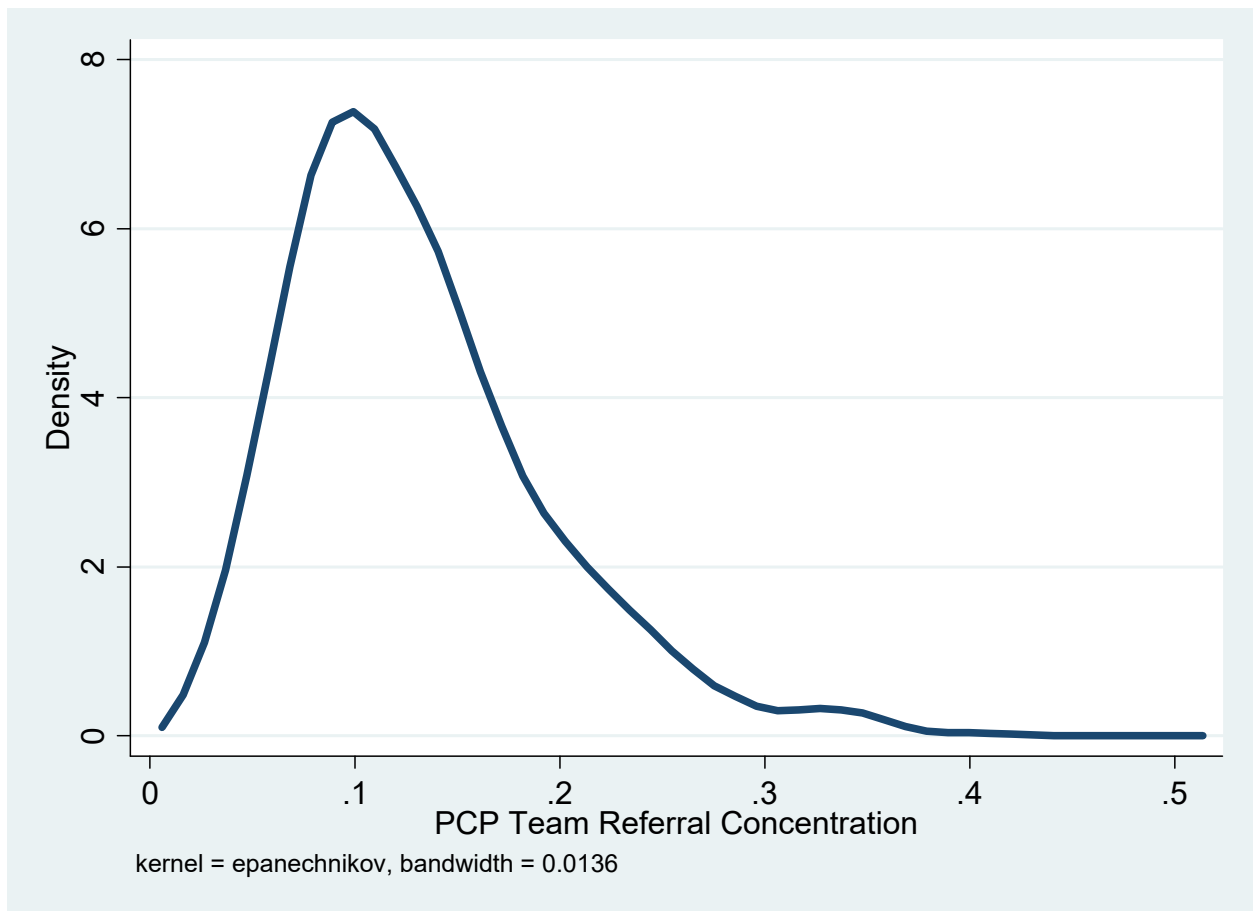
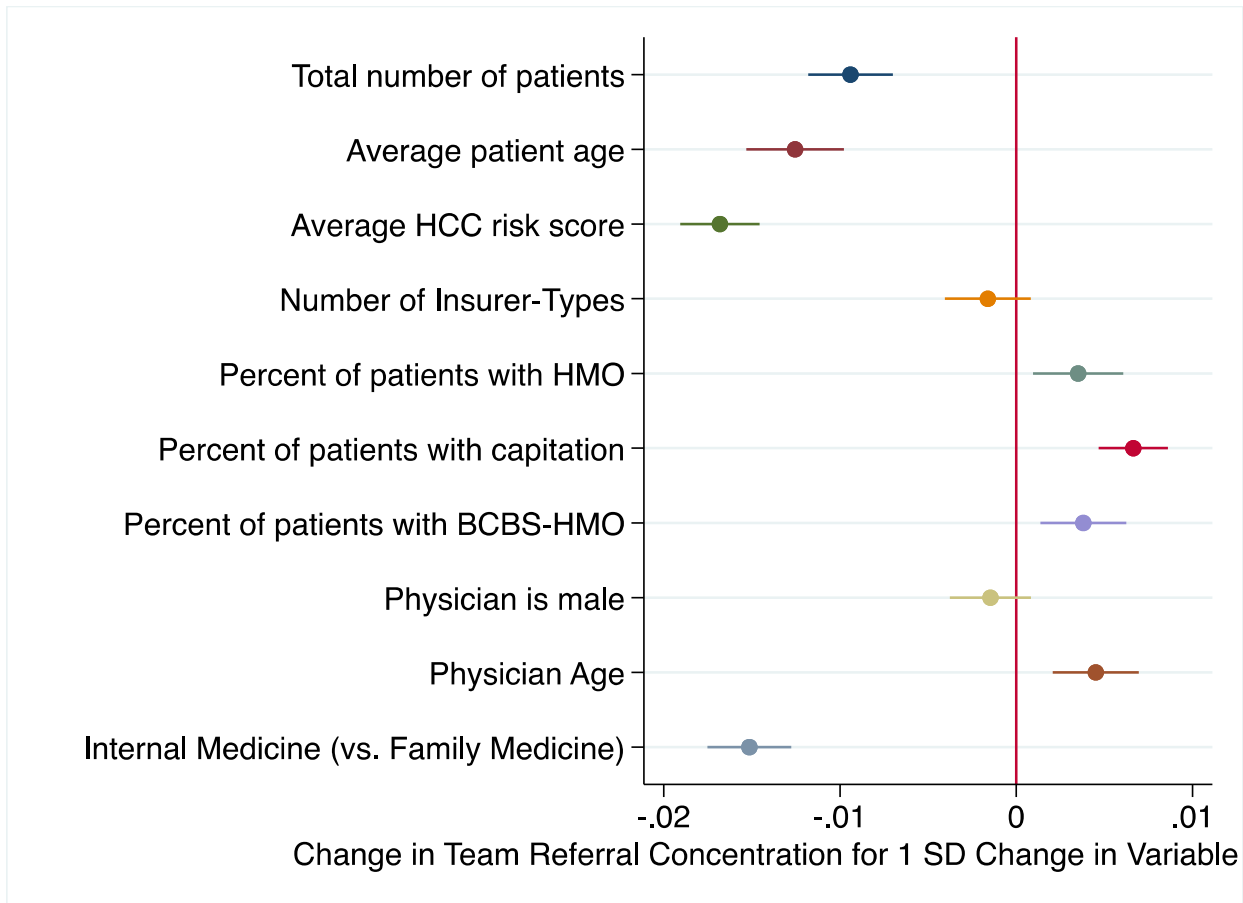


Figure 2. Distribution of PCP Team Referral Concentration.



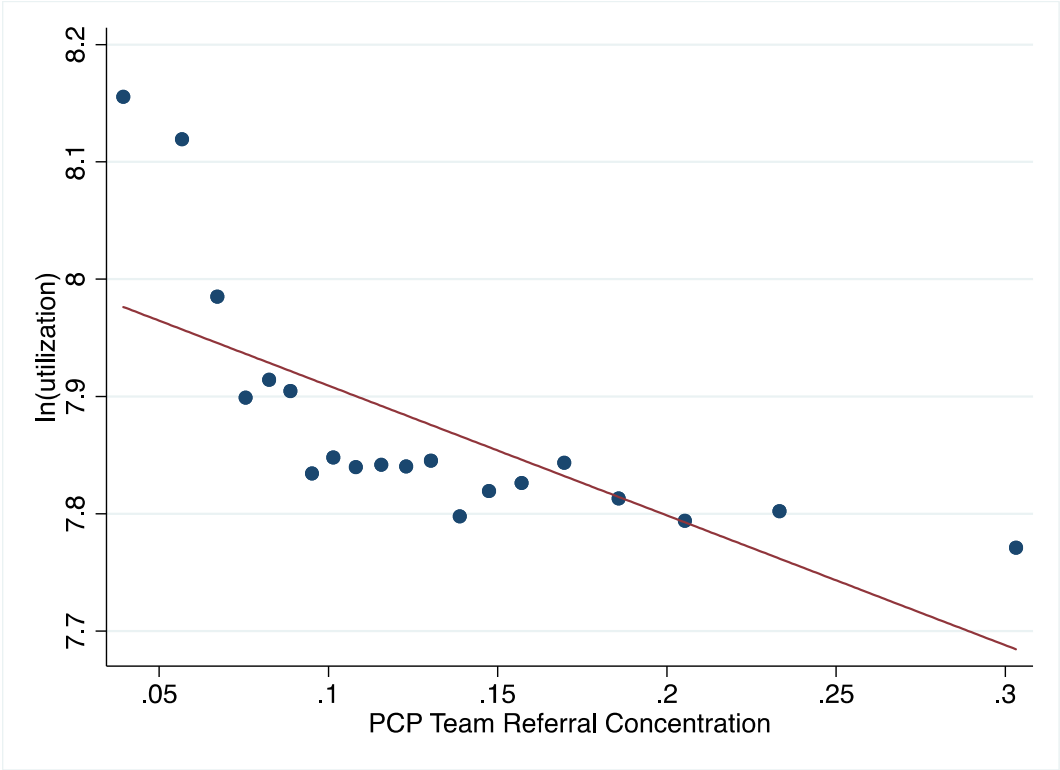
Notes: Data from Massachusetts All Payer Claims Data chronic illness analysis sample. One observation per patient.

Figure 3: Bivariate Associations of Team Referral Concentration with PCP Characteristics



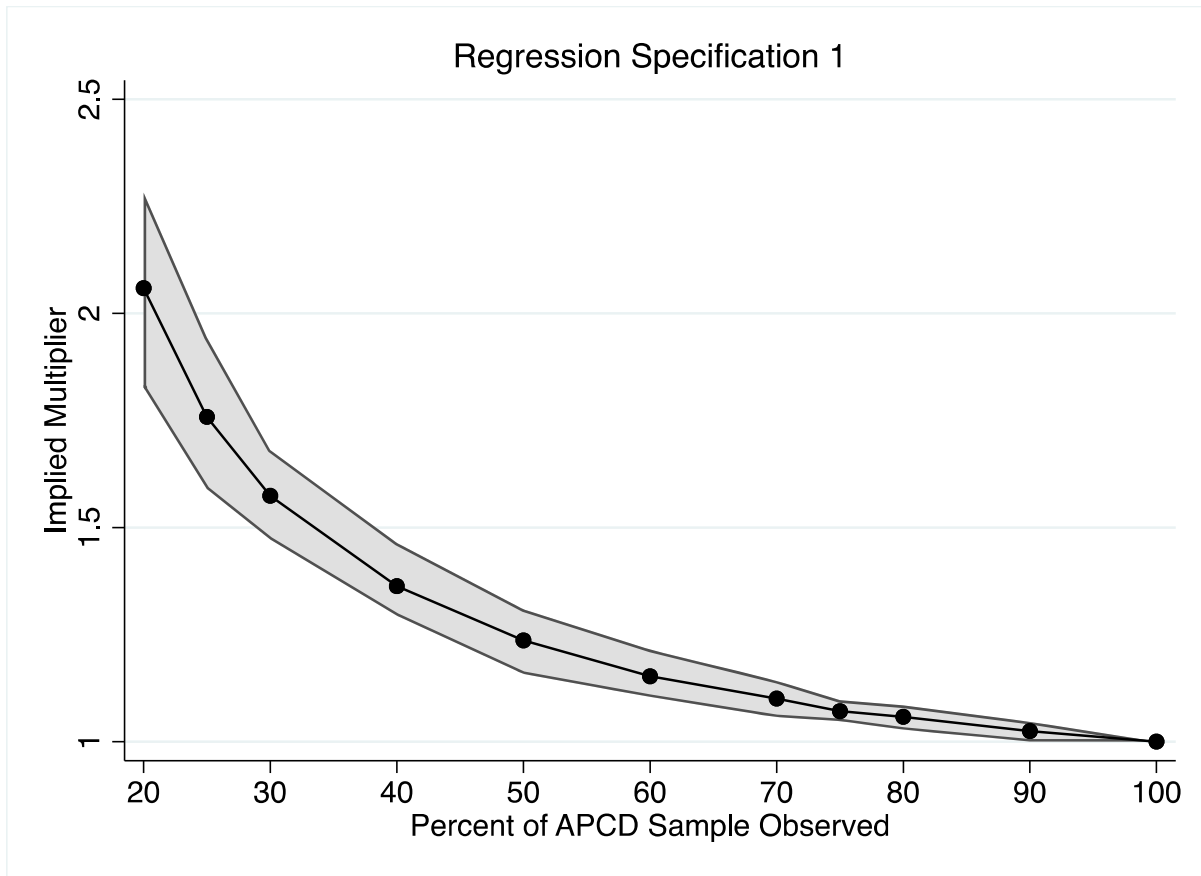
Note: Data from Massachusetts All Payer Claims Database. Limited to PCPs treating at least 50 patients, one observation per PCP. Figure plots coefficients and 95% confidence intervals from regressions of team referral concentration on each covariate separately, normalized to show the change in team referral concentration associated with 1 standard deviation change in each variable. Total number of patients, number of insurer-types, percent of patients with HMO, percent of patients with capitation, and percent of patients with BCBS-HMO are calculated for the broad sample. Average HCC risk score and average age are calculated for continuously eligible.

Figure 4: Binned Scatterplot of PCP Team Referral Concentration and Patient Utilization



Notes: Sample: MA APCD chronic illness 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. Best fit linear line plotted.

Figure 5. Measurement Error Multiplier Simulations Using Subsamples of APCD Data



Note: Plots mean multiplier. Shaded area shows 5th, and 95th percentiles bootstrapped from 50 random samples per percent subsample. Follows the regression specification for Column 1 of Table 3.

Table 1. Descriptive Statistics of the Massachusetts APCD Analysis Sample

	PCP is Below Median Team Referral Concentration	PCP is Above Median Team Referral Concentration
<u>Patient characteristics:</u>		
Mean Spending (\$)	8155	6407
Median Spending (\$)	2952	2450
Age	48.9	49.1
Male	0.49	0.49
Pr(Any Inpatient Admission)	0.08	0.06
Patient care continuity HHI	0.44	0.47
Pr(Diabetes)	0.11	0.12
Pr(Heart Condition)	0.14	0.12
Pr(Bipolar and Major Depressive)	0.07	0.07
Pr(Asthma)	0.10	0.10
<u>PCP characteristics:</u>		
PCP's Team Referral Concentration	0.08	0.18
Pr(Internal Medicine)	0.75	0.71
Fraction Capitated Encounters	0.04	0.08
Fraction HMO Patients	0.60	0.62
PCP is Male	0.61	0.60
N Patients	157,360	157,318

Notes: Data from Massachusetts All Payer Claims Data, chronically ill analysis sample. Conditions 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.

Table 2. Average Team Referral Concentration by Physician Contracting Network, MA APCD

	Average Team Referral Concentration	Std. Dev. Of Team Referral Concentration	N PCPs
<i>Physician Contracting Network</i>			
Beth Israel Deaconess P.O.	0.098	0.065	296
Lahey Clinic	0.103	0.072	91
UMass Memorial Health Care	0.110	0.071	233
Partners Community Health Care	0.111	0.069	909
New England Quality Care Alliance	0.124	0.059	226
Baycare Health Partners	0.130	0.070	136
Southcoast Physicians Network	0.134	0.038	53
Atrius Health	0.136	0.052	253
Caritas Christi Network Service	0.151	0.075	203
Fallon Clinic	0.189	0.064	94
No Physician Contracting Network	0.169	0.104	1974

Notes: Physician Contracting Network information obtained from the 2010 Massachusetts Provider Database. Unit of Observation is a PCP.

Table 3. Referral Concentration and Healthcare Utilization and Spending

Dependent Variable	Massachusetts All Payer Claims Data						Medicare 20% Sample
	(1) ln(utilize)	(2) ln(utilize)	(3) ln(utilize)	(4) ln(spend)	(5) ln(spend)	(6) ln(spend)	(7) ln(utilize)
PCP Team Referral Concentration	-0.626*** (0.056)	-0.369*** (0.055)	-0.420*** (0.0692)	-1.160*** (0.069)	-0.900*** (0.069)	-0.816*** (0.0858)	-0.286*** (0.011)
Patient Care Continuity HHI	No	Yes	Yes	No	Yes	Yes	No
PCP controls	No	No	Yes	No	No	Yes	No
Patient controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Insurer controls	Yes	Yes	Yes	Yes	Yes	Yes	NA
N patients	314,678	314,678	314,678	314,678	314,678	314,678	1,848,071

Notes: PCP's team referral concentration is jackknifed (a patient's own visits are excluded from the calculation for their PCP). Standard errors are clustered at the PCP level. Dependent variables are ln(utilization) or ln(spending). Utilization is measured using standardized prices as described in the text. Insurer controls are not applicable for Medicare; for patients in Massachusetts they are a fixed effect for each payer and a fixed effect for each of 13 types of insurance plan (i.e. HMO, PPO, EPO, indemnity, etc.). Patient controls are patient ZIP code fixed effects, sex (male/female), age (included as a 5-knot spline), and comorbidity controls. The patient comorbidity control in the Massachusetts data is the HCC risk score included as a 5-knot spline; in the Medicare data, it is a vector of comorbidity fixed effects for each of the 27 conditions recorded in the Chronic Condition Warehouse. Massachusetts PCP controls include the average HCC risk score of the PCP's commercial patients (included as a 5-knot spline), total number of patients (included as 5-knot spline), indicator for PCP sex, an indicator for whether the PCP's specialty is Internal Medicine or Family Medicine, and PCP modal ZIP Code fixed effects. Medicare PCP controls include PCP specialty (i.e., family medicine or internal medicine), gender, a 5-knot spline in number of patients, and the PCP-level mean value of each of the 27 Chronic Condition Warehouse comorbidities. Data: Columns 1-6, 2012 APCD analysis sample: commercially insured Massachusetts residents with chronic illness. Data for Column 7: 2012 Medicare beneficiaries in 20% sample. See discussion of measurement error in text for interpretation.

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

	Dependent variable: ln(utilization)			Dependent variable: ln(spending)		
	(1)	(2)	(3)	(4)	(5)	(6)
PCP Team Referral Concentration in relevant specialty	-0.297*** (0.032)	-0.143*** (0.038)	-0.192*** (0.040)	-0.403*** (0.036)	-0.115*** (0.039)	-0.268*** (0.041)
PCP fixed effect	No	Yes	No	No	Yes	No
Specialist fixed effect	No	No	Yes	No	No	Yes
Patient and insurer controls	Yes	Yes	Yes	Yes	Yes	Yes

Note: PCP's team referral concentration is. Standard errors are clustered at the patient level. Sample: N = 108,442. Sample is restricted to patients who saw at least one specialist in exactly one specialty: cardiology, orthopedics, endocrinology, dermatology, or OB/GYN. All specifications include fixed effects for the specialty consulted by the patient. Patient and insurer controls are the same as in Table 3, column 1. Column 1 specification replicates Table 3, column 1 on the restricted sample with the specialty-specific measure of referral concentration. Column 2 includes PCP fixed effects. Column 3 includes specialist fixed effects.

Table 5. Utilization-Based Quality Outcomes in Massachusetts APCD Sample (OLS Results)

	(1)	(2)	(3)	N of obs.	Dep. var. mean
Dependent variable:					
Any emergency department visit	-0.076*** (0.017)	-0.016 (0.018)	-0.120*** (0.025)	314,678	0.215
Any inpatient visit (excluding pregnancy)	-0.071*** (0.0106)	-0.063*** (0.014)	-0.020 (0.014)	314,678	0.070
Any hospitalization for ambulatory care sensitive condition	-0.005 (0.003)	-0.005 (0.003)	0.005 (0.005)	314,678	0.008
Any duplicate imaging	-0.007 (0.007)	-0.003 (0.007)	0.014 (0.009)	314,678	0.040
Miles from patient ZIP to plurality specialist ZIP	-15.94*** (1.10)	-15.70*** (1.10)	-3.863*** (1.02)	157,168	11.046
Patient controls	Yes	Yes	Yes		
Patient care continuity HHI	No	Yes	Yes		
PCP controls	No	No	Yes		

Note: PCP's team referral concentration is jackknifed. 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. Our measure is adapted from Lammers et al. (2014), and includes 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, 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. Standard errors are clustered at the PCP level for Panel A. Columns 1-3 mirror specifications from Table 3 Columns 1-3.

Table 6. Difference-in-Differences and Instrumental Variables Results with Patient Movers in Medicare

	(1)	(2)	(3)
<i>A. Difference-in-Differences</i>			
Dependent variable: ln(utilization)			
(Δ Team Referral Concentration)*Post	-0.613*** (0.176)	-0.434*** (0.136)	-0.448*** (0.131)
<i>B. First Stage</i>			
Dependent variable: (Δ Team Referral Concentration)*Post			
Pre-move Referral Concentration*Post	-0.859*** (0.041)	-0.858*** (0.041)	-0.851*** (0.043)
F statistic of excluded instrument	441	438	393
<i>C. Two Stage Least Squares</i>			
Dependent variable: ln(utilization)			
(Δ Team Referral Concentration)*Post	-0.584** (0.237)	-0.445** (0.186)	-0.468** (0.180)
Patient fixed effect	Yes	Yes	Yes
Patient care continuity HHI	No	Yes	Yes
PCP controls	No	No	Yes
Number of individual patients	1639	1639	1639
Number of observations (patient X year)	6230	6230	6230

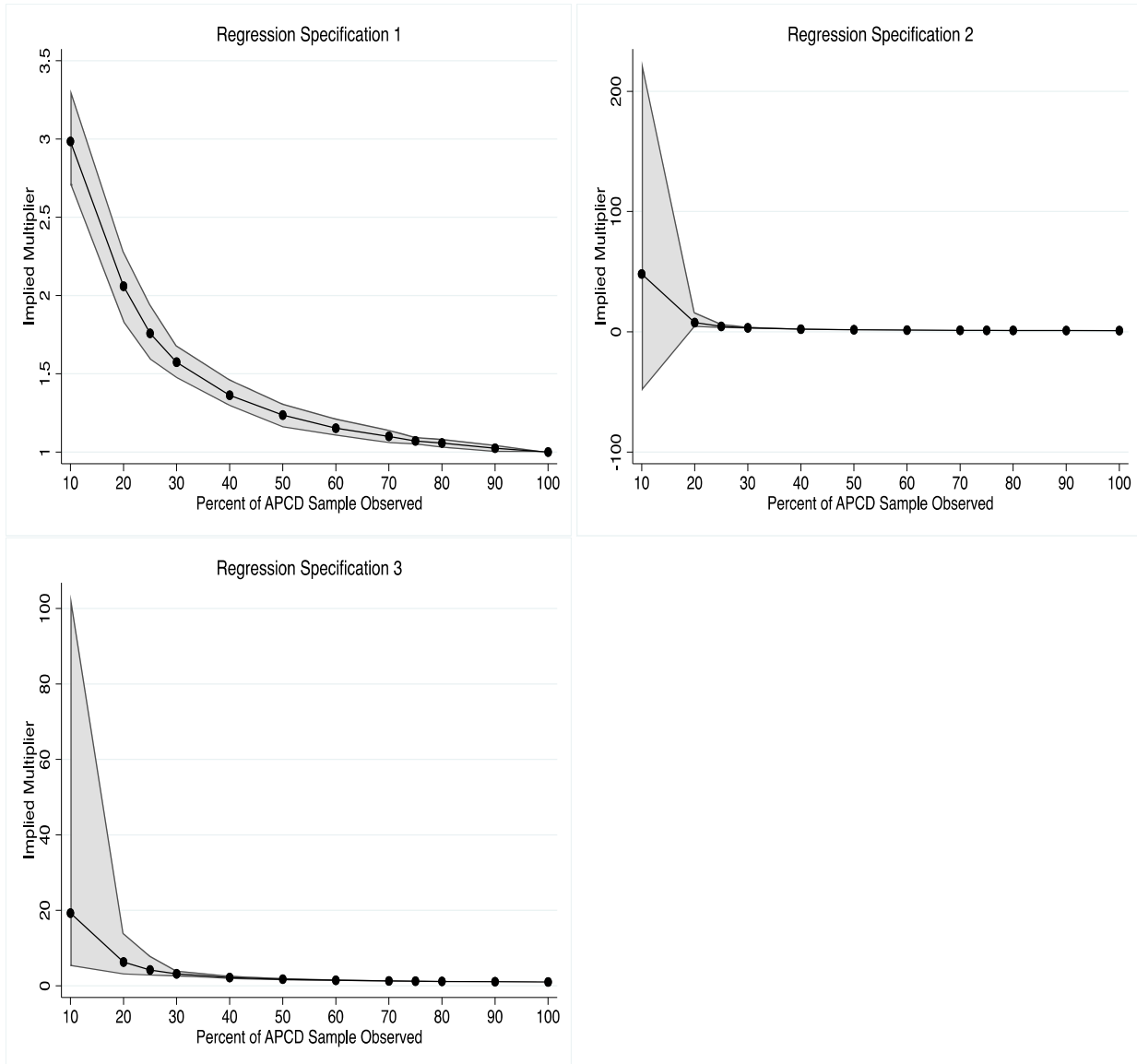
Note: PCP's team referral concentration is jackknifed. Standard errors are clustered at the patient level. Data: patients in Medicare 20% sample who move across regions during 2007-2012. In panels B and C, the PCP's team referral concentration in the origin region is used as an instrumental variable for the change in team referral concentration experienced after the move, exploiting mean reversion in PCP team referral concentration at the time of a switch in PCP. All regressions include an individual patient fixed effect, year fixed effect, and a series of indicators for event year relative to the move. The regression reported in column (2) adds a time varying measure of the patient's care continuity HHI. The regression reported in column (3) adds time varying PCP controls: specialty (internal medicine vs. family medicine), gender, and 5-knot spline of PCP patient volume. Recall that these specifications, particularly columns 2 and 3, are likely to suffer from substantial attenuation bias due to measurement error.

Online Appendix

For Agha et al., "Team Formation and Performance: Evidence from Healthcare Referral Networks"

Appendix A: Additional Empirical Results

Figure A1. Measurement Error Multiplier Simulations Using Subsamples of APCD Data



Note: Plots mean multiplier, 5th, and 95th percentiles bootstrapped from 50 random samples per percent subsample. Follows the regression specifications for Column 1, 2, and 3 respectively of Table 3.

Table A1. PCP Team Referral Concentration and Spending and Utilization, With Physician Contracting Network Fixed Effects

	All Physician Contracting Networks	Partners Healthcare Only
	Ln(Utilization)	Ln(Utilization)
	(1)	(2)
PCP Team Referral Concentration	-0.132** (0.060)	-0.233** (0.114)
Patient controls	Yes	Yes
Patient care continuity HHI	Yes	Yes
PCP controls	Yes	Yes
N patients	314,678	67,815

Note: PCP's team referral concentration is jackknifed (a patient's own visits are excluded from the calculation for their PCP). Standard errors are clustered at the PCP level. Utilization is measured using standardized prices as described in the text. Insurer controls are a fixed effect for each payer and a fixed effect for each of 13 types of insurance plan (i.e. HMO, PPO, EPO, indemnity, etc.). Patient controls are patient ZIP code fixed effects, sex (male/female), age (included as a 5-knot spline), and comorbidity controls. The patient comorbidity control in the Massachusetts data is the HCC risk score included as a 5-knot spline. Massachusetts PCP controls include the average HCC risk score of the PCP's commercial patients (included as a 5-knot spline), total number of patients (included as 5-knot spline), indicator for PCP sex, and an indicator for whether the PCP's specialty is Internal Medicine or Family Medicine. PCP controls do not include PCP modal ZIP Code fixed effects, as most ZIP codes do not have multiple contracting networks. A fixed effect for the physician contracting network is included, including an indicator of not being in a physician contracting network. Column 2 is limited to only PCPs affiliated with Partners Healthcare.

Table A2. Medicare Sample Descriptive Statistics

	Below Median PCP Team Referral Concentration (2012)	Above Median PCP Team Referral Concentration (2012)	Sample of moving beneficiaries (2007-2012)
<i>Patient characteristics:</i>			
Age	77.1	76.8	79.3
Male	0.41	0.41	0.34
Mean Spending (\$)	13,127	10,786	10,917
Median Spending (\$)	4543	3647	4098
Pr(Any Inpatient Admission)	0.27	0.23	0.28
Patient care continuity HHI	0.35	0.39	0.38
Pr(Diabetes)	0.30	0.30	0.28
Pr(Heart Condition)	0.45	0.42	0.50
Pr(Depression)	0.16	0.15	0.18
Pr(Asthma)	0.06	0.05	0.05
<i>PCP characteristics:</i>			
PCP's Team Referral Concentration	0.21	0.46	0.33
Pr(Internal Medicine)	0.69	0.56	0.66
PCP is Male	0.79	0.77	0.79
Number of observations	925,681	925,661	6230

Notes: Similar to our findings in Massachusetts, we find that 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: below the median, the average referral concentration is 0.21 versus 0.46 above the median.

Table A3. PCP Team Referral Concentration and Medicare Spending, With Measurement Error

	Ln(utilization)		
	(1)	(2)	(3)
PCP Team Referral Concentration	-0.286*** (0.011)	-0.139*** (0.009)	-0.035*** (0.008)
Patient controls	Yes	Yes	Yes
Patient care continuity HHI	No	Yes	Yes
PCP controls	No	No	Yes
N patients	1,851,342	1,851,342	1,851,342

Note: Specifications likely suffer from substantial attenuation bias due to measurement error; simulations suggest measurement error is particularly acute in columns (2) and (3) after additional controls are incorporated. PCP's team referral concentration is jackknifed (a patient's own visits are excluded from the calculation for their PCP). Standard errors are clustered at the PCP level. Data: 2012 Medicare beneficiaries in our 20% sample. Specifications parallel Columns 1-3 of Table 3. However, there are no insurer controls, since all patients are enrolled in traditional Medicare. Further, patient HCC risk scores are replaced with a simple vector of comorbidity fixed effects for each of the 27 conditions recorded in the Chronic Condition Warehouse. Column 3 includes PCP controls: PCP specialty (family medicine or internal medicine), gender, a 5-knot spline in number of patients, and the PCP-level mean value of each of the 27 Chronic Condition Warehouse comorbidities.

Table A4. Decomposition of Medicare Spending

<i>Dependent variable:</i>	<i>Independent variable:</i> PCP team referral concentration			<i>N</i>
	(1)	(2)	(3)	
ln(Inpatient claims)	-0.186*** (0.012)	-0.144*** (0.011)	-0.0527*** (0.011)	468,611
ln(Outpatient claims)	-0.111*** (0.028)	0.114*** (0.027)	0.0852*** (0.027)	1,851,342
ln(Provider submitted claims)	-0.151*** (0.008)	-0.034*** (0.007)	0.030*** (0.007)	1,851,342
Any Inpatient spending	-0.096*** (0.004)	-0.077*** (0.004)	-0.038*** (0.003)	1,851,342
Patient controls	Yes	Yes	Yes	
Patient care continuity HHI	No	Yes	Yes	
PCP controls	No	No	Yes	

Note: Specifications likely suffer from substantial attenuation bias due to measurement error; simulations suggest measurement error is particularly acute in columns (2) and (3) after additional controls are incorporated. PCP's team referral concentration is jackknifed (a patient's own visits are excluded from the calculation for their PCP). Standard errors are clustered at the PCP level. Data: 2012 Medicare beneficiaries in our 20% sample. However, there are no insurer controls, since all patients are enrolled in traditional Medicare. Further, patient risk scores are replaced with a simple vector of comorbidity fixed effects for each of the 27 conditions recorded in the Chronic Condition Warehouse. Column 3 includes PCP controls: PCP specialty (family medicine or internal medicine), gender, a 5-knot spline in number of patients, and the PCP-level mean value of each of the 27 Chronic Condition Warehouse comorbidities.

Table A5. Underpowered Utilization-Based Quality Outcomes in Medicare Mover Sample (IV Results)

	(1)	(2)	(3)	N of obs.	Dep. var. mean
Dependent variable:					
Any emergency department visit	-0.007 (0.063)	0.031 (0.060)	0.017 (0.061)	6,232	0.393
Any inpatient visit (excluding pregnancy)	-0.089 (0.062)	-0.054 (0.055)	-0.071 (0.055)	6,232	0.286
Any hospitalization for ambulatory care sensitive condition	-0.010 (0.042)	-0.001 (0.042)	-0.007 (0.043)	6,232	0.085
Patient fixed effects	Yes	Yes	Yes		
Patient care continuity HHI	No	Yes	Yes		
PCP controls	No	No	Yes		

Note: PCP's team referral concentration is jackknifed. 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). 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. Standard errors are clustered at the patient level. Columns 1-3 mirror specifications from Table 6 Panel C, Columns 1- 3.

Appendix B. Chronic Illness Definitions for Massachusetts Population

We use the definition of Chronic Illness from Frandsen et al. (2015). A patient is included for having a chronic illness if they received an ICD-9 diagnostic code 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

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

Classification	Taxonomy Code	Physician Specialty
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
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
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

Classification	Taxonomy Code	Physician Specialty
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

Appendix D. Proofs of Results in Section 3

Result 1: The PCP's optimal team structure has a number of specialists N^* . The PCP invests the same amount of effort r_s^* in developing team-specific capital with each specialist and refers to each specialist with probability $\frac{1}{N^*}$.

Proof: The first order condition for each choice of effort is given by $r_s = \frac{1}{\sum_{s=1}^N e^{V_{is}}} e^{V_{is}} (\omega + \theta)$. It is easy to see that the symmetric solution satisfies the vector of first order conditions. However, $\ln[\sum_{s=1}^N e^{V_{is}}]$ is convex in the vector of efforts r_s . To guarantee this symmetric solution uniquely satisfies the first order conditions, we need the Hessian matrix of second order derivatives to be negative definite everywhere. The assumption $\omega + \theta < \sqrt{2}$ guarantees that.

Result 2: Investment in team-specific capital r_s^* is inversely proportional to N^* , the number of specialists the PCP refers to. A PCP with a higher fixed cost φ of working with an additional specialist will work with fewer specialists, invest more in team-specific capital with each specialist, and have lower expected healthcare costs for their patients.

Proof:

From the first order condition for the choice of effort, we have $r_s^ = \frac{1}{N^*} (\omega + \theta)$. Moreover the PCP's choice of N^* is chosen to $\max_{N, \{r_s\}_{s \in N}} \ln[\sum_{s=1}^N e^{V_{is}}] - \varphi N - \sum_{s=1}^N [\frac{1}{2} r_s^2]$.*

A higher value of φ leads to a weakly lower N^ being chosen. Note that φ only directly affects the choice of N^* , and the first order condition for r_s^* determines the level of investment in team-specific capital. This decreases in N^* . Finally, Assumption 1 gives that higher r_s^* leads to lower healthcare costs.*

Appendix E. 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 1 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 1 referred patient, we would be able to perfectly calculate his TRC=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 OLS 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 OLS specification (cf. Section 7A).

A similar logic extends to the instrumental variable (IV) case, and the net effect of measurement error is theoretically ambiguous. We have non-classical measurement error, since measurement error in our

instrument (pre-move team referral concentration) is mechanically correlated with measurement error in the endogenous variable (change in team referral concentration).

To compare attenuation bias in the difference-in-differences or IV specifications to attenuation bias in the OLS Medicare results (e.g. coefficients reported in Table 3, column 7), 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.