UNIVERSITY of York

HEDG

HEALTH, ECONOMETRICS AND DATA GROUP

WP 17/19

Two-Sided Matching in Physician-Insurer Networks: Evidence from Medicare Advantage

Kathleen Nosal

August 2017

http://www.york.ac.uk/economics/postgrad/herc/hedg/wps/

Two-Sided Matching in Physician-Insurer Networks: Evidence from Medicare Advantage

Kathleen Nosal*
University of Mannheim

June 3, 2016

Abstract

Many health insurance plans in the U.S. restrict enrollees to choose from a set of providers the insurer has contracted with. These provider networks are formed via bilateral bargaining between insurers and providers. Provider networks are an important tool for product differentiation and cost containment for insurers and also put real restrictions on consumers' choice of providers. In this paper, I analyze matching between insurers offering Medicare Advantage Plans and physicians, using a unique data set consisting of all insurer-physician links in several counties. I estimate parameters of a two-sided, many-to-many matching model which describes formation of provider networks, using the Maximum Score estimator of Fox (2010). This method uses implications of a pairwise stability condition to estimate a joint surplus function which depends on insurer-physician links. The surplus function accounts for the role of physician and insurer characteristics in determining their match values, and also for interactions between physicians linked to the same insurer, whose services may be complements or substitutes. The results indicate that insurers prefer on the margin to link with physicians who increase the specialty concentration of their network and who are located near other physicians in the network. Physicians are negatively affected by having a broader referral network, as defined by having a larger set of physicians with whom they have insurer links in common. Finally, compared with regional insurers, nationally active insurers benefit more from matching with physicians with U.S. medical degree. Preliminary counterfactual analyses suggest that insurers and physicians would be collectively better off if all physicians were matched to all insurers– that is, if selective contracting were eliminated entirely.

1 Introduction

Health insurance plans featuring limited provider networks are increasingly common in the U.S. health insurance industry. Despite ample academic and media attention on the

*Email: knosal@mail.uni-mannheim.de. I am grateful to Greg Crawford, Jeremey Fox, Michelle Sovinsky and Hidenori Takahashi for helpful comments and suggestions. I also received useful feedback at EARIE Munich, the University of Zurich, MaCCI Competition and Regulation Day, WinE at EEA Mannheim, and the New Empirical Developments in Health and Labor Markets workshop at KU Leuven.

effects of provider networks on consumers,[1] little has been done to understand the other side of the equation: the incentives of insurers and providers. Why do they choose to organize their interactions in this way? How do the structure of provider networks affect insurer and provider profitability?

In this paper, I analyze matching of insurers and physicians in provider networks. The aim is to understand how the network structure and the interaction of insurer and physician characteristics generate surplus for insurers and physicians. I estimate parameters of a two-sided, many-to-many matching model which explains formation of the provider networks. In the model, profitability depends on the geographic scope of the network, the breadth and depth of physician specialty coverage, and interactions between physician and insurer characteristics, while allowing for monetary transfers between insurers and physicians. I estimate the model using a unique data set linking physicians and insurers participating in Medicare Advantage in several U.S. counties. Medicare Advantage is a program which allows Medicare eligibles to opt into subsidized plans offered through private insurers in place of regular Medicare coverage. Restricting attention to Medicare Advantage allows me to focus on a group of insurers competing for the same group of consumers in well-defined markets.

Estimation is based on the maximum score estimator proposed by Fox (2010). The equilibrium concept is pairwise stability, which requires that no coalition of one physician and one insurer can deviate to do better than the observed matching. Following Fox (2010), inequalities are formed which follow from pairwise stability of the observed matching. The maximum score estimator selects the parameters which maximize the number of these inequalities that hold.

Provider networks limit the ability of insured patients to freely chose a primary care physician, and to access a full range of specialists. At the same time, these networks are an important tool used by insurers to cut costs and manage care. Despite the importance of provider networks in the healthcare system, surprisingly little is known about the formation of these links between physicians and insurers. The aim of this paper is to quantitatively study the incentives of physicians and insurers which affect the formation of provider networks. In the next few paragraphs, I discuss the trade-offs that insurers and physicians face in forming links and preview how these trade-offs are reflected in the structural surplus function that I estimate. The estimated parameters of the surplus function pin down the relative importance of different features of the network in determining physician and insurer profit.

Insurers receive a risk-adjusted, county-specific capitation payment from Medicare for each enrollee in one of their Medicare Advantage plans, and may also charge a premium. Since Medicare eligibles choose among plans, insurers compete for Medicare Advantage enrollees. Medicare Advantage plans are differentiated products, which differ in coverage, cost sharing, and the provider networks. Insurers may therefore use the quality of the provider network as a way to attract enrollees. Perceived quality of the network may depend in part on the number of physicians, their specialties and locations, and other characteristics. The trade-off faced by insurers is that offering a higher quality network attracts enrollees, increasing revenue, but may also increase costs. There are two channels for the cost increase. First, in order to add more or better physicians, the insurer might have to offer better financial terms to these physicians. (In the model, this is captured

[1] For example, the Washington Post ran an article entitled "How a narrow network can really mess with your choice of doctors" in June 2015. The New York Times has also published several pieces on provider networks. I survey the academic literature on provider networks later in this Introduction.

by monetary transfer payments from insurers to physicians.) Second, the greater access to care afforded by a better network may encourage greater utilization of healthcare by patients. Indeed, a central tenet of managed care is that insurers use restricted access to providers as a way to reduce the cost of care. Having too broad of a network undermines the cost savings of managed care.

How do insurers manage the conflicting goals of using the network to attract consumers and to contain costs? To measure the net effects of features of the network on insurer profits, I include functions of physician specialties and locations in the surplus function. Do insurers profit from having broad networks, with a few physicians in almost every specialty, or do insurers benefit from a network more concentrated on a few more heavily used specialties? Do insurers try to cover the market with physicians located geographically close to every potential enrollee, or do they focus on recruiting physicians in a few popular locations? These factors are captured in the surplus function with terms based on the breadth and concentration of specialty offerings and physician locations.

Physicians join an insurer's network in order to access the insurer's patient pool. The physician and insurer must agree on a schedule of payments for services the physician provides to the insurer's enrollees, which typically entails a discount from the physician's regular rates.[2] Thus, physicians must determine whether it is profitable to serve the insurer's patients at the proposed rates. Another consideration for physicians is the way that network membership affects cooperation and competition with other physicians. Referrals generally occur within a network. Consequently, a physician's referral network is limited to the other physicians with whom she has an insurer link in common. On the competition side, network overlap would also determine which physicians are directly in competition for the same patients.

In order to capture the net effect of the provider networks on the physicians' interaction with other physicians, the physician surplus function includes a term which depends on the size of the physician's implied referral network. The referral network is defined as the set of physicians with whom the physician has at least one insurer link in common. This is the relevant set of physicians both for defining from whom potential referrals can come, and also who is a direct competitor for patients.

The surplus function also contains interactions between physician and insurer characteristics. The idea is to determine whether there are complementarities between certain types of insurers and physicians. For example, it could be the case that high quality physicians tend to match with high quality insurers. The characteristics that are interacted are intended as proxies for quality. These characteristics are an indicator for whether the insurer operates nationally, an indicator for whether the physician graduated from a US medical school, and the physician's years of experience (defined as years since graduation from medical school.)

Most of the existing literature on provider networks focuses on *hospital* insurer networks. Papers in this literature estimate the value of hospitals to consumers (Ho, 2006, Ericson and Starc, 2014) or consider bargaining between hospitals and insurers (e.g., Ho, 2009, Ho and Lee, 2013 and Gowrisankaran, Nevo and Town, 2015). A few papers address physicians in insurer networks. Some of these are mostly descriptive, such as Chernew et al. (2004) and Gaynor and Marks(2002). Others measure the effect of physician networks on consumer demand for health plans and consumer welfare (Dahl and Forbes, 2014) or

[2]Other arrangements are possible. For example, some insurers offer a captitation based system, where physicians are compensated based on the number of patients covered, regardless of the actual services used.

on insurer costs (Gruber and McKnight, 2014) but take the network as given. The current paper is the first to consider the formation of physician-insurer networks. Physicians and hospitals are both important, and, in fact, complementary, components of provider networks and of the healthcare system more broadly. This paper fills in a gap in the literature on provider network formation, which has ignored the role of physicians. In 2012, only 9% of Americans had an overnight hospital stay, but 82% had some contact with a healthcare provider (Center for Disease Control, 2013 and 2014). Modeling only access to hospitals and not to physician office visits misses the way that most patients encounter the healthcare system in a given year, and the most typical use of insurance coverage. Furthermore, literature on continuity of care emphasizes the importance to patient health of access to the same physicians over time (e.g. Pouret et al, 2015). The continuity of care literature provides a channel through which access to physicians in providers networks could influence health. For these reasons, it is important to also consider the role of physicians in provider networks.

The Fox (2010) maximum score estimator has been applied to matching problems in numerous industries: matching of bidders to licenses in spectrum auctions (Fox and Bajari, 2013), matching of marketing rights for pharmaceutical products to firms (Levine, 2009), matching between banks and firms in commercial loan markets (Chen and Song, 2013), and matching between professional athletes and teams (Shi, Yang and Goldfarb, 2009), among other applications. Most of these are examples of one-to-many matching. My application is a case of many-to-many matching, and thus fully exploits that methodology in Fox (2010), which was originally applied to many-to-many matching between auto assemblers and parts suppliers.

In the results, I find that insurers prefer to add physicians who increase the specialty concentration of the network, rather than add additional specialties, and whose practices are physically located near those of other physicians in the network, rather than branching out to new areas. Compared with regional insurers, nationally active insurers have more exclusive links with physicians, and match with more physicians with US medical degrees, suggesting positive assortative matching. Finally, physicians seem to be negatively affected by having a broader referral network, as defined by having a larger set of physicians with whom they have insurer links in common.

Provider networks are an increasingly important arena for product differentiation by insurers. A trend in the regulation of health insurance is to require standardization of plan benefits and cost sharing.[3] For example, plans offered through the health care marketplaces established by the Affordable Care Act must provide a standardardized set of benefits, and must implement one of four pre-defined levels of cost sharing. In contrast, insurers have considerable freedom in what form provider networks can take. As a result, provider networks remain a way that insurers can vertically or horizontally differentiate plans in an environment where other dimensions of plan quality are regulated. Given these developments, a question of particular interest is whether plan quality and the quality of the provider network are complements or substitutes. If they are substitutes, then insurers may respond to regulation requiring higher plan quality by offering lower quality, more restricted networks. If they are complements, regulating plan benefits alone may be sufficient to also encourage provision of high quality networks. Preliminary results in this paper suggest that the latter is more likely.

[3] Starc and Ericson (2013) study standardization in the context of the Massachussetts health insurance exchange.

2 Data

2.1 Constructing the Data

The novel data set created for this paper constructs the complete matching between physicians and Medicare Advantage plans in several counties. Each insurer's network was observed through provider directories from the insurers. The network data was matched to comprehensive administrative data on licensed physicians. For each physician-level observation, the data indicates which insurers the physician is linked to and also includes a variety of physician characteristics.

The data covers Medicare Advantage plans offered in six counties in Arizona and Nevada in 2014.[4] Data on the insurers offering Medicare Advantage plans came from public sources made available by the Center for Medicare and Medicaid Services (CMS).[5] The annual Contract Service Area data set from the CMS lists every health plan in the Medicare Advantage program and the counties and states in which each plan is available. The data specifies the type of plan (HMO, PPO, etc.) and the identity of the insurer offering the plan. Other data sets from CMS feature specific plan characteristics such as premiums, benefits, and copayment levels. In order to focus on the Medicare Advantage plans available for all Medicare eligibles, I removed "Special Needs Plans", which are subject to additional requirements, such as having a particular chronic condition, and plans that are restricted to employees or former employees of particular employers.

Data on the universe of physicians practicing in the relevant counties comes from files provided by the state medical boards of Arizona and Nevada. (Throughout this section, I refer to these data sets as the "medical board data.") The data lists all physicians licensed in the state, along with their business address, specialties, medical school, graduation year, and residencies completed. The physicians are assigned to a county based on the business address.

To establish the matching between physicians and insurers, I collected provider directories for all the plans in the data set. Provider directories are documents intended for consumers enrolled in the plan, which list network providers, along with contact, location and specialty information for each provider. The Center for Medicare and Medicaid Services requires that insurers participating in Medicare Advantage periodically send hard copies of the directories to enrollees. Insurers typically make the provider directories available on their websites as well, often as a PDF file. In addition, many insurers provide an interactive search tool to search a database of providers on their website. For the majority of the plans and insurers included in the data, PDF directories were available, and I obtained the list of providers from the index section of the directory. For the remaining insurers, I used the search tool to extract the plan's full list of providers.[6]

Many insurers offer multiple Medicare Advantage plans in the same county. In some cases, all plans have the same network, but in other cases there are different networks for different plans. For example, an insurer may offer a basic plan with a narrow network

[4] The five counties are: Maricopa County, Arizona; Pima County, Arizona; Yavapai County, Arizona; Washoe County, Nevada; and Clark County, Nevada.

[5] Data on Medicare Advantage plans can be downloaded at https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/index.html and https://www.medicare.gov/download/downloaddb.asp.

[6] Provider directories for the insurer Wellcare are missing from the data. Wellcare entered Pima and Maricopa counties in 2014, and exited in 2015. Because of its short tenure, Wellcare is unlikely to have been an important player in these markets.

and a deluxe plan with a larger network. This within-insurer network differentiation is interesting, but it is not the main focus here. Furthermore, there is not always a clear mapping between directories and plans. A directory may be for one or several plans in one or several counties.[7] It is always clear, however, which insurer the directory corresponds to, and which insurers are present in each county. To abstract from plans, I create a composite network for each insurer in each county from all of the provider directories for that insurer's plans. In other words, I assume that networks are at the insurer level, and aggregate the data accordingly.

From the provider directories, I extracted the names of network physicians, and linked them to the physician names in the medical board data. I built indicators for membership in each insurer's network based on the directories a physician appears in. The result is a data set with both physician characteristics and network inclusion information.

To prepare for the join between the provider directory data and the medical board data, the text data from the provider directories had to be filtered to produce a list of physician names. The provider directories contain listings of many types of medical professionals and facilities. Entries which have the structure of human names and contain an MD (Medical Doctor) title were kept.[8] In joining the physician names with the board data, I employed fuzzy matching techniques to allow for formatting differences and spelling discrepancies in names. Spurious duplicate matches were reduced by subjecting duplicate pairs to a stricter matching criteria in a second pass. Despite the inherent difficulties in matching names, this procedure worked quite well: for most insurers, a match was found for more than 95% of physicians, and many duplicate matches were successfully eliminated on the second pass. Physicians for whom no match was found in the medical board data were discarded.[9]

Specialty indicators were constructed from the specialty fields in the state medical board data. The goal is to sort physicians into specialties, and further determine who is a primary care physician and who is a specialist. This categorization does not follow immediately from the medical board data, because multiple specialties can be listed for each physician, and the coding of specialty names is inconsistent. To standardize specialties, I adopted the list of specialties used in the Medicare Physician Compare tool.[10] After combining a few closely related specialties, 57 categories remained. I then mapped the 363 distinct specialty names that appear in the Arizona medical board data and the 179 in the Nevada data to these 57 categories. There are eight specialty descriptions that I mapped to primary care: Primary Care, Family Practice, General Practice, Geriatric Medicine, Internal Medicine, Gynecology, Pediatrics[11] and Preventive Medicine. If only

[7]Also, it is not always clear how to define what a plan is when some insurers offer many nominally distinct plans that have the same network and nearly identical benefits, while others offer few plans or only one.

[8]The restriction to entries with an MD title eliminates some providers who could arguably be called physicians. For example, Doctors of Osteopathic Medicine (associated with an OD title) and psychologists or optometrists who have a PhD but not an MD are excluded. This narrow definition allows for a closer match to the medical board data.

[9]For a small number of these discarded physicians, I investigated further with Google. In some cases, these physicians were reported to have died or moved to another state, and apparently remained in the provider directory erroneously, despite being no longer listed as an active physician in the medical board data.

[10]Available at https://www.medicare.gov/physiciancompare/.

[11]Since Medicare is for the elderly, one would not expect many pediatricians in the networks! While there are some, pediatricians are indeed underrepresented in the Medicare Advantage plan networks compared to the overall share of licensed physicians that are pediatricians.

one specialty is given for a physician, I simply applied the mapping to that specialty. When multiple specialties are specified for a physician, I used the following procedure: any physician whose first two specialties are both in the primary care group are considered primary care.[12] If one is primary care and the other is a proper specialty, the non-primary-care specialty is used. If a physician has two specialties, neither of which is primary care, the first of the two is taken.

The state board data also contains each physician's office address, medical school, and medical school graduation date. These items required further processing to be transformed into usable variables. From the addresses, I pulled out the postal code as a proxy for the physician's office location. The postal codes were linked to their respective populations. The postal code populations give a measure of how many prospective patients live near the physician's office. The medical school graduation years were used to calculate each physician's time since graduation. This can be interpreted as a "years of experience" variable, with the caveat that time spent in additional training after medical school, such as a residency, is also included, so it does not perfectly measure years in medical practice. Finally, to make a rough classification of medical schools, I divided them into medical schools located inside and outside of the United States. This division is intended as a proxy for perceived medical school quality.[13]

Some insurers operate regionally, while others offer Medicare Advantage plans nationwide. This distinction is important because national insurers may benefit from economies of scale, allowing them to offer higher quality insurance. I constructed a indicator for national operation based on the number of states in which an insurer offers health plans. Those offering plans in 18 or more states were coded as national.[14]

The data set is a (nearly) complete picture of the matching between physicians and insurers within the Medicare Advantage market. However, there are health insurance plans outside of Medicare Advantage that also have provider networks. Many insurers who offer Medicare Advantage plans also participate in other markets, offering employer group insurance, individual plans, or Medicaid Managed Care plans. In addition, there are insurers who participate in these other markets, but are not considered here at all because they do not offer Medicare Advantage plans. Links between physicians and insurers through these other programs could have spillover effects on the matching decisions within the Medicare Advantage market. Therefore, the ideal data set would also include these other links. Unfortunately, it is infeasible to collect such a data set from publically available sources. Employer sponsored health plans in particular typically do not have publically available network information. Given this difficulty, the focus was on collecting as complete data as possible on Medicare Advantage networks. On the bright side, it might be possible to learn something about provider networks in these other programs from the Medicare Advantage data. Within Medicare Advantage, insurers who offer

[12]The use of two specialties to determine primary care status is because there were many physicians who had, for example, Internal Medicine as a first specialty, and a more specific specialty, like cardiology, as the second. This type of Internal Medicine physician is probably not a generalist and should not be classified as primary care.

[13]For example, a number of non-U.S. medical schools in the data were located in India and the Caribbean. Patients would plausibly perceive physicians from these medical schools to have received a lower quality of education compared to physicians from well-known U.S. medical schools (regardless of whether this is true.) It would be ideal to incorporate detailed medical school ranking data instead of this very rough proxy.

[14]The cutoff of 18 is arbitrary, but represented a natural gap in the data: the next lowest data point after 18 was 6.

Table 1: Physician and Insurer Summary Statistics by County

County	Ins	Phys	MA Phys	MA Prime	MA Spec	Specialties
Maricopa, AZ	9	9029	4735	1495	3240	46
Pima, AZ	6	2852	1463	544	919	43
Yavapai, AZ	3	380	200	87	113	21
Washoe, NV	3	1174	560	176	384	42
Clark, NV	4	3460	1520	584	936	46

Table Notes: Columns "Ins" and "Phys" give the total number of Medicare Advantage insurers and the total number of licensed physicians in the county. "MA Phys," "MA Prime," and "MA Spec" give the total number of physicians, primary care physicians, and specialist physicians in the county who participate in Medicare Advantage (i.e., are in the network for at least one Medicare Advantage insurer.). "Specialties" is the number of distinct specialties represented among the MA specialists.

multiple plans tend to have similar networks across plans. If this trend is also true for plans offered by the same insurer across different programs, then the insights from this paper are perhaps not only applicable to Medicare Advantage.

2.2 Summary Statistics

Table 1 provides summary statistics about the physicians and insurers in each of the five counties in the data. The number of insurers offering Medicare Advantage plans ranges from three (in Yavapai and Washoe counties) to nine (in Maricopa county). The number of licensed physicians in the counties also has a wide range, with 380 in Yavapai county and 9029 in Maricopa county.[15] In every county, around half of the licensed physicians participate in Medicare Advantage, meaning that they are in the network of at least one Medicare Advantage plan. The table further breaks down the Medicare Advantage participants into primary care physicians and specialists, and indicates how many distinct specialties are represented among the specialists.

Drilling down to the within-county, across-insurer level, Table 2 contains summary statistics about the networks for each of the nine MA insurers in Maricopa county. There is considerable variation across insurers in the number of physicians in the insurer's network and also the distribution of the physicians across the specialty and primary care categories. Insurers also differ in the number of specialties and zip codes represented by the physicians, and in the percentage of physicians who are exclusive to that insurer. From this table, it is clear that networks are differentiated.

Figure 1 shows the distribution of physicians' number of network memberships within each county. These distributions demonstrate a great deal of heterogeneity across physicians. In each county, some physicians are linked to just one insurer, some to every insurer, and some are in each bin in between. Therefore, there is no general rule about whether physicians "single-home" or "multi-home" with insurers. Both possibilities appear to happen in equilibrium.

[15]The variation in number of physicians in part reflects the different population of the counties. Maricopa county has over 4 million residents, while Yavapai has about 200,000.

Table 2: Insurer Summary Statistics for Maricopa County

Insurer	Nat'l	# Phys	# Spec	# Prime	Specs	Zips	Excl.	Mean exp	US Med
BCBS	Yes	802	547	255	39	83	1.2%	23.8	59.1%
United	Yes	2652	1667	985	44	118	3.1%	23.5	66.2%
Humana	Yes	2582	1721	861	45	116	6.6%	23.1	64.6%
Aetna	Yes	2964	2044	920	45	120	4.1%	23.7	66.2%
Healthnet	Yes	834	501	333	38	94	7.8%	22.8	63.5%
Cigna	Yes	3000	2520	480	46	111	22.6%	23.9	70.3%
Caremore	No	482	419	63	35	67	0.2%	23.2	60.4%
Phoenix	No	1779	1183	596	43	110	2.6%	23.2	60.6%
Scan	No	2131	1459	672	43	103	2.0%	23.4	62.8%

Table Notes: This table presents summary statistics about the networks of the nine MA insurers in Maricopa county. The columns contain: indicator for national insurer, total number of physicians, number of specialist physicians, number of primary care physicians, number of distinct specialties represented, number of distinct zip codes represented in physician locations, percentage of physicians who are exclusive to the insurer, mean years of experience, percentage of physicians with a US Medical degree.

3 Model

In this section, I describe a model of two-sided, many-to-many matching between insurers and physicians.

Consider a market, m. Let I_m be the set of insurers in the market, and J_m be the set of physicians in the market. Insurers are indexed by $i = 1, 2, ...|I_m|$ and physicians are indexed by $j = 1, 2, ...|J_m|$.

The outcome of a matching game consists of a *matching* and a vector of *transfers*. A matching μ is a subset of $I_m \times J_m$ such that $(i, j) \in \mu$ if and only if i and j are linked– that is, if physician j is in the network of insurer i. A *match* refers to an ordered pair that is an element of a matching. Many-to-many matching is allowed, meaning that each physician may be linked to multiple insurers and each insurer to multiple physicians in a matching. Some agents may be unmatched. An *unmatched* agent is not a member of any of the ordered pairs in μ, because the agent is not linked to any agent on the other side of the market.[16] A transfer is a monetary amount t_{ij} paid from insurer i to physician j, where $(i, j) \in \mu$. Let T_μ be the vector of transfers associated with matching μ. Transfers are not restricted to be positive.

A given matching maps to a profit level for each insurer and physician. Let $\pi_i(\mu)$ be the profit, net of transfers, to insurer i from matching μ. Likewise, let $\pi_j(\mu)$ be the profit, net of transfer, to physician j from matching μ. The profit is allowed to depend on the whole matching, μ, rather than on just a single match (i, j), to allow for complementarities or other types of interactions between matches. Since utility can be transferred between

[16] In practice, many physicians in the data are unmatched. I do not introduce a special notation for unmatched agents, because they do not show up in the inequalities used in estimation. The methods used here were developed specifically for the case where unmatched agents are part of the theoretical model, but are unobservable in the data. Thus, the estimation inequalities are designed to identify the parameters of interest without reference to unmatched agents. Fox, Hsu and Yang (2015) introduce an alternate estimator that does exploit information on unmatched agents.

agents through the monetary transfers T_μ, an important quantity is the joint profit of two matched agents, $\pi_i(\mu) + \pi_j(\mu)$, which constitutes the total surplus from matching which can be divided between them.

The equilibrium concept is pairwise stability. Pairwise stability requires that no pair of agents can collectively do better by forming a new match (which requires bilateral consent) and possibly breaking one or more existing matches that one of the agents is involved in (which requires only unilateral consent).[17] The following is an implication of pairwise stability of matching μ:

$$\pi_i(\mu) + t_{ij} \geq \pi_i((\mu \setminus \{(i,j),(i',j')\} \cup (i,j')) - \tilde{t}_{ij'} \tag{1}$$

where:

$$\begin{aligned}&(i,j),(i',j') \in \mu\\&\tilde{t}_{ij'} = \pi_{j'}(\mu) + t_{i'j'} - \pi_{j'}(\mu \setminus \{(i,j),(i',j')\} \cup (i,j'))\end{aligned}$$

The inequality compares insurer i's profit at the matching μ and a modified matching that breaks the matches (i,j) and (i',j') and adds the match (i,j'), accounting for transfers. The transfer $\tilde{t}_{ij'}$ is constructed to make physician j' indifferent between the original and modified matchings. If insurer i were strictly better off with the modified matching, then insurer i and physician j' could block the matching, and it would not satisfy pairwise stability. Therefore, the left side of (1) must be weakly greater for the condition to hold.[18]

These inequalities are slightly different than those in Fox (2010) because I allow agents to care about matches that they are not involved in– for example, insurer i might be affected by whether insurer i' is matched to j or not. Many applications, such as Chen and Song (2013), assume additive separability of profit across a firm's matches.[19] In that case, agent i gets the same profit from being matched with agent j regardless of what other matches i is involved in. Fox (2010) allows a firm to care about its portfolio of matches. This is motivated by his application, in which upstream auto parts manufacturers may benefit from specialization based on the set of downstream auto assemblers it is matched with. Similarly, in the spectrum auction application in Fox and Bajari (2013), there are complementarities across licenses, so firm profits depend on the package of licenses won, and profits are not additively separable across licenses. My application requires even more generality. As I will exposit later in the detailed section on the profit functions, an insurer's profit may depend on how exclusive its relationship with a physician is. Therefore, the insurer cares about *other* matches of the physicians it is matched with.

Since data on transfers is not available, the estimation must be based on conditions that do not involve transfers. Following Fox (2010), I manipulate condition (1) to eliminate the transfers. First, the definition of $\tilde{t}_{ij'}$ is plugged into (1):

$$\pi_i(\mu)+t_{ij} \geq \pi_i((\mu\setminus\{(i,j),(i',j')\}\cup(i,j'))-\pi_{j'}(\mu)-t_{i'j'}+\pi_{j'}(\mu\setminus\{(i,j),(i',j')\}\cup(i,j')) \tag{2}$$

[17]This is the definition of pairwise stability used by Fox(2010) and traditionally in the matching literature such as Roth and Sotomayor(1990). It is stronger than the definition used in the literature on network formation (e.g. Bloch and Jackson(2006) and Jackson and Wolinsky (1996)), which only considers single deviations and therefore rules out that links may be simultaneously added and deleted to block a matching. Bloch and Jackson refer to the stronger notion as "Pairwise Nash Stability."

[18]Note that while there is a vector of transfers T_μ associated with the matching μ, the inequality refers only to transfer t_{ij} and $\tilde{t}_{ij}$. This is because the other transfers are the same on both sides of the inequality, and therefore cancel.

[19]Chen and Song (2013) also differ in that they study many-to-one matching rather than many-to-many matching. The issues discussed here are slightly simplified in that case because the notion of a "portfolio" of matches is only relevant for one side of the market. The same is true of Fox and Bajari (2013).

Then, the analogous condition is formed for insurer i':

$$\pi_{i'}(\mu)+t_{i'j'} \geq \pi_{i'}((\mu\backslash\{(i,j),(i',j')\}\cup(i,j'))-\pi_j(\mu)-t_{ij}+\pi_j(\mu\backslash\{(i,j),(i',j')\}\cup(i',j)) \quad (3)$$

Adding inequality (2) to inequality (3) generates an inequality without transfers:

$$\begin{aligned}\pi_i(\mu) + \pi_{i'}(\mu) + \pi_j(\mu) + \pi_{j'}(\mu) \geq \\ \pi_i((\mu \setminus \{(i,j),(i',j')\} \cup (i,j')) + \pi_{i'}((\mu \setminus \{(i,j),(i',j')\} \cup (i,j')) \\ + \pi_j(\mu \setminus \{(i,j),(i',j')\} \cup (i',j)) + \pi_{j'}(\mu \setminus \{(i,j),(i',j')\} \cup (i,j'))\end{aligned} \quad (4)$$

The inequality (4) is known as a *sum of revenues inequalities*. Sum of revenues inequalities are used in the estimation to determine the extent to which pairwise stability is satisfied. Such an inequality can be generated for every two matches (i,j) and (i',j') observed in the data.

3.1 Profit Functions

In this section, I provide more details about the profit functions π_i and π_j. In section 3.1.1 I discuss three assumptions on the profit functions which facilitate estimation and identification. In section 3.1.2 I present the empirical specification of the profit function that is used in the estimation.

3.1.1 Key Assumptions and General Form of Profit Function

Let x_i and y_j be vectors of observed characteristics of insurer i and physician j respectively, and let X_m and Y_m be matrices of these characteristics for all agents in market m. The profit functions consist of a deterministic term, which is a parametric function of the observed characteristics and the matching, and a stochastic term, whose distribution may also depend on the matching. The profit functions are written as:

$$\begin{aligned}\pi_i(\mu; X_m, Y_m, \alpha) = \bar{\pi}_i(\mu, X_m, Y_m, \alpha) + \epsilon_i(\mu) \\ \pi_j(\mu; X_m, Y_m, \beta) = \bar{\pi}_j(\mu, X_m, Y_m, \beta) + u_j(\mu)\end{aligned} \quad (5)$$

where α and β are parameters to estimate. To simplify notation, I will sometimes suppress the arguments X_m, Y_m, α and β.

Assumption 1 *Consider a matching μ. Suppose $(i,j') \notin \mu$. Then, for any i' s.t. $(i',j') \in \mu$, $\bar{\pi}_i(\mu) = \bar{\pi}_i(\mu \setminus (i',j'))$ and for any i'' s.t. $(i'',j') \notin \mu$, $\bar{\pi}_i(\mu) = \bar{\pi}_i(\mu \cup (i'',j'))$. Likewise, if $(i',j) \notin \mu$, then for any j' s.t. $(i',j') \in \mu$, $\bar{\pi}_j(\mu) = \bar{\pi}_j(\mu \setminus (i',j'))$ and for any j'' s.t. $(i',j'') \notin \mu$, $\bar{\pi}_j(\mu) = \bar{\pi}_j(\mu \cup (i',j''))$.*

The assumption says that if insurer i is not matched with physician j', then i is indifferent (according to the deterministic part of the profit function) to j' adding or deleting a match, and j' is indifferent to i adding or deleting a match. This assumption is weaker than what is typical in the literature, because it allows for agents to be affected by its own matches' *other* matches. What is ruled out, is an agent caring about the matches of agent on the other side of the market, with whom it is not matched.

Writing the sum of revenue inequality for a matching μ with respect to the deterministic part of the profit function yields:

$$\begin{aligned}&\bar{\pi}_i(\mu)+\bar{\pi}_{i'}(\mu)+\bar{\pi}_j(\mu)+\bar{\pi}_{j'}(\mu)\geq\\&\qquad\bar{\pi}_i((\mu\setminus\{(i,j),(i',j')\}\cup(i,j'))+\bar{\pi}_{i'}((\mu\setminus\{(i,j),(i',j')\}\cup(i',j))\\&\qquad\quad+\bar{\pi}_j(\mu\setminus\{(i,j),(i',j')\}\cup(i',j))+\bar{\pi}_{j'}(\mu\setminus\{(i,j),(i',j')\}\cup(i,j'))\end{aligned}\tag{6}$$

Under Assumption (1), the sum of revenues inequality (6) can be rewritten as follows:

$$\begin{aligned}&\bar{\pi}_i(\mu)+\bar{\pi}_{i'}(\mu)+\bar{\pi}_j(\mu)+\bar{\pi}_{j'}(\mu)\geq\\&\quad\bar{\pi}_i((\mu\setminus\{(i,j),(i',j')\}\cup\{(i,j'),(i',j)\})+\bar{\pi}_{i'}((\mu\setminus\{(i,j),(i',j')\}\cup\{(i,j'),(i',j)\})\\&+\bar{\pi}_j(\mu\setminus\{(i,j),(i',j')\}\cup\{(i',j),(i,j')\})+\bar{\pi}_{j'}(\mu\setminus\{(i,j),(i',j')\}\cup\{(i',j),(i,j')\})\end{aligned}\tag{7}$$

Because of the substitutions allowed by Assumption (1), the profit terms on the right side are all a function of the same set, $\mu\setminus\{(i,j),(i',j')\}\cup\{(i,j'),(i',j)\}$. Therefore, the left and right sides just evaluate the profit functions of the same four agents at different matchings.

The sum of revenues inequality (7) involves only the deterministic profit function, but profit also includes the error terms ϵ_i and u_j. Under an additional assumption on the error terms, the rank order condition given below, it is possible to work directly with the deterministic version of the sum of revenues inequality given in (7).

Assumption 2 ***Rank Order Condition*** *Let $Pr(\mu)$ be the probability that matching μ occurs (where ϵ_i and u_j are the source of the uncertainty.) Suppose μ_1 and μ_2 are two matchings such that $(i,j),(i',j')\in\mu_1$, $(i,j'),(i',j)\notin\mu_1$ and $\mu_2=\mu_1/\{(i,j),(i',j')\}\cup(i',j),(j',i))$. Then, the following equivalence holds:*
$Pr(\mu_1)>Pr(\mu_2)\iff\bar{\pi}_i(\mu_1)+\bar{\pi}_{i'}(\mu_1)+\bar{\pi}_j(\mu_1)+\bar{\pi}_{j'}(\mu_1)>\bar{\pi}_i(\mu_2)+\bar{\pi}_{i'}(\mu_2)+\bar{\pi}_j(\mu_2)+\bar{\pi}_{j'}(\mu_2)$

The inequality on the right side is the sum of revenues inequality for the deterministic part of the profit function. The condition states that matchings for which the sum of revenue inequality holds for the deterministic profit are more likely than those for which it does not hold. The estimator relies on evaluating the deterministic inequalities, and exploits this property for consistency.

The unobservables $\epsilon_i(\mu)$ and $u_j(\mu)$ are left unrestricted other than the rank order condition. As an example, $\epsilon_i(\mu)$ could be a sum of match-specific logit errors ϵ_{ij} for every j such that (i,j) is in μ. However, much more general unobservables are possible, as there is no requirement of additive separability across matches or any particular functional form.

Finally, I make a linear functional form assumption about the profit functions.

Assumption 3 ***Linearity*** *Parameters enter the profit function linearly. That is:*

$$\bar{\pi}_i(\mu,X_m,Y_m,\alpha)=\alpha z_i(\mu,X_m,Y_m)$$

and

$$\bar{\pi}_j(\mu,X_m,Y_m,\beta)=\beta z_j(\mu,X_m,Y_m)$$

where $z_i(\mu,X_m,Y_m)$ and $z_j(\mu,X_m,Y_m)$ are vectors of functions, and α and β are vectors of parameters.

This assumption is common in the literature, and not very restrictive since the functions z_i and z_j may themselves be non-linear.

3.1.2 Profit Function Specifics

In this section, I specify the insurer and physician profit functions used in estimation, after a discussion of some related identification issues.

There are limits to what can be identified in the profit function. Any variable that enters the profit function that is just a physician characteristic, or just an insurer characteristic will have the same value on both sides of the sum of revenues inequality and will cancel. The coefficients on such characteristics are therefore not identified. The same is true for a constant term. On the other hand, variables consisting of *interactions* between physician and insurer characteristics do not cancel. The coefficients on these interactions are identified. Variables that depend on the structure of the network are implicitly interactions of indicators for agents identities, and also fall in this category. For brevity, in what follows I will include only the identified interaction terms in the profit functions, anticipating that any non-interacted characteristics simply cancel later. However, the intention is not to imply that these characteristics do not contribute to profit, but rather that they are not of primary interest since their effect cannot be identified in this setting. In part because of the issue of these canceling terms, estimated profit does not have a cardinal interpretation and requires a normalization. Estimated coefficients reveal the relative importance of contributors to profit, but no absolute dollar value can be inferred.

Insurer Profit Function

Insurers receive revenue from premiums paid by consumers and, in the case of Medicare Advantage, capitation payments paid by Medicare. Revenue therefore depends on the number of consumers who enroll[20] in the insurer's plans, which in turn depends on how appealing the plans are to consumers compared to other plans. An important component of quality is the quality of the insurer's physician network. As detailed below, the quality of the network will be measured in terms of the geographic coverage of the network and the breadth of specialties offered. Of course, the plans' premiums, benefits, and co-payments are another important component of insurer quality. Since these are examples of constant insurer characteristics (conditional on a fixed portfolio of plans) these are not included in the profit function specification.

Insurer costs reflect the expense of providing health care services to enrollees. Since a major justification for insurers offering limited networks is to control costs, features of the physician network also affect costs. A principle of managed care is to use primary care physicians as gatekeepers to regulate patients' access to specialists. Thus, the specialty composition of the network also influences cost.

Based on these observations about insurer revenue and cost, I posit the following form for insurer profit:

$$\bar{\pi}_i(\mu) = \beta_1 zip\_count_i(\mu) + \beta_2 zip\_pop_i(\mu) + \beta_3 spec\_count_i(\mu) + \beta_4 spec\_conc_i(\mu) \qquad (8)$$

The variables zip_count_i and zip_pop_i measure the geographic coverage of the physician network associated with insurer i. Gruber and McKnight (2014) emphasize that a consequence of limited network plans may be that patients travel further to visit providers. Since patients generally prefer closer providers, all else equal,[21] the location of physicians

[20]In addition to the number of consumers enrolling, the demographic composition of enrollees also matters because capitation payments are risk adjusted. I don't directly model anything related to patient demographics.

[21]For example, in the context of hospitals, Gowrisankaran, Lucarelli, Schmidt-Dengler and Town (2011) show that distance has a significant effect on patient choice.

with respect to patients could be an important way that networks are differentiated. Both variables are based on the network physicians' office postal codes. The variable zip_count_i is the number of unique postal codes represented by the physicians in the insurers network. It measures the network's geographic breadth. The variable zip_pop_i is meant to capture the extent to which physicians in the network are located near patients. I take the total population of each postal code, to measure the population of "potential patients" in the postal code.[22] Then, for insurer i, the populations of the postal codes for each physician linked to plan i are summed and divided by the number of physicians in the network to get the zip_pop_i measure. Therefore, an insurer with a higher value of zip_pop_i has physicians located in more populated areas on average.

The $spec\_count_i$ and $spec\_conc_i$ variables measure the specialty composition of the insurer's network. The variable $spec\_count_i$ is a count of the distinct physician specialties in the insurer's network. The variable $spec\_conc_i$ measures how physicians in the insurer's network are spread across specialties. It is calculated like an HHI. A share is computed for each specialty, by dividing the number of network physicians in that specialty by the total number of network physicians. The $spec\_conc_i$ variable is formed by taking the sum of the squares of these shares.

Physician Profit Function

The benefit to physicians of belonging to insurers' networks is access to a pool of patients. Physicians are compensated by the insurer for treating patients either through a fixed capitation payment per patient, or a fee-for-service scheme (accounted for by the transfers), plus a co-payments from patients depending on the specifics of the insurance contract. Physicians in the same network can be either substitutes or compliments. Substitutability arises from patients choosing one physician over another. Two physicians of the same specialty would generally be substitutes. Complementarity comes from a physician generating business for another physician, such as through referrals. For example, a primary care physician may be a complement to a oncologist if the primary care physician screens patients for cancer then refers those with positive results to the oncologist.

Interaction terms between physician and insurer characteristics are also included in the profit function. For interactions of a single physician and single insurer characteristic, what is identified is the contribution to joint surplus. Therefore, the choice to make these part of the physician profit function instead of the insurer profit function is for expositional and notational convenience, and this should be kept in mind when interpreting the resulting coefficients. An interaction indexed by j is the sum of the interaction terms for physician j for all of the insurers i that j is matched with.

The specification of the profit function for physician j is:

$$\bar{\pi}_j(\mu) = \alpha_1 ref\_net_j(\mu) + \alpha_2 med\_inter_j(\mu) + \alpha_3 years\_inter_j(\mu) + \alpha_4 years\_inter\_sqr_j(\mu) + \alpha_5 excl\_inter_j(\mu) \tag{9}$$

The variable ref_net_j measures the size of physician j's "referral network." It is defined as the total number of physicians with whom physician j has at least one insurer in common. This is theoretically the set of physicians from whom physician j could receive referrals of Medicare Advantage patients.

[22] These are total population measures, not numbers of Medicare eligibles, so not all members of the population are potential enrollees in Medicare Advantage. Nonetheless, the Medicare eligible population likely correlates strongly with total population.

On the insurer side, the variable entering into the interaction terms is the indicator for whether the insurer operates nationally. As discussed earlier, national insurers are inherently different because they benefit from economies of scale that regional insurer cannot access. National insurers may therefore be able to offer higher quality for a given level of cost. Each of the interaction terms in the physician profit function consists of a physician characteristic interacted with this indicator.

On the physician side, the first three interacted characteristics are the physician's years of experience, years of experience squared, and indicator for attending a U.S. medical school. These variables are meant to capture aspects of physician quality. Interactions between these three variables and the national insurer indicator make up the terms med_inter_j, $years\_inter_j$ and $years\_inter\_sqr_j$.

The last physician-side characteristic is a measure of how exclusive the physician is– that is, whether the physician links to one or many insurers. It is constructed as a count of the number of links that the physician has (so that a lower number indicates greater exclusivity.) The exclusivity index functions like a fixed physician characteristic because it never varies for a given physician in the sum of revenue inequalities. As matches are hypothetically swapped, the number of matches that a given agent has is always kept constant. Exclusivity is important for the value of a physician to an insurer. Insurers may be better able to control costs when they have exclusive contracts with physicians. Gaynor and Mark (2002) point out that physicians are more likely to change their practice style to conform to incentives provided by an insurer when that insurer's patients comprise a large portion of the physician's total income. The term $excl\_inter_j$ is the exclusivity measure interacted with the national insurer indicator.

4 Estimation

Estimation is carried out using the Maximum Score estimator of Fox (2010).

Data from multiple markets is used in the estimation. Let M be the set of all markets and let $m = 1, 2, ...|M|$ index markets. Let μ_m be the observed matching in market m. Let I_m be the set of all (i, j, i', j') such that $(i, j), (i', j') \in \mu_m$ and $(i', j), (i, j') \notin \mu_m$. Then, one valid sum of revenues inequality (for the deterministic part of profit, as in (7)) can be formed for each element of I_m. In practice, only a randomly chosen subset of the theoretically valid inequalities are used in the estimation. Let $\hat{I}_m \subseteq I_m$ be that subset.

The basic idea of the estimation is to find the parameter values that maximize the number of theoretically valid sum of revenues inequalities that hold for the observed matching. Recall that under assumption (2), matchings that satisfy the corresponding sum of revenue inequalities are more likely to be observed than those which do not. Intuitively, parameters that lead to more of the inequalities being satisfied are more likely to have generated the observed matching.[23] The objective function is:

$$\begin{aligned}
&Q_M(\alpha, \beta) = \\
&\quad \frac{1}{|M|} \sum_{m \in M} \sum_{(i,j,i'j') \in \hat{I}_m} 1[\bar{\pi}_i(\mu_m, \alpha) + \bar{\pi}_{i'}(\mu_m, \alpha) + \bar{\pi}_j(\mu_m, \beta) + \bar{\pi}_{j'}(\mu_m, \beta) \geq \\
&\qquad \bar{\pi}_i((\mu_m^{(i,j,i',j')}, \alpha) + \bar{\pi}_{i'}(\mu_m^{(i,j,i',j')}, \alpha) + \bar{\pi}_j(\mu_m^{(i,j,i',j')}, \beta) + \bar{\pi}_{j'}(\mu_m^{(i,j,i',j')}, \beta)] \qquad (10)
\end{aligned}$$

[23]Fox (2010) discusses formal consistency of the estimator.

where:

$$\mu_m^{(i,j,i',j')} = \mu \setminus \{(i,j),(i',j')\} \cup \{(i,j'),(i',j)\}$$

The estimator is:

$$\alpha^*, \beta^* = \underset{\alpha,\beta}{\operatorname{argmax}}\; Q_M(\alpha,\beta) \tag{11}$$

The estimator is the parameter vector (α^*, β^*) which maximizes the number of sum of revenue inequalities which hold.

In practice, estimation proceeds as follows:

1. For each market m, construct the set I_m which indexes all theoretically valid sum of revenues inequalities in the market. The set consists of all (i,j,i',j') such that $(i,j),(i',j') \in \mu_m$ and $(i',j),(i,j') \notin \mu_m$.

2. Randomly select a subset $\hat{I}_m \subseteq I_m$ to use in the estimation. Within the market, each inequality should be chosen with equal probability.

3. For every element in $\hat{I}_m$, construct the vectors of regressors, $z_i(\mu_m, X_m, Y_m)$, $z_{i'}(\mu_m, X_m, Y_m)$, $z_j(\mu_m, X_m, Y_m)$, $z_{j'}(\mu_m, X_m, Y_m)$, $z_i(\mu_m^{(i,j,i',j')}, X_m, Y_m)$, $z_{i'}(\mu_m^{(i,j,i',j')}, X_m, Y_m)$, $z_j(\mu_m^{(i,j,i',j')}, X_m, Y_m)$, and $z_{j'}(\mu_m^{(i,j,i',j')}, X_m, Y_m)$.

4. For a given candidate parameter vector (α, β) evaluate $Q_H(\alpha, \beta)$. Doing so requires evaluating the sum of revenues inequality corresponding to every element in each $\hat{I}_m$. For a given element (i,j,i',j') in market m:

   a. Calculate $\bar{\pi}_i(\mu_m, \alpha)$ as $\alpha z_i(\mu_m, X_m, Y_m)$, $\bar{\pi}_j(\mu_m, \beta)$ as $\beta z_j(\mu_m, X_m, Y_m)$ and so forth for all terms in sum of revenues inequality.

   b. Evaluate the sum of revenues inequality. If it holds, the sum in (10) increments by 1. If it doesn't hold, the sum doesn't increment.

   c. Repeat for next (i,j,i',j') in $\hat{I}_m$.

5. Maximize the objective function $Q_M(\alpha, \beta)$ over α and β (with one coefficient set to ± 1 to establish scale). The previous step is repeated for each candidate (α, β).

4.1 Implementation

Because it is a sum of indicator functions, the the maximum score objective function is not smooth in the parameters. Therefore, optimization routines that rely on smoothness, such as derivative-based methods, will not reliably find an optimum in this case. To overcome this difficulty, I use the Genetic Algorithm (GA) optimization routine available in Matlab. GA is intended for potentially poorly behaved functions, and uses a stochastic search algorithm to find a global optimum. In practice, GA does not always find the global optimum of the maximum score function in a single run. To reduce optimization error, I run the optimization 20 times, and take the parameters corresponding to the best objective function value.[24]

[24] The stochastic nature of the optimization routine allows for different outcomes across otherwise identical runs. The 20 runs consist of 10 for the positive value of the normalized parameter, and 10 for the negative value of the normalized parameter.

Table 3: Parameter Estimates for Profit Function

variable	coefficient	90% conf. interval
zip_count	-5.59*	(-8.88,-0.015)
zip_pop	-0.489	(-4.04,5.83)
spec_concentration	4.03*	(1.83,12.12)
spec_count	-9.18*	(-9.53,-3.57)
ref_net	-1	-
med_inter	3.02*	(2.20,8.63)
yrs_inter	0.47	(-0.45,1.79)
yrs_inter_sqr	-0.009	(-0.034,-0.015)
excl_inter	7.48*	(2.54,13.01)

Table notes: Parameter estimates are from Maximum Score estimation. The 90% confidence intervals were generated using subsampling.

In the estimation, data is pooled from five markets. In total, nearly 20 million theoretically valid sum of revenues inequalities are generated. To ease computation, a random 5% sample of the inequalities are taken to be used in the objective function. According to Fox (2010), the estimator is still consistent when only a randomly selected subset of the theoretically valid inequalities are used.[25]

4.2 Confidence Intervals

Confidence intervals for the parameters are calculated using a method called *Subsampling.*[26] Similar to the bootstrap, subsampling entails drawing random samples from the data and using the empirical distribution of parameters across samples. In subsampling, sampling is without replacement, and each subsample is smaller than the full data set. Here, subsampling is done on the level of physicians. For each subsample, 30% of physicians in the data are randomly selected. Inequalities involving physicians who both belong to the selected subset are kept and the rest discarded. The estimation is then run on this subset of inequalities. This procedure is repeated 20 times, generating 20 vectors of parameters, each from a different subsample of physicians. Ninty percent confidence intervals are then computed from the empirical distribution of these parameters.

5 Results

Estimation results are presented in Table 3. The percentage of inequalities satisfied at the optimal parameters is a measure of goodness of fit of the model. Here, 62.7% were satisfied, indicating that the model has some explanatory power for the matching patterns, but there is still substantial room for improvement. Some other papers using the matching maximum score estimator have reported 80% or more of the inequalities satisfied.

[25]In practice I do not find the estimates or confidence intervals to qualitatively change across different realizations of the sample or different sample sizes.

[26]See Politis et al. 2012 for more on subsampling. Fox (2010) notes that subsampling is consistent for the Maximum Score Estimator while the Bootstrap is not.

The magnitude of the coefficient on ref_net, the physician's referral network size, was normalized. While the magnitude was restricted to be 1, the sign was determined through estimation, and is therefore meaningful. The estimation produced a negative sign, so the reported coefficient is -1. Because of the normalization, no standard errors are reported for this coefficient. The negative sign of the coefficient is surprising. It indicates that physicians benefit from having a smaller referral network. One explanation is that the effect of competition dominates. Competing with a larger group of physicians for an insurer's patients could outweigh the benefits of having more colleagues from whom to receive referrals. Furthermore, what is really being measured here is the *breadth* of the referral network: the number of unique individuals with whom the physician has links in common. An alternate possibility is that physicians prefer to have *deeper* referral networks, in the sense of having multiple insurers in common with each colleague. A deeper network might facilitate repeated interactions and better knowledge about others in the referral network.

The two variables describing geographic features of the insurer's network, *zip_count* and *zip_pop* both have negative coefficients, though only the coefficient on *zip_count* is statistically significant at the 90% level. *Zip_count* is a count of the distinct zip codes within the county covered by offices of physicians in the insurer's network. The negative sign indicates that on the margin insurers benefit more from having a physician in an already covered zip code than from having a physician in a new zip code. In other words, insurers tend to geographically specialize. The underlying cause could be an agglomeration-type effect, where there are positive spillovers between physicians who are near each other and linked to the same insurer. On the other hand, it could indicate that physicians who are located in more remote areas of the county are less profitable for all insurers. In either event, this geographic specialization means that a given insurer's network will not be equally appealing for all consumers. Those who live near an insurer's "cluster" will have a better selection of conveniently located physicians than those who do not.

The variables *spec_concentration* and *spec_count* are related to the range of specialties offered in an insurer's network. The estimated coefficient on *spec_concentration* is positive and the coefficient on *spec_count* is negative. Both are statistically significant. Together, these coefficients suggest that insurers benefit more from having additional physicians in specialties that are already well represented in the network, compared to adding additional specialties or adding physicians to the least covered specialties currently in the network. In other words, insurers seem to prefer having many physicians in a few specialties to having a few physicians in every specialty. Forming the network in this way would likely be good for many consumers, while harmful for a few consumers. Insurers are incentivized to contract with plenty of physicians in the most common specialties, like primary care, anesthesiology and cardiology. Patients who need the most common services therefore have plenty of options. However, patients who need the services provided by more obscure specialty types might find the networks to be very restrictive in those specialties.

The remaining terms in the surplus function are interactions between physician and insurer characteristics. The interaction between the indicator for a physician with a U.S. medical degree and a national insurer, *med_inter* has a positive and significant coefficient. The coefficient on the interaction between a physician with a single, exclusive link and a national insurer, *excl_inter*, is also positive and significant. Based on these results, physicians with a U.S. Medical degree and physicians who contract selectively with insurers are more likely to match with national insurers. Insofar as these variables are measures of physician and insurer quality, these results point towards positive assortative matching between physicians and insurers. On the other hand, the interaction of physician years of experience (and years of experience squared) with the national insurance indicator, *yrs_inter* (respectively *yrs_inter_sqr*) led to statistically insignificant coefficient estimates.

6 Pseudo-Counterfactuals

In this section, I discuss exercises in which surplus is compared across different matching configurations. I calculate total deterministic surplus across all agents in the model for the observed matching using the estimated parameters. Then, I repeat this calculation at different matchings. These include a full networks matching, where all physicians are linked to all insurers, and a maximum exclusivity matching where each physician is randomly assigned to exactly one insurer.

Consider a matching μ. Let $(\hat{\alpha})$ and $(\hat{\beta})$ be the estimated parameters of the surplus functions. The total deterministic surplus from the matching at the estimated parameter values is simply the sum of each agent's surplus evaluated at the given matching and parameter values:

$$\hat{\bar{\Pi}}^{total}(\mu, \hat{\alpha}, \hat{\beta}) = \sum_{i \in I_m} \bar{\pi}_i(\mu, \hat{\alpha}) + \sum_{j \in J_m} \bar{\pi}_j(\mu, \hat{\beta}) \tag{12}$$

This total surplus expression can be evaluated at any matching.

I call these exercises "psuedo-counterfactuals" because they differ from the typical counterfactuals that accompany structural estimation in two ways. First, I do not claim that these matchings represent an equilibrium outcome. Since the game is not fully specified, and the estimation procedure avoids having to fully solve a game, there is no way to solve for new equilibria without imposing additional assumptions. With the pseudo-counterfactuals, I take a different tact, evaluating total surplus at matchings chosen for their policy interest which are probably not equilibria. Since there is no guarantee that joint surplus would be maximized at the observed equilibria, it is theoretically possible to find other matchings with greater total surplus, and it is informative to see how different matchings affect total surplus.

The second limitation of the pseudo-counterfactuals is that only the deterministic and not the stochastic part of total surplus can be evaluated. This limitation comes about because the stochastic terms $\epsilon(\mu)$ and $u(\mu)$ are not estimated. For the sake of the estimation procedure, as long as these terms are such that the rank order condition is satisfied, their exact values are unimportant. In the case of the counterfactuals, the stochastic terms should be included in the total surplus calculations to get a full accounting of surplus. On the other hand, since the rank order condition in some sense limits how influential the stochastic terms can be in relation to the deterministic part of surplus, one can at least hope that omitting the stochastic terms will not change the ranking of total surplus across the different matchings in the pseudo-counterfactuals.

This way of doing counterfactuals follows ideas from Bajari and Fox (2013). They study allocation of licenses in the FCC Spectrum Auctions as a matching game. They define a measure of efficiency for license allocations that depends on the deterministic part of a surplus function for portfolios of licenses, which corresponds to the common value component of the license valuation. The omitted stochastic terms are the bidders

Table 4: Pseudo-Counterfactuals

County	Baseline Matching	Full Matching	Exclusive Matching
Maricopa, AZ	691,703	2,439,641	75,578
Pima, AZ	141,097	282,571	17,406
Yavapai, AZ	11,726	17,291	3,447
Clark, NV	102,606	217,501	27,026
Washoe, NV	26,401	46,819	9,388

Table notes: All numbers are in arbitrary units of surplus and should only be compared ordinally. Surplus is calculated at various matchings using the estimated parameters. "Baseline matching" is the matching observed in the data. "Full matching" matches all physicians to all insurers within each market. "Exclusive matching" randomly assigns each physician to exactly one insurer, putting equally probability on each insurer. The reported surplus the average over 10 such random assignments.

unobserved private values. They rank the efficiency of allocations based on common value only. While in my case the neat interpretation from the auction setting does not apply, the concept is the same.

A final caveat is that the surplus numbers only account for physician and insurer surplus. Since consumer choices are not explicitly modeled, there is no way to evaluate consumer welfare separately. Matchings that make physicians and insurers better off could do so by making consumers worse off.

Results from the pseudo-counterfactuals are in table (4). The 'Baseline Matching' column reports the total surplus at the original matching observed in the data in each county. The 'Full Matching' column reports total surplus when every physician is linked with every insurer in the county. The 'Exclusive Matching' column reports surplus when each physician is linked to exactly one insurer. The insurer each physician is linked to is chosen randomly from the insurers in the county with equal probability. Ten random matchings with exclusivity are generated in this way, and the number reported is the mean surplus across the matchings.

The total surplus numbers are in arbitrary units and should be interpreted only ordinally.[27] What is meaningful is the ranking by total surplus of the matchings within a county. In all counties, total surplus is the greatest in the full matching, followed by the baseline matching, followed by the exclusive matching. It is unsurprising that the exclusive matching performs poorly. Since matches are made randomly, many physician-insurer complementarities from the baseline matching are not realized. Furthermore, the geographic and specialty concentration that benefits total surplus is diluted in these matchings. The full matching performs best in every county. Compared to the baseline matching, specialty and geographic concentration goes down, and the referral network measure increases. These effects are surplus decreasing, but given the parameter values the impact is small. Since every physician is matched to every insurer, all possible complementarities from interactions are realized, which increases total surplus dramatically. This result is striking, but many robustness checks are necessary to verify that it is not

[27] To get a rough idea of what surplus in this industry is like in dollar amounts, consider that according to the CMS total national expenditures on private health insurance was $991 billion in 2014. This paper considers only five counties, and only the insurers participating in Medicare Advantage, but given the overall scale of the insurance industry, total profits could be quite large.

an artifact of the particular way of specifying the surplus functions.

The dominance of the full matching over the baseline matching in terms of total surplus has policy implications. According to these results, insurers and physicians would be collectively better off under regulations that prevented limited networks. While the baseline matching is presumably an equilibrium outcome and individual insurers would not have an incentive to unilaterally deviate and move to full networks, if all insurers and physicians change to this model total industry surplus would increase. Of course, the missing piece in the welfare analysis is consumers. All else equal, greater access to providers should be welfare increasing for consumers, but if insurers can choose different prices and plan benefits at different network configurations, it becomes less clear what is best for consumers. In related work (Nosal 2012), I show that there are very large switching costs (over $4000) associated with changing Medicare Advantage plans, which may be driven by network differences across plans that imply that consumers switching plans are often forced to switch providers. Others find that consumers have high willingness to pay for more generous networks (Gruber and McKnight, 2014 and Ericson and Starc 2014). In light of the high valuation of increased access to providers, it is plausible that consumers would be better off with less restricted networks, even if prices are higher. Therefore, policies which encourage full (or at least broader) networks may increase total welfare.

7 Planned Extensions

I plan to improve the specification of the surplus function by including the following elements:

- *Hospitals.* Hospitals are an important complement to physicians. I have collected data on the hospitals in each insurer's network. The hospitals in the network can be treated as a fixed insurer characteristic (assuming, for example, that hospital-insurer contracts are negotiated for a longer term than physician-insurer contracts.) The number and quality of hospitals that an insurer is connected to is a measure of insurer quality. A physician may care about insurer connections to hospitals geographically near the physician's office, or hospitals where the physician has admitting privileges. Any of these elements could be incorporated in the surplus functions.

- *Plan Ratings.* CMS releases annual ratings for Medicare Advantage plans. Some components of the ratings are related to health outcomes of enrollees, and would just as easily be a consequence of the network configuration as a cause of it. However, other measures like ratings of the plans' customer service can be interpreted as pure insurer quality measures. These could be included in the surplus function interacted with physician characteristics to further test for assortative matching.

- *Physician Practices.* While inclusion in insurer networks is technically on the physician level, physicians may make correlated decisions within practices because of administrative efficiencies from accepting the same plans. Data on practice membership would make it possible to allow for complementarities when physicians from the same practice match with the same insurer.

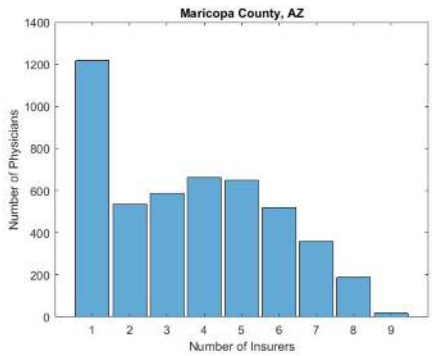
- *Referral Networks.* The current specification of the referral network function does not allow for asymmetries in how physicians interact based on specialties. For example, a specialist may benefit more from sharing insurer connections with primary care physicians than with other specialists, especially those of the same specialty. A more flexible referral network specification can allow for these asymmetries.

8 Conclusion

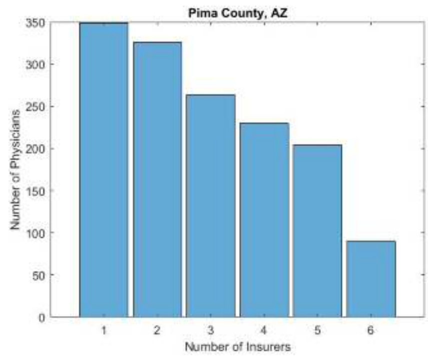
Two themes are apparent in the estimation results. First, the geography and specialty coefficient estimates point to a tendency towards specialization. Insurers do not include as many specialties and office locations as possible in the network, but rather prefer to offer more physicians in a smaller set of specialties and locations. Since consumer preferences are not modeled, there is no way to directly quantify the welfare implications for consumers. However, there likely would be both winners and losers. Consumers whose needs coincide with the geographic and specialty areas that the insurers focus on may perceive an abundance of choice of physicians. On the other hand, consumers who live in more remote parts of the county or who need specialty care from less common specialty types may not be able to find a Medicare Advantage plan with a network that suits their needs. Of course, such consumers can always choose original Medicare, which puts fewer restrictions on choice of providers. The second theme is that there is some evidence of positive assortative matching based on physician and insurer characteristics. If this result applies more broadly, in the sense that insurers that offer plans with better benefits offer better networks, this is good news for consumers. In this case, consumers would not have to trade off choosing a high quality plan and one that offers an appealing provider network.

The pseudo-counterfactual results imply that physicians and insurers would be collectively better off if networks were complete, in the sense that all physicians match with all insurers. This finding suggests that legislation requiring broader networks may actually help the industry as a whole. Assuming broader networks are also good for consumers, total welfare could be increased with such a policy.

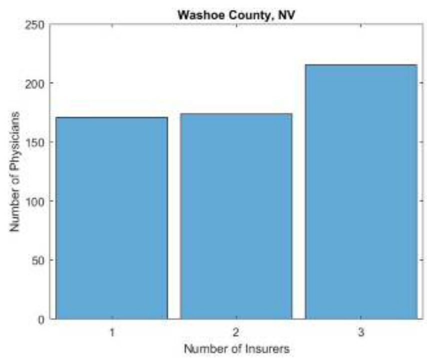
Figure 1: Distribution of Physicians' Number of Networks



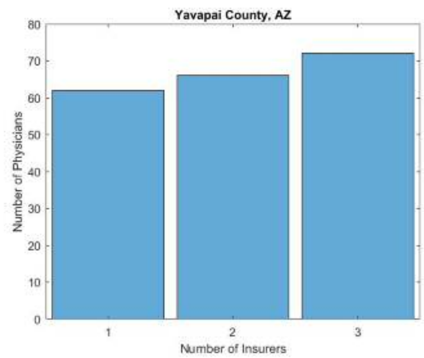
(a) Maricopa County, AZ



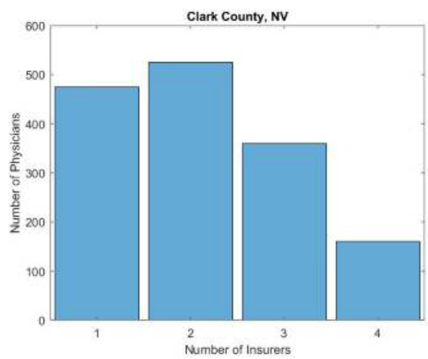
(b) Pima County, AZ



(c) Washoe County, NV



(d) Yavapai County, AZ



(e) Clark County, NV

Figure notes: In the graphs, each bar represents a number of insurers, ranging from one to the total number of MA insurers in the county. The height of the bar indicates the number of physicians who are in the network of exactly that number of insurers.

References

[1] Bloch, F., and Jackson, M. O. (2006). "Definitions of equilibrium in network formation games." *International Journal of Game Theory*, 34(3), 305-318.

[2] Center for Disease Control. (2013). *Summary Health Statistics for U.S. Population: National Health Interview Survey, 2012.* Vital and Health Statistics, Series 10, Number 259. Retrieved from http://www.cdc.gov/nchs/data/series/sr_10/sr10_259.pdf

[3] Center for Disease Control. (2014). *Summary Health Statistics for U.S. Adults: National Health Interview Survey, 2012.* Vital and Health Statistics, Series 10, Number 260. Retrieved from http://www.cdc.gov/nchs/data/series/sr_10/sr10_260.pdf.

[4] Chen, J., and Song, K. (2013). "Two-sided matching in the loan market." *International Journal of Industrial Organization*, 31(2), 145-152.

[5] Chernew, M. E., Wodchis, W. P., Scanlon, D. P., and McLaughlin, C. G. (2004). "Overlap in HMO physician networks." *Health Affairs*, 23(2), 91-101.

[6] Dahl, G. B., and Forbes, S. J. (2014) "Doctor Switching Costs in Health Insurance." Working Paper.

[7] Ericson, K. M., and Starc, A. (2014). "Measuring Consumer Valuation of Limited Provider Networks" Working Paper no. w20812, National Bureau of Economic Research.

[8] Ericson, K. M. M. and Starc, A. (2013). "How Product Standardization Affects Choice: Evidence from the Massachusetts Health Insurance Exchange" Working paper no. w19527, National Bureau of Economic Research.

[9] Fox, Jeremy (2010). "Estimating Matching Games with Transfers." NBER Working Paper.

[10] Fox, Jeremy, David Hsu and Chenyu Yang. (2015). "Unobserved Heterogeneity in Matching Games with an Application to Venture Capital. " Working Paper, Rice University.

[11] Fox, Jeremy T., and Patrick Bajari. (2013) "Measuring the efficiency of an FCC spectrum auction." *American Economic Journal: Microeconomics* 5(1): 100-146.

[12] Gaynor, Martin and Tami Mark. (2002). "Physician Contracting with Health Plans: A Survey of the Literature" In Teh-Wei Hu and Chee-Ruey Hsieh (eds.), *The Economics of Health Care in Asia-Pacific Countries*, Cheltenham, UK: Edward Elgar, 2002.

[13] Gowrisankaran, G., Lucarelli, C., Schmidt-Dengler, P., and Town, R. (2011). "The impact of the Medicare Rural Hospital Flexibility Program on patient choice." *International Journal of Industrial Organization*, 29(3), 342-344.

[14] Gowrisankaran, Gautam, Aviv Nevo, and Robert Town (2015). "Mergers When Prices Are Negotiated: Evidence from the Hospital Industry." *American Economic Review* 105.1: 172-203.

[15] Gruber, J., and McKnight, R. (2014). "Controlling health care costs through limited network insurance plans: Evidence from Massachusetts state employees" Working paper w20462, National Bureau of Economic Research.

[16] Ho, Katherine. (2006) "The welfare effects of restricted hospital choice in the US medical care market." *Journal of Applied Econometrics* 21.7: 1039-1079.

[17] Ho, Katherine. (2009). "Insurer Provider Networks in the Medical Care Market." *American Economic Review*, 99(1), 393-430.

[18] Ho, Katherine and Robin Lee (2013). "Insurer competition and Negotiated Hospital Prices" Working Paper w19401, National Bureau of Economic Research.

[19] Jackson, Matthew O. and Asher Wolinsky. (1996) "A strategic model of social and economic networks." *Journal of economic theory* 71(1): 44-74.

[20] Levine, A. A. (2009). "Licensing and scale economies in the biotechnology pharmaceutical industry." Working paper.

[21] Nosal, K. (2012) "Estimating Switching Costs for Medicare Advantage Plans." Working paper.

[22] Politis, Dimitris N., Joseph P. Romano, Michael Wolf. (2012). *Subsampling.* Springer Series in Statistics. New York, New York: Springer Science and Business Media.

[23] Pourat, N., Davis, A. C., Chen, X., Vrungos, S., and Kominski, G. F. (2015). "In California, Primary Care Continuity Was Associated With Reduced Emergency Department Use And Fewer Hospitalizations." *Health Affairs*, 34(7), 1113-1120.

[24] Roth, A. E., and Sotomayor, M. A. O. (1992). *Two-sided matching: A study in game-theoretic modeling and analysis.* Econometric Society Monographs. Cambridge University Press.

[25] Town, Robert, Robert Feldman and John Kralewski. (2011). "Market Power and Contract Form: Evidence from Physician Group Practices." *International Journal of Health Care Finance and Economics*, 11, 115-132.

[26] Yang, Y., Shi, M., and Goldfarb, A. (2009). "Estimating the value of brand alliances in professional team sports." *Marketing Science*, 28(6), 1095-1111.