WORKING PAPER: COMMENTS WELCOME

The Local Influence of Principal Investigators on Technology Adoption

Evidence from New Cancer Drugs

Leila Agha, Boston University
David Molitor, Stanford University *

May 1, 2013

Abstract

Adoption of new health care technologies is widely considered to be a key driver of both rising health costs and improved outcomes in the United States. This paper explores the diffusion of new cancer drugs by testing the influence of physician investigators who lead clinical trials. In particular, we exploit variation across drugs in the location of clinical trials to test whether geographic proximity to a principal investigator influences the speed of technology adoption. Using original data on clinical trial study authors and sites for 21 new cancer drugs along with Medicare claims data from 1998-2008, we estimate that patients are 30% more likely to receive treatment with a new drug if they seek care in the hospital referral region where the drugÕs principal investigator (PI) practices. This effect, which is estimated in the first two years following initial FDA approval, fades over time until there is no apparent difference in utilization after four years. The increased local utilization is not driven solely by physicians practicing at the PI's hospital, but is concentrated within PI's metropolitan area. The evidence suggests that early information about a new drug can lead to large regional differences in initial adoption rates, but does not drive persistent specialization in treatment with the new technology.

*We would like to thank Amitabh Chandra, Joseph Doyle, Amy Finkelstein, and Jonathan Gruber, as well as seminar participants at MIT for helpful comments and suggestions. This research was supported by the National Institute on Aging, grant number T32-AG000186.

1 Introduction

There is broad consensus that the adoption of new medical technologies is a key factor underlying both rising health costs and improved outcomes in the United States (Newhouse, 1992; Cutler, 1995; CBO, 2008; Smith, Newhouse, and Freeland, 2009). This potential role of technology adoption to serve both as an engine of productivity growth in the health sector as well as a source of pressure on public and private budgets through increased costs highlights the importance of understanding the process by which medical technologies diffuse across medical providers. Despite its importance, surprisingly little is known about what actually drives some providers to adopt innovations more quickly than others. Existing evidence has largely focused on provider characteristics associated with adoption behavior, as in the pioneering work by Coleman, Katz, and Menzel (1957) documenting that physicians with greater social connections were faster to adopt a new antibiotic. Yet providers who are quick to adopt new technologies may likely differ from slow adopters in many correlated and potentially unobserved respects, making it difficult to tease out causal adoption influences from associations alone.

In this paper, we study the effect of geographic proximity to principal investigators (PIs) of new medical technologies on the speed of adoption of those technologies. The central idea behind this spatial mechanism is that when uncertainty surrounds the underlying benefits of a new medical technology, proximity to the PI may increase a potential adopter's informational advantage. This information could spread through face-to-face social interactions (à la "see one, do one, teach one"), or it could result from a salience effect such as if physicians tend to pay special attention to the activities of local opinion leaders.

While existing evidence on the role of geography on medical technology adoption is limited, a few empirical studies have found evidence consistent with this possibility. For example, technology adoption appears to cluster at local levels such as the hospital (Escarce, 1996) or region (Baicker and Chandra, 2010), suggesting a potential role for localized knowledge spillovers. Perhaps the clearest existing evidence comes from Harpaz-Rotem and Rosenheck (2009) who find that an off-label drug treatment for posttraumatic stress disorder developed at a Veterans Affairs medical center diffused with a steeply negative geographic gradient across other VA centers, despite apparently minimal financial or organizational barriers.

The context of our study is the adoption of new cancer drugs by physicians. Specifically, we examine whether new cancer drugs approved by the Food and Drug Administration (FDA) are

adopted more rapidly in the geographic region containing the principal investigator of the pivotal clinical study used in the FDA review process. A key feature of the empirical context is that there may be substantial uncertainty about the efficacy and appropriate applications of a newly introduced drug; the PI's detailed knowledge of drug mechanisms, patient responses and side effects may put him at a significant informational advantage in the early stages of a drug's diffusion. Many new cancer drugs are approved based on promising evidence for a narrowly defined indication. For example, many clinical trials are conducted on patients whose cancers have relapsed after initial treatment, so the efficacy of the new drug as an initial treatment is not yet established upon drug approval. In addition, many cancer drugs come with side effects that range from temporary but severely uncomfortable (e.g. nausea, fever, pain) to serious or life-threatening (e.g. kidney failure, lung damage, nerve damage, secondary cancers). A host of other drugs and additional monitoring may be required to mitigate these side effects, and physicians may develop expertise in this management over time.

A central challenge to identifying the proximity channel empirically is that innovative activity may tend to take place in regions in which new technologies are adopted quickly for reasons other than proximity. For example, if rates of adoption of a new drug with a Boston-based PI are especially high in the Boston region, we would like to interpret this as a result of proximity to the PI. However, an alternative reason for this association could be that Boston physicians simply tend to adopt all new drugs—or at least a class of drugs in which this drug falls—much faster than the national average. To address this issue, we adopt a controlled approach in the spirit of Jaffe, Trajtenberg, and Henderson (1993). Specifically, we construct for each geographic region the average "control" rate of adoption of a class of new drugs relative to the national rate of adoption, excluding any drugs for which the PI was based in that region. Our test for proximity effects amounts to asking whether the rate of adoption of a new drug is higher in the PI's region than the control rate.

While the context of our study is medical care, our work contributes to a broader literature on the role of geography and the diffusion of innovations. It has long been argued that geographic proximity facilitates the spread of knowledge and innovations across individuals or firms (Marshall, 1890). Existing empirical studies of the role of geographic spillovers have primarily focused on the creative process of new ideas and technologies, such as the tendency for inventors to cite patents developed in their geographic region (Jaffe, Trajtenberg, and Henderson, 1993), or for academic

citation patterns to follow migrant scientists (Azoulay, Zivin, and Sampat, 2012).[1] Our study thus expands on this work to shed light on the geographic connection between innovative activity and the subsequent adoption of resulting technologies. Moreover, to the best of our knowledge this is the first study to exploit the diffusion of multiple similar technologies to identify geographic spillovers in technology diffusion.

Our analysis is based on a novel data collection effort that identifies the principal investigators of the pivotal clinical trials for 21 new cancer drugs, and matches the locations of these PIs to adoption patterns of the drugs using Medicare claims records from over 600,000 patient cancer care episode from 1998-2008. The key finding from our baseline analysis is that patients treated in the hospital referral region where the PI is located within the first two years following a drug's FDA approval increases the propensity to receive the new drug by nearly 4 percentage points, from a mean of roughly 12 percentage points (or a 33% increase in propensity). This proximity effect fades over time so that there is no discernible effect after four years following a drug's approval. Together, these results provide evidence that geographic proximity to a PI speeds the rate adoption of new cancer drugs but does not drive differences in longer-term levels of adoption.

We probe the geographic scope of these findings, demonstrating that the markedly faster adoption of a drug within the PI's HRR is robust to excluding the prescribing patterns of the PI himself. However, neighboring HRRs do not experience faster rates of adoption, suggesting a high degree of spillover-localization. Finally, we attempt to separate two distinct channels through which the proximity effect could operate. The first is an "adoption" effect, whereby providers in the PI's region have an increased propensity to treat a fixed population of patients with the new drug. The second channel is patient sorting, as might occur if patients especially suitable for the drug sorted into regions specializing in the use of that drug. Using an instrumental variables strategy based on whether patients reside in the PI's HRR, we find evidence that roughly 70% of the proximity effect is driven by the adoption effect, while the remaining 30% of the proximity effect comes from patients sorting into the PI's region.

At a broader level, our results can also be interpreted in light of common models of regional disparities in medical treatment choices and technology adoption. First, the Roy model of medical treatment choice by Chandra and Staiger (2007) implies that in the presence of productivity spillovers, greater physician experience with a particular treatment may lead to steady-state specialization in that treatment relative to alternatives. While not strictly a rejection of this model,

[1] Audretsch and Feldman (2004) provide an extensive review.

our finding that early experience with new cancer drugs due to PI proximity does not lead to long-run differences in utilization suggests that productivity spillovers are not a key factor in explaining regional disparities in cancer drug utilization.

Second, the evidence our results shed on why some regions adopt medical technologies faster than others may provide insight into what underlies large regional variation in health care productivity, both across U.S. regions as well as across countries. As shown by Parente and Prescott (1994), even small barriers to technology adoption can result in large and persistent productivity disparities across regions. Skinner and Staiger (2009) draw on this model to show that in the case of heart attack management, speeds of adoption by hospitals of highly effective technologies explain large differences in productivity gains across hospitals. One loose interpretation of our results is that regions with higher levels of investigational activity remain as much as four years ahead of the curve as new technologies continue to be introduced, suggesting that funding clinical research can reduce the barrier to medical technology adoption. Whether this lower barrier increases productivity growth as in the Parente and Prescott (1994) remains an important question for future research.

The organization of the remainder of the paper is as follows. Section 2 describes the empirical context and key data elements. Section 3 lays out the primary empirical strategies and results. Section 4 probes the extent of patient sorting, and Section 5 concludes.

2 Setting and Data

There are two key data elements necessary for our analysis: the utilization of new chemotherapy drugs across regions over time, and the location of principal investigators who lead the pivotal clinical trials on which each drug's initial FDA approval was based.

2.1 Measuring chemotherapy use

We measure chemotherapy diffusion using Medicare Part B reimbursement claims over the 11-year period 1998-2008. Over the period of our study, 21 new chemotherapy agents covered by traditional Medicare were approved by the FDA. The diffusion of these drugs forms the basis of our analysis.

While traditional Medicare does not pay for most outpatient prescription drugs, an exception is made for drugs that are not typically self-administered, including chemotherapy drugs administered intravenously or intramuscularly. These expenses have comprised a rising proportion of Medicare

spending in recent years. In 2004, Medicare Part B spent $11 billion on drugs, a category dominated by chemotherapy expenses; these costs rose 267% in the 7-year period since 1997, as compared to a 47% rise in total Medicare spending (Bach, 2009). The growth has been driven by both the rapid growth in the prices charged for new chemotherapeutic agents and rising rates of drug use.

We analyze drug use at the level of a Hospital Referral Region (HRR), as defined by the Dartmouth Atlas for Healthcare, which partitions U.S. zip code areas into 306 distinct regions. Regions are defined by where the majority of the population in each zip code are referred for tertiary health care services, and are commonly used as the unit of analysis for cancer care (see e.g Fisher et al., 2003a,b; Onega et al., 2008). For some analyses, we also consider Hospital Service Areas (HSA), which partitions zip codes more finely into 3,436 regions based on where patients in each zip are typically hospitalized. The key thing to note is that over 80% of patients treated for cancer within an HRR are treated within a single hospital service area containing the area's large tertiary care center, so much of the within-region diffusion occurs in a single hospital.

To track the adoption and use of the 21 chemotherapy agents, we analyze a 100% sample of Medicare outpatient claims as well as a 20% sample of Medicare physician carrier claims. For each provider region, we estimate the fraction of appropriate patients who receive a given new chemotherapy drug. To do this, we first define a patient chemotherapy treatment episode to include all chemotherapy claims for a patient within a given provider region (e.g. HRR) and calendar year. For each patient episode, we observe any cancers for which the patient received treatment, as well as the chemotherapy drugs administered to the patient. Our data comprise 3.7 million chemotherapy care episodes in our main specification, with provider regions defined as HRRs.

2.2 Principal investigator locations

In addition to the Medicare claims data, we also collected a new data set linking chemotherapy drugs to the clinical trial that provided the primary support for FDA approval. The data were collected through review of FDA approval history documentation and the relevant academic medical literature. While each drug application typically cites several studies from various stages of drug development, the applicant must pre-specify a "pivotal trial," which is typically a randomized trial that provides the most comprehensive evidence to date on the efficacy of the drug.[2] By matching the pivotal trial information in the FDA application to the authors of the academic article reporting

[2]Detailed information on the development and approval process for new drugs can be found on the FDA's website at http://www.fda.gov/Drugs/DevelopmentApprovalProcess.

the trial's findings, we are able to identify the researchers who were primarily responsible for the trial.

In our analysis, we exploit the convenient fact that the first author in medical journal articles represents the major contributing author (Baerlocher et al., 2007). In the case of clinical trial reports, the trial's principal investigator (PI) typically receives the first-author position; hence, for the discussion that follows we will use the author position to determine the study's PI, and use the terms PI and first author interchangeably. We categorize authors as "first," "other" and "last," and we record each author's location based on the zip code of the author's institution at the time of the article's publishing. Using this location information, provider regions can be categorized based on their geographic proximity to authors of the clinical trial. In the discussion that follows, the phrase "PI HRR" refers to the HRR where the PI for the particular drug under analysis practiced.

2.3 Drug Information and Summary Statistics

For this analysis, we identified 21 new chemotherapy drugs that may be covered by Medicare Part B.[3] Of these drugs, 17 of them had clinical trials led by researchers in the United States, and thus may be used to identify the impact of proximity to a PI on drug diffusion. The remaining four drugs are included in the sample to improve the precision of coefficients on other control variables.

The twenty-one drugs in our study are listed in Table 1, sorted by order of their FDA approval dates. These drugs target a variety of cancer types, including common carcinomas such as breast, lung, and colon cancer, as well as hematologic and urologic cancers. All drugs were FDA approved between 1998-2007. A majority of the pivotal clinical trials (13/21) were published in the Journal of Clinical Oncology; the New England Journal of Medicine was the next most frequent publication venue (4/21).

Table 1 column 8 reports the number of observed patient chemotherapy care episodes for the targeted diagnosis in the two calendar years following FDA approval. There is substantial heterogeneity in target population size due to variation in disease prevalence, ranging from 800 observed episodes of the relatively rare cutaneous T-cell lymphoma to 84,900 episodes of lung cancer. As described below, most regression results are weighted by the size of the target population to improve precision; however, note that the regression results are qualitatively similar when each drug is weighted equally, rather than by target population size.

[3] We began with a list of 26 new chemotherapy drugs that we obtained from Bach (2009). Of these drugs, five of them were billed fewer than 10 times in our sample over the first two years after approval. Given that we were not able to observe any measurable diffusion for these six agents, they were excluded from the analysis.

The first author on these trials practice at a wide set of academic medical centers that together span all four U.S. census regions. The most frequent first author locations within our sample are: Houston, Texas, (four trials); Chicago, Illinois, (three trials); Durham, North Carolina, (two trials); and New York, New York (two trials). There are 11 unique HRRs that contain a PI for at least one drug, 55 HRRs that contain a non-lead author for at least one drug (but never contain a PI), and 248 remaining HRRs that never contain any author.

Summary statistics are reported in Table 2. Over the first two years following drug approval, the average utilization rate of the new drug for indicated patient episodes ranges from 16% in regions where the PI practices to 10% in regions that never contain any investigators.

Among the regions that do not contain any investigators for the observed drug, those regions that contain PIs for other in-sample drugs are the most intensive adopters; regions that contain another drug's PI are over 20% more likely to use a new chemotherapy agent compared to regions that never contain any investigators, as reported in columns 3 and 5. Thus, PIs tend to be located in regions that have a high degree of enthusiasm, expertise, or patient demand for new chemotherapy agents in general.

Within the set of regions that contain a PI for at least one in-sample drug, the PI HRR has 40% greater utilization rates on average as compared to utilization in regions that contain no authors for that particular drug (cf. columns 1 and 3). Thus, despite the overall higher rates of new drug use in regions that contain a PI for at least one drug, utilization is even greater when the lead researcher of the particular observed drug is in the area.

In the third row of Table 2, we can observe that regions containing PIs have more chemotherapy episodes treating the targeted disease than regions with secondary authors or regions without any authors. However, within the set of regions that contain a PI for at least one drug, comparing columns 1 and 3, there are slightly larger observed target populations in those regions that do *not* contain the PI for a given drug. This suggests that PIs tend to reside in more populous regions (or regions with particularly high rates of cancer), but that the drug trials do not appear to be further targeted to areas with idiosyncratically larger appropriate populations amongst this set of regions that contain at least one drug's PI.

3 Empirical Evidence

3.1 Empirical Strategy

Our central idea is to exploit variation in the geographic location of lead study authors across multiple new chemotherapy drugs to identify the impact of geographic proximity to a PI. If the location of study authors were randomly assigned across the country, then we could simply compare drug utilization across locations and infer that any increased propensity to use the drug was due to the extra experience or information about the drug that the first author has and may communicate to others in his geographic region. In the absence of random assignment, we endeavor to control richly for regional propensities to use new chemotherapy agents and for changing attitudes about new drugs, in order to isolate the impact of PI location on drug usage.

Our empirical framework is analogous to a difference-in-differences setting, where we compare drug utilization in PI and non-PI regions, controlling for baseline differences in the region's propensity to use new chemotherapeutic agents as well as controlling for time trends in the needs for and desirability of each drug. Since we observe the diffusion of twenty-one newly introduced drugs, we exploit regions' usage of other new drugs (for which the area contains no authors) to establish its propensity to adopt when it does not contain a study author. In addition, we use the time path of drug usage in non-author regions to establish how the drug usage evolved outside the PI HRR.

Our baseline regression specification takes the following form:

$$\begin{aligned}(drug)_{ijtd} = &\{HRR \times \textit{disease-group FEs}\}_{ijd} + \{drug \times year\ FEs\}_{dt} \\ &+ \beta_f \mathbf{1}(PI\ HRR)_{jd} + \varepsilon_{ijtd}\end{aligned} \tag{1}$$

An observation is a patient-drug episode (patient i treated in provider region j, t years after drug d was approved), limited to episodes for which drug d is indicated based on patient diagnoses. The regression is estimated over patient-drug episodes that fall within three years following FDA approval of the drug.

The first term in the regression above is a vector of fixed effects measuring each HRR's propensity to use new chemotherapy drugs for each of three cancer disease types. Targeted diseases are grouped based on the cancer subtype: hematologic cancers (leukemias and lymphomas), urologic cancers (kidney and bladder cancer), and other carcinomas (brain, breast, colon, and lung cancer). This allows regions to differ in their enthusiasm and patient suitability for new chemotherapy

treatments of each disease group. The second term in the regression allows each drug to face an idiosyncratic yearly shock to its popularity that is constant across regions. Lastly, the key independent variable of interest is an indicator for whether the first author of drug d's pivotal clinical trial is located in region j. The coefficient β_f on this indicator describes how much more likely a cancer patient will receive a new chemo drug if treated in the HRR where the PI of the drug's pivotal clinical trial is located.

The primary threat to the validity of this approach would stem from the possibility that PI regions are systematically more likely to use the new drug (for reasons not driven by author status) than their utilization of other new drugs for this cancer type and the national utilization of this particular new drug would predict. This could occur if, for example, clinical trials were located in areas with idiosyncratically high latent demand for that particular drug. We deal with this threat to validity in a number of ways, outlined in further detail below, including limiting the analysis to regions that ever contain a first author and controlling for HRR-by-disease fixed effects with finer groupings of disease types.

In the first set of results discussed below and presented in Section 3.2, we match patients to provider regions on the basis of where patient care is actually delivered. Thus, it is important to recognize that any effect PI status has on a region's propensity to prescribe a new chemotherapy agent could be driven by two separate channels: (a) a partial-equilibrium effect in which providers in the PI region have an increased propensity to treat a fixed population of patients with the new drug; and (b) a general equilibrium effect in which patients suitable for particular treatments sort to providers who specialize in those treatments.[4] For example, an increased number of suitable patients may travel into the PI author region for treatment, or suitable resident patients may be more likely to stay within the region for their care. (In this context, patient suitability could encompass both clinical appropriateness and the patient's demand for a new drug.)

To exemplify this point, suppose $f_d(\theta, p)$ denotes the fraction of cancer patients treated with drug d in a given region, where p measures the drug's suitability for patients treated in the region, and θ indexes the propensity of physicians in the region to administer the drug to a standard patient. The essential point is that both θ and p may respond to a change in a region's author

[4]Note that both channels are present even under true random assignment of principal investigators to geographic regions.

status τ, and thus the aggregate effect of author status on regional treatment intensity is given by

$$\frac{df_d}{d\tau} = \frac{\partial f_d}{\partial \theta}\frac{d\theta}{d\tau} + \frac{\partial f_d}{\partial p}\frac{dp}{d\tau}$$

The first term on the right-hand side of this equation measures the effect that PI status has on a region's propensity to use a drug holding fixed patient suitability, while the second term measures the increased usage of the drug due to patient sorting. Our baseline specification (1) measures the aggregate impact of first-author status on drug utilization, but does not disentangle the two mechanisms. Because these two channels have very different implications for policy, we present in section 4 an instrumental variables approach that isolates the change in utilization driven by channel (a), i.e. the increased propensity of first author regions to treat any particular patient.

3.2 Baseline Proximity Effects

We begin by presenting evidence on whether geographic proximity to a new chemotherapy drug's principal investigator impacts a physician's propensity to prescribe that drug for indicated patients, as estimated by our baseline specification (1) described above. The results from this estimation are reported in Table 3. As shown in column (1), we find that patients with the targeted diagnosis who receive chemotherapy treatment in the PI's HRR are 3.8 percentage points more likely to receive the new drug. This result is calculated over the first two years following the year of the drug's FDA approval. To provide a useful benchmark, this PI impact is a 37% increase over the 10.3% average utilization rate of these new chemotherapy drugs within regions that contain a PI for at least one other chemotherapy drug in our sample but do not contain any author for that particular drug (see Table 2, column (3)).

Next we explore the extent to which drug utilization is influenced by geographic proximity to other authors of a drug's pivotal trial. To do this, we first estimate a slightly modified version of specification (1) in which the indicator for whether patient care is delivered in the PI's HRR is replaced by an indicator for whether care is delivered in the HRR of any investigator on the drug's pivotal trial. As reported in Table 3 columns (2) and (4), the average effect that proximity to any investigator has on drug utilization is substantially smaller at 1.3pp, suggesting that there is important heterogeneity in the role of proximity to authors based on their level of contribution to the clinical trial.

We explore this heterogeneity by re-estimating specifications reported in columns (2) and (4)

with the addition of indicators for proximity to first and last authors. The results, reported in columns 3 and 5, estimate proximity to any author to have a positive (but not statistically significant) impact of 0.7pp. Proximity to the last author has a slightly smaller but not significantly different effect. However, the effect of proximity to the first author is 3.4pp higher than is proximity to any author (statistically significant at the 1% level). Given that most of the proximity effect on drug utilization appears to be driven by the first author, we focus the remainder of our analysis and discussion primarily on proximity to first authors, although we report any-author specifications as well.

In Table 3, Panel C, we further investigate the impact of proximity to a drug investigator, here limiting the sample only to the first year following initial FDA approval of the drug. As we discuss in more detail in the next section below, the impact of proximity is greatest immediately following drug approval and fades out over the study period. In column (8), we find that the regions containing any investigator (who is neither the first nor the last author) are 2.1pp more likely to prescribe the new chemotherapy drug, significant at the 5% level. This effect is double the magnitude estimated in column (6) when the second event year following FDA approval is included in the sample. This suggests that there is an initial positive relationship between new drug utilization and proximity to other study investigators, but that the effect fades out more quickly than the effect of PI proximity and is significantly smaller 2-years post FDA approval.

3.3 The Evolution of Proximity Effects

An important question we turn to next is how the effect of geographic proximity to a principal investigator evolves over time. With a Roy model of productivity spillovers, we may find geographic specialization in the use of medical treatments as described by Chandra and Staiger (2007). High-use areas develop expertise in the technology and have higher returns to its usage, and so they continue to use it more frequently in the steady state than low-use areas that do not develop a similar expertise. Under this model of productivity spillovers, we might expect to find long-run differences in the use of new chemotherapy agents across PI HRRs and other regions. An alternative model such as Phelps (2000) where information asymmetries are the reason for delayed adoption amongst non-PI regions would predict convergence in practice patterns as information about the new treatments reaches each physician.

To measure the evolution of the PI proximity effect, we estimate a modified version of specification (1) in which the PI HRR indicator is interacted with a full set of event-year dummies $\mathbf{1}(s = t)$

ranging from 1-4 calendar years following drug approval (0 corresponds to FDA approval year).[5] We also flexibly allow each HRR to follow separate diffusion rates by cancer disease type. The regression equation we estimate is given by

$$(drug)_{ijtd} = \{HRR \times disease\text{-}group \times event\text{-}year\ FEs\}_{ijd} + \{drug \times year\ FEs\}_{dt} + \sum_{s=1}^{4} \left[\beta_t \mathbf{1}(PI\ HRR)_{jd} \mathbf{1}(s = t)\right] + \varepsilon_{ijtd} \quad (2)$$

The primary coefficients of interest in equation (2) are the β_t, which describe the effect of geographic proximity to drug d's principal investigator on utilization of that drug t years after FDA approval. A modified version of (2) in which the PI dummy $\mathbf{1}(PI\ HRR)_{jd}$ is replaced by an indicator for whether any of a drug's pivotal trial study authors is located in the HRR estimates the evolution of the average effect of proximity to any investigator.

Panel A of Figure 1 plots how the estimated effect of proximity to a drug's principal investigator on drug utilization evolves over time, while Panel B plots the analogous result with proximity measured with respect to any of the drug's pivotal clinical trial authors. The time pattern of proximity effects traced out in these graphs reveals a number of insights. First, recently approved cancer drugs are used much more intensively in regions a study investigator, an effect even stronger in the PI's region. These patterns correspond to the results in Table 3 discussed previously. The second pattern highlighted by Figure 1 is that the proximity effect fades over time, though it is roughly twice as persistent in the PI's region compared with the average investigator region. In both cases, any proximity effect on drug utilization vanishes within 4 years after drug approval.

Taken as a whole, these estimates suggest that proximity to a pioneer investigator drives more-intensive take-up of new drugs, an effect which is stronger and more persistent for principal investigators than for other early investigators of the drug. Yet despite the initial eagerness to use the drug, this difference in intensity between investigator and non-investigator regions converges within a few years.

3.4 Geographic Extent of Investigator Influence

In the previous section, we documented that patients treated in regions (HRRs) that contain pioneer investigators of a new cancer drug are significantly more likely to receive that new drug than are

[5] Medicare drug codes are not introduced until the calendar year following FDA approval for the large majority of drugs in our sample, limiting our ability to measure diffusion prior to the first calendar year.

patients treated in other regions. While this implies that geography plays an important role in the adoption of new drugs, it does not address the extent of geographic influence. On the one hand, the effects of proximity to a drug's investigator may extend beyond the HRR in which that investigator is located to neighboring HRRs as well, or perhaps even further. On the other hand, the HRR may be too broad, for example if influence only extended to doctors within the investigator's hospital or physician's group. In the extreme, all geographic influence could be driven by investigators themselves being more likely to use the drug, with no effect on other physicians in the area.

We next test the geographic extent of investigator influence. We do this for two primary reasons. First, it sheds light on the degree of "friction" in the geographic network through which investigator influence operates, which in turn can help distinguish between potential mechanisms. Second, understanding the breadth of influence also impacts the interpretation of our baseline estimates from specification (1). There, we estimated the wedge between investigator HRRs relative to non-investigator HRRs. But if proximity effects are concentrated within, for example, only the PI's own hospital, then the region may be a misleading unit of analysis and the estimated proximity effects will be too small, as they would average over both treated and non-treated regions. On the other hand, if proximity effects extend more broadly than an investigator's own HRR, some of the comparison non-investigator regions are themselves influenced by the treatment, resulting in proximity effects that are also too small.

In principle, if oncologists were sufficiently dispersed geographically, it would be feasible to non-parametrically identify precisely how the influence of proximity to an investigator changes as distance from the investigator grows. However, cancer care in the U.S. is highly specialized and major cancer care sites are often geographically dispersed (Onega et al. 2008). To the extent that cancer care is primarily delivered in distinct geographic clusters, testing the extent of investigator influence amounts to testing how intensity across cancer care sites varies based on distance to the investigator's site. If influence is confined to the investigator's site, then even adjacent regions will show no increased use of the new drug.

In the remainder of this section, we explore the geographic extent of investigator influence on regional propensities to use new drugs. We begin by testing whether the increased drug usage is driven only by the PI himself or only at the hospital at which the PI practices, and then extend to testing whether higher usage extends outside of the PI's metropolitan area, or even further afield to neighboring HRRs.

3.4.1 Effect beyond investigator direct effect

As noted above, the evidence on geographic proximity we have presented so far leaves open the possibility that the increased use of a new drug in an investigator's HRR is entirely due to that investigator's propensity to use the drug. For example, if the principal investigator of a colon cancer drug treats 4% of colon cancer patients in his or her region and treats all of them with the new drug while other physicians use the drug at the national rate, this would lead to a nearly 4 percentage point increase in that region's use of the drug.

To evaluate whether the increase in drug use in the principal investigator's region is driven by other physicians practicing in that region, we simply re-estimate specification (1) after dropping out patients who are treated by the principal investigators themselves. If cancer patients uniquely matched to a single oncologist, this exercise would be straightforward. However, cancer patients are often treated by teams of physicians, making it difficult to disentangle the impact any one of the physicians played in the patient's treatment. Our approach is conservative in that it drops out patients episodes in which the principal investigator ever administered cancer care, even if a different physician delivered the new chemotherapy drug. Thus, this approach will not pick up proximity effects channeled through the influence the investigator may exert on other physicians while they are working together on the same patient, providing a lower-bound of the full proximity effect on nearby physicians.

The results of this estimation are shown in Table 4, column (2). Roughly 1,000 patient episodes receiving treatment from principal investigators of the drugs in our sample are dropped from the baseline specification (repeated in the table's first column for reference). Despite this restriction, the estimated proximity effect drops only very slightly (and insignificantly) from 3.9 to 3.6 percentage points, indicating that the proximity effect measured in the baseline specification is not a direct effect of the PI's own treatment choices but rather impacts the choices of physicians in the broader community.

3.5 Effect beyond investigator's hospital

Having demonstrated that patients treated by the PI himself are not driving the increased utilization within the PI's HRR, we now turn to investigating the precise geographic scope of the PI's influence. The next test is whether the PI's influence extends beyond the hospital where he practices, distinguishing intra- from inter-organization spillovers. In particular, we investigate whether

being treated in the PI's HRR, but not at the hospital where the PI practices, is associated with higher utilization of a new cancer drug. 26% of patients treated in the PI HRR are also treated by a physician practicing at the PI's hospital; these patients would need to have *much* higher propensity to receive the drug than their immediate neighbors, if they are to account for the entire increase in drug diffusion.

These specifications require each physician to be matched to the hospital where he typically delivers cancer care; the Carrier file does not indicate the precise location at which the service is performed, so this requires observing the physician to be associated with a bill in the Outpatient file. We drop the 14% of patients for whom we cannot observe the hospital affiliation of their physician. A second limitation of our data is that if physicians practice at multiple hospitals within a region, we may not observe the full set of physician-hospital affiliations. In particular, if we do not observe Outpatient cancer drug bills linked to the provider from *every* hospital with which the doctor is affiliated, then we may undercount the number of physicians who are practicing at the PI's hospital. Due to this data limitation, the results described below should be interpreted as suggestive of the PI's influence beyond his hospital, but we cannot fully rule out interaction of physicians at the PI's hospital as a mechanism for this effect.

As a final caveat to the interpretation of these results, as we analyze smaller geographic units, bias from patient sorting may become more acute. Our estimates of the being treated in a PI hospital as compared to a neighboring hospital may be a lower bound on any true spillovers, if suitable patients sort to the PI's hospital, leaving the population treated at nearby centers less appropriate for treatment with the new drug.

Results are reported in Table 4, column (3). We find that patients treated in the PI HRR but outside the PI's hospital are 3.6 percentage points more likely to receive treatment with the drug, suggesting that the PI's local influence extends beyond his own hospital. In fact, the point estimate suggests that patients treated at the PI's hospital are only 1 percentage point more likely to receive treatment than those patients treated at neighboring facilities, a result that is not statistically significant. However, the estimate is somewhat imprecise, so we cannot rule out the possibility that patients treated at the PI's hospital are, for example, twice as likely to receive the new drug as others treated in the same HRR.

3.6 Effect beyond investigator's HSA

To shed further light on the geographic extent of investigator influence, we consider using Hospital Service Areas as an alternative and finer definition of provider geography than HRRs. Because 45% of patients cross HSA boundaries to receive cancer treatment (as compared to only 20% crossing HRR boundaries), we prefer the HRR analysis for our main specification to reduce bias from patient sorting. HSAs roughly correspond to metropolitan areas across which patients travel for lower acuity hospital care.

Because the investigators in our sample practice primarily in large urban areas, it is important to keep in mind that the PI's HSA will often contain more than one tertiary care center. As a result, the HSA analysis is distinct from the hospital-level analysis. Also, 80% of observed chemotherapy care in the PI's HRR also occurs within the PI's HSA. Because cancer care is highly concentrated at major tertiary care centers which are often collocated within cities, relatively few patients in our sample seek treatment at the smaller hospitals within the PI's HRR but outside his HSA.

To measure whether investigator influence extends beyond an investigator's HSA, for each drug we identify the "neighbor" HSAs that lie geographically adjacent to the HSA in which the drug's principal investigator is located. We then measure the degree to which investigator regions and their neighbors increase drug utilization relative to non-neighbor regions, by estimating

$$\begin{aligned}(drug)_{ijtd} = \{HRR \times \textit{disease-group FEs}\}_{ijd} + \{drug \times year\ FEs\}_{dt} \\ + \beta_f \mathbf{1}(PI\ HRR)_{jd} + \beta_n \mathbf{1}(neighbor\ HRR)_{jd} + \varepsilon_{ijtd}\end{aligned} \tag{3}$$

As shown in Table 4 columns (4) and (7), the proximity effects are concentrated in the HSA containing the investigator. Patients treated in neighboring HSAs are not significantly more likely to receive treatment with the drug than patients who are further afield, with the point estimate suggesting a 0.7 percentage point higher drug utilization. Although it is important to note that the estimate is not very precise, due to the small numbers of patients treated for cancer in HSAs neighboring the PI HSA, and upper end of the 95% confidence interval corresponds to a substantial 2.9 p.p. higher utilization.

These results suggest that while drug investigators exert a strong influence on the propensity for nearby physicians to adopt and utilize those drugs, this influence may be highly concentrated geographically. If patients who are appropriate for treatment with the new chemotherapy drug would tend to be referred to the large tertiary care center within their HRR, then patient sorting

may drive the estimated coefficient on neighboring HSA's drug adoption towards zero. This possibility motivates using the geographically broader HRR as the unit of geography, as we do in the following section.

3.6.1 Effect on neighboring HRRs

Lastly, we estimate the analogous version of specification (3) where all variables and patient episodes are redefined at the HRR level rather than at the HSA. Recall that this analysis provides an important test since if the PI's influence extends to neighboring HRRs, our estimates of the extent of the PI's impact on drug utilization will be attenuated towards zero due to contamination of the "control" HRRs.

Table 4 shows in column (5) that while PI HRRs have a 3.9 pp increase in their propensity to use the new drug, there is no observed increase in propensity in neighboring HRRs. The point estimate suggests a 0.1 pp increase in new drug utilization in neighboring HRRs, which is small in magnitude and statistically not distinguishable from zero. The 95% confidence interval is relatively tight, with the upper bound consistent with at most 1.3 percentage point higher drug utilization. When investigator regions and their neighbor regions are defined based on the presence of any author on the drug's pivotal trial, the results are qualitatively the same (column 8). Based on these estimates, we find no evidence suggesting the influence of investigators extends beyond the investigators' own HRRs.

The estimates in this section demonstrate that the HRR unit provides a reasonable geographic area over which to study cancer drug diffusion, balancing concerns about patient sorting with the importance of identifying the precise region over which drug diffusion occurs. We have demonstrated that the increased utilization is not driven by patients treated by the PI himself, nor only by physicians practicing at the PI's own hospital. Although most of the effect appears to be concentrated within the HSA that contains the PI and where most cancer patients in the HRR receive treatment, we cannot rule out the possibility of substantial treatment effects in neighboring HSAs. However, there is no evidence of the PI's influence extending beyond his HRR into the larger neighboring regions, and these results are precise enough to rule out effect sizes that are one third as large as the impact of being treated in the PI's HRR.

3.7 Drug-Specific Investigator Influence

While the evidence above measures the average influence investigators exert on nearby physicians' drug adoption rates, the degree of this influence may vary across drugs, for example due to differences in the degree of the investigator's "connectedness" in the local community or the level of enthusiasm for the new drug based on trial results. This section seeks to flesh out this heterogeneity, by measuring how investigator influence varies over the drugs in our sample.

To do this, we estimate our baseline specification (1), with the modification that the investigator region dummy is fully interacted with a set of K indicators for each drug with a U.S. investigator. This new specification is given by

$$\begin{aligned}(drug)_{ijtd} = \{HRR \times \textit{disease-group FEs}\}_{ijd} + \{drug \times year\ FEs\}_{dt} \\ + \sum_{k=1}^{K} \left[\beta_d \mathbf{1}(PI\ HRR)_{jd} \mathbf{1}(drug\ d = k)\right] + \varepsilon_{ijtd}\end{aligned} \tag{4}$$

The coefficient of interest β_d describes how much more (or less) intensively drug d was used in the principal investigator's region relative to other regions, after controlling for each region's typical adoption rate for similar new drugs.

The results of this regression are plotted in Figure 2, with the drugs sorted in increasing order of β_d. For reference, the figure also plots the mean use of each drug among the HRRs that contain a principal investigator for any drug in our sample. As the figure shows, there is significant heterogeneity in the proximity effect across drugs. Two drugs in the sample show statistically negative proximity effects, while 7 are statistically positive.

To break these results down even further, we calculate the (residualized) utilization rates of each of these drugs across the 11 HRRs that contain a PI for any drug in our sample (controlling for the fixed effects in (4)). The results are shown in Figure 3. Note that Figure 2 derives from the results in Figure 3 by taking, for each drug, the (regression-weighted) average difference between the drug utilization rate in the PI region and the average utilization of that drug across the non-PI regions. By exposing the data, Figure 3 shows not only the variation in utilization of each new drugs across regions, but also shows the PI region's usage rank relative to the control regions. In fact, the PI region is the most intensive region for 5 (30%) of the 17 drugs in this sample, and is ranked in the top half for all but 3 of the drugs.

4 Patient Sorting and IV Estimates

As discussed above, there are two possible mechanisms for the observed increased propensity to prescribe new drugs in PI regions: an increased propensity to use the drug on a fixed set of patients, and a change in patient sorting such that the PI regions see patients with higher latent demand. In this section, we test for changes in patient sorting, and then use an instrumental variables strategy to identify the differences in drug utilization that occur holding the set of patients fixed.

In Table 5, we test whether patients with the targeted diagnosis who seek treatment in the PI's HRR are more likely to have traveled from a different HRR of residence. This would occur if, for example, savvy patients travel into the PI's HRR for treatment in order to gain access to the new chemotherapy drugs. In columns (1) through (4), the regression specification mirrors that in the main specification described in equation (1) above, but the outcome variable has been replaced with an indicator variable for whether the patient has traveled in from a different HRR.

In the baseline specification, including all observations and testing for travel to the PI's HRR, we find a positive point estimate suggestive of increased travel, but the effect is not statistically distinguishable from zero. Once we narrow the sample to the set of HRR's that ever contain a PI for any new drug, the estimated effect increases and becomes statistically significant at the 5% level. In particular, we find that 4 pp more of the patients treated in the PI's HRR have traveled from an outside HRR, as reported in column (2). In columns (3) and (4), we investigate the propensity to travel into a region containing any study author and continue to find an increased proportion of traveling patients, with approximately 3 pp more patients traveling from outside the HRR in author regions, as compared to non-investigator regions, significant at the 1% level.

This evidence suggests that some patients are aware of new centers of expertise for the new chemotherapy agent and are willing to travel further to improve their access to the drug. If these patients who are newly traveling into the PI's HRR are either more clinically appropriate for the new drug or have higher demand for trying the new technology, then part of the increased levels of drug utilization may be driven by the changing patient composition.

In Table 5, columns (6) and (8), we test whether the patients who travel from outside HRRs differ in their propensity to receive the new drug. In these columns, we report results from a regression that again mirrors equation (1), maintaining new drug utilization as the outcome variable of interest, but adding interactions with whether the patient has traveled from outside the HRR to

every term in the regression. The regression takes the following form:

$$(drug)_{ijtd} = \{HRR \times disease\text{-}group \times traveler\ FEs\}_{ijd} + \{drug \times year \times traveler\ FEs\}_{idt} + \beta_f \mathbf{1}(PI\ HRR)_{jd} + \beta_{ft} \mathbf{1}(PI\ HRR{*}traveler)_{ijd} + \varepsilon_{ijtd} \quad (5)$$

where *traveler* is a binary indicator for whether the observed patient is seeking care outside his home HRR.

We find that patients traveling to the PI's HRR are indeed more likely to receive treatm ent with the new drug than patients treated in the PI HRR who also reside within that HRR. These traveling patients are 2.6pp more likely to receive the new treatment, significant at the 5% level. Together, the findings in this table suggest that not only are more patients with the targeted diagnosis traveling into the PI's region to seek out the new treatment, but that these patients are indeed more likely to receive treatment than those who were already residing in the region. This evidence suggests strongly that the overall 3.9 pp higher new drug use in PI regions is driven at least in part by changing patient composition, and not solely a higher propensity to use the drug on a fixed set of patients.

To isolate whether the PI regions are indeed more likely than other regions to use the drug on any given patient, we pursue an instrumental variables (IV) strategy. In particular, we use each patient's HRR of residence as an instrumental variable to predict whether they will seek treatment in an HRR that contains the PI for the relevant drug. This instrument mitigates the concern that patient sorting renders the patients treated in the PI region more suitable to the new chemotherapy treatment. The exclusion restriction requires that, after conditioning on the included fixed effects, where a patient lives is uncorrelated with his suitability or demand for treatment with the new chemotherapy drug.

The first stage equation of the IV model takes this form:

$$(treated\ in\ author\ HRR)_{ijtd} = \{HRR \times disease\text{-}group\ FEs\}_{ijd} + \{drug \times year\ FEs\}_{dt} + \gamma_1 \mathbf{1}(residence\ in\ PI\ HRR)_{ijd} + \varepsilon_{ijtd} \quad (6)$$

The reduced form follows:

$$(drug)_{ijtd} = \{HRR \times disease\text{-}group\ FEs\}_{ijd} + \{drug \times year\ FEs\}_{dt} + \gamma_2 \mathbf{1}(residence\ in\ PI\ HRR)_{ijd} + \varepsilon_{ijtd} \quad (7)$$

The IV estimate is the simple ratio of γ_2/γ_1. Note that paralleling the baseline regression specification, we include fixed effects for HRR by disease group and for drug by year. We also report results from an enriched IV specification where in addition to using residence in the PI HRR as an instrument, we also include two additional instruments: (1) residence in a neighboring HRR (i.e. a region that shares a border with the PI HRR); and (2) residence in a neighbor-of-a-neighbor HRR (i.e. a region that is separated from the PI HRR by one neighbor).

The exclusion restriction could be violated under a few conditions. One possibility is that patients with the targeted cancer who reside in the PI region could have idiosyncratically high demand for the drug; this could occur if, for example, the drug targets a particular sub-type of colon cancer that has a higher-than-typical prevalence in the PI's region, so that a larger fraction of colon cancer patients in the region are appropriate for treatment. Second, the instrument would be invalid if patients change their HRR of residence in response to the availability of new chemotherapy drugs.

Under the IV framework, the exclusion restriction is not directly testable, but it seems plausible that the fraction of targeted cancer patients suitable for treatment with the new drug would not vary systematically across regions and that elderly Medicare patients would be very unlikely to move across regions within a three-year period in response to the location of a new chemotherapy trial.

Results from the IV regressions are reported in Table 6. Column 1, row 1, reports that patients who live in the PI's HRR are nearly 80 pp more likely than non-resident patients to receive treatment within the PI HRR, indicating a strong correlation between the instrumental variable and the endogenous regressor in the first stage. The reduced form results find that patients residing in the PI's HRR are 2 percentage points more likely to receive treatment with the new drug. The Wald estimate reported in the final row of results rescales the reduced form estimate and estimates that providers in the PI's region are 2.8 pp more likely to prescribe the new drug compared to other providers, significant at the 5% level. The finding is robust to restricting the sample only to patients residing in HRRs that contain a PI for any drug, as reported in Column 2. Adding

the additional instrumental variables for residence in neighbor HRRs to the model also does not substantially change the estimated IV coefficient.

The results are similar in the specification investigating the impact of being treated in a region with any study author, but the magnitude of the reduced form and two-stage-least-squares estimates are markedly smaller and less statistically significant. Overall, the results suggest a 1 percentage point increase in a targeted patient's propensity to receive the new drug when being treated in an author's region, significant at the 10% level.

The IV results suggest that physicians in the PI's region of residence are approximately 2.8 pp more likely to use the new drug on a fixed set of patients, which is a somewhat smaller effect than the baseline regression results reporting a 3.9 pp increase in drug utilization across all patients treated in the PI region. Suitable patients sorting into the region may have contributed to the high observed point estimate in the baseline regression, but the result remains large and statistically significant in the IV specification.

Taken together, the patient traveling results and the IV regressions find strong support for both hypothesized channels by which the presence of a PI may affect care in his region. Patients with high latent demand for the drug seem to seek out care in areas with expertise in the new technology. In addition, doctors in the PI's region are more likely to use the new drug, holding fixed the population of patients seeking treatment. The two results work in tandem, since if PI regions had no greater expertise with a new drug, it would be surprising to find increasing numbers of appropriate patients undertaking costly travel to seek treatment in the region.

5 Conclusion

The results presented above suggest a significant role for informational frictions affecting the adoption of new chemotherapy agents in the first few years after drug introduction. Regions that contained a new drug's PI were 2.8 percentage points more likely to prescribe the new drug in the preferred IV specification presented in Table 6. This effect is primarily driven by physicians practicing within the same health service area as the PI, i.e. doctors who are in regular contact with the area's large tertiary care center, but it is not driven by the prescribing patterns of the PI himself. Patients outside the PI's HRR are increasingly likely to travel into the PI's region from a different HRR of residence to seek care after the new drug is introduced, suggesting that some patients are aware of the differences in drug adoption or expertise across regions and willing to

endure greater travel costs to gain access to the new treatment.

Despite these marked differences in early adoption of new chemotherapy drugs, there is no evidence that early expertise with a drug drives higher rates of long-term utilization. PI HRRs are no more likely to specialize in treatment with the new drug than other regions by the fourth year following drug introduction. Thus, the information frictions that may hamper early adoption seem to ease over time and the utilization of PI and other HRRs converges within a four-year period. However, even if there are no observed long-run differences in utilization of the studied drugs, research tends to cluster at particular prestigious academic hospitals. Although long-run prescription patterns for any particular drug may converge across regions, overall treatment patterns for cancer patients may persistently differ as the research-intense regions continue to house investigators for each wave of new treatment innovations.

References

Audretsch, David B. and Maryann P. Feldman. 2004. "Knowledge Spillovers and the Geography of Innovation." *Handbook of Regional and Urban Economics* 4:2713–2739.

Azoulay, Pierre, Joshua S. Graff Zivin, and Bhaven N. Sampat. 2012. *The Diffusion of Scientific Knowledge Across Time and Space: Evidence from Professional Transitions for the Superstars of Medicine*. University of Chicago Press, 107–155. URL http://www.nber.org/chapters/c12350.

Bach, Peter B. 2009. "Limits on Medicare's Ability to Control Rising Spending on Cancer Drugs." *New England Journal of Medicine* 360 (6):626–633.

Baerlocher, Mark Otto, Marshall Newton, Tina Gautam, George Tomlinson, and Allan S. Detsky. 2007. "The Meaning of Author Order in Medical Research." *Journal of Investigative Medicine* 55 (4):174.

Baicker, Katherine and Amitabh Chandra. 2010. "Understanding Agglomerations in Health Care." In *Agglomeration Economics*, edited by Edward L. Glaeser. The University of Chicago Press, 211 – 236.

CBO, United States. 2008. *Technological Change and the Growth of Health Care Spending*. Congressional Budget Office paper. United States Congressional Budget Office.

Chandra, Amitabh and Douglas O. Staiger. 2007. "Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks." *The Journal of Political Economy* 115 (1):103–140.

Coleman, James, Elihu Katz, and Herbert Menzel. 1957. "The Diffusion of an Innovation Among Physicians." *Sociometry* 20 (4):253–270.

Cutler, David M. 1995. "Technology, Health Costs, and the NIH." In *National Institutes of Health Roundtable on the Economics of Biomedical Research*.

Escarce, Jose J. 1996. "Externalities in Hospitals and Physician Adoption of a New Surgical Technology: An Exploratory Analysis." *Journal of Health Economics* 15 (6):715–734.

Fisher, Elliott S., David E. Wennberg, Thérèse A. Stukel, Daniel J. Gottlieb, F.L. Lucas, and Étoile L. Pinder. 2003a. "The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care." *Annals of Internal Medicine* 138 (4):273–287.

———. 2003b. "The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care." *Annals of Internal Medicine* 138 (4):288–298.

Harpaz-Rotem, Ilan and Robert A. Rosenheck. 2009. "Tracing the Flow of Knowledge: Geographic Variability in the Diffusion of Prazosin Use for the Treatment of Posttraumatic Stress Disorder Nationally in the Department of Veterans Affairs." *Archives of General Psychiatry* 66 (4):417.

Jaffe, Adam B., Manuel Trajtenberg, and Rebecca Henderson. 1993. "Geographic Localization of Knowledge Spillovers as Evidenced by Patent Citations." *the Quarterly Journal of Economics* 108 (3):577–598.

Newhouse, Joseph P. 1992. "Medical care costs: how much welfare loss?" *The Journal of Economic Perspectives* 6 (3):3–21.

Onega, Tracy, Eric J. Duell, Sun Shi, Dongmei Wang, Eugene Demidenko, and David Goodman. 2008. "Geographic Access to Cancer Care in the US." *Cancer* 112 (4):909–918.

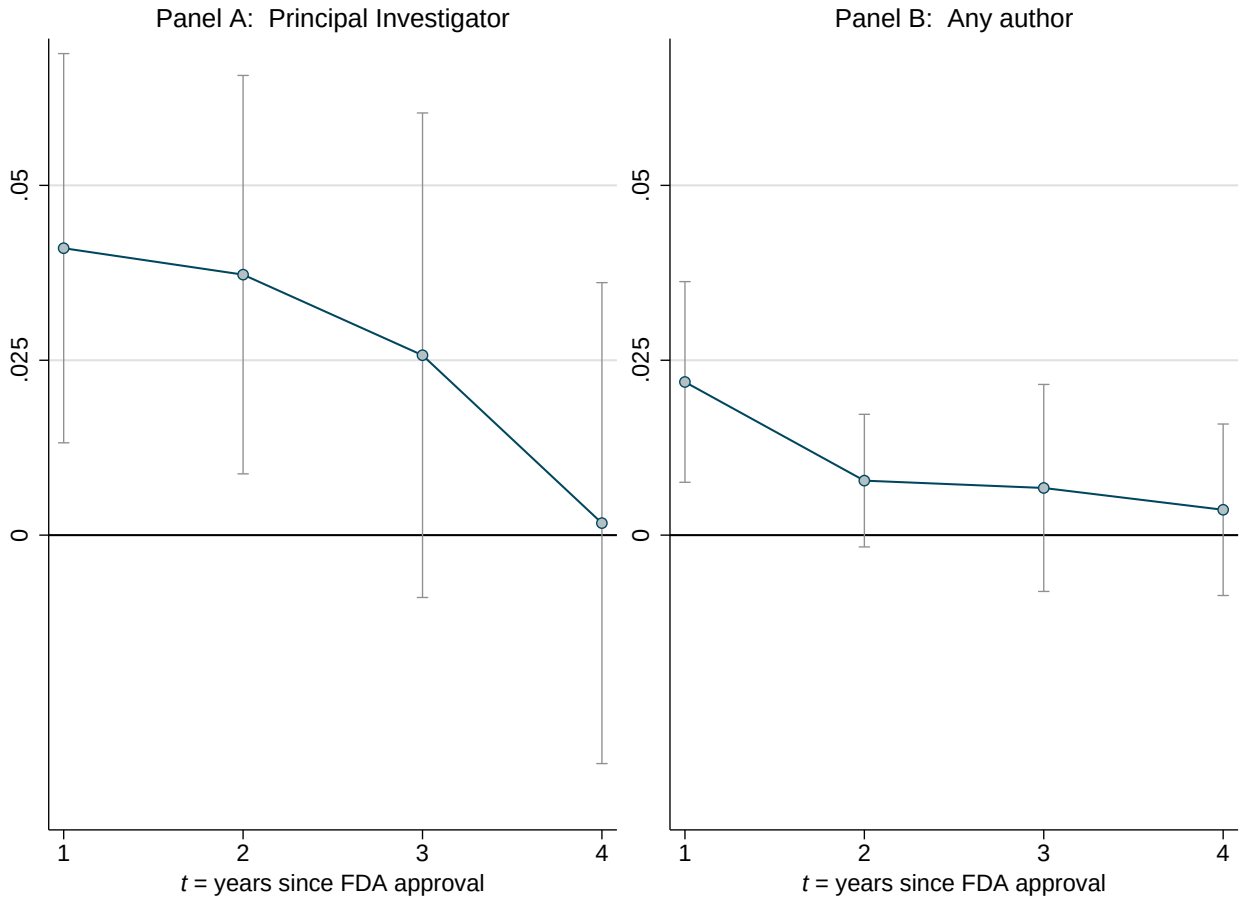
Parente, Stephen L. and Edward C. Prescott. 1994. "Barriers to technology adoption and development." *Journal of Political Economy* :298–321.

Phelps, Charles E. 2000. "Information Diffusion and Best Practice Adoption." *Handbook of Health Economics* 1:223–264.

Skinner, Jonathan and Douglas Staiger. 2009. "Technology Diffusion and Productivity Growth in Health Care." Working Paper 14865, National Bureau of Economic Research.

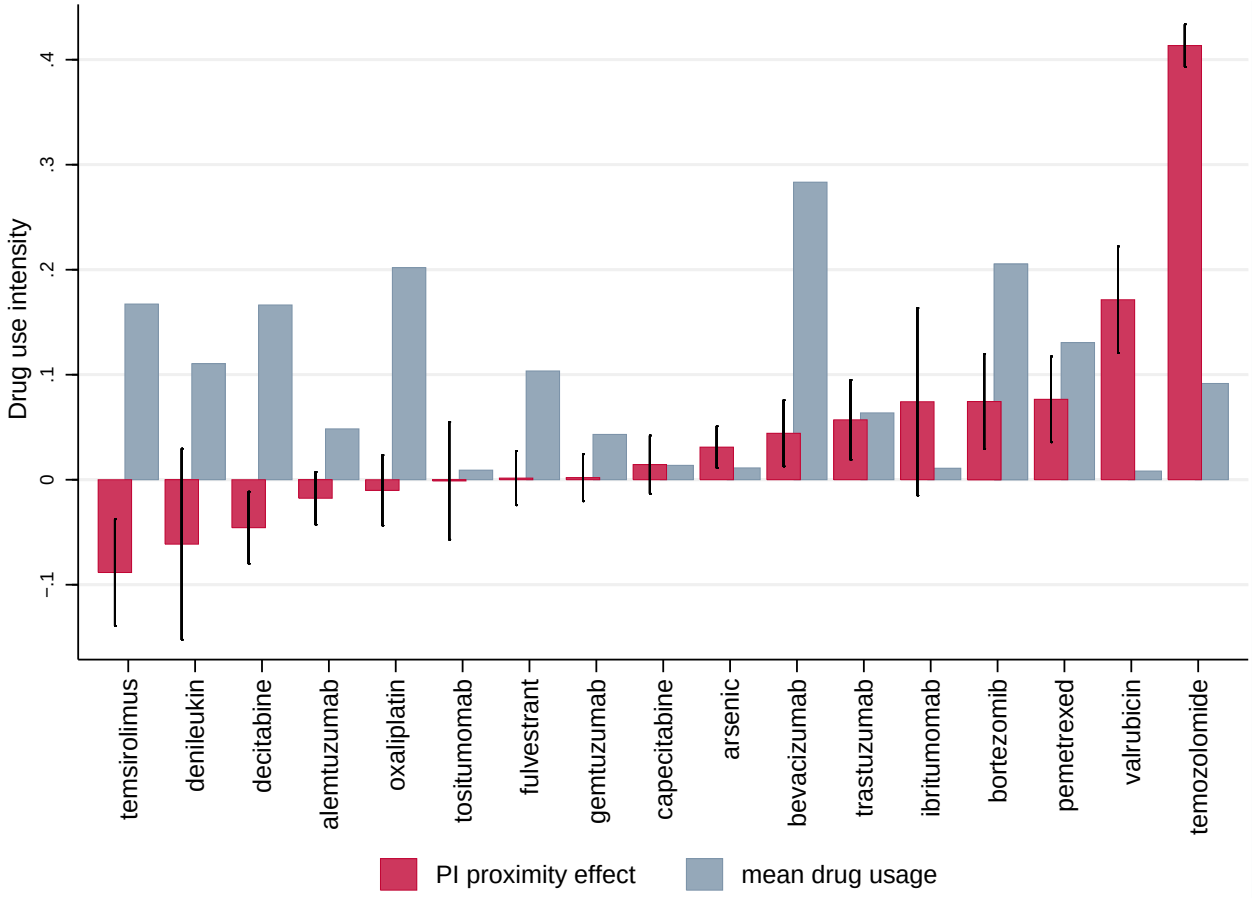
Smith, Sheila, Joseph P. Newhouse, and Mark S. Freeland. 2009. "Income, Insurance, and Technology: Why Does Health Spending Outpace Economic Growth?" *Health Affairs* 28 (5):1276–1284.

Figure 1: Influence of Author Proximity on Drug Use



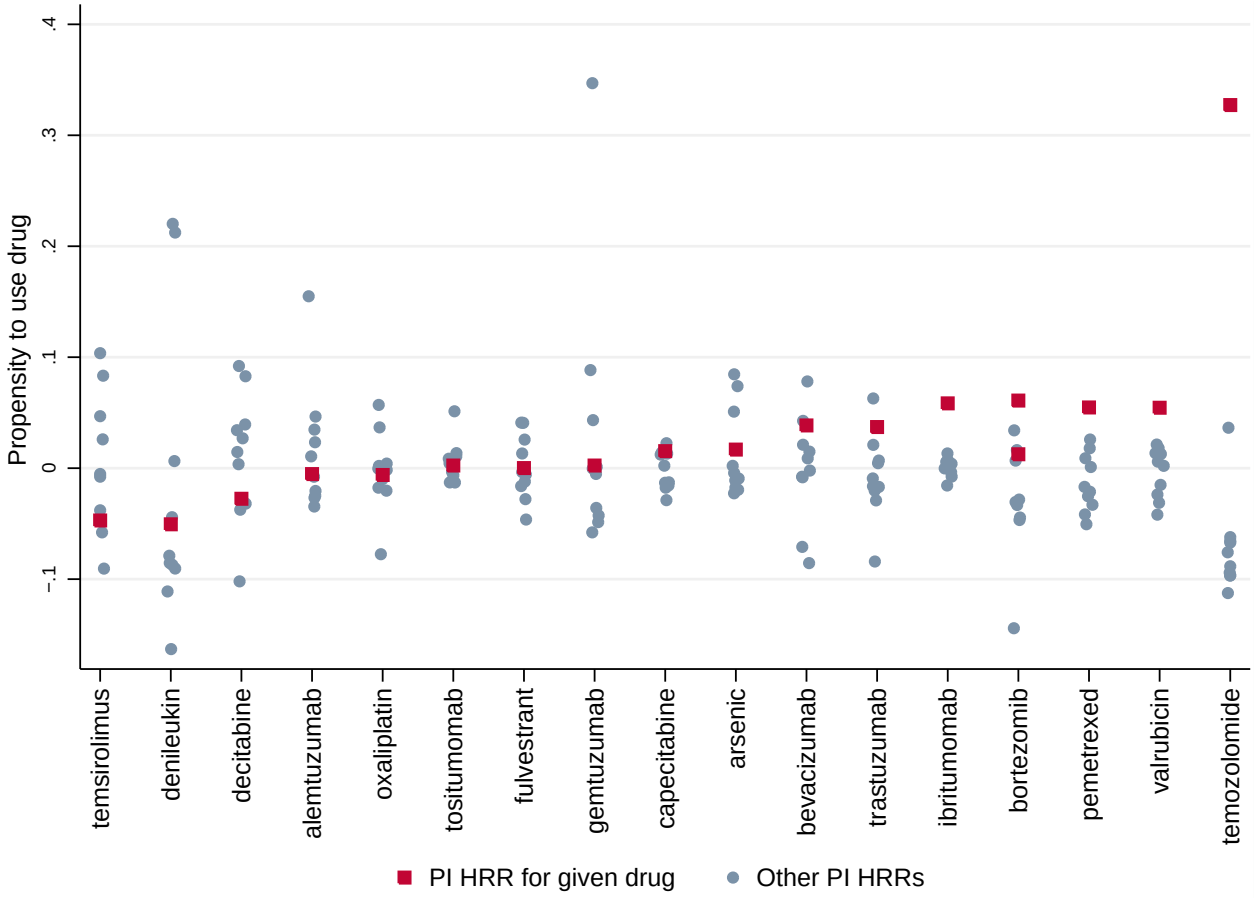
Notes: Graphs plot estimates of the effect pioneer investigator proximity has on drug utilization, t years since the corresponding chemotherapy drug became FDA approved. Bands indicate 95% confidence intervals constructed from standard errors clustered at the provider HRR-drug level.

Figure 2: Principal Investigator Influence on Drug Utilization, by Drug



Notes: Red bars plot estimates of the principal investigator proximity effect on drug utilization, by drug. Blue bars plot fraction of indicated patients receiving drug within 2 years of FDA approval. Bars on the proximity effects indicate 95% confidence intervals constructed from standard errors clustered at the provider HRR-drug level. All estimates calculated using patient episodes occurring in HRRs that ever contain a principal investigator for a drug in our sample.

Figure 3: Distribution of Drug Utilization across PI HRRs, by Drug



Notes: Each observation in the plot represents the average drug use (after adjusting for drug-eventyear and HRR-cancertype fixed effects) over indicated patients within an HRR, limited to HRRs that ever contain a principal investigator for any drug in the sample.

Table 1: List of studied chemotherapy drugs

Generic drug name	Trade name	FDA approval date	Target disease	1st author city	Year of pivotal trial publication	Journal of pivotal trial	Size of target population	No. of authors on pivotal trial
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Capecitabine	Xeloda	4/30/1998	breast cancer	Dallas, TX	1999	Journal of Clinical Oncology	26,410	10
Trastuzumab	Herceptin	9/25/1998	breast cancer	Chicago, IL	1999	Journal of Clinical Oncology	26,410	11
Valrubicin	Valstar	9/25/1998	bladder cancer	Chicago, IL	2000	Journal of Urology	13,557	6
Denileukin diftitox	Ontak	2/5/1999	cutaneous T-cell lymphoma	Durham, NC	2001	Journal of Clinical Oncology	819	26
Temozolomide	Temodar	8/11/1999	brain cancer	Houston, TX	2000	British Journal of Cancer	1,797	22
Epirubicin hydrochloride	Ellence	9/15/1999	breast cancer	Canada	1998	Journal of Clinical Oncology	53,762	18
Gemtuzumab ozogamicin	Mylotarg	5/17/2000	acute myeloid leukemia	Seattle, WA	2001	Journal of Clinical Oncology	2,192	17
Arsenic trioxide	Trisenox	9/25/2000	acute myeloid leukemia	New York, NY	2001	Journal of Clinical Oncology	1,079	15
Alemtuzumab	Campath	5/7/2001	chronic lymphocytic leukemia	Houston, TX	2002	Blood	12,027	11
Zoledronic acid	Zometa	8/20/2001	hypercalcemia of malignancy	Canada	2001	Journal of Clinical Oncology	2,694	11
Ibritumomab tiuxetan[1]	Zevalin	2/19/2002	non-Hodgkin's lymphoma	Rochester, MN	2002	Journal of Clinical Oncology	51,042	13
Fulvestrant	Faslodex	4/25/2002	breast cancer	Houston, TX	2002	Journal of Clinical Oncology	64,045	14
Oxaliplatin	Eloxatin	8/9/2002	colon cancer	Nashville, TN	2003	Journal of Clinical Oncology	52,778	8
Bortezomib[2]	Velcade	5/13/2003	multiple myeloma	Boston, MA	2003	New England Journal of Medicine	23,819	21
Tositumomab-I 131	Bexxar	6/27/2003	non-Hodgkin's lymphoma	Stanford, CA	2005	Journal of Clinical Oncology	54,275	7
Pemetrexed	Alimta	2/4/2004	lung cancer	Chicago, IL	2003	Journal of Clinical Oncology	84,918	13
Cetuximab	Erbitux	2/12/2004	colon cancer	United Kingdom	2004	New England Journal of Medicine	55,528	12
Bevacizumab	Avastin	2/26/2004	colon cancer	Durham, NC	2004	New England Journal of Medicine	55,528	15
Decitabine	Dacogen	5/2/2006	myelodysplastic syndromes	Houston, TX	2006	Cancer	15,460	16
Panitumumab	Arranon	9/27/2006	colon cancer	Belgium	2007	Journal of Clinical Oncology	59,028	12
Temsirolimus	Torisel	5/30/2007	kidney cancer	Philadelphia, PA	2007	New England Journal of Medicine	3,794	19

[1]There were two pivotal trials for Ibritumomab tiuxetan; the second trial had 1 total authors, with the same first author also in Rochester, MN, and was also published in the Journal of Clinical Oncology.

[2]There were two pivotal trials for Bortezomib; the second trial had a 15 total authors, with the first author in New York, NY, and was published in the British Medical Journal 2004.

Table 2: Drug use summary stastics

	PI HRR for observed drug (1)	Non-PI investigator HRR for obs. drug (2)	PI HRR for another drug, no investigator for observed drug (3)	Non-PI investigator HRR for another drug, no investigator for obs. drug (4)	HRR with no pivotal investigators for any drug in sample (5)
Variables:					
Drug utilization rate	0.156	0.098	0.117	0.100	0.096
No. of patient-year-HRR-drug obs.	6,988	28,359	52,206	130,663	271,734
Avg. no. of patients per HRR-drug pair	388	229	424	183	69
No. of HRR-drug pairs	18	124	123	713	3923
No. of unique HRRs	11	55	11	58	248

Notes: Each column represents mutually exclusive sets of observations. Also, to maintain the same set of drugs represented in each column, drug observations with non-US based pivotal trials are exluded.

Table 3: Author Proximity Effect on Drug Utilization

Dependent variable: (drug)_id in {0,1}, indicates receipt of new cancer drug d by patient i

	Panel A: All HRRs			Panel B: Author HRRs only			Panel C: All HRRs, 1st event year		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
(PI in HRR)	**0.039*** **		**0.035*** **	**0.035*** **		**0.033*** **	**0.041*** **		**0.035** **
	(0.013)		(0.013)	(0.011)		(0.012)	(0.014)		(0.015)
(any author in HRR)		**0.014*** **	**0.009**		**0.014*** **	**0.009**		**0.022*** **	**0.021** **
		(0.005)	(0.006)		(0.005)	(0.006)		(0.007)	(0.008)
(last author in HRR)			**-0.007**			**-0.008**			**-0.025***
			(0.012)			(0.011)			(0.015)
drug-eventyr FEs	X	X	X	X	X	X	X	X	X
HRR-cancer FEs	X	X	X	X	X	X	X	X	X
N	660,962	660,962	660,962	96,577	293,128	293,128	289,775	289,775	289,775

Notes: Each observation is a patient-drug episode (patient episodes may indicate multiple drugs). Author HRRs based on location of authors of each drug's pivotal clinical trial on which initial FDA approval was based. HRR-cancer fixed effects defined using three categories of cancer drugs: urologic, hematologic, and other (including breast, colon, lung, and brain). Standard errors clustered at the HRR-drug level shown in parentheses. *: p<0.10; **: p<0.05; ***: p<0.01.

(1-3) All HRRs and patient episodes within 1-2 calendar years after drug's FDA approval

(4) Sample limited to HRRs which ever contain a Principal Investigator (PI)

(5-6) Sample limited to HRRs which ever contain any pivotal trial author

(7-9) All HRRs and patient episodes within the calendar year immediately following drug's FDA approval

Table 4: Geographic Extent of Investigator Influence

Dependent variable: (drug)_id in {0,1}, indicates receipt of new cancer drug *d* by patient *i*

	Principal Investigator (PI)					Any Author		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(PI in HRR/HSA)	**0.039*** **	**0.036*** **	**0.036** **	**0.044*** **	**0.039*** **			
	(0.013)	(0.013)	(0.015)	(0.016)	(0.013)			
(treated at PI's hospital)			**0.010**					
			(0.023)					
(PI neighbor)				**0.007**	**0.001**			
				(0.011)	(0.006)			
(any author in HRR/HSA)						**0.014*** **	**0.020*** **	**0.014*** **
						(0.005)	(0.007)	(0.005)
(any author neighbor)							**0.002**	**-0.001**
							(0.005)	(0.003)
provider region	HRR	HRR	HRR	HSA	HRR	HRR	HSA	HRR
drug-eventyr FEs	X	X	X	X	X	X	X	X
region-cancer FEs	X	X	X	X	X	X	X	X
N	660,962	659,971	660,962	678,671	550,761	660,962	678,671	660,962

Notes: Each observation is a patient-drug episode (patient episodes may indicate multiple drugs). Author HRR/HSAs based on location of authors of each drug's pivotal clinical trial on which initial FDA approval was based. HRR/HSA-cancer fixed effects defined using three categories of cancer drugs: urologic, hematologic, and other (including breast, colon, lung, and brain). Standard errors clustered at the HRR-drug level shown in parentheses. *: p<0.10; **: p<0.05; ***: p<0.01.

(1) From Table 3, column (1), for comparison.

(2) Patients ever treated by a Principal Investigator (of any pivotal trial for drugs in sample) excluded from estimation.

(3) Specification adds indicator for "neighbor" HRRs located geographically adjacent to the HRR in which the PI is located.

(4) Specification adds indicator for being treated by a physician who practices at the same hospital as the PI. Excludes observations with patients treated by doctors who cannot be linked to a hospital practice.

(5) Specification adds indicator for "neighbor" HSAs located geographically adjacent to the HAS in which the PI is located.

(6) From Table 3, column (2), for comparison.

(7) Specification adds indicator for "neighbor" HRRs located geographically adjacent to an HRR containing any study author.

(8) Specification adds indicator for "neighbor" HSAs located geographically adjacent to an HSA containing any study author.

Table 5: Patient Travel and Proximity Effects

Dependent variables
(1-2): (travel)_id in {0,1}, indicates treatment outside patient *i* 's HRR of residence
(3-8): (drug)_id in {0,1}, indicates receipt of new cancer drug *d* by patient *i*

	Travel				Drug Utilization			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(PI in HRR)	**0.026**	**0.041****			**0.039*****	**0.030****		
	(0.019)	(0.016)			(0.013)	(0.012)		
(PI HRR)*traveler						**0.026****		
						(0.010)		
(any author in HRR)			**0.032*****	**0.032*****			**0.014*****	**0.012****
			(0.009)	(0.009)			(0.005)	(0.005)
(any auth HRR)*traveler								**0.011**
								(0.007)
HRR definition	provider	provider	provider	provider	provider	provider	provider	provider
drug-eventyr FEs	X	X	X	X	X		X	
drug-eventyr-traveler FEs						X		X
HRR-cancer FEs					X		X	
HRR-cancer-traveler FEs						X		X
N	660,962	96,577	660,962	293,128	660,962	660,962	660,962	660,962

Notes: Each observation is a patient-drug episode (patient episodes may indicate multiple drugs). Author HRRs based on location of authors of each drug's pivotal clinical trial on which initial FDA approval was based. HRR-cancer fixed effects defined using three categories of cancer drugs: urologic, hematologic, and other (including breast, colon, lung, and brain). Standard errors clustered at the HRR-drug level shown in parentheses. *: p<0.10; **: p<0.05; ***: p<0.01.
(1-4) Dependent variable indicates whether patient received care outside the patient's HRR of residence.
(2,4) Same as (1,3), except estimated over author HRRs only.
(5,7) From Table 3, columns (1,2), for comparison.
(6,8) Estimates the differential proximity effect for travelers, defined as patients residing outside provider HRR.

Table 6: IV Estimates of Proximity Effect on Drug Utilization

	Principal Investigator (PI)			Any Author		
	(1)	(2)	(3)	(4)	(5)	(6)
first stage: treated in author HRR effect						
residence in author HRR	**0.796*** **	**0.876*** **	**0.806*** **	**0.739*** **	**0.828*** **	**0.760*** **
	(0.034)	(0.033)	(0.005)	(0.002)	(0.020)	(0.002)
residence in author neighbor HRR			**0.131*** **			**0.103*** **
			(0.003)			(0.001)
residence in neighbor of neighbor HRR			**0.030*** **			**0.016*** **
			(0.001)			(0.001)
reduced form: drug receipt effect						
residence in author HRR	**0.022** **	**0.022** **	**0.023** **	**0.007**	**0.009** **	**0.007**
	(0.010)	(0.009)	(0.010)	(0.004)	(0.004)	(0.004)
residence in author neighbor HRR			**0.013**			**0.002**
			(0.007)			(0.003)
residence in neighbor of neighbor HRR			**-0.003**			**-0.002**
			(0.004)			(0.002)
Wald/2SLS						
provider in author HRR	**0.028** **	**0.025** **	**0.031*** **	**0.010**	**0.011** **	**0.010**
	(0.012)	(0.010)	(0.012)	(0.005)	(0.005)	(0.005)
drug-eventyr FEs	X	X	X	X	X	X
provider HRR-cancer FEs	X	X	X	X	X	X
Sample						
residence in any HRR	X		X	X		X
residence in an author HRR		X			X	
N	660,359	76,100	660,359	660,359	267,287	660,359

Notes: Each observation is a patient-drug episode (patient episodes may indicate multiple drugs). Author HRRs based on location of authors of each drug's pivotal clinical trial on which initial FDA approval was based. HRR-cancer fixed effects defined using three categories of cancer drugs: urologic, hematologic, and other (including breast, colon, lung, and brain). Standard errors clustered at the HRR-drug level shown in parentheses. *: p<0.10; **: p<0.05; ***: p<0.01.

Appendices

Table A1: Author Proximity Effect on Drug Utilization for Indicated/All Cancer Patients

Dependent variable: (drug)_id in {0,1}, indicates receipt of new cancer drug *d* by patient *i*

	Indicated Patients				All Cancer Patients			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
(1st author in HRR)	**0.039*** **	**0.035*** **			**0.0025** **	**0.0035*** **		
	(0.013)	(0.011)			(0.0012)	(0.013)		
(any author in HRR)			**0.014*** **	**0.014*** **			**0.0007**	**0.0008**
			(0.005)	(0.005)			(0.0007)	(0.0007)
drug-eventyr FEs	X	X	X	X	X	X	X	X
HRR-cancer FEs	X	X	X	X	X	X	X	X
N	660,962	96,577	660,962	293,128	13,672,002	1,944,314	13,672,002	5,992,456

Notes: Each observation is a patient-drug episode. Columns (1-4) include only drug-episodes recording a diagnosis for the drug's initial indication. Columns (5-8) include all drug-episodes, regardless of whether the patient was diagnosed with the drug's initial indication. Author HRRs based on location of authors of each drug's pivotal clinical trial on which initial FDA approval was based. HRR-cancer fixed effects defined using three categories of cancer drugs: urologic, hematologic, and other (including breast, colon, lung, and brain). Standard errors clustered at the HRR-drug level shown in parentheses. *: p<0.10; **: p<0.05; ***: p<0.01.

(1) From Table 3, column (1). All indicated patient episodes, all HRRs.

(2) From Table 3, column (4). All indicated patient episodes, limited to HRRs ever containing a first author.

(3) From Table 3, column (2). All indicated patient episodes, all HRRs.

(4) From Table 3, column (5). All indicated patient episodes, limited to HRRs ever containing any author.

(5) All patient episodes (indicated and non-indicated), all HRRs.

(6) All patient episodes (indicated and non-indicated), limited to HRRs ever containing a first author.

(7) All patient episodes (indicated and non-indicated), all HRRs.

(8) All patient episodes (indicated and non-indicated), limited to HRRs ever containing any author.

Table A2: Proximity Effect Robustness to Cancer Type Categories

Dependent variable: (drug)_id in {0,1}, indicates receipt of new cancer drug *d* by patient *i*

	Principal Investigator (PI)			Any Author		
	(1)	(2)	(3)	(4)	(5)	(6)
(PI in HRR/HSA)	**0.039*** **	**0.026*** **	**0.024** **			
	(0.013)	(0.010)	(0.011)			
(any author in HRR/HSA)				**0.014*** **	**0.014*** **	**0.015*** **
				(0.005)	(0.004)	(0.004)
drug-eventyr FEs	X	X	X	X	X	X
HRR-cancer (3 classes) FEs	X	X		X	X	
HRR-cancer (7 classes) FEs			X			X
N	660,962	571,553	660,962	660,962	571,553	660,962

Notes: Each observation is a patient-drug episode. Author HRRs based on location of authors of each drug's pivotal clinical trial on which initial FDA approval was based. Three categories of cancer drugs: urologic, hematologic, and other (including breast, colon, lung, and brain). Seven categories of cancer drugs: urologic, hematologic, breast, colon, lung, brain, and hypercalcemia; the lung, brain, and hypercalcemia categories contain only one drug per class. Standard errors clustered at the HRR-drug level shown in parentheses. *: p<0.10; **: p<0.05; ***: p<0.01.

(1,3) From Table 3, columns (1,2), for comparison.

(2,5) Drug categories with only one drug (lung, brain, and hypercalcemia) excluded.

(3,6) HRR-cancer fixed effects defined using seven categories of cancer-related drugs.