NBER WORKING PAPER SERIES

THE LOCAL INFLUENCE OF PIONEER INVESTIGATORS ON TECHNOLOGY ADOPTION: EVIDENCE FROM NEW CANCER DRUGS

Leila Agha
David Molitor

Working Paper 20878
http://www.nber.org/papers/w20878

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
January 2015

We thank Joshua Aronson, Brittany Bychkovsky, Amitabh Chandra, Joseph Doyle, Amy Finkelstein, Jonathan Gruber, Nolan Miller, Jim Rebitzer, and Laila Saied, as well as seminar participants at Boston University, Georgia State University, MIT, NBER, and Northwestern Kellogg for helpful comments and suggestions. We are also grateful to Jean Roth and Mohan Ramanujan for assistance obtaining and managing the data. This research was supported by the National Institute on Aging, grant number T32-AG000186. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

The Local Influence of Pioneer Investigators on Technology Adoption: Evidence from New Cancer Drugs

Leila Agha and David Molitor

NBER Working Paper No. 20878
January 2015, Revised February 2017
JEL No. I11,O33

ABSTRACT

Local opinion leaders may play a key role in easing information frictions associated with technology adoption. This paper analyzes the influence of physician investigators who lead clinical trials for new cancer drugs. By comparing diffusion patterns across 21 new cancer drugs, we separate correlated regional demand for new technology from information spillovers. Patients in the lead investigator's region are initially 36% more likely to receive the new drug, but utilization converges within four years. We also find that “superstar ” physician authors, measured by trial role or citation history, have broader influence than less prominent authors.

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

David Molitor
University of Illinois at Urbana-Champaign
340 Wohlers Hall
1206 S. Sixth Street
Champaign, IL 61820
and NBER
dmolitor@illinois.edu

1 Introduction

Across many industries, technology adoption exhibits spatial clustering (Comin, Dmitriev, and Rossi-Hansberg, 2012). In medicine, technology diffusion shows similar patterns, with notable clustering within interpersonal networks (Coleman, Katz, and Menzel, 1957), hospitals (Escarce, 1996), and geographic regions (Baicker and Chandra, 2010). The potential for geographic proximity to facilitate the spread of knowledge and innovations across individuals or firms has been recognized at least since Marshall (1890). It has long been argued that local opinion leaders play a key role in this diffusion process (Rogers, 1962), yet direct empirical evidence has been limited. If local opinion leaders have significant sway, the value of well-informed, prominent physicians could extend beyond their contributions to the scientific literature or the treatment of their own patients; influential doctors may shape the practice of medicine across their region.

The limited evidence on the role of opinion leaders is in part due to the challenges of empirical identification. Clustered technology diffusion within a network, organization, or region could be driven by common local demand or capacity for technology adoption rather than information spillovers. Furthermore, it is difficult to isolate the role of opinion leaders as a specific pathway for knowledge spillovers. This study takes on both of these challenges, analyzing whether new cancer drugs approved by the Food and Drug Administration (FDA) are adopted more rapidly in the geographic regions containing study authors of the pivotal clinical study used in the FDA review process. First, by comparing diffusion patterns across many new drugs with different locations of study authors, we are able to separate local demand or taste for technology from the role of knowledge spillovers. Second, by comparing the influence of more and less prominent physicians as measured by academic citations and clinical trial role, we demonstrate that "superstar" authors have a substantially broader reach than their less prominent peers.

Understanding the determinants of technology adoption is of particular interest in the health care context, where new technologies are a key factor underlying both rising costs of care and improved health outcomes in the United States (Newhouse, 1992; Cutler, 1995; CBO, 2008; Smith, Newhouse, and Freeland, 2009). Against the backdrop of this aggregate growth is substantial heterogeneity across regions in the extent and speed of new technology adoption (Fisher et al., 2003a,b; Skinner and Staiger, 2005), and large geographic disparities in access to new cancer treatments (Nattinger et al., 1992; Farrow, Hunt, and Samet, 1992; Fairfield et al., 2010; Bristow et al., 2014).

While previous research has documented an extensive role of social learning in determining technology adoption in developing countries (Conley and Udry, 2010; Adhvaryu, 2014), we may expect little role for local information frictions in this setting where adoption decisions are made by expert physicians with high human capital and ready access to scientific information. On the other hand, it is a setting with substantial uncertainty about the efficacy and appropriate applications of newly introduced drugs. As a result, the clinical trial authors' detailed knowledge of drug mechanisms, patient responses, and side effects may put them and their peers at an informational advantage in the early stages of a drug's diffusion.

Existing empirical studies on 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] We expand on this work to investigate the geographic connection between research activity and the subsequent adoption of resulting technologies. Further, there is relatively little evidence on how the prominence or connectedness of an opinion leader affects his influence. Banerjee et al. (2013) show that the take-up of a new microfinance product in Indian villages depends on the network centrality of the village leader who is initially informed about the product. Our study sheds light on whether opinion leaders continue to matter in a context where technology adoption decisions are made by highly specialized experts.

Our analysis is based on a novel data collection effort that identified the study authors of the pivotal clinical trials for 21 new cancer drugs and matched the locations of these authors to adoption patterns of the drugs using Medicare claims records from over 1.4 million patient cancer care episode from 1998-2008. For scientific publications of pivotal clinical drug trials, the principal investigator is typically credited with the first author position. The last author is often a research scientist affiliated with the sponsoring drug company and is not a practicing physician for most of the drugs in our sample. We restrict attention to study authors who are also practicing clinicians.

The key finding from our baseline analysis is that patients treated in the hospital market where the first author is located are 36% more likely to receive treatment with the new drug within the first two years following a drug's FDA approval. This increased use is driven by higher rates of adoption by physicians both within and outside the author's practice group, suggesting the first author's influence extends beyond the boundaries of his organization.

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

By contrast, other physician study authors boost utilization only within their own physician group; while they share the first author's enthusiasm for the drug, they have a narrower sphere of influence and have smaller effects on regional patterns of care. We show that alternative definitions of "superstar" authors based not on authorship position but on citation counts to previous publications yield similar results; across a variety of definitions of "superstar" status, proximity to high-profile authors is associated with larger increases in local drug utilization than proximity to their lower-profile coauthors.

While initial proximity effects are large, the effects fade over time so that there is no discernible effect four years after a drug's approval. Despite this eventual convergence, initial differences in new drug use have significant implications for access to care, which we explore in Section 4 of this paper. This pattern of convergence also sheds light on the underlying driver of regional disparities in our context. The Roy model of medical treatment choice by Chandra and Staiger (2007) demonstrates 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, but our findings suggest that experience-related productivity spillovers are not a key factor in explaining regional disparities in cancer treatments.

Next, we explore heterogeneity in investigator influence across regions and find that investigator proximity has the largest impact on regions with the lowest levels of adoption of other new drugs, suggesting that information frictions are particularly acute in less technology-intensive areas. From a policy perspective, this also suggests that total utilization of a new drug may be influenced by the choice of investigator locations.

We also examine the impact that author location may have on where patients seek care. We document that appropriate patients appear more likely to travel into a study author's region after a new drug is approved, suggesting that patients may benefit from access to broad provider networks. An instrumental variables strategy based on whether patients reside in an author's hospital market reveals that differential patient sorting accounts for about one third of our main finding.

Finally, we probe the welfare implications of our findings by studying the survival improvements associated with the adoption of new cancer drugs. By comparing regions with fast and slow drug diffusion tendencies before and after the introduction of a new drug, we estimate that fast-diffusing regions show evidence of higher returns to new drug use. The survival improvements are so large that they appear unlikely to be driven solely by the greater fraction of patients receiving the new treatment in the fast-adopting regions. Rather, they point to larger average treatment effects

which could result from better physician selection of patients for treatment or improved dosing. This evidence provides further support for the idea that the local information environment may be a key determinant of both adoption and returns to new drug use, and suggests that policies that boost utilization without changing the quality of local information may fail to realize the full potential benefits of the new technology.

The organization 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 investigates the role and extent of patient travel and selective sorting. Section 5 describes evidence on the survival benefits of new cancer drug adoption and Section 6 concludes.

2 Setting and Data

In the United States, prescription drugs are regulated by the U.S. Food and Drug Administration (FDA), which between 2004-2013 granted approval for 26 new drugs per year, on average (FDA, 2014). In order to receive approval, new drugs undergo an extensive review process, in coordination with both the drug's sponsor (the manufacturer) and the FDA.[2] This process begins with the submission of an Investigational New Drug application, which includes a proposal for testing the drug on human subjects through clinical trials. FDA regulations place primary responsibility on the drug sponsor to select clinical investigators and research sites, and on a qualified institutional review board to review and approve each clinical investigator's qualifications before participation in the investigation (21 CFR § 312.53).

While each drug application may cite several studies from various stages of drug development, the applicant must pre-specify a "pivotal trial," which is typically a randomized controlled trial that provides the most comprehensive evidence to date on the efficacy of the drug. For establishing generalizable efficacy of treatment effects, FDA industry guidance states that drug sponsors may wish to invoke a multicenter trial design since the results arise from a broader patient population and multiple clinical settings (FDA, 1998). Drug sponsors may further minimize the risk of an unsuccessful trial by employing a design in which the investigator enrolls a small fraction of the total number of subjects, especially in cases where the investigator has a disclosable financial interest in the study (21 CFR § 54.5(c)). As we show below, the pivotal trials for all the drugs we study utilized a multicenter trial design.

Once clinical testing is complete, the drug sponsor submits the results as part of a New Drug

[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.

Application. If approval is granted, an official prescription drug label is written describing the indications for which the drug may be legally marketed. The FDA publicly releases this label as part of a detailed approval package describing the basis for approval and the pivotal clinical trial that provided the primary support for the approval decision.

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, in which case the efficacy of the new drug as an initial treatment is not 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.

In this study, we investigate the adoption and utilization of new cancer drugs in the years following initial approval by the FDA. There are two key data elements necessary for our analysis: the utilization of new cancer drugs across regions over time, and the location of physician study authors who lead the pivotal clinical trials on which each drug's initial FDA approval was based.

2.1 Measuring cancer drug use

We measure the diffusion of cancer drugs using Medicare Part B reimbursement claims over the 11-year period 1998-2008. During the study period, 21 new cancer drug agents covered by Medicare Part B were approved by the FDA.[3] The diffusion of these drugs forms the basis of our analysis.

While Medicare Parts A and B do not pay for most outpatient prescription drugs, an exception is made for drugs that are not typically self-administered, including cancer drugs administered intravenously or intramuscularly. These payments 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 cancer drug expenses; these costs rose 267% in the 7-year period since 1997, as compared to a 47% rise in total Medicare spending (Bach, 2009). Medicare Part B drug spending also comprises a significant share of total Medicare drug spending. As of 2010, spending on Part B drugs totaled $19.5 billion, compared to the $61.7 billion spent on Part D drugs which are typically self-administered (U. S. Government Accountability Office, 2012).

[3]While more than 21 new cancer agents were introduced over this period, only drugs that are typically administered in a doctor's office are billable through Medicare Part B. Because we can only track drug utilization with Medicare Part B claims, this restricts our sample to this group of 21 agents.

We analyze drug use at the level of Hospital Referral Regions (HRR), as defined by the Dartmouth Atlas of Health Care, which partition 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). The map in Figure 1 shows the boundaries of each HRR.

To track the adoption and use of new cancer drugs, we analyze a 100% sample of Medicare outpatient claims as well as a 20% sample of Medicare physician claims. For each drug in our sample, we study a cohort of patients diagnosed and treated for the targeted cancer type (e.g. colon, lung, etc.) for up to four years after the drug's initial approval. The unit of observation is the patient-year episode of cancer care: we analyze all claims associated with a cancer diagnosis for that patient within a calendar year. Our data comprise 1.4 million cancer care episodes within the first four years following drug introduction, of which 659,000 occur within the first two years after drug introduction. These data allow us to track the utilization rate of each new cancer drug among indicated patients across HRRs and over time following drug approval.

For this analysis, we identified 21 new cancer drugs that were covered by Medicare Part B and FDA approved between 1998-2007.[4] 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 first author on drug diffusion. The remaining four drugs are included in the sample to improve the precision of coefficients on other control variables.

Summary information for the 21 drugs in our study and their pivotal clinical trials is listed in Table 1, sorted in order of drug FDA approval date. 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. Nearly all the pivotal trials for drugs in our study were large, multi-center trials. On average, each trial enrolled 299 patients across 56 trial sites. 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 (11) reports the number of indicated patient cancer care episodes observed 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,

[4]We began with a list of 26 new cancer 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 agents, they were excluded from the analysis.

an observation in our regressions is a patient cancer episode, effectively weighting our regressions by the size of the target patient population; however, the results from our baseline specifications are qualitatively unchanged when each drug is given equal weight.

2.2 Author roles and locations

In addition to the Medicare claims data, we also collected a new data set linking cancer drugs to the pivotal 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. By matching the pivotal trial information in the FDA application to the authors of the academic article reporting the trial's findings, we are able to identify the researchers who were primarily responsible for the trial.

There were an average of 14 authors per paper in our sample, ranging from 6 to 26 (Table 1 column (9)). We restricted our analysis to studying the influence of authors who are also practicing clinicians, excluding from analysis the drug company employed scientists who often co-author clinical trials.[5] We categorized authors as "first" and "other" according to the order listed on the clinical trial, and we recorded each author's location based on the ZIP code of the author's institution at the time of the article's publishing. As indicated by Table 1, there are many more trial sites than authors for most drugs in our sample; authorship typically signals intellectual involvement with the research process that goes beyond facilitating a clinical trial site.

Our analysis exploits the convenient fact that authorship order is a strong signal of author contribution and involvement. In contrast to other types of biomedical publications where the last author is often the principal investigator, in large clinical trials, the first author is typically a senior physician who was leading the trial effort as the principal investigator (Baerlocher et al., 2007). The first author was a practicing clinician for all of the studies in our sample, in contrast to the last author position which was only held by a practicing clinican for 7 out of 17 drugs. (The last author is frequently an employee of the sponsoring drug company.) Furthermore, in 8 out of 17 cases, the first author was also the single most highly cited clinical author on the trial, but this is only true for 1 of the 7 last author clinicians.[6]

[5] In no cases was the first author of the study affiliated with the drug company. Often, the drug company employees' contributions were credited with a middle or final author slot.

[6] As further evidence of the prominent role of first authors in this setting, a search of clinical trials listed on ClinicalTrials.gov found that for over 80% of the drugs in our sample, the first author is registered as an investigator for at least one trial involving that drug. By contrast, only 24% of last authors are listed as investigators for that drug. We do not use the ClinicalTrials.gov database to assign author roles because many of the clinical trials in our sample pre-date the formation of this registry, so it is not possible to match to the pivotal trial to registry information, even if subsequent trials involving the same drugs and investigators are part of the ClinicalTrials.gov database.

In addition to using authorship position to determine a physician's trial role and prominence, we also develop a second measure of each author's prominence within academic medicine using data on publication and citation histories from Web of Science. We rank authors based on citation counts accruing to publications in the relevant medical field published in the 10 years leading up to FDA drug approval. Specifically, we find all research articles matching each author's last name and initials over a 10 year period leading up to the year of FDA approval. Using keywords coded within Web of Science, we restrict to articles only within the relevant field. For all authors in our sample, this includes oncology articles. Dermatology, neurology, hematology, urology, and nephrology are added to the definition of matched articles for drugs targeting those specific cancer types. These field restrictions provide a more targeted measure of prominence within the relevant medical field, as well as help disambiguate authors with common names. Next we count all citations that have accrued to those publications to the present date. Finally, we define the top 10% (or 50%) authors as the top-cited author on that drug trial plus any other author whose citation count places them in the top 10% (50%) of all authors on the same drug's trial.

This citation-based measure of "superstar" status will capture authors' academic prominence in their field, tagging the most highly cited authors on each trial. To disentangle "superstar" influence from any differences across drugs or subspecialties in rates of drug take-up and prominence of investigators, we prefer this relative measure of "superstar status" which allows us to compare the influence of authors within the set of researchers associated with each drug.

Hospital referral regions (HRRs) can be categorized based on their geographic proximity to authors of the clinical trial. In the discussion that follows, the phrase "first author HRR" refers to the HRR where the first author for the particular drug under analysis practiced. For our main analysis, we create non-overlapping definitions of author regions, so that a region cannot simultaneously be coded as a first and other author region, with first author designation taking precedence. Similarly, we code regions that contain "superstar" authors by the citation count metric.

The first authors on these trials practice at a wide set of academic medical centers that together span all four U.S. census regions. Author locations are pictured on a map in Figure 1, with first author locations marked by circles and other author locations by triangles. The most frequent first author locations within our sample are: Houston, Texas, (four first authors); Chicago, Illinois, (three first authors); Durham, North Carolina, (two first authors); and New York, New York (two first authors). There are 11 unique HRRs that contain a first author for at least one drug; 54

HRRs that contain a non-lead author for at least one drug, but never contain a first author; and 252 remaining HRRs that never contain any author (see Table 2). For our baseline results, we match patients to HRRs on the basis of where they received cancer care.

Within the HRRs that contain study authors, we further separate patients treated by doctors in the study authors' own practice groups from patients treated by other doctors in the region. To accomplish this, we group together all physicians who bill to the Medicare Carrier files using the same tax identification number as a clinical trial author. If a patient has at least one bill that year from a physician who is linked to a trial author's tax ID, then we code the patient as treated within an author's practice group. Because academic oncologists typically work as part of large group practices, this allows us to separate the author's influence within his own organization from his influence on outside physicians.

2.3 Summary statistics

The difference between a drug's utilization in the first author's region versus other regions—and how this difference changes over time since initial approval—is perhaps most easily seen in Figure 2. This figure plots the fraction of indicated patients treated in each HRR type who receive the new drug.[7] The solid line plots raw drug utilization rates in each drug's respective first author region (HRR), while the dashed line plots the average drug utilization in all other regions.

Two of the main results from our empirical analysis that follows are suggested by the raw data plotted in Figure 2. First, the figure shows that when a new cancer drug is introduced, indicated patients treated in the region containing the drug's first author are more likely to receive the drug than are patients in other regions. Second, this gap closes over time, so that drug utilization in the first author's HRR is no more intensive than in other regions after four years. This convergence occurs primarily because other regions increase their new drug utilization until it is similar to the first author region's use, and not due to substantial decreases in use in the first author's region. If physicians are learning about the drug's value over time, it appears they are learning that the drug is a valuable addition to their practice. We do observe a slight decrease in average utilization rates in first author regions between years 3 and 4, from 18.0% to 17.2%, which could reflect new competing drugs entering the market.

More detailed summary statistics are reported in Table 2. Over the first two years following drug approval, the average utilization rate of new drugs for indicated patient episodes ranges from 15.6%

[7] To make rates over time comparable, the drugs used to create this graph are restricted to the balanced sample (67%) of drugs in our sample for which four years of usage rates are available.

in regions where the first author practices to 8.6% in regions that never contain any investigators.

Among the regions that do not contain any investigators for a given drug (columns 3-4), those regions that contain authors for other in-sample drugs (column 3) are the more intensive adopters, with 9.2% of patients receiving new drugs compared to 8.6% in regions that never contain any investigators. This suggests that authors tend to be located in regions that have a high degree of enthusiasm, expertise, or patient demand for new cancer drugs in general.

Within the set of regions that contain an author for at least one in-sample drug, the first author HRR has 70% greater utilization rates on average (cf. columns 1 and 3). Thus, despite the overall higher rates of new drug use in regions that contain an author, utilization is even greater when the lead researcher of the particular observed drug is in the area.

The second row of Table 2 reports the fraction of patients treated in each region type who are ever treated by a physician in an author's practice group. 53% of patients treated in the first author's region receive treatment from a doctor in the same practice group as a study author; only 36% of patients in a middle or last author's region receive treatment from a doctor in an author's practice group. Given that the authors' practice groups do not have complete regional penetration for the targeted cancers, we can compare drug utilization within and outside the authors' own practice groups to test whether their influence extends beyond their own organization.

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 cancer drugs to identify the impact of geographic proximity to these investigators on drug utilization. If the location of study investigators 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 in an investigator's geographic region was due to this proximity. However, study authors are not chosen randomly, and their locations are likely to be correlated with other regional factors (e.g. co-location with innovation-loving physicians) that influence the rate of new drug adoption. As a result, a naïve comparison of the utilization rate of a new drug in a study author's region versus other regions does not separately identify the impact of proximity to a drug's pivotal study authors on the adoption rate of that drug.

The key methodological innovation in our analysis is to identify the effect of proximity to a drug's pivotal study authors by implementing an empirical design analogous to a difference-in-

differences approach. Specifically, we compare drug utilization in study author and non-author regions, controlling for baseline differences in each region's propensity to use new cancer drugs as well as controlling for time variation in drug utilization to capture the demand for each drug. Since we observe the diffusion of 21 newly introduced drugs, our strategy allows us to exploit each region's usage of other new drugs to establish its propensity to adopt a new drug when the region 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 in the absence of author influence.

Our baseline regression specification takes the following form:

$$\begin{aligned}(drug)_{ijtd} = \beta_f \mathbf{1}(\textit{first author HRR})_{jd} + \beta_o \mathbf{1}(\textit{other author HRR})_{jd} \\ + \{HRR \times \textit{disease-group FEs}\}_{ijd} + \{\textit{drug} \times \textit{year FEs}\}_{dt} + X_{it} + \varepsilon_{ijtd}\end{aligned} \tag{1}$$

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 two years following FDA approval of the drug.

The first two terms in the regression above are the key independent variables of interest: indicators for whether a study author of drug d's pivotal clinical trial is located in region j. To reflect the possibility that proximity effects may differ for these lead authors, we split our author proximity measures into separate indicators for whether a patient is treated in the first author HRR versus in a region containing any other author (but not the first author). The coefficients on these indicators describe how much more likely it is for a cancer patient to receive a new drug if treated in an HRR where an author of the drug's pivotal clinical trial is located. As discussed previously in Section 2.2, the first author on large clinical trial publications is typically a senior physician leading the trial effort as the principal investigator.

The third term in this regression is a vector of fixed effects measuring each HRR's propensity to use new cancer 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).[8] This allows

[8]Data limitations prevent us from controlling for even finer divisions of cancer types; for each separate cancer type, we must observe at least two new drugs (with different author locations), so that we can establish each region's counterfactual enthusiasm for new drug adoption within this diagnosis type. We chose to separate urologic cancers because they are typically treated by urologists who have completed a different type of fellowship training program (urology rather than hematology and oncology). It is common for hematologists to specialize in blood disorders,

regions to differ in their enthusiasm and patient suitability for new cancer treatments within each disease group. The fourth term in the regression allows each drug to face an idiosyncratic yearly shock to utilization that is common across regions. Finally, we include patient characteristics X_i which include indicators for patient sex, race, age (in 5-year bins), and whether this is a new cancer treatment episode (i.e. patient had no cancer claims in the previous calendar year).

The primary threat to the validity of this approach stems from the possibility that study author regions are systematically more likely to use the new drug (for reasons not driven by author proximity) 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. As outlined in further detail below, we also address this potential threat to validity in a number of other ways, including limiting the analysis to regions that ever contain a study author and studying the persistence of our estimated proximity effect. To preview our findings, the fact that the measured proximity effect converges within four years suggests that there are no permanent differences in patient appropriateness or latent demand for the new drug in first author regions.

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 delivered. Thus, any effect author status has on a region's propensity to prescribe a new cancer drug could be driven by two separate channels: (a) a *prescribing* effect in which providers in the author region have an increased propensity to treat a fixed population of patients with the new drug; and (b) a *sorting* effect in which patients suitable for particular treatments sort to providers who specialize in those treatments.[9] For example, an increased number of suitable patients may travel into an 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

including leukemias and lymphomas, and not treat solid tumors. While oncologists may also subspecialize in treating particular types of solid tumors, their fellowship training spans cancer types and it is not uncommon for the same therapeutic agent to find applications to other solid tumors besides the type for which it first gained approval. As a result, we group the solid tumors together for reasons of statistical power.

[9]Note that both channels would be present even under true random assignment of study authors to geographic regions.

patient. The essential point is that both θ and p may respond to a change in a region's author 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 first author 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 in equation (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, Section 4 applies an instrumental variables approach to isolate the change in utilization driven by prescribing, i.e. the increased propensity of first author regions to treat a given patient population.

3.2 Baseline Proximity Effects

Effects by Geographic Proximity

We begin by presenting evidence on whether geographic proximity to a new cancer drug's pivotal study authors impacts a physician's propensity to prescribe that drug for indicated patients. We begin by examining how the effect of geographic proximity to a study author 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 cancer drugs across author HRRs and other regions. An alternative model such as Phelps (2000) where information asymmetries are the reason for delayed adoption amongst non-first author regions would predict convergence in practice patterns as information about the new treatments reaches each physician.

To measure the evolution of the author proximity effect, we estimate a modified version of specification (1) in which the author HRR indicators are interacted with a full set of event-year dummies ranging from 1-4 calendar years following drug approval (0 corresponds to FDA approval year).[10] The coefficients on these interactions describe the corresponding proximity effect separately

[10]Medicare drug codes are not introduced until the calendar year following FDA approval for the large majority of

for each year.

The left panel of Figure 3 plots how the estimated effect of proximity to a drug's first author on drug utilization evolves over time, while the right panel plots the effect of proximity to any of the drug's other 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 more intensively on average in regions containing a study investigator, an effect that is much stronger in the first author's region. The second pattern highlighted by Figure 3 is that the proximity effect fades over time, so that any proximity effect on drug utilization vanishes within 4 years after drug approval.

This figure provides a novel way to benchmark the speed of technology adoption. Prior measures of the speed of technology adoption have primarily focused on absolute rates of take-up, such as the length of time since invention for an individual to adopt (e.g. Comin and Hobijn, 2010), or rate of acceptance (e.g. Griliches, 1957). However, these measures can be inappropriate in settings where the "optimal" level of adoption is difficult to ascertain (e.g. due to informational uncertainty) and may even change over time, as competing technologies are introduced and the scientific undertanding evolves. In contrast, our measure of convergence describes how quickly regions conform to a benchmark adoption pattern set in regions containing the experts involved in the technology's development.

Taken as a whole, these estimates suggest that proximity to a pivotal study trial investigator drives higher take-up of new drugs, an effect which is stronger and more persistent for first authors than for other early investigators of the drug. Yet despite the initial eagerness to use the drug, this difference in diffusion between investigator and non-investigator regions converges within a few years. This convergence provides further support for the econometric assumption that the drug's first author is not located in a geographic area with idiosyncratically higher latent demand for that particular cancer drug; if this were the case, we may expect to see persistent differences in drug use across the first author and non-first author regions.

Table 3 shows results from our baseline specification in Equation (1). Because the main proximity effects were found in Figure 3 to be concentrated in the first two years following FDA approval, we focus the remainder of our regressions on this period. As shown in column (1) of Table 3, we estimate that patients who receive cancer treatment in the first author's HRR are 4.0 percentage points more likely to receive the new drug, significant at the 1% level. To provide a useful benchmark, this first author impact is a 36% increase over the 11.1% average utilization rate in regions

drugs in our sample, limiting our ability to measure diffusion prior to the first calendar year.

that contain a first author for a different in-sample drug with US-based first authors. Patients who receive treatment in a middle or last author's region, by contrast, are only 0.69 percentage points more likely to receive the new drug, an estimate which is positive but not statistically significant. The difference between utilization in first author and other author HRRs is statistically significant at the 5% level.

Extent of author influence: results by physician group

The average effect a study author has on prescribing behavior within the HRR may obscure important heterogeneity in regional utilization. The smaller average effect of "other" study authors relative to the first author could result from a narrower sphere of influence or from less enthusiastic adoption even within the other author's practice group. To explore this possibility, we use physician group tax IDs from a 20% random sample of all Medicare physician claims between 1998-2008 to measure which physicians practice in the same organization as a drug's trial authors.

Patients indicated for a given drug and treated in an author HRR are marked as treated by an author's group if the patient was treated by a physician practicing in the same group as a trial author at any point during the year. As reported in Table 2, 53% of indicated patients treated in the first author's region receive care from a physician who is part of a study author's practice group; 36% of patients in other author regions are treated by a physician from a study author's practice group.

To estimate proximity effects separately within and outside an author's physician group, we estimate a modified version of equation (1) where the indicators for being treated in an author HRR are interacted with indicators for being treated by the author's physician group. Column (2) of Table 3 shows the results of proximity separately by author group status. Among patients treated in the first author's HRR, patients treated within an author's physician group are 4.21 percentage points more likely to receive the new drug ($P < 0.01$) compared to patients treated in non-author regions, while patients treated outside the author group are 4.16 percentage points more likely to receive the drug. The point estimate suggests that the first author appears broadly influential, increasing new drug adoption in his region by almost equal amounts within and outside his practice group. However, the statistical significance of the utilization boost outside the first author's practice group is sensitive to the inference method, as discussed at greater length in the Robustness section which follows.

The results are quite different in other author regions. In those regions, patients treated within

an author's practice group are 2.8 percentage points more likely to receive the new drug ($P < 0.01$). However, there is no estimated increase in drug utilization outside the author group, and the 95% confidence interval is bounded above by 0.7 percentage points. Despite being enthusiastic adopters of new drugs within their own practice group, middle and last authors do not appear to influence practice patterns in neighboring physician groups.

An important consideration for interpreting these results is that because patients are not randomly assigned to doctors, some of the increased utilization found within author group could be driven by a compositional shift in which patients are treated by author group physicians. To the extent that patients most appropriate for the new drug sort into the author group, some of the increase in prescribing rates within the physician group may not correspond to a net increase in propensity to treat a given patient with the new drug. The region level results in column (1), however, indicate that overall prescribing does in fact increase at the region level.

Taken as a whole, the results from columns (1-2) provide evidence of important proximity effects within author regions. The higher rates of drug utilization in a first author region compared to other study author regions appear to be driven primarily by the first author's broader sphere influence. While both types of study authors boost utilization within their own practice group, the evidence suggests that only first authors substantially increase drug adoption by doctors who are not part of the author's firm.

Robustness

To probe the robustness of the baseline proximity effects reported in Panel A of Table 3, we run an identical analysis in a restricted sample that only includes regions with at least one study author for a drug in our sample. Restricting the sample mitigates the concern that non-research-intensive regions provide a poor counterfactual for a new drug's popularity in author regions. Results are reported in Table 3, Panel B, with estimates very similar to those found in the full sample reported in Panel A. We continue to find approximately 4 percentage point higher new drug use in the first author region ($P < 0.01$) and a smaller, insignificant aggregate effect on other author regions. Notably, restricting the sample in this way does not substantially attenuate the estimated effects.

The interpretation we emphasize for our baseline results is that increased utilization in author HRRs results from knowledge spillovers from proximity to lead investigators. However, a competing explanation is that heightened levels of utilization among patients in author HRRs might simply reflect the continuation of treatment for patients enrolled in the trial itself. The most direct and

perhaps ideal approach to test the continuation hypothesis would be to identify patients in our analysis who were enrolled in the pivotal trial, and to remove them from the analysis in a robustness check. While the Medicare data do not permit identifying patients enrolled in specific clinical trials, we re-estimate our baseline results in a sample restricted to new cancer patients, defined to be those with no cancer treatment observed in the previous calendar year. This sample restriction plausibly removes patients enrolled in the pivotal study and eligible for Medicare who survive and continue to receive the new drug following FDA approval. Estimates based on new cancer patients alone, reported in Panel C of Table 3, are nearly identical to our baseline estimates in Panel A of Table 3 that include all cancer patients. We find these results to support the interpretation that our baseline effects are driven by informational spillovers, rather than a continuation mechanism alone.

There are several other reasons why it is unlikely that the continuation mechanism drives the results we observe. First, while the largest trial sites tend to be located in an author HRR, they are not systematically located in the lead author's HRR.[11] Second, even for a large site with many patients enrolled, it is likely that few of the patients assigned to receive the study drug would show up in our Medicare analysis sample.

To give a back of the envelope sense of the magnitudes that could be explained by continuation effects, we approximate the number of potentially continuing patients in our sample, using information from FDA medical reviews. As a plausible upper bound, we allow the lead author trial sites to enroll 33 patients on average, the average size of the largest site in our analysis of FDA disclosures.[12] Of these 33 patients, assume 16.5 patients are randomized to the treatment arm receiving the study drug, 8.25 (50%) of these patients survive until after FDA approval, and 2.9 (36%) of the survivors are eligible for Medicare. Even if all 2.9 of these patients continued with the new

[11] We examined FDA medical reviews and clinical trial publications for each of the 21 drugs in our sample to determine, when possible, the total number of patients enrolled in the pivotal clinical trial, the number of trial centers enrolling patients in the trial, the number of patients enrolled in author regions, and the age distribution and median survival of patients in the trial. On average, the pivotal trials for the drugs in our sample enrolled 299 patients across 56 trial sites. While neither the FDA medical reviews nor clinical trial publications systematically report on enrollment by site, select information about trial sites with high enrollment was available for 12 (57%) of the drugs in our study. The average trial size for these twelve drugs was 300 patients, nearly identical to the average trial size of 297 patients for the remaining nine drugs. For the 11 drugs for which we could determine the author location of the largest trial site in the study, 8 (73%) were not located in the lead author's HRR.

[12] Except where otherwise stated, the parameters for this calculation derive from the FDA medical reviews described in the previous footnote. Across the 12 drugs for which site enrollment information was available, the trial site enrolling the most patients enrolled 33 patients, on average. The average age of enrolled patients was approximately 58 years; only 6 of the trials reported the fraction of elderly patients, and in all cases fewer than 36% of patients were 65 years or older. We therefore take 36% as our estimated fraction of trial patients eligible for Medicare. For the 13 drugs reporting median survival in the trial publication, median overall survival in the most favorable treatment arm reported was 14.3 months. This suggests that fewer than half (50%) of individuals enrolled in the pivotal study would have survived to the years following FDA approval of the drug, the baseline years of our analysis.

drug following FDA approval and were counted in our analysis, this would only account for 0.7% of the average base population of 388 indicated cancer patients per drug in the lead author region (Table 2, Column 1, Row 4), only a small fraction of the 4 percentage point increase in new drug utilization we observe in lead author HRRs in the first two years following FDA approval. This calculation further suggests that the continuation mechanism is not likely to explain a significant share of the increase in utilization we estimate in lead author regions.

Finally, we explore robustness of our findings to bootstrap methods of statistical inference, with results reported in Appendix Table A1. Our main results use conventional cluster robust variance estimators, supported by the fact that we have many clusters (6,086 region-drug pairs) of which 18 clusters are "treated" (i.e. first author region for given drug). To account for possible bias due to a modest number of unequally sized treated clusters, we apply the wild cluster bootstrap developed by Cameron, Gelbach, and Miller (2008) and Cameron and Miller (2015). We also vary the cluster definition, estimating p-values with clusters at the HRR-drug level as in our baseline results, as well as at the broader HRR level.

Note that when there are a small number of treated clusters, results from MacKinnon and Webb (2016) suggest that conventional clustered standard errors are likely to over-reject, but the wild cluster bootstrap tends to under-reject. As a result, the wild cluster bootstrap estimates should be conservative, with the true p-values lying between the values given by each approach.

Our main findings remain statistically significant across all four methods of inference reported in Appendix Table A1. In particular, we consistently find that first author regions have significantly higher drug use than non-author regions (p-value ranges from 0.002 to 0.012, depending on clustering method). We also find that the impact of being treated by the author's group in a region that contains a middle or last author consistently significant across inference approaches ($P \leq 0.002$ with all methods).

One result is sensitive to the inference method: our finding that patients treated in the first author's region but outside the first author's group are significantly more likely to get the new drug than patients in non-author regions. The p-value is marginally significant with conventional CRVE ($P = 0.048$) and becomes not statistically distinguishable from zero with the wild cluster bootstrap ($P = 0.168$ or 0.278 depending on cluster level). While the point estimates for utilization in the first author region within and outside the author's practice group are very close, we caution that the estimate of utilization rates outside the author's group is imprecisely measured.

Mechanisms of estimated author influence

The greater impact of proximity to first authors could be driven by two potential factors: (a) first authors take on more responsibility for analyzing and writing the paper, and thus are better informed about the new drug's value; or (b) even if all authors had the same quality of information about a new drug, first authors may be more influential due to their greater professional stature. Both channels have potential *a priori* support.

The pivotal clinical trials for drugs in our study have an average of 14 authors per paper (Table 1). In the publication of these trials, the first author often takes the lead role in trial design and preparing the manuscript (cf. Hudes et al., 2007; DeAngelo et al., 2007), suggesting he may also have the most detailed, comprehensive view of the drug's efficacy. On the other hand, the first author is also likely to be one of the highest profile physicians involved with the research; he is the single most highly cited clinical author for 8 of the 17 drugs in our sample with US based trials. The findings reported in the previous section on authors' scope of influence outside their practice group lends support to the idea that the first author's status as a local opinion leader may primarily drive these differences. In the next section, we will further explore whether identifying "superstar" authors based on citation histories rather than authorship sequence leads to similar findings.

Another complementary explanation for the observed impact of first authors is that they may have stronger ties to the sponsoring drug companies and be more actively involved in drug promotion efforts. Of the 21 drugs in our sample, 9 of the published trials report disclosure statements detailing which authors have financial relationships with drug companies. For these drugs, an average of 52% of all clinical authors report financial ties to the sponsoring drug company, compared to 67% of first authors. These financial ties include consulting fees, lecture fees, research support, expert testimony, and stock ownership.

While drug companies are only 1.3 times more likely to have financial ties to a first author compared to a middle or last author, the estimated impact of being treated in the first author's region on new drug use is over 5.5 times larger than the estimated impact of being treated in another author's region. If each disclosed financial tie indicated an equal amount of funding support and financial ties were the only driver of our observed effect, then we would expect differences in the frequncy of financial relationships to scale linearly with the estimated effect of author proximity. Hence, the observed frequency of financial ties between drug companies and clinical authors would not lead us to predict the first author's apparent outsized influence on regional drug utilization.

Unfortunately, our assessment of these differences is limited by the fact that many drug trials did not report disclosure statements over this period, and even trials with disclosure statements do not list the amounts of money exchanged. As a result, we cannot rule out the possibility that drug companies have stronger relationships or expend more resources supporting the first author's drug promotion efforts compared to other authors. Crucially, if drug companies were investing more in the first author, this would suggest that they perceived a higher return to the first author's potential promotion efforts; in that case, the drug company's investment is complentary to the superior information or professional stature that the first author already offers.

A related issue is that because the geographic location of a study author is also typically the location a trial site (though one of many), it is difficult to disentangle whether local physicians primarily learn about a new drug because of proximity to study author or proximity to trial activity. To explore this issue, we first consider whether our contrasting results on first author and other author regions could be driven by differences in the scale of the trial site. As described previously, for the 11 drugs in our study for which we could determine the author location of the largest trial site in the study, 8 (73%) were not located in the lead author's HRR. Thus, it does not appear that the differential effect of first author proximity compared to other authors can be easily explained by substantial differences in the size of the trial site by author type. However, to the extent that local trial activities such as patient enrollment and clinical care patterns are themselves an intermediate outcome influenced by the local investigator leading that trial site, trial activity may be an important channel through which the investigator proximity effect works.

Outside of drug company sponsored events, there are many other opportunities for oncologists to meet with their local peers and share ideas. Within a given oncology practice group, formal mechanisms may include the establishment of internal drug treatment protocols and "tumor board" meetings where treatment options for new cancer cases are often discussed with a broad team of care providers. Across separate practice groups, opportunities for sharing ideas include invited "grand rounds" seminars, local and regional professional society meetings, contact through shared patients and patient referrals, as well as casual interpersonal networks.

In personal communications with oncologists, physicians described significant barriers to the adoption of new cancer agents. Because trial participants are often selected for being in more stable health than many cancer patients, physicians cited significant concern about the risk of severe side effects and uncertainty about optimal dosing regimens. Oncologists are also aware of the potential importance of heterogeneous responses to treatment; hearing about individual

successfully treated patients may be more compelling than reading about modest average response rates. The expertise of a prominent physician in the community on a new drug's applications and efficacy could substantially lessen these barriers to new drug adoption.

Finally, another explanation for the observed regional differences in drug utilization is differences in drug prices. There are two relevant prices to consider in this context: the reimbursement that physicians are paid for prescribing a drug, and the cost to the physician to purchase the drug. If drug-specific reimbursements were higher or purchasing costs lower in first author regions than in other regions, prices could potentially explain our proximity findings. However, we think neither of these price effects is a likely explanation in our context. Because we are studying utilization amongst Medicare patients where reimbursement rates are set by administrative rule, there is effectively no scope for reimbursements to be idiosyncratically higher in a given region for first author drugs. In addition, while it is conceivable that drug manufacturers could offer drug discounts to specific physicians or groups and might specifically target physician's practicing at the first author's hospital, our high estimated utilization rate for local physicians who do not practice in the first author's physician group makes it seem unlikely that drug discounts could explain our results.

3.3 Superstar Proximity Effects

Our baseline results in Section 3.2 suggest that clinical trial authors with greater expertise or prominence (as captured by first author status) have a greater impact on drug utilization in their region. To further explore the differential effect of "superstar" physicians, we apply our citation-based measure of author prominence which identifies the top 10% (or 50%) authors as the top-cited author on that particular drug trial plus any other author whose citation count places them in the top 10% (50%) of all authors on the same drug's trial. Because we want to make comparisons of doctor influence within the set of physicians on a particular drug trial, this relative measure provides a clear comparison that will ensure our ability to separate the relative prominence of authors on each trial. Note that because trial authors are a selected population, it is likely that these citation ranks would be even more favorable if compared to the overall population of physicians in their field.

For any measure of superstar status, our baseline regression in Equation (1) is easily modified to estimate the differential impact of proximity to a superstar author. For these regressions, we allow author proximity effects to vary by drug, and then estimate the differential impact of proximity to

a superstar author. Our superstar regression takes the form:

$$(drug)_{ijtd} = \beta_s \mathbf{1}(superstar\ author\ HRR)_{jd} + \{author\ HRR \times drug\ FEs\}_{jd} + \{HRR \times disease\text{-}group\ FEs\}_{ijd} + \{drug \times year\ FEs\}_{dt} + X_{it} + \varepsilon_{ijtd}. \quad (2)$$

The key coefficient of interest is β_s, which describes how much more a new drug is used in a superstar HRR relative to other author HRRs, on average. Thus, if $\beta_s = 0$, utilization in a superstar region is no more intensive than in other author HRRs, while $\beta_s > 0$ corresponds to higher utilization in superstar regions. The second term in this regression allows the effect of author proximity to vary by drug, and the last three terms are the same as in Equation (1).

The results from Regression (2) are shown in Table 4. Column (1) shows that drug utilization in first author HRRs is 3.00 percentage points higher on average than in middle or last author HRRs for the same drug, which closely matches the result obtained by differencing the first and other author HRR results in Table 3. From columns (2-3), new drug utilization is 2.27 percentage points higher in a top 50% cited author HRR (column 2), and 2.32 percentage points higher in a top 10% cited author HRR.

Columns (4-6) run horse races between these three measures of superstar status. Column (4), which includes superstar indicators for both first author and top 50% cited author shows that both indicators correspond to higher utilization; the coefficient on top 50% cited author is marginally significant, with $P = 0.056$. Column (5) includes first author and top 10% cited author indicators. In this case, both coefficients point to higher utilization, although only the first author corresponds to a statistically significant increase. Similarly, in column (6) which includes all three measures of superstar status, all coefficients are positive, but only significant for the first author.

Each of the author regions contains a physician investigator who is well informed about the new drug, but among those regions, those with the most prominent authors are the ones that experience the most substantial increases in new drug use. These findings suggest an important role of local opinion leaders even in the context of drug adoption by highly expert decision makers with access to clinical trial findings. Regional information frictions may dissipate within 3-4 years, but during the initial two years after drug introduction, local opinion leaders have substantial influence on adoption rates in their region. A subtle but important note is that these results do not estimate the causal impact of increasing an individual author's citation history or authorship order *ceteris paribus*; principal investigators and authors with high citation counts are likely to be exceptional

along other unmeasured dimensions as well.

3.4 Extent of Investigator Influence

In this section, we probe the extent of drug authors' influence. In particular, we test which types of regions are most heavily influenced by investigator proximity; whether study authors affect utilization in neighboring regions; and lastly, whether study authors affect off-label drug use.

First, we test whether regions that are typically slow to adopt new cancer drugs experience a greater boost in utilization when a study author is located in the region. There may be greater scope for the study author to affect practice patterns in slower-adopting regions that are not already very high users of new cancer drugs.

We develop a measure of each region's speed of new cancer drug adoption by looking at the average rates of new drug use when no author is present in the region. In particular, we regress an indicator for new drug use on a series of region dummy variables, controlling for drug by year fixed effects, patient demographic characteristics and whether this is a new cancer spell. To avoid including the direct impact of author proximity in this measure of regional adoption speed, we exclude from the sample any observations where an author for the relevant drug was located in that region. This regression includes only observations in the first two years following initial FDA approval. The region fixed effects from this regression form the basis of our measure of regional technology adoption speed.

For ease of interpretation, we standardize this measure of technology adoption speed by de-meaning the variable and dividing each fixed effect by the standard deviation. As a result, the average regional technology adoption speed index is 0, and a value of 1 corresponds to a region with average new drug use 1 standard deviation above the national mean.

For estimation, we augment our baseline estimating Equation (1) to include interaction terms between whether the first (or other) author is located in the region and the region's technology adoption speed index. Results are reported in Table 5, columns (1-2). We find that the first author's influence is greatest in regions that are typically slower to adopt new drugs; the interaction term is negative and statistically significant at the 5% level. For regions near the mean of the adoption speed index, patients are 4.1 percentage points more likely to receive treatment with the new drug when the first author is located in that region. The impact of being treated in a first author HRR increases to 6.2 percentage points for regions that are typically 1 standard deviation slower to adopt than the average region; the effect falls to 1.9 percentage points for regions that are typically 1

standard deviation faster to adopt.

We continue to find no significant effect of being treated in a middle or last author's region. The coefficient on the interaction between other author HRR and regional adoption speed index similarly suggests that slower-adopting regions experience a greater effect of proximity to other study authors; however, the result is not statistically significant.

The first author's influence boosts regional use more for regions that tend to be technology adoption laggards. From a policy perspective, this suggests that investment in clinical research may yield the greatest spillovers to medical practice in regions that are not already among the fastest adopters of new technologies. It should be noted that this effect is estimated only using variation in adoption speed within the set of regions that contain a first author for at least one in-sample drug.

Next, we test the geographic extent of investigator influence. This analysis not only addresses the geographic extent of the authors' reach, but also impacts the interpretation of our baseline estimates from specification (1). There, we estimated the wedge between investigator HRRs relative to non-investigator HRRs. 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 estimated proximity effects that are too small.

To measure whether investigator influence extends beyond his own HRR, for each drug we identify the "neighbor" HRRs that share a border with the HRR in which the drug's first author is located. We augment our baseline estimating Equation (1) to include two indicator variables: one for patients treated in a region that neighbors a first author region, and a second for patients treated in a region that neighbors a middle or last author region.

Table 5 shows in columns (3-4) that while first author HRRs have a 4.0 percentage point increase in their propensity to use the new drug, there is no observed increase in drug use of neighboring HRRs. The point estimate suggests a less than 0.2 percentage point increase in new drug utilization in neighboring HRRs, which is small in magnitude and statistically not distinguishable from zero. This null effect is quite precise, with a 95% confidence interval that excludes effect sizes that are one third as large as the impact of being treated in the first author's HRR. There is a similarly small, insignificant effect estimated for neighbors of other author HRRs. Although the first author's influence may extend beyond physicians in his own practice group to other physicians practicing in the same region, there is no evidence that his influence raises utilization in neighboring regions.

Finally, we investigate whether study authors influence the use of new drugs for applications

not covered by the initial FDA approval label. While drug labels typically provide relatively narrow indications for application, physicians have wide latitude in determining how they will prescribe the drug.[13] Across our 21 drugs, 22% of utilization within the first two years was for patients with diseases not indicated on the FDA label, which we will call "far" off-label drug use. In columns (5-6) of Table 5, we estimate whether the study authors' influence increases the use of new drug for other applications. The sample is restricted to patients who do *not* have the broad cancer type covered by the initial FDA label, and the estimating equation mirrors the specification in Equation (1).

We find no evidence of higher use of the drug for off-label patients in the authors' regions, suggesting the authors' influence is largely local to the cancer type on the initial label. A limitation of this analysis is that it is relatively unusual for any given off-label cancer patient to receive treatment with a particular new drug; mean utilization is 0.37 percentage points in regions that ever contain a study author. The point estimate from column (5) suggests that off-label utilization increases by 0.06 percentage points when the first author is in the region; the 95% confidence interval bounds the effect as no larger than 0.16 percentage points.

If increased use in the first author's region was driven by a pure "advertising" or "salience" effect boosting awareness or enthusiasm of the new drug, we might have expected greater spillovers to off-label applications. On the other hand, the authors' expertise may be local to the indications studied in the clinical trial they led; they may not have a strong informational advantage when it comes to applications of the drug beyond that population. Taken as a whole, the evidence in this section suggests that first authors boost on-label drug use within their own region, especially when they are located in a region of relative technology laggards, but that authors have little measurable influence on use in neighboring regions or on applications of the drug to other populations.

4 Patient Travel and Selective Sorting

As discussed in Section 3.1, there are two possible channels through which the observed increased propensity to prescribe new drugs in first author regions may occur: first, an increased propensity to use the drug on a fixed set of patients; and second, a change in patient sorting such that the first

[13] For example, capecitabine was initially approved in 1998 for the treatment of metastatic breast cancer that had already proved resistant to both paclitaxel and an anthracycline-containing chemotherapy regimen. In the preceding analysis, we analyzed the use of capecitabine across all breast cancer patients, since our data do not allow us to capture the more specific clinical conditions defining the label indications. However, in the first two years after capecitabine's introduction, a full 39% of its use was on colon cancer patients; colon cancer and breast cancer may be biologically similar, but robust clinical trial evidence was not yet available for the application to colon cancer. The FDA eventually added colon cancer to the label in 2001–after the two-year period covered by our main analysis.

author regions see patients with higher latent demand. In this section, we test directly for changes in patient sorting as indicated by patient travel patterns, and then use an instrumental variables strategy to identify the differences in drug utilization that occur over a fixed set of patients.

In Table 6, we begin by testing whether patients with the targeted diagnosis who seek treatment in the first author's HRR are more likely to have traveled from a different HRR of residence. This would occur if, for example, savvy patients travel into author regions for treatment in order to gain access to the new cancer drugs. In columns (1-2), 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 travel, defined by whether the patient's HRR of residence is not the same as the HRR where he receives care.

In the baseline specification reported in column (1), we find a 3.3 percentage point increase in the fraction patients treated in the first author's HRR who do not live in the region, significant at the 10% level; on average 22.2% of patients treated in an author region reside outside the region. There is a similar 3.0 percentage point increase in the fraction of traveling patients treated in other author's HRRs, significant at the 1% level. Restricting to the set of regions that ever contain a study author in column (2) yields similar results.

This evidence suggests that some patients are aware of new centers of expertise for the new cancer drug (perhaps due to physician referral) and are willing to travel further to improve their access to the drug. Note that these findings do not necessarily require patients to cross large distances for care; 66% of travelers are being treated in a neighboring region that shares a border with the patient's HRR of residence. In an unreported regression, we found that these travel effects are driven by patients seeking care in an author's practice group; excluding patients treated by an author group leads to negative, insignificant coefficients in the travel regression. Taken together, these findings are consistent with the possibility that some patients are on the margin between seeking care at their local tertiary care center or traveling to a neighboring, prestigious academic medical center for treatment. If the center has recently participated in a clinical trial for a new drug treating the patient's cancer, the patient may be more likely to travel for treatment.

If these patients who are newly traveling into an author's practice group 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 in the author's region may be driven by the changing patient composition. In Table 6, columns (3-4), we test whether the patients who travel from outside HRRs differ in their propensity to receive the new drug relative to non-movers. In these columns,

we report results from a regression that augments our baseline Equation (1) by interacting the author proximity indicators with a binary indicator *traveler* for whether the observed patient is seeking care outside his HRR of residence. Based on this regression, we estimate that patients traveling to the first author's HRR are 2.2 percentage points more likely to receive treatment with the new drug than patients treated in the first author HRR who also reside within that HRR, although the result is not statistically significant at conventional levels. Travelers to other author regions are only 0.1 percentage points more likely to get the drug, which is also not significant. The confidence intervals on these travel estimates allow the possibility that the overall 4.0 percentage point higher new drug use in first author regions may be driven at least in part by changing patient composition, and not solely by a higher propensity to use the drug on a fixed set of patients.

These estimates on patient travel are particularly relevant considering the new attention to provider networks available on the Affordable Care Act's health insurance exchanges. A common feature of these new insurance plans is restricted provider networks (Hancock, 2013), with consumers facing much higher prices for out-of-network care. Our findings on travel suggest that severely ill patients, such as the cancer patients in our study, may travel strategically to improve access to providers with additional expertise in new treatments.

Instrumental variable analysis

To isolate whether trial author regions are indeed more likely than other regions to use the drug on a given set of patients, we pursue an instrumental variables (IV) strategy. In particular, we use indicators for whether the patient resides in the first or other author's HRR as an instrumental variable to predict whether they will seek treatment in an HRR that contains the first or other author for the relevant drug. This instrumentation strategy mitigates the concern that patient sorting renders the patients treated in the first author region more suitable to treatment with the new cancer drug.

The reduced form equation of the IV model takes the following form:

$$\begin{aligned}(drug)_{ijtd} = \gamma_1 \mathbf{1}(reside\ in\ first\ author\ HRR)_{ijd} + \gamma_2 \mathbf{1}(reside\ in\ other\ author\ HRR)_{ijd} \\ + \{HRR \times disease\text{-}group\ FEs\}_{ijd} + \{drug \times year\ FEs\}_{dt} + X_{it} + \varepsilon_{ijtd}\end{aligned} \tag{3}$$

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 the two indicators for residence in an author HRR as instrumental variables, we also include

two additional instruments: (1) residence in a first author's neighboring HRR, i.e. a region that shares a border with the first author HRR; and (2) residence in an other-author's neighboring HRR, i.e. a region that shares a border with an other-author HRR.

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 cancer drug. For example, because we include region by drug class fixed effects, this allows regions to vary in their latent demand for drug classes, but does not allow author regions to have higher latent demand for the author's drug compared to other drugs in the same class.

The exclusion restriction could be violated under a few conditions. One possibility is that patients with the targeted cancer who reside in the first author 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 first author'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 cancer 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 cancer drug trial. This assumption is further bolstered by the observed convergence in drug usage across first author and non-first author regions, as reported in Figure 3, suggesting no permanent differences in patient eligibility for treatment in the first author's region.

Results from the IV regressions are reported in Table 7. The reduced form results show that patients residing in the first author's HRR are 2.3 percentage points more likely to receive treatment with the new drug, significant at the 5% level; there is no significant increase in use for patients residing in other author regions. The IV estimate reported in the final rows of the table rescale the reduced form estimate and shows that providers in the first author's region are 2.9 percentage points 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 first author for any drug, as reported in column (2). Adding the instrumental variables for residence in neighbor HRRs to the model also does not substantially change the estimated IV coefficient (cf. column 3).

The IV results suggest that over 70% of the baseline effect reported in section 3.2 is due to the increased propensity of physicians in the first author's region to prescribe the drug on a given set of patients. While the differential sorting of high-demand or high-appropriateness patients to author regions explains some of observed the boost in drug utilization, most of the effect is driven by differences in physician behavior, not patient sorting. Under a local average treatment effect interpretation, the IV result implies that doctors in the first author's region are 2.9 percentage points more likely to use the new drug on a given set of patients, namely those for whom location of residence determines location of care.

Taken together, the patient traveling results and the IV regressions find support for both hypothesized channels by which the presence of a first author may affect care in his region. Patients with high latent demand for the drug seem to seek out care in areas with high expertise in the new technology. In addition, doctors in the first author's region are more likely to use the new drug, holding fixed the population of patients seeking treatment, with this channel representing the primary driver behind our baseline result.

Extrapolating from the IV regression result, if all providers behaved like those in the first author's region, approximately 2,500 additional Medicare Fee for Service (FFS) patients would be treated with each new cancer drug in the first two years after initial drug approval. This amounts to an estimated 53,000 Medicare FFS patients in total over the 11 years of our sample who did not receive treatment with one of the 21 drugs under study due to the lower patterns of initial usage in areas that didn't contain the study's first author.

5 Welfare implications of early drug diffusion

The evidence presented thus far suggests that clinical trial investigators influence local drug use in the first few years following drug introduction. A key remaining question is whether faster drug diffusion into late adopting regions would be welfare enhancing. Are non-author regions slower to adopt new drugs because they are realizing more modest survival benefits?

Contrasting the survival benefits to new drug adoption in author and non-author regions is not statistically powered in our setting, given the modest number of author regions and the limited impact of most new drugs on survival. If we maintain the identification strategy specified in Equation (1) but replace the outcome variable with 1-year mortality, we estimate that 1-year mortality is 0.65 percentage points lower in first author regions, with the 95% confidence interval running from -2.6 to 1.3 percentage points. The implied Wald instrumental variable estimate for

the effect of a new drug is similarly imprecise, with the point estimate suggesting a statistically insignificant 15 percentage point reduction in 1-year mortality.

Due to the lack of statistical power for this analysis, we turn to a broader comparison of cancer mortality in fast and slow adopting regions to identify whether the survival gains from new cancer drugs are related to observed adoption patterns in our data. Details of the difference in differences estimation are described in Appendix Figures A1 and A2, as well as accompanying notes. The key source of identification comes from comparing survival rates across fast- and slow-diffusion regions, before and after the introduction of a new cancer drug. For this strategy, we measure a region's enthusiasm for adopting a new drug using a leave-out estimate of enthusiasm for other new drugs, excluding the region's realized utilization for that particular new drug.

We find that new drug use is 3.0 percentage points higher in fast-diffusing regions in the first four years following FDA approval ($P < 0.01$). Further, regions that are fast adopters experience a 1 percentage point greater reduction in 1-year mortality for the relevant cancer diagnosis than slow adopting regions ($P < 0.05$). The effect is even larger for 18-month mortality, with fast adopters experiencing a 1.5 percentage point greater reduction than slow adopters ($P < 0.01$). If we interpret the effect on drug utilization as the first stage of an instrumental variables regression to calculate the local average treatment effect of new drug use, the Wald instrumental variable estimate would suggest that new drug use is associated with a 51 percentage point reduction in 18-month mortality, among the marginal treated patients in high-diffusing regions, with a 95% confidence interval of $[-0.83, -0.19]$.

However, the implied instrumental variable estimate of the drug treatment effect stretches credibility, considering that the average 18-month mortality rate is 35.8% in our sample. It could be that the marginal patient who would be treated in a fast adopting region but not in a slow adopting region has much higher than average latent mortality. However, a more compelling explanation may be that fast-diffusion regions have higher returns to new drug use than slow-diffusion regions even for infra-marginal patients. Due to superior dosing, design of drug cocktails, management of side effects, or patient selection, these regions may achieve higher average returns to new drug use among all treated patients.[14]

Considered in the context of our earlier results, these findings suggest that policies designed

[14]Oncologists have documented dramatic variation in the return to drug use corresponding to variation in information regarding appropriate use of the drug. For example, when trastuzumab was first introduced, breast cancer patients taking the drug survived for a median of 25 months. Now, median survival has increased to 41 months for patients treated with the same drug, which expert oncologists have attributed to longer dosing schedules and better side effect management (Pollack, 2014).

to increase drug utilization without improving the surrounding information environment on appropriate application and management of the new drug may not achieve substantial survival gains. Information frictions may be a key limiting factor not only for drug adoption but also for the medical benefits of new drug use. On the other hand, interventions that improve local information about a new drug, such as through closer contact with experts, may result in higher use and improved patient welfare by facilitating the effective application and targeting of the new technology.

6 Conclusion

The results above suggest that information frictions significantly limit the adoption of new cancer drugs in the first few years after drug introduction. Prominent physicians who are well-informed about a new drug in the early stages of diffusion may play a key role in easing these frictions. Our results show that new drug utilization is 36% higher over the two years following FDA approval in the hospital market where the lead physician investigator on the drug's pivotal clinical trial practices. Despite the marked regional differences in early adoption of new cancer drugs, there is no evidence that early expertise with a drug drives higher rates of long-term utilization. Author 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 as utilization rates in first author and other HRRs converge within a four-year period.

While both first authors and other authors increase new drug utilization within their physician practice group, the influence of first authors is larger. Further, we show that relying on publication and citation history rather than authorship order to identify "superstar" study authors yields similar results, suggesting that investigators with greater professional status are more influential in increasing drug adoption in their region. We also find that investigator proximity has the largest impact on regions with the lowest levels of adoption of other new drugs, pointing to greater adoption frictions in less technology-intensive areas.

Regions that are quick to adopt new cancer drugs also see greater improvements in patient survival following new drug introduction. The large magnitude of these survival gains in fast-adopting regions suggest they may be driven not only by the greater fraction of patients receiving treatment with the new drug in these regions, but also by higher average returns to new drug use.

This study has important limitations that may point the way to informative avenues for future research. While the results suggest a complementarity between the opinion leader's professional

prominence and superior information about new drug, we cannot fully disentangle how information and prominence each contribute to local diffusion. Future work should assess the role of local opinion leaders outside of the context of research involvement to better understand the potential policy mechanisms that could improve local dissemination of information about medical innovations. Finally, a more detailed understanding the factors shaping the heterogeneous estimated returns to new technology adoption will be important to understanding the welfare consequences of variation in early diffusion of medical treatment.

The evidence in this paper suggests that the local information environment may be a key determinant of both the adoption of new cancer drugs and the realized survival benefits. Merely increasing new drug use in slower-adopting regions may not lead to substantial improvements in patient survival, but improving the local information environment to support new drug adoption could have significant welfare benefits both by increasing appropriate utilization and improving survival rates for treated patients. Contact with prominent physicians with expertise on the new drug may increase local drug use by reducing information frictions.

References

Adhvaryu, Achyuta. 2014. "Learning, Misallocation, and Technology Adoption: Evidence from New Malaria Therapy in Tanzania." *The Review of Economic Studies* 81 (4):1331–1365.

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.

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.

Banerjee, Abhijit, Arun G Chandrasekhar, Esther Duflo, and Matthew O Jackson. 2013. "The Diffusion of Microfinance." *Science* 341 (6144).

Bristow, Robert E, Jenny Chang, Argyrios Ziogas, Hoda Anton-Culver, and Veronica M Vieira. 2014. "Spatial analysis of adherence to treatment guidelines for advanced-stage ovarian cancer and the impact of race and socioeconomic status." *Gynecologic oncology* 134 (1):60–67.

Cameron, A Colin, Jonah B Gelbach, and Douglas L Miller. 2008. "Bootstrap-based improvements for inference with clustered errors." *The Review of Economics and Statistics* 90 (3):414–427.

Cameron, A Colin and Douglas L Miller. 2015. "A practitioner's guide to cluster robust inference." *Journal of Human Resources* 50 (2):317–372.

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.

Comin, Diego and Bart Hobijn. 2010. "An Exploration of Technology Diffusion." *American Economic Review* 100 (5):2031–2059.

Comin, Diego A, Mikhail Dmitriev, and Esteban Rossi-Hansberg. 2012. "The spatial diffusion of technology." Working Paper 18534, National Bureau of Economic Research.

Conley, Timothy G. and Christopher R. Udry. 2010. "Learning about a New Technology: Pineapple in Ghana." *American Economic Review* 100 (1):35–69.

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

DeAngelo, Daniel J, Daohai Yu, Jeffrey L Johnson, Steven E Coutre, Richard M Stone, Alison T Stopeck, Jon P Gockerman, Beverly S Mitchell, Frederick R Appelbaum, and Richard A Larson. 2007. "Nelarabine induces complete remissions in adults with relapsed or refractory T-lineage acute lymphoblastic leukemia or lymphoblastic lymphoma: Cancer and Leukemia Group B study 19801." *Blood* 109 (12):5136–5142.

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.

Fairfield, Kathleen M., F. Lee Lucas, Craig C. Earle, Laurie Small, Edward L. Trimble, and Joan L. Warren. 2010. "Regional variation in cancer-directed surgery and mortality among women with epithelial ovarian cancer in the medicare population." *Cancer* 116 (20):4840–4848.

Farrow, Diana C, William C Hunt, and Jonathan M Samet. 1992. "Geographic variation in the treatment of localized breast cancer." *New England Journal of Medicine* 326 (17):1097–1101.

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.

Griliches, Zvi. 1957. "Hybrid corn: An Exploration in the Economics of Technological Change." *Econometrica* 25:501–522.

Hancock, Jay. 2013. "'Narrow Networks' Trigger Push-Back from State Officials." Kaiser Health News. URL http://www.kaiserhealthnews.org/stories/2013/november/25/states-balk-at-narrow-networks.aspx.

Hudes, Gary, Michael Carducci, Piotr Tomczak, Janice Dutcher, Robert Figlin, Anil Kapoor, Elzbieta Staroslawska, Jeffrey Sosman, David McDermott, István Bodrogi et al. 2007. "Temsirolimus, interferon alfa, or both for advanced renal-cell carcinoma." *New England Journal of Medicine* 356 (22):2271–2281.

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.

MacKinnon, James G and Matthew D Webb. 2016. "Wild bootstrap inference for wildly different cluster sizes." *Journal of Applied Econometrics* .

Marshall, Alfred. 1890. *Principles of Economics*. MacMillan, London.

Nattinger, Ann Butler, Mark S Gottlieb, Judith Veum, David Yahnke, and James S Goodwin. 1992. "Geographic variation in the use of breast-conserving treatment for breast cancer." *New England Journal of Medicine* 326 (17):1102–1107.

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.

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

Pollack, Andrew. 2014. "Roche Breast Cancer Drug Perjeta Appears to Greatly Extend Patients' Lives." *New York Times* September 28.

Rogers, Everett M. 1962. *Diffusion of Innovations*. New York: Free Press.

Skinner, Jonathan and Douglas Staiger. 2005. "Technology Adoption From Hybrid Corn to Beta Blockers." Working Paper 11251, National Bureau of Economic Research.

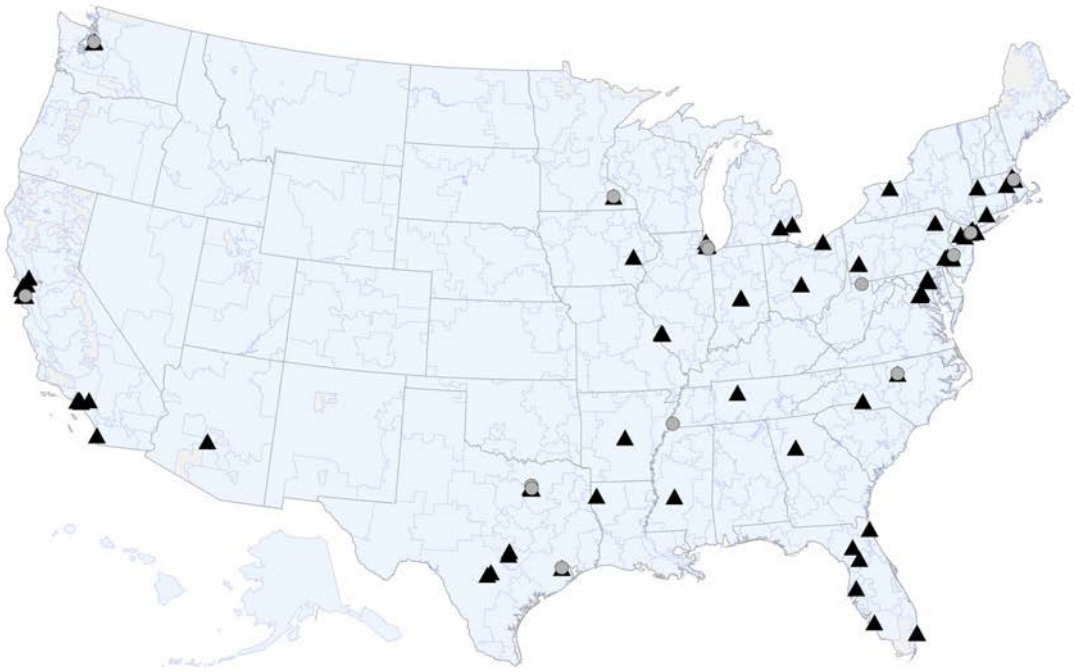
Smith, Sheila, Joseph P. Newhouse, and Mark S. Freeland. 2009. "Income, Insurance, and Technology: Why Does Health Spending Outpace Economic Growth?" *Health Affairs* 28 (5):1276–1284.

U. S. Government Accountability Office. 2012. "Medicare: High-Expenditure Part B Drugs."

U.S. Food and Drug Administration. 1998. "Guidance for Industry: E9 Statistical Principles for Clinical Trials."

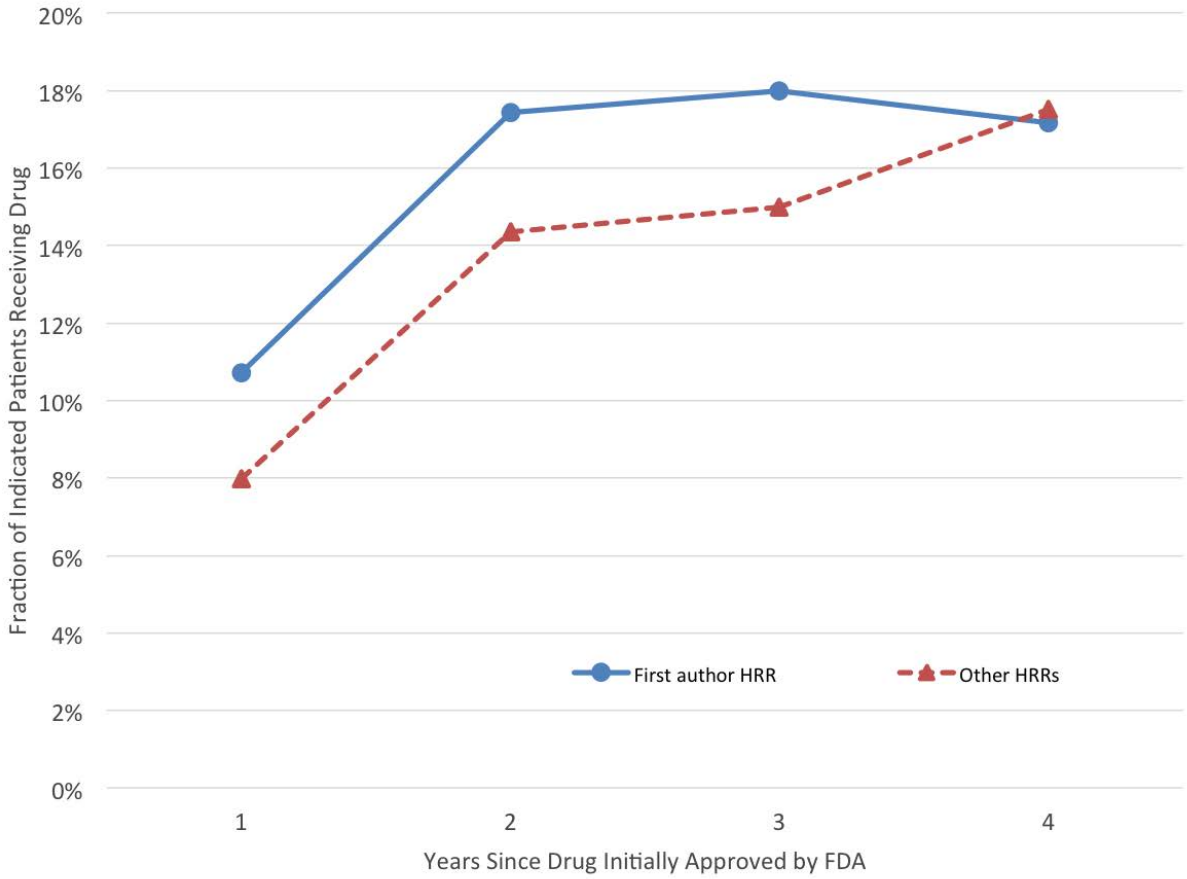
———. 2014. "Novel New Drugs 2013 Summary."

Figure 1: Locations of Drug Pivotal Study Authors



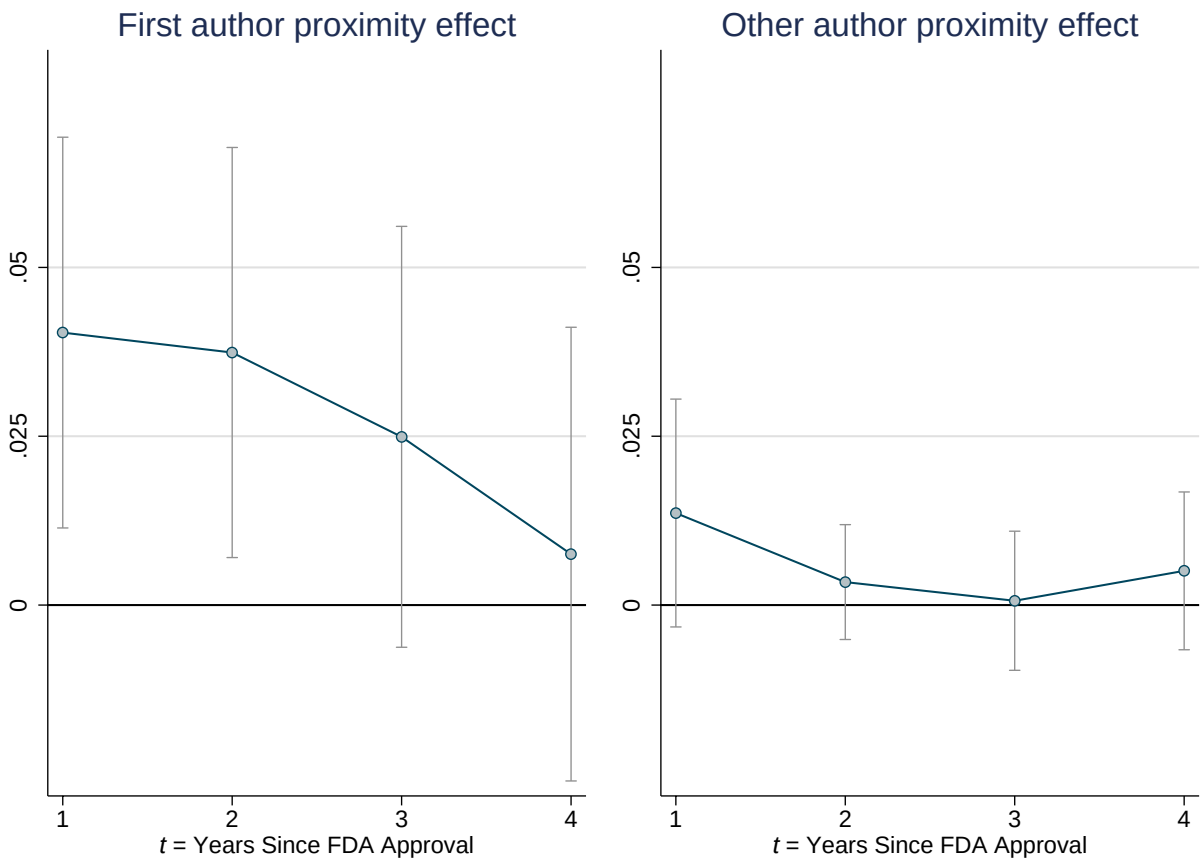
Notes: Circles mark the Hospital Referral Regions (HRR) that contain a first author for the pivotal clinical trial of a cancer drug in our sample. Some regions are the site of multiple first authors: Houston, Texas, (four first authors); Chicago, Illinois, (three first authors); Durham, North Carolina, (two first authors); and New York, New York (two first authors). Triangles mark HRRs that contain other authors.

Figure 2: New Cancer Drug Utilization Rates, by Years Since FDA Approval



Notes: The figure plots raw usage for new drugs in the first author's HRR versus all other HRRs. The "First author HRR" and "Other HRRs" rates are generated by averaging over the drug-specific first author HRR and non-first author HRR rates calculated separately for each drug in each year. To make rates over time comparable, we restrict to the 67% of drugs in our sample for which four years of usage rates are available.

Figure 3: 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 cancer drug became FDA approved. Bands indicate 95% confidence intervals constructed from standard errors clustered at the provider HRR-drug level.

Table 1: List of Studied Cancer Drugs

Generic name	Trade name	FDA approval date	Target disease	Pivotal clinical trial						Size of target population
				Trial centers	Patients enrolled	Journal of publication	Year of publication	No. of authors on pivotal trial	1st author city	
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
Capecitabine	Xeloda	4/30/1998	breast cancer	25	163	Journal of Clinical Oncology	1999	10	Dallas, TX	26,410
Trastuzumab	Herceptin	9/25/1998	breast cancer	54	222	Journal of Clinical Oncology	1999	11	Chicago, IL	26,410
Valrubicin	Valstar	9/25/1998	bladder cancer	41	90	Journal of Urology	2000	6	Chicago, IL	13,557
Denileukin diftitox	Ontak	2/5/1999	cutaneous T-cell lymphoma	25	71	Journal of Clinical Oncology	2001	26	Durham, NC	819
Temozolomide	Temodar	8/11/1999	brain cancer	21	225	British Journal of Cancer	2000	22	Houston, TX	1,797
Epirubicin hydrochloride	Ellence	9/15/1999	breast cancer	39	710	Journal of Clinical Oncology	1998	18	Canada	53,762
Gemtuzumab ozogamicin	Mylotarg	5/17/2000	acute myeloid leukemia	32	104	Journal of Clinical Oncology	2001	17	Seattle, WA	2,192
Arsenic trioxide	Trisenox	9/25/2000	acute myeloid leukemia	9	40	Journal of Clinical Oncology	2001	15	New York, NY	1,079
Alemtuzumab	Campath	5/7/2001	chronic lymphocytic leukemia	21	93	Blood	2002	11	Houston, TX	12,027
Zoledronic acid	Zometa	8/20/2001	hypercalcemia of malignancy	87	287	Journal of Clinical Oncology	2001	11	Canada	2,694
Ibritumomab tiuxetan[1]	Zevalin	2/19/2002	non-Hodgkin's lymphoma	45	200	Journal of Clinical Oncology	2002	13	Rochester, MN	51,042
Fulvestrant	Faslodex	4/25/2002	breast cancer	83	400	Journal of Clinical Oncology	2002	14	Houston, TX	64,045
Oxaliplatin	Eloxatin	8/9/2002	colon cancer	120	463	Journal of Clinical Oncology	2003	8	Nashville, TN	52,778
Bortezomib[2]	Velcade	5/13/2003	multiple myeloma	14	202	New England Journal of Medicine	2003	21	Boston, MA	23,819
Tositumomab-I 131	Bexxar	6/27/2003	non-Hodgkin's lymphoma	3	40	Journal of Clinical Oncology	2005	7	Stanford, CA	54,275
Pemetrexed	Alimta	2/4/2004	lung cancer	88	456	Journal of Clinical Oncology	2003	13	Chicago, IL	84,918
Cetuximab	Erbitux	2/12/2004	colon cancer	56	329	New England Journal of Medicine	2004	12	United Kingdom	55,528
Bevacizumab	Avastin	2/26/2004	colon cancer	164	923	New England Journal of Medicine	2004	15	Durham, NC	55,528
Decitabine	Dacogen	5/2/2006	myelodysplastic syndromes	23	170	Cancer	2006	16	Houston, TX	15,460
Panitumumab	Arranon	9/27/2006	colon cancer	81	463	Journal of Clinical Oncology	2007	12	Belgium	59,028
Temsirolimus	Torisel	5/30/2007	kidney cancer	148	626	New England Journal of Medicine	2007	19	Philadelphia, PA	3,794

[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 Statistics

Variables:	First author HRR (1)	Other author HRR (2)	Author HRR for different drug (3)	HRR with no authors (4)
Drug utilization rate	0.156	0.097	0.092	0.086
Fraction treated in author's group	0.534	0.357	0.000	0.000
Number of observations	6,985	29,322	250,330	372,831
Avg. no. of patients per HRR per drug	388	236	254	75
No. of HRR-drug pairs	18	124	986	4958
No. of unique HRRs	11	54	54	252

Notes: Regions are defined by the 306 Dartmouth Atlas Hospital Referral Regions (HRRs). For each drug in the sample, regions are partitioned into four groups based on geographic proximity to authors of the pivotal trial, corresponding to the four columns in the table. Statistics are then reported for each column by aggregating over the set of drugs in the sample. Reported statistics reflect drug utilization over the first two years following initital introduction. Data on drug utilization come from Medicare claims 1998-2008.

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

Proximity Measures	Panel A: All HRRs		Panel B: Author HRRs only		Panel C: New Cancer Patients	
	(1)	(2)	(3)	(4)	(5)	(6)
First author HRR	**0.0404*** **		**0.0383*** **		**0.0399*** **	
	(0.0131)		(0.0122)		(0.0154)	
Other author HRR	**0.0069**		**0.0068**		**0.0059**	
	(0.0048)		(0.0051)		(0.0053)	
First author HRR & in author group		**0.0421*** **		**0.0417** **		**0.0421*** **
		(0.0125)		(0.0124)		(0.0162)
First author HRR & non-author group		**0.0416** **		**0.0392** **		**0.0409** **
		(0.0211)		(0.0195)		(0.0203)
Other author HRR & in author group		**0.0276*** **		**0.0286*** **		**0.0271*** **
		(0.0074)		(0.0073)		(0.0082)
Other author HRR & non-author group		**-0.0031**		**-0.0033**		**-0.0036**
		(0.0054)		(0.0058)		(0.0057)
Number of observations	659,468	659,468	286,637	286,637	393,618	393,618

Notes: This table reports results from 6 separate regressions, where the dependent variable is an indicator that equals 1 if the patient is treated with the new drug over the observed episode of care. Each observation is a patient-episode, i.e. a 1 year episode of care during which the patient receives treatment for the indicated cancer type and so may be eligible for the new drug. "First author HRR" (hospital referral region) is an indicator variable that equals 1 if the patient is treated in the same HRR as the trial's first author. "Other author HRR" indicates treatment in the same HRR as another trial author (but not the first author). Author physician group indicators equal 1 if the patient was at any point during the year treated by a physician practicing in the same group as a trial author. Author HRR, non-author group indicators equal 1 if the patient is treated in the author's HRR but never by a doctor practicing in the same group as a trial author. All regressions include drug-year fixed effects; HRR-cancer type fixed effects defined using three categories of cancer drugs: urologic, hematologic, and other (including breast, colon, lung, and brain); and indicators for patient age, race, sex, and new cancer treatment episode. Standard errors clustered at the HRR-drug level shown in parentheses. *: p<0.10; **: p<0.05; ***: p<0.01.

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

(3-4) Sample limited to HRRs which ever contain any pivotal trial author

(5-6) All HRRs, but limited to new cancer patients, defined as patients with no cancer treatment in previous calendar year

Table 4: Superstar Author Proximity Effect on Drug Utilization

Dependent variable: (drug)_id in {0,1}, indicates receipt of new cancer drug d by patient i						
Independent variables:	(1)	(2)	(3)	(4)	(5)	(6)
First author HRR	**0.0300****			**0.0246****	**0.0242****	**0.0230****
	(0.0124)			(0.0119)	(0.0111)	(0.0110)
Top 50% cited author HRR		**0.0227****		**0.0154***		**0.0144**
		(0.0090)		(0.0081)		(0.0090)
Top 10% cited author HRR			**0.0232****		**0.0100**	**0.0034**
			(0.0122)		(0.0108)	(0.0117)
Number of observations	659,468	659,468	659,468	659,468	659,468	659,468

Notes: These regressions test whether "superstar" authors are more influential than other study authors for the same drug; to that end, the baseline regression specification is augmented to include a vector of (drug)*(any author HRR) fixed effects. Reported coefficients describe whether regions with authors of the noted type have higher new drug use compared to the rest of the author regions for the same drug. Top 50% and top 10% authors are defined as the most prominent academic authors for each drug, as measured by citation counts accruing to publications produced over the 10 years leading up to FDA drug approval in the relevant field. All regressions include drug-year fixed effects; HRR-cancer type fixed effects defined using three categories of cancer drugs: urologic, hematologic, and other (including breast, colon, lung, and brain); and indicators for patient age, race, sex, and new cancer treatment episode. See notes to Table 3 for further details. *: p<0.10; **: p<0.05; ***: p<0.01.

Table 5: Scope of Author Influence

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

	Regional technology intensity		Neighbor region drug adoption		Off-label drug use	
Independent variables:	(1)	(2)	(3)	(4)	(5)	(6)
First author HRR	**0.0409*** **	**0.0385*** **	**0.0402*** **	**0.0402*** **	**0.0006**	**0.0007**
	(0.0112)	(0.0105)	(0.0131)	(0.0130)	(0.0005)	(0.0005)
First author HRR * Fast adoption index	**-0.0215****	**-0.0197*** **				
	(0.0084)	(0.0077)				
Neighbor of first author HRR			**0.0017**	**0.0014**		
			(0.0066)	(0.0066)		
Other author HRR	**0.0064**	**0.0063**	**0.0066**	**0.0067**	**0.0002**	**0.0003**
	(0.0047)	(0.0050)	(0.0048)	(0.0049)	(0.0004)	(0.0004)
Other author HRR * Fast adoption index	**-0.0042**	**-0.0035**				
	(0.0052)	(0.0053)				
Neighbor of other author HRR			**-0.0031**	**-0.0030**		
			(0.0030)	(0.0031)		
Sample						
Restricted sample?	No	Yes	No	Yes	No	Yes
Number of observations	659,468	286,637	659,468	547,256	7,712,248	3,063,237

Notes: This table reports coefficients and standard errors from 6 separate regressions. See notes to Table 3 for further details. Columns 2 and 6 restrict the sample to only include regions that contain at least one study author. Column 4 restricts the sample to include only regions that contain a study author or border an author region. p<0.10; **: p<0.05; ***: p<0.01.

(1-2) Fast adoption index summarizes regional utilization rates of new drugs when there is no author present in the region; the variable is normed to be mean zero and have a standard deviation of 1. Higher values correspond to regions that are quicker to adopt new cancer drugs.

(3-4) Neighbor regions are defined as those that share a border with author regions.

(5-6) Sample is changed to include all cancer care episodes treating patients who do not have the indicated disease type for the observed drug (e.g. testing use of a colon cancer drug in patients with other cancer types, such as breast cancer or lung cancer).

Table 6: Patient Travel and Proximity Effects

	Dependent variables:			
	Travel		New drug use	
Independent variables:	(1)	(2)	(3)	(4)
First author HRR	**0.0329***	**0.0309***	**0.0327*****	**0.0311*****
	(0.0192)	(0.0197)	(0.0124)	(0.0116)
Traveler to first author HRR			**0.0224**	**0.0226**
			(0.0140)	(0.0138)
Other author HRR	**0.0295*****	**0.0285*****	**0.0066**	**0.0064**
	(0.0094)	(0.0089)	(0.0052)	(0.0056)
Traveler to other author HRR			**0.0012**	**0.0011**
			(0.0061)	(0.0064)
Sample				
Author HRRs only?	No	Yes	No	Yes
Number of observations	659,468	286,637	659,468	286,637

Notes: Columns 1 and 2 report results from regressions where the dependent variable indicates whether patient received care outside the patient's HRR of residence. Columns 3 and 4 report results from regressions where the dependent variable indicates whether the patient received treatment with the new drug. Columns 2 and 4 restrict the sample to regions that contain a study author for at least one drug in our sample. See notes to Table 3. *: p<0.10; **: p<0.05; ***: p<0.01.

Table 7: IV Estimates of Proximity Effect on Drug Utilization

	Outcome: new drug use		
	(1)	(2)	(3)
A. Reduced form: drug receipt effect			
Residence in first author HRR	**0.0228****	**0.0217****	**0.0228****
	(0.0099)	(0.0099)	(0.0099)
Residence in other author HRR	**0.0038**	**0.0044**	**0.0046**
	(0.0044)	(0.0050)	(0.0061)
Residence in first author's neighbor HRR			**0.0035**
			(0.0045)
Residence in other author's neighbor HRR			**-0.0028**
			(0.0028)
B. Two Stage Least Squares			
Provider in first author HRR	**0.0293****	**0.0259****	**0.0263****
	(0.0127)	(0.0114)	(0.0115)
Provider in other author HRR	**0.0056**	**0.0060**	**0.0065**
	(0.0058)	(0.0061)	(0.0061)
Sample			
Author HRR only?	No	Yes	No
Number of observations	659,468	286,637	659,468

Notes: Reduced form results report coefficients from 3 regressions where the outcome variable is new drug use and the key explanatory variables are indicators for whether a given patient resides in the same region as the study author (or in column 3, in a neighboring region). In Panel B, the Two Stage Least Squares results use patient residence variables as instrumental variables for whether the patient is treated in author region. All regressions include drug-year fixed effects, and HRR-cancer type 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: P-values from Alternate Approaches to Inference

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

Proximity Measures	Panel A: Standard cluster at HRR-drug		Panel B: Standard cluster at HRR		Panel C: Wild cluster bootstrap at HRR-drug		Panel D: Wild cluster bootstrap at HRR	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
First author HRR	0.0404***		0.0404***		0.0404***		0.0404**	
	(0.002)		(0.010)		(0.010)		(0.012)	
Other author HRR	0.0069		0.0069		0.0069		0.0069	
	(0.149)		(0.155)		(0.212)		(0.180)	
First author HRR & in author group		0.0421***		0.0421***		0.0421***		0.0421***
		(0.001)		(0.000)		(0.002)		(0.002)
First author HRR & non-author group		0.0416**		0.0416*		0.0416		0.0416
		(0.048)		(0.076)		(0.168)		(0.278)
Other author HRR & in author group		0.0276***		0.0276***		0.0276***		0.0276***
		(0.000)		(0.002)		(0.002)		(0.002)
Other author HRR & non-author group		-0.0031		-0.0031		-0.0031		-0.0031
		(0.564)		(0.549)		(0.652)		(0.626)
Number of observations	659,468	659,468	659,468	659,468	659,468	659,468	659,468	659,468

Notes: This table explores robustness of our main findings to alternative methods of statistical inference to account for clustering. In columns 1 and 2, we replicate the analysis from Table 3 Panel A, but we now report p-values associated with each coefficient in parentheses. Each subsequent panel reports the same regression results, but applies a different methodology for inference. Panel A uses the usual cluster robust standard error with clusters defined at the hospital referral region (HRR) by drug level. Panel B reports p-values using the usual cluster robust standard error with clusters defined at the HRR. Panel C reports p-values from a wild cluster bootstrap, accounting for clusters at the HRR by drug level. Panel D reports p-values from a wild cluster bootstrap, accounting for clusters at the HRR level.

Table A2: Author Proximity Effect on Drug Utilization: Finer divisions of author role

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: New Cancer Patients	
	(1)	(2)	(3)	(4)	(5)	(6)
First author HRR	**0.0406*** **		**0.0385*** **		**0.0403*** **	
	(0.0131)		(0.0122)		(0.0154)	
Middle author HRR	**0.0077**		**0.0074**		**0.0076**	
	(0.0050)		(0.0054)		(0.0055)	
Last author HRR	**0.0006**		**0.0028**		**-0.0064**	
	(0.0140)		(0.0127)		(0.0175)	
First author physician group		**0.0422*** **		**0.0407*** **		**0.0422*** **
		(0.0125)		(0.0126)		(0.0162)
First author HRR, non-author group		**0.0417** **		**0.0395** **		**0.0411** **
		(0.0211)		(0.0196)		(0.0203)
Middle author physician group		**0.0273*** **		**0.0273*** **		**0.0259** **
		(0.0074)		(0.0074)		(0.0082)
Middle author HRR, non-author group		**-0.0018**		**-0.0023**		**-0.0007**
		(0.0059)		(0.0064)		(0.0060)
Last author physician group		**0.0324**		**0.0314**		**0.0438**
		(0.0324)		(0.0319)		(0.0329)
Last author HRR, non-author group		**-0.0137**		**-0.0101**		**-0.0275*** **
		(0.0016)		(0.0097)		(0.0087)
Number of observations	659,468	659,468	281,253	281,253	393,618	393,618

Notes: Each observation is a patient-episode, i.e. a 1 year episode of care during which the patient has the indicated cancer type and so may be eligible for treatment with the new drug. First author HRR (hospital referral region) is an indicator variable that equals 1 if the patient is treated in the same region as the trial's first author. Other author HRR indicates treatment in the same region as another trial author (besides the first author). Author physician group indicators equal 1 if the patient was at any point treated by a physician practicing in the same group as the trial author. HRR, non-author group indicators equal 1 if the patient is treated in the author's HRR but never by a doctor practicing in the same group as a trial author. All regressions include drug-year fixed effects, and 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.

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

(3-4) Sample limited to HRRs which ever contain any pivotal trial author

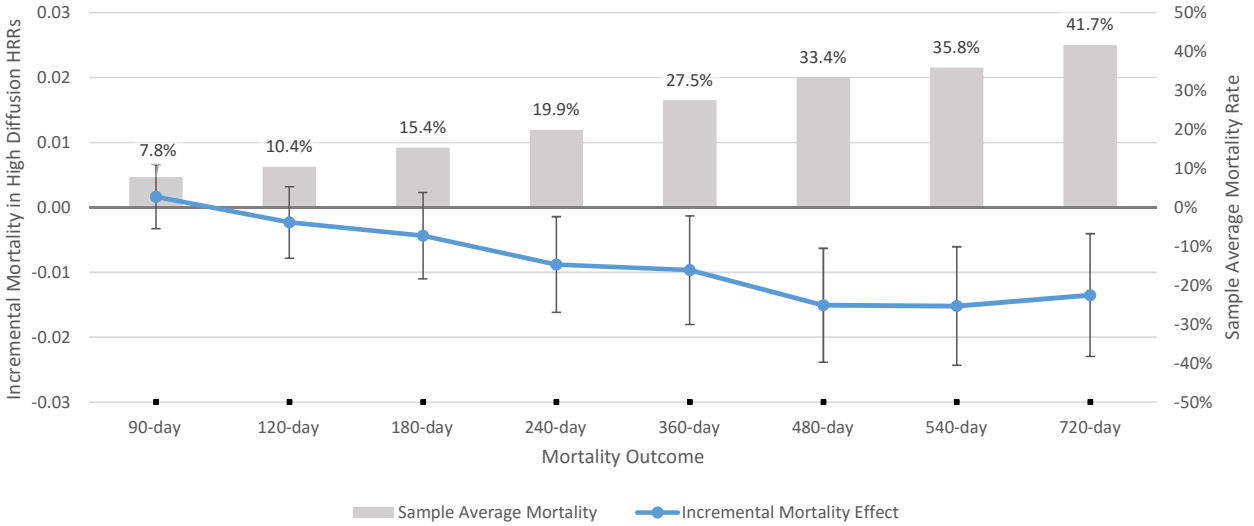
(5-6) All HRRs, but limited to new cancer patients, defined as patients with no cancer treatment in previous calendar year

Notes on Appendix Table 2

This table reports results from 6 separate regressions that mirror our baseline specification (1), but further subdivide other authors into middle authors and last authors on the basis of authorship order on the academic publication. For new drug clinical trials, the first author is typically the principal investigator and the last author is often a scientist employed by the sponsoring drug company. In fact, only seven drugs in our sample have a last author who is a practicing clinician, whereas all 18 drugs with US-based trials have a practicing clinician as the first author. Note that we only investigate the role of practicing physicians; for the many drugs with a non-clinical final author, there is no last author region coded.

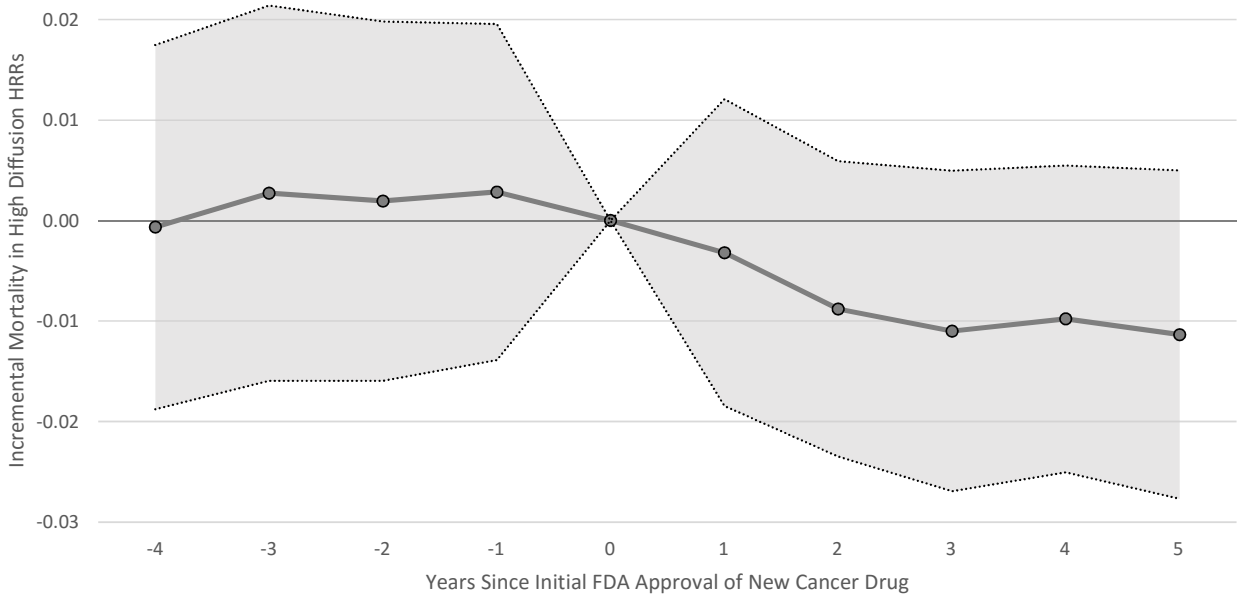
Results reported here find that patients treated in last author regions are not significantly more likely to receive the new drugs compared to other regions; however, the estimates are relatively imprecise due to the small number of drugs with clinicians in the last author position. In particular, the 95% confidence interval includes an up to 2.8 percentage point higher utilization in last author regions, reported from the column (1) baseline specification. Point estimates are suggestive of 3.2 percentage point higher use within the last author's physician group, as reported in column (2), but the findings are not statistically significant.

Figure A1: Mortality Effect of Regional Adoption Speed



Notes: On the primary axis, this solid blue line plots difference-in-differences estimates from eight separate regressions (described in Equation (4)) where the mortality outcome varies from 90-day to 720-day mortality. Each mortality outcome is an indicator for whether a patient died within the specified period from the initial date of the cancer treatment episode. The regression is estimated over new patient-drug episodes that fall within four years before and after FDA approval of the drug. Bands indicate 95% confidence intervals constructed from standard errors clustered at the HRR-drug level. Each difference-in-differences coefficient describes how the mortality rate changes following new cancer drug introductions for indicated cancer patients residing in fast-diffusing regions, compared to mortality rate changes in slow diffusing regions. Regional drug diffusion speed is measured by a leave-out index which is defined as the fraction of cases among all other drugs in our sample for which new drug utilization was higher than the national average over the first four years following FDA approval. This index is a continuous measure ranging from 0 to 1. Thus the plotted coefficients describe the mortality impact of a new drug introduction in places which are above-average early adopters of all other new cancer drugs in our sample (diffusion index = 1) relative to a region which has below-average early utilization of all other sample drugs (index = 0). The secondary axis shows average mortality rates among the regression sample.

Figure A2: Event Study Mortality Effect of Regional Adoption Speed



Notes: This graph plots the coefficients from the difference-in-differences 360-day mortality specification (see Equation (4) and notes to Figure A1), but with the binary "post" indicator replaced by a series of indicators for the number of years relative to the drug's initial FDA approval. The omitted event year is 0. Bands indicate 95% confidence intervals constructed from standard errors clustered at the HRR-drug level. Each coefficient describes how changes in mortality relative to year zero differs across regions with a high versus low propensity to adopt new drugs, where this propensity is measured by a leave-out diffusion index defined as the fraction of cases among all other drugs in our sample for which new drug utilization was higher than the national average over the first four years following FDA approval. Thus estimates of zero in the negative event years prior to drug FDA approval are consistent with the parallel trends assumption that mortality between high- and low-adoption regions would not have changed absent the new drug. Negative mortality estimates in event years following FDA approval suggest larger mortality reductions in fast-diffusion regions compared to slow-diffusion regions.

Notes on Appendix Figures A1 and A2

In this section, we discuss the mortality findings summarized in Section 5 at greater length. First, it is useful to note a few key features of the identification problem for studying mortality effects. First, our analysis will rely on comparing survival patterns across fast- and slow-diffusion regions. Differences in survival across these regions could be driven either by differences in take-up of the new drug or by differing returns to drug use among infra-marginal treated patients.

Second, to understand the marginal returns to the additional drug use driven by investigator influence, we would ideally compare changes in survival rates across author and non-author regions, echoing our identification strategy from the earlier analysis. The problem is that such comparisons are underpowered for estimating mortality effects: we are relying on moderate variation in drug use from a handful of regions to measure changes in a noisy outcome variable (i.e. mortality rates). Any effect on mortality would be smaller than the total change in drug utilization, and thus the regression would likely have less statistical power than the main specifications.

Studying mortality rates rather than drug take-up confers one crucial benefit which contributes to identification: we can observe mortality rates for each cancer diagnosis in each region both before and after drug approval in order to construct difference-in-differences estimates of survival effects. Controlling for pre-period differences in mortality allows us to remove regional variation in survival that may be driven by patient health status and isolate survival gains that may be plausibly related to drug introduction. By contrast, our earlier analysis could not exploit such an approach since a new drug's measurable use is zero for all regions in the years before FDA approval and is therefore uninformative about the region's enthusiasm for new technology. For our analysis of drug take-up, we instead relied on regional adoption rates of other new drugs to construct counterfactual adoption behavior.

In light of these considerations, we proceed with a difference-in-differences estimation strategy that compares mortality rates before and after new drug introduction, and across regions with high and low propensities to adopt new cancer drugs. This approach exploits wider variation in regional drug use, while still controlling for any baseline differences in survival by region and cancer diagnosis. Our primary regression specification takes the following form:

$$\begin{aligned}\mathbf{1}(\textit{mortality})_{ijtd} = \beta(\textit{HRR diffusion index})_{jd} \times \mathbf{1}(\textit{post FDA approval})_{jtd} \\ + \{\textit{HRR} \times \textit{drug FEs}\}_{jd} + \{\textit{drug} \times \textit{year FEs}\}_{dt} + \epsilon_{ijtd}. \qquad (4)\end{aligned}$$

The mortality outcome $\mathbf{1}(\textit{mortality})$ is an indicator for whether a patient died within a specified period (e.g. 1-year) of the initial date of the cancer treatment episode. To avoid estimating this relationship using multiple observations from surviving patients, we include only new cancer treatment episodes (i.e. patients with no cancer claims in the previous calendar year). The regression is estimated over patient-drug episodes that fall within four years before and after FDA approval of the drug.

The key coefficient of interest is β which describes how the mortality rate changes after drug introduction in regions that are likely to be fast adopters of the new drug, among patients treated for the relevant cancer type. The variable (*HRR diffusion index*) codes a leave-out estimate of the region's enthusiasm for new cancer drugs, excluding the region's utilization for the particular drug under analysis. To remove potential bias from changes to patient sorting after drug introduction, we match patients to regions based on their HRR of residence. The indicator variable $\mathbf{1}(\textit{post FDA approval})$ equals 1 in the years following initial FDA approval for drug j and 0

otherwise.

More specifically, $(\textit{HRR diffusion index})_{jd}$ codes for region j the fraction of cases among all drugs in our sample excluding drug d for which new drug utilization in region j was higher than the national average over the first four years following FDA approval. A value of 0 would indicate the region is never an above-average adopter of any other new cancer drug in our sample over the first four years; a value of 1 would indicate the region is an above-average adopter of all other new cancer drugs in our sample. Because we use a leave-out estimate of diffusion speed, this measure should be uncorrelated with the idiosyncratic latent demand for this particular cancer drug in each region, which in turn could be related to differences in health status.

The regression also controls for region by drug (i.e. drug-specific diagnoses) fixed effects. These fixed effects provide a fine level of control for baseline disease-specific regional mortality rates; they are identified in this regression by the inclusion of pre-period years prior to drug introduction. Finally, we control for drug by year fixed effects, which capture national trends in survival for each drug-specific set of cancer diagnoses in our sample.

The identifying assumption for interpreting β as the causal impact of differences in drug utilization on patient survival depends on the usual parallel trends assumption. In this case, the assumption is that fast and slow diffusion HRRs would have experienced parallel changes in mortality rates if no new drug had been introduced. To explore the validity of this assumption, we show the event study plot reporting regression results that replace the binary $\mathbf{1}(\textit{post FDA approval})$ variable with a series of dummy variables for each year relative to FDA approval. An absence of pre-trends in this plot would support the validity of the difference-in-differences results.

Our baseline results from estimating equation (4) are reported in Figure A1. Each point in the figure represents the coefficient β estimated from a separate regression, where the outcome variable is mortality over the indicated time window, ranging from 90 days to 720 days. Beginning with the outcome of 240 day mortality rates, all regressions show statistically significant mortality reductions in fast-adopting regions relative to slow-adopting regions, significant at the 5% level. For example, the point estimate implies a 1 percentage point greater reduction in 1-year mortality for the relevant cancer diagnosis for a region that is an above-average adopter of all other in-sample drugs compared to a region that is a below-average adopter of all other drugs, from a base mortality rate of 27.5%.

To investigate the validity of the parallel trends assumption, we create an event study plot showing the year-by-year differences in 1-year mortality rates across fast- and slow-adopting regions. As Figure A2 shows, the differences in cancer mortality were stable across regions in the pre-period years; after a new drug is introduced, the fast-adopting regions experience steep declines in patient mortality, consistent with the basic difference-in-differences results. These findings corroborate the assumption that fast- and slow-adopting regions were on similar trends in cancer mortality rates before new drug introduction.

To further interpret the magnitude of these results, we examine how our leave-out estimate of regional diffusion speed relates to average new drug utilization rates. To proceed, we use an analogous specification as in (4) above, with pre-period utilization of a new drug before set to zero prior to FDA approval. We find that new drug use is 3.0 percentage points (standard error of 0.36) higher in fast-diffusing regions in the first four years following FDA approval. This change corresponds to a 37% increase in drug utilization over the average rate of 8.1% in the new cancer patient sample. If we interpret the effect on drug utilization as the first stage of an instrumental variables regression to calculate the local average treatment effect of new drug use, the Wald instrumental variable estimate would suggest that new drug use is associated with a 33 percentage point reduction in 1-year mortality, among the marginal treated patients in high-diffusing regions.

Results of this analysis are interpreted in section 5.