The value of randomized trials for efficacy in oncology: Lessons from off-label prescribing

Rebecca McKibbin*

20 November 2017

Job Market Paper

Click here for the latest version

Abstract

This paper investigates the role of efficacy evidence produced in randomized controlled trials (RCTs) on demand for cancer drugs. I use the unique setting of off-label prescribing, where it is possible to observe demand for a drug for a particular cancer both before and after a trial result is released. Using a new data set that combines information from scientific articles, FDA data, and Medicare claims, I estimate the effect of the release of results from an RCT on demand for off-label uses of a drug. By exploiting variation in the timing of the release of results across off-label uses of drugs, I find that off-label prescriptions of drugs increases on average by 150% if the RCT finds a statistically significant increase in overall survival. In contrast, there is no change in the absolute level of demand if the trial result is inconclusive. I estimate the effect of market and drug characteristics on the firm’s decision to invest in an RCT for a potential off-label use. The results suggest that trials are not selected based on efficacy. A back of the envelope welfare calculation suggests that firms have an incentive to invest in trials that are not socially efficient on the margin. However, on average the RCTs conducted for off-label uses of drugs are worth more to consumers than they cost the firm. A key lesson for drug regulation is that efficacy data produced in RCTs plays an important role in the adoption of new drugs into medical practice.

*I thank my advisors, Fiona Scott Morton, Jason Abaluck, and Amanda Kowalski for their guidance and support. I also thank Joseph Altonji, Steven Berry, Philip Haile, Costas Meghir, Joseph Shapiro, Noriko Amano, Giovanni Compiani, Benjamin Friedrich, Sharat Ganapati, Hugo Gerard, Jennifer Gippel, Matthew Grant, Matthew Hom, Corina Mommaerts, Ljubica Ristovska, and seminar participants at Yale for helpful comments and discussions. I am also grateful to Jean Roth and Mohan Ramanujan for assistance obtaining and managing the data. Funding from the National Institute on Aging, through Grant Number R24AG058049 to the National Bureau of Economic Research, and by the Economic and Social Research Council, through Grant Number ES/M008673/1 to the Institute for Fiscal Studies, is gratefully acknowledged. The content is solely the responsibility of the author and does not necessarily represent the official views of the NIH, NBER or IFS.
1 Introduction

Should firms be required to produce randomized controlled trials demonstrating the efficacy of their cancer drugs? In the US, drug companies are required to demonstrate that new cancer drugs are effective for one indication through a randomized controlled trial (RCT). However, once this requirement has been met, the drugs can be prescribed for other untested indications. As a result, many cancer drugs are currently prescribed for purposes where there is limited or no evidence about the efficacy from clinical trials. Whether or not the regulation is too strict for the first uses of drugs, or too lax for additional uses of drugs is greatly debated (see Cavalla, 2015). When regulation was first introduced in 1962, the evidence-based medicine paradigm was in its infancy. Although the importance of efficacy evidence in medicine was understood in academic and government circles, it had not yet made its way into mainstream medical practice (Bothwell and Podolsky, 2016). At that time, firms generally had not been providing any evidence from clinical trials that demonstrated the effectiveness of their drugs (Davies and Kerimani, 2008). This may have been due to market failures or a lack of demand for evidence. Today, doctors have greater awareness of the importance of efficacy evidence, as well as greater access to information on the available evidence for a drug. Moreover, fifty years of regulation has resulted in a pharmaceutical market with numerous drugs with an existing evidence base. This suggests a need to re-evaluate the evidentiary requirement in the context of the market and state of medicine today. Research into the benefits of setting the evidentiary standard as an RCT has been limited by lack of variation in policy. In this paper I revisit the question of regulatory requirements for new cancer drugs by using a new data set to examine the importance of efficacy evidence produced in RCTs, relative to only providing evidence from a non-randomized method.

Drug prescriptions are considered to be off-label when they are prescribed for a purpose that has not been approved by the US Food and Drug Administration (FDA). This is common practice in oncology and it has been suggested that as much as 75% of cancer therapy usage in the US might be off-label (Pfister, 2012). The focus of this paper is uses of drugs that are off-label because they have been approved for one cancer, based on organ site of origin, and are being prescribed for treating a cancer originating from another organ site. For example, a lung cancer drug being used to treat breast cancer. The unique data set compiled for this study shows that approximately 23% of prescriptions for cancer drugs included in the sample are for an off-label use of this type. Off-label prescribing presents a unique opportunity to study the trade-offs resulting from requiring RCTs. This is because we can observe drugs being prescribed before and after randomized evidence becomes available. There is no requirement that off-label uses of drugs be tested in any kind of clinical trial. However, to be covered by insurance, and therefore accessible to the majority of patients, there is typically evidence on the efficacy of the drugs for the off-label uses from a non-
randomized trial. Firms then have the option to provide additional evidence on efficacy from an RCT, which they an incentive to do because of differences in advertising rules.

The paper proceeds in three parts. First, I develop a model of drug demand and estimate the direct effect of the release of positive trial results from an RCT and compare this to non-positive results. This is important because it is not known to what extent RCTs directly affect demand. Data on off-label prescribing is not readily available in existing data sets, and first uses of drugs already have evidence from an RCT when prescribing begins. A new data set is compiled for this paper, which allows off-label uses of drugs to be identified. The data set is compiled from scientific articles and FDA data, which were merged with Medicare claims data. I estimate the effect of the release of the results of an RCT for off-label uses of drugs that firms choose to test during the sample period (1999-2013). This gives an estimate of the effect of the treatment on the treated. The identification strategy is to exploit variation within the drug (before and after the trial) and across drugs (in the timing of the trial for the disease relative to when the drug was first released). The latter controls for take-up of the drug that may have occurred regardless of the trial. The results show that on average an RCT with a statistically significant increase in overall survival increases demand for an off-label use of a drug by an additional 140 patients per quarter (150%) taking up the drug. On average there is no change in the absolute level of demand if the trial result is inconclusive (statistically insignificant increase in overall survival).

The second part of the paper examines the firm’s decision to conduct an RCT for a potential off-label use of a drug. This is important because the previous estimate is a treatment on the treated effect. If greater or fewer trials were run due to a policy change, the effect on the margin might be different from this estimate. The main finding is that firms are more likely to choose to invest in RCTs for off-label uses of drugs in cancer markets where there are more approved drugs and less prescribing of off-label drugs. Market size has no effect; however, this may be a result of measurement error. To determine whether or not the results of the trial are predictable based on observable characteristics of the drugs and market, I estimate a Heckman selection model. None of the observable characteristics are statistically significant predictors of the result of the trial. The coefficient on the selection correction is also statistically insignificant. However, the magnitude is reasonably large. I further examine the selection issue by plotting the trial outcome against the predicted probability of the trial and the number of drugs in the market. Again, it does not appear that there is selection.

Finally, I examine the welfare implications of the estimates and interpret the findings in terms of broader lessons for drug regulation policy in cancer treatment. Requiring RCTs to demonstrate the efficacy of a drug has both efficiency and distributional welfare implications. From an efficiency perspective, the question is whether or not the benefit of reducing uncertainty around efficacy is outweighed by the cost of the trials because different levels of evidence quality have different costs.
It is unknown to what extent, and in what direction, private firms will provide an inefficient number of trials. Trials are a fixed cost and firms may over or under-invest in them relative to the social optimum. It depends on whether or not the trials create new valuable information or simply draw market share from competitors. In addition, prices are distorted by insurance and the need for firms to set a single price for a drug, rather than different prices based on the efficacy of the drug for each type of cancer. To examine this I outline a conceptual framework of the decision of a firm and social planner to invest in a trial. I then provide empirical evidence on their relative incentives to conduct trials. I also compute the social value of the trials, under some assumptions, and compare this with an estimate of the cost of conducting an RCT, which I draw from the literature. I find that while on average the trials that firms choose to conduct for off-label uses of drugs are worth more to society than what they likely cost, on the margin firms have an incentive to invest in trials that are not socially efficient because drug reimbursement levels are so high.

From a distributional perspective there are winners and losers. When more evidence is required, fewer drugs will be available but there will more information on those drugs. Greater information should lead to improved treatment choices for patients. However, since there is heterogeneity in how people respond to drugs some people will lose from the reduction in variety. Moreover, RCTs take more time to complete and during the study period it is necessary to place some restrictions on access to the drug in order to provide an incentive to participate in research. This means that people who are very sick now and may potentially benefit from the drug are not able to access it. While quantifying the distributional effects is outside the scope of this paper, I show that the demand for untested uses of drugs is low. This means that few people benefit from having access to untested drugs, but prescription of untested drugs also contribute little to health care costs.

This paper provides a first step in evaluating the policy of requiring RCTs demonstrating efficacy for cancer drugs. In favor of reduced entry restrictions on new drugs, this study shows that evidence has an important role in demand for drugs. The majority of prescribing is for uses of drugs that have been tested. Untested new drugs are only a small proportion of the market. This suggests that new untested drugs would not be widely adopted if they were made available under less strict regulations, at least in the short-run. Although the results are based on keeping current prices fixed, the anticipated effect on prices from reducing the evidentiary requirement, if there were any change, would be downwards. Lower prices are likely (if anything) to lead to less prescribing due to physician agency because patients are typically fully insured but physicians are paid in proportion to the price of the drug. I also find that firms have an incentive to run trials voluntarily, possibly exceeding the social optimum, because demand increases in response to RCTs that show an improvement in overall survival and because prices are so high. Against reduced entry restriction, or in favor of greater regulation around off-label prescribing, this study finds that demand for drugs is persistent in the face of evidence that could be considered negative.
Doctors do not respond to a failure to find an increase in overall survival as if it were a negative finding and demand persists among drugs with a trial even if the trial result was not favorable. The policy implication of this finding is unclear. On the one hand, spending occurs on drugs that would not have been available if an entry rule were in place because the firm has failed to demonstrate efficacy. On the other hand, allowing the drug to be prescribed provides an increased incentive for firms to run the trial in the first place.

This paper is one of very few that has examined the effect of requiring RCTs because it is difficult to find sources of variation. Peltzman (1973) studies the effect on consumer surplus of the introduction of the 1962 amendments to drug regulation laws. He finds that the benefits generated by reducing the cost of understanding the effectiveness of drugs, by making firms produce trial information, is far outweighed by the loss of new drugs. Peltzman is able to compute a measure of the consumer surplus because consumers face prices – during this period insurance coverage for prescription drugs was low (around 30% of the population (Reed and Carr, 1969)). More recently Grennan and Town (2015) examined the regulatory trade-off for medical devices. They used variation in regulatory rules between the EU and the US to estimate the effect of trials on demand for stents (a medical device). They find that testing is essential to the adoption of the medical device. The difference in these results might be explained by the increased emphasis on evidence in medicine in the decades since the 1962 regulations were introduced. This paper extends this literature by examining cancer drug markets in the evidence-based medicine era using a new source of variation. My results are consistent with the finding of Grennan and Town (2015) i.e information is essential to the adoption of a new medical product.

This paper also contributes to the literature on the role of information in drug prescribing, including new drug adoption. Although no papers have causally estimated the effect of the first RCT on prescribing, which is the most similar to the regulatory policy margin, previous work has examined how other types of information about drugs affects demand and the take-up of new drugs. This includes information from post-market clinical trials, FDA updates, advertising and learning from experience and from other doctors. For example, Arrow et al. (2017) shows that physicians with access to a pharmaceutical reference database begin prescribing newly approved drugs sooner than other doctors. Agha and Molitor (2015) and Coleman, Katz and Menzel (1966) look at the role of medical opinion leaders in the diffusion of new drugs. Chintagunta, Jiang and Jin (2009) examine the relative importance of learning from patients, news articles, academic articles and FDA updates on the sales of Cox-2 inhibitors. There is a large literature on the role of advertising on demand for drugs (for example, Chintagunta, Goettler and Kim, 2012; Ching and Ishihara, 2010; Ching et al., 2015), including the relative importance of scientific information from clinical trials versus persuasive advertising (Azoulay, 2002; Sood, Kappe and Stremersch, 2014).

This paper also complements the literature on the incentives of firms to invest in drug research
and development (R&D). Many papers have examined the relationship between potential market size and investment in pharmaceutical R&D for first uses of drugs or generic entry (Acemoglu and Linn, 2004; Lichtenberg and Waldfogel, 2003; Finkelstein, 2004; Berndt and Trusheim, 2012; Dubois et al., 2015; Blume-Kohout and Sood, 2013). Using cancer as an application, Budish, Roin and Williams (2015) show that firms underinvest in long term research, because fixed patent terms make short term research more profitable. My paper examines the firms incentives to invest in research into secondary uses of drugs.

Finally, this paper provides an estimate of off-label prescribing in the market for cancer drugs in Medicare and considers the role of off-label prescribing and its evidence base in the market as a whole. Previous medical literature has examined how much off-label prescribing occurs and whether or not it is justified based on current evidence for a subset of drugs, cancers or for other diseases (Conti et al., 2013; Radley, Finkelstein and Stafford, 2006; Jung et al., 2014; Hamel et al., 2015; Eaton, Sima and Panageas, 2016; Bradford, Turner and Williams, 2014; Graziul, Gibbons and Alexander, 2012). The economic literature has considered how policy can affect the amount of off-label prescribing. Shapiro (2017) looks at the effect of detailing on off-label prescribing and Smieliauskas et al. (2017) at the effect of state mandates on private health insurance coverage on the amount of off-label prescribing. This paper considers off-label prescribing and its role in the market in the broader context of the regulatory system.

The paper proceeds as follows. Section 2 provides further background information about drug regulation and off-label prescribing in the US. Section 3 describes the data set and Section 4 presents key facts about both on and off-label prescribing of cancer drugs in Medicare. Section 5 describes the demand model and the estimation of the effect of RCTs on demand. Section 6 describes the firms investment decision and shows empirically which market and drug characteristics explain the decision. Section 7 describes the welfare issues in requiring trials and describes the computation and results of the back of the envelope welfare calculation. Section 8 summarizes what can be learned from off-label prescribing for drug regulation for more generally.

2 Drug Regulation & Off-label Prescribing in the US

2.1 Drug Regulation & Clinical Trials

Drug companies have been required to produce evidence on the effectiveness of a drug for its intended use, using controlled trials, since the Kefauver-Harris Drug Amendments to the Federal FD&C Act were passed in 1962. The amendments were introduced in response to the observation that either poor quality or no clinical research was being undertaken by drug companies (Davies and Kerimani, 2008). In addition to requiring controlled trials on efficacy, they also introduced a
pre-market approval process. Previously, drug companies had only been required to demonstrate the safety of the drug and they received approval automatically if the FDA did not act on the information within a certain timeframe (Food & Drug Administration, 2006).

The current drug approval process involves pharmaceutical companies conducting a series of trials on the safety and efficacy of the drug relative to a control. First, pharmaceutical companies must obtain permission to test the drug in humans. This is granted on the basis of pre-clinical research conducted in laboratories. After obtaining permission for human testing, firms need to demonstrate efficacy and safety of their drug in three phases. The first phase is designed to determine the safety and optimal dosage of the drug. This phase typically includes small non-randomized studies (Food & Drug Administration, 2015). In the case of cancer drugs, 15-30 people is a typical sample size during this phase (National Cancer Institute, n.d.). Phase 2 trials are designed to test the efficacy and side effects of the drug. These trials can be randomized or non-randomized and may contain hundreds of people. In cancer trials it is typical to use a non-randomized trial design during phase 2, with a much smaller sample size (on average about 50 people) relative to what would be used for non-cancer drugs (Gan et al., 2010; Michaelis and Ratain, 2007). Finally, Phase 3 trials are intended to further study the efficacy of the drug and the adverse reactions in a much larger sample over a longer period of time. For cancer drugs, this is usually the first randomized trial. Although there are no specific rules about the control group for the drug, cancer drugs are typically tested against the current conventional treatment (National Cancer Institute, n.d.).

Once the trials are complete, the drug company submits the results to the FDA for review. If the FDA finds that the safety and efficacy profile of the drug represents a reasonable trade-off, the drug is approved for marketing in the US. Uses of the drug that have been reviewed and approved by the FDA are referred to as on-label. This is because these uses are listed as intended uses of the drug on the drugs package insert (or label). Drug companies can choose to conduct trials on the effectiveness of the drug for other indications and submit the results for review and approval by the FDA. The label would then be amended to include these new approved indications.

### 2.2 Off-label prescribing

Off-label prescribing is the administration of a drug in any manner that has not been approved by the FDA. This could include: prescribing a drug for a different indication (for example, a cancer drug given to treat arthritis); a different dosage; in combination with other drugs; or prescribing for a group of people (children instead of adults) that is different from what is described on the drugs' label. Off-label prescribing is legal because FDA regulations surrounding pharmaceuticals pertain to the firms activities: marketing and disseminating information about the drug. Doctors
are able to prescribe drugs as they see medically appropriate. The only restrictions are those imposed by malpractice law. That the drug is off-label is insufficient grounds for malpractice. It is the medical justification for prescribing the drug that is considered in malpractice cases (Cavalla, 2015).

Off-label prescribing is common, particularly in cancer, psychiatric diseases and pediatrics. Previous literature has found estimates of off-label prescribing ranging from 21%-39% of prescriptions (Bradford, Turner and Williams, 2014; Radley, Finkelstein and Stafford, 2006). In cancer specifically, Molitor and Agha (2012) found that 22% of Medicare prescriptions made between 1998-2008 for cancer drugs were for off-label uses. Conti et al. (2013) found that 30% of the top-10 on-patent cancer drugs were prescribed for an off-label use in 2009. Off-label prescribing is common in cancer for several reasons. Firstly, the regulatory approval process is slow and costly for the drug company. This means that the approved uses of drugs do not necessarily reflect the latest scientific literature and medical thinking (Gupta and Nayak, 2014). Second, the severity of cancer means that patients may be willing to accept the greater safety risk posed by an unapproved drug than they would with a less severe disease. Third, many cancers are rare diseases and do not have many (or any) approved treatments. However, it is known that cancers share biology and sensitivity to therapeutic agents, making it a reasonable strategy to consider unapproved treatments.

2.3 Insurance Coverage

Off-label uses of cancer drugs are well insured in the US. Under the 1993 Omnibus Budget Reconciliation Act, Medicare is required to cover off-label uses of drugs in oncology if the use is one listed as supported in one of five designated compendia or if there is evidence published in any of twenty-six peer-reviewed journals (American Society of Clinical Oncologists, 2009). The drug compendia are independently operated and there are no criteria governing their system of reviews or recommendations. Abernethy et al. (2009) conducted a review of the compendia and found that the recommendations were based on very little of the available evidence and were not based on the most recent or highest quality studies. Moreover, across compendia, there are substantial differences in the evidence cited and recommendations made. This has led to Medicare coverage of off-label indications for cancer that only have evidence of efficacy from a small non-randomized phase 2 trial.

Private insurance also covers off-label uses of cancer drugs. In thirty-six states, legislation has been enacted that mandates coverage by private insurers (based on population, these mandates affect 74% of the US population (Bach, 2009)). Plans covered by the Employee Retirement Income Security Act of 1974 (ERISA) are not subject to these mandates. Nevertheless, empirical work

---

1Medicaid coverage rules are similar, but have a restricted list of compendia and journals.
suggests that private insurers do provide generous off-label coverage for cancer drugs. Smieliauskas et al. (2017) find that off-label prescribing rates for chemotherapy in a sample of private claims data are 41%, larger than estimates observed in Medicare. Bradford, Turner and Williams (2014) examine off-label prescribing across insurance types (across all types of diseases) and finds similar rates across insurance types.

2.4 Advertising

Regulation around the information that can be distributed about drugs provides an important incentive for firms to undertake clinical trials. In addition to regulating the entry of new drugs onto the market, the FDA also regulates the information that firms can provide about the drugs. All advertising to doctors and patients as well as promotional materials must contain truthful information about the safety and effectiveness of the drug. That is, the information must be supported by evidence from clinical trials. This leads to an important difference between on and off-label uses of drugs – off-label uses cannot be advertised or promoted. However, it is permitted for the drug company to distribute published scientific studies (Ventola, 2009). Moreover, there are a number of commercial drug compendia, which provide physicians with a synthesis and evaluation of drugs, based on the current literature, for both on and off-label uses. The compendia are not subject to regulation and are permitted to make recommendations about off-label prescribing, and synthesize scientific findings in accordance with their own policies.

2.5 Clinical Trial Registration

A final important component of drug regulation is trial registration, which affects the drug company’s ability to conceal the existence of trials. The Federal Drug Administration Modernization Act in 1997 introduced compulsory registration of clinical trials conducted in the US for ‘serious or life-threatening conditions’. Initially the low number of registered trials suggested that compliance was low. In 2005, the International Committee of Medical Journal Editors (ICMJE) introduced a rule that publications of clinical trials must be indexed to a clinical trial registry. Since then, registration has improved, although it is difficult to determine the compliance rate. In addition to registering the existence of a trial, reporting of the results has been compulsory since 2008. Compliance with timely reporting of results has found to be poor (Gill, 2012; Prayle, Hurley and Smyth, 2012). The implication of this is that it is difficult for firms to hide the existence of trials. However, as will be discussed in Section 4.2, it is not necessary for them to do so. In fact, firms

---

2Although off-label promotion is illegal, and civil and criminal penalties are substantial, firms engage in the practice. Between 2004 and 2013, eighteen cases were settled with fines above $300 million, with some companies being implicated more than once (Cavalla, 2015).
openly report cancelling trials where drugs are failing to perform as hoped (Pak, Rodriguez and Roth, 2015). As such, it should be possible to observe all trials ever begun in the data set, at least in the later years of the sample.

3 Data

The main empirical goal of the paper is to identify how RCTs affect demand for off-label uses of cancer drugs. To do this, I require data on off-label prescriptions of cancer drugs and data on RCTs that have been conducted for these off-label uses. Data on cancer drug prescriptions is obtained from Medicare claims data. I classify the prescriptions as on or off-label using a crosswalk I developed to match approved drug indications with the diagnosis codes included in the claims data. I then match the claims data with a data set containing RCTs conducted for off-label uses of cancer drugs, which I compiled from the scientific literature. This section details the construction of the data.

3.1 Cancer Drug Prescriptions

Cancer drug prescriptions are obtained from Medicare administrative claims data. Medicare is the US government national insurance program for people aged 65 or older. There are four parts (A,B,C,D), which cover different types of services. Cancer drugs that are administered by a doctor (for example because they are injected or intravenously administered), are covered under Part B. Part B also covers some oral drugs, mainly ones that were approved before the introduction of Medicare Part D in 2006.

The files I use are the 100% Outpatient and the 20% Carrier Claims research identifiable files for 1999-2013. These files contain every claim made by institutional providers (such as hospital outpatient facilities) and 20% of claims made by non-institutional providers (such as physicians in private practice) on behalf of beneficiaries covered under Medicare Fee-For-Service (FFS). Medicare FFS (traditional Medicare) accounted for 82% of beneficiaries in 1999, but has declined over time and by 2013 accounted for 69% (Kaiser Family Foundation, 2017). The data is a sample of approximately a third of Medicare FFS claims.3 Medicare likely covers more than half of the total cancer population in the US. Fifty-three percent of new cases of cancer are diagnosed in people aged 65 and over (National Cancer Institute, n.d.a). A further 25% are diagnosed at age 55-64, some of whom will still be in treatment by age 65 and some of whom may qualify for Medicare early due to disability. Unfortunately, no data exists on the prevalence of cancers that are actively

---

3The 100% Outpatient sample has approximately the same number of claims as the 20% Carrier claims for the drugs included in this sample.
being treated.

To identify claims for cancer drugs I begin by compiling a list of drugs used to treat cancer. This is obtained from two sources. First, the National Cancer Institute provides a list of drugs currently used to treat cancer. Second, I supplement this with a list of approved drugs that was obtained from a historical version of the FDA website to ensure that I include drugs that may have become obsolete during my sample period. I then used the drugs@FDA database (Food & Drug Administration, 2015b) to find the date of first approval, approved indications and the approval dates of supplementary indications for each of the drugs. I excluded drugs that were approved after December 31 2013, as they would not be observed in the claims data, and drugs that were for the treatment of symptoms (such as pain, nausea or osteoporosis).

Drugs are identifiable in Medicare claims by their Health Care Common Procedure Coding System (HCPCS) codes. I use the HCPCS Release and Code Sets from 1999-2013, published by the Center for Medicare & Medicaid Services, to match the list of drugs with their HCPCS codes in each year. Some drugs do not have a unique HCPCS code or are not covered by Medicare Part B (for example because they are administered orally and are covered under Part D). There are 80 drugs that are uniquely identifiable in the Medicare claims data.

### 3.2 Classifying Prescriptions as On & Off-label

The Medicare claims data does not provide a record of whether or not the claim is for an on or off-label purpose. However, each claim has several ICD-9CM diagnosis codes attached, which I use to determine the approval status of the claimed indication. This allows me to identify off-label prescribing of the type where a cancer drug approved for treatment of one type of cancer is used to treat another type of cancer, where the type is defined based on organ site. This method for identifying off-label prescriptions has been used by other researchers in different data sets and diseases (Radley, Finkelstein and Stafford, 2006; Graziul, Gibbons and Alexander, 2012; Smieliauskas et al., 2017). It provides a lower bound on the amount of off-label prescribing because approved indications are actually defined more narrowly than organ site.

In order to classify the prescriptions as on or off-label I create a cross walk between the approved indications listed on the drug labels and these ICD-9CM codes. The indications are taken from the drug labels, which were obtained from drugs@FDA. I define cancer as ICD-9CM codes ranging between 140 and 239, which is the three digit level for neoplasms. Within cancer, I define 25 types of cancer based on organ site. Eighteen of these are for solid cancers and 7 are blood cancers.

---

4This list can be obtained from the National Cancer Institute webpage “A-Z List of Cancer Drugs” at www.cancer.gov/about-cancer/treatment/drugs. I used the list as it was on November 4 2015.

5Historical versions of the FDA website can be accessed using the wayback machine at http://archive.org/web/.
To create these groupings I first take the anatomical region groups described in Skeel and Khleif (2011). I divide some of these regional groups into subgroups considering the groupings suggested by (Weiner et al., 2003) as well as the prevalence of the subtypes (see National Cancer Institute, 2017) and the cancer types mentioned on the drug labels.

3.3 Sample Construction: Demand

There are 26.1 million claims made in the Medicare data between 1999-2013 for the 80 cancer drugs of interest. After excluding beneficiaries with multiple types of cancers, there are 23.3 million claims for these drugs with a cancer diagnosis attached. These come from 1.765 million Medicare beneficiaries treated at least once for cancer with cancer drugs over the period. I define a treatment episode for a patient as all prescriptions for the same drug, taken without a break of more than sixty days or twenty percent of the median treatment break. There are 4.961 million treatment episodes in the data.

Table 1 shows the number of people in the sample with each type of cancer who have at least one claim for at least one drug that is included in the sample. In 2013, the prevalence of cancer patients receiving drug treatment varies from 244 people with male genitourinary cancer to 41,348 people with lung cancer.

3.4 Clinical Trial Data

To study the effect of RCTs on demand, I focus on drugs covered by Medicare Part B that were first released to the US market between 1996 and 2010. The drugs included are primarily intravenous drugs, however, also included are oral drugs that have an individual claim code that is covered by Part B. I focus on new drugs for two reasons. Firstly, it allows us to observe the effect of the first randomized trial. Secondly, drugs that are still on patent are more likely to be tested by drug companies, which is the testing margin of interest. The date range was chosen to maximize the sample size while allowing a reasonable length of time to observe demand in the post-trial period. Since the last year of data is in 2013, only drugs launched before 2010 are included, which leaves a period of three years to observe post-launch demand. Although the first year of data is 1999, drugs launched in the previous three years are included so as to maximize the sample size. Based on the dates of approval for cancer drugs described in Section 3.1, a total of 39 drugs were released during this period. However, only 27 are covered by Medicare Part B with a unique identifying code (some orals that are covered by Part B do not have a unique claim code and some orals are covered under Part D).

I compile data on RCTs that have been conducted for off-label uses of these forty drugs. To do this I first query the PubMed database for each of the molecule names. I then perform a text
Table 1: Medicare beneficiaries receiving at least one drug treatment: by cancer site 1999-2013

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Bladder</td>
<td>3,944</td>
<td>6,983</td>
<td>9,970</td>
<td>13,127</td>
<td>15,792</td>
<td>18,847</td>
</tr>
<tr>
<td>Brain</td>
<td>506</td>
<td>393</td>
<td>530</td>
<td>1,310</td>
<td>2,332</td>
<td>2,686</td>
</tr>
<tr>
<td>Breast</td>
<td>14,046</td>
<td>17,844</td>
<td>17,726</td>
<td>23,323</td>
<td>29,315</td>
<td>33,575</td>
</tr>
<tr>
<td>Cervical</td>
<td>485</td>
<td>606</td>
<td>692</td>
<td>847</td>
<td>1,097</td>
<td>1,279</td>
</tr>
<tr>
<td>Colorectal</td>
<td>17,073</td>
<td>17,582</td>
<td>18,484</td>
<td>19,994</td>
<td>21,909</td>
<td>23,690</td>
</tr>
<tr>
<td>Gastroesophageal</td>
<td>2,595</td>
<td>3,480</td>
<td>4,393</td>
<td>5,278</td>
<td>6,817</td>
<td>8,160</td>
</tr>
<tr>
<td>Gynecologic</td>
<td>1,125</td>
<td>1,508</td>
<td>2,204</td>
<td>3,255</td>
<td>4,927</td>
<td>6,619</td>
</tr>
<tr>
<td>Head &amp; Neck</td>
<td>1,274</td>
<td>1,982</td>
<td>2,980</td>
<td>4,656</td>
<td>5,982</td>
<td>7,129</td>
</tr>
<tr>
<td>Hodgkin's Lymphoma</td>
<td>509</td>
<td>541</td>
<td>667</td>
<td>780</td>
<td>910</td>
<td>1,129</td>
</tr>
<tr>
<td>Kidney</td>
<td>660</td>
<td>944</td>
<td>1,154</td>
<td>1,605</td>
<td>2,002</td>
<td>2,113</td>
</tr>
<tr>
<td>Leukemia (ALL)</td>
<td>149</td>
<td>176</td>
<td>221</td>
<td>296</td>
<td>419</td>
<td>490</td>
</tr>
<tr>
<td>Leukemia (AML)</td>
<td>210</td>
<td>310</td>
<td>491</td>
<td>843</td>
<td>1,465</td>
<td>1,923</td>
</tr>
<tr>
<td>Leukemia (CLL)</td>
<td>1,875</td>
<td>2,949</td>
<td>4,155</td>
<td>4,784</td>
<td>5,855</td>
<td>6,769</td>
</tr>
<tr>
<td>Leukemia (CML)</td>
<td>267</td>
<td>158</td>
<td>129</td>
<td>175</td>
<td>271</td>
<td>347</td>
</tr>
<tr>
<td>Liver</td>
<td>575</td>
<td>601</td>
<td>1,083</td>
<td>1,546</td>
<td>3,034</td>
<td>3,707</td>
</tr>
<tr>
<td>Lung</td>
<td>19,036</td>
<td>24,125</td>
<td>29,394</td>
<td>33,205</td>
<td>37,882</td>
<td>41,348</td>
</tr>
<tr>
<td>Male Genitourinary</td>
<td>66</td>
<td>92</td>
<td>96</td>
<td>136</td>
<td>192</td>
<td>244</td>
</tr>
<tr>
<td>Melanoma</td>
<td>731</td>
<td>857</td>
<td>1,086</td>
<td>1,504</td>
<td>2,192</td>
<td>3,334</td>
</tr>
<tr>
<td>Multiple Myeloma</td>
<td>2,264</td>
<td>3,925</td>
<td>3,969</td>
<td>5,392</td>
<td>9,455</td>
<td>12,978</td>
</tr>
<tr>
<td>Non-Hodgkin’s Lymphoma</td>
<td>10,486</td>
<td>14,119</td>
<td>18,717</td>
<td>21,160</td>
<td>23,919</td>
<td>26,379</td>
</tr>
<tr>
<td>Ovarian</td>
<td>6,420</td>
<td>7,626</td>
<td>8,386</td>
<td>9,255</td>
<td>10,613</td>
<td>11,767</td>
</tr>
<tr>
<td>Pancreas</td>
<td>2,585</td>
<td>3,396</td>
<td>4,346</td>
<td>5,979</td>
<td>7,730</td>
<td>9,330</td>
</tr>
<tr>
<td>Prostate</td>
<td>4,579</td>
<td>5,846</td>
<td>8,650</td>
<td>9,876</td>
<td>11,841</td>
<td>10,518</td>
</tr>
<tr>
<td>Thyroid</td>
<td>143</td>
<td>115</td>
<td>171</td>
<td>297</td>
<td>502</td>
<td>678</td>
</tr>
<tr>
<td>Urinary</td>
<td>76</td>
<td>103</td>
<td>146</td>
<td>252</td>
<td>341</td>
<td>427</td>
</tr>
<tr>
<td>Other</td>
<td>4,762</td>
<td>5,816</td>
<td>6,332</td>
<td>7,735</td>
<td>10,778</td>
<td>11,445</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>96,441</td>
<td>122,077</td>
<td>146,172</td>
<td>176,610</td>
<td>217,572</td>
<td>246,911</td>
</tr>
</tbody>
</table>

Notes: This table shows the counts of cancer patients with each type of cancer who are receiving treatment with at least one drug included in the sample. These counts differ from national prevalence estimates because they do not include patients who receive no drug treatment (such as those only receiving radiation or surgery during the calendar year and those who are not being treated). The counts also differ from the Medicare population because they have not been weighted to account for the fact that the data is a sample (of approximately one third).
search of the article title and the abstract and determine the phase of the trial and the possible
cancer types. For each cancer type, the abstracts of all articles classified as phase 3 trials or phase
2 trials that contain a variant of the word “random” are manually checked to determine if: (1) the
disease and drug classification were correct; (2) this was an RCT for the drug of interest, designed
demonstrate efficacy. It is possible that trials were conducted that were never published/have not
been published yet. To check this I also search the clinicaltrials.gov database for each of the
molecules of interest. The clinicaltrials.gov database has a primary indication variable, which I use
to determine whether or not the trial is for an off-label use.

Using these publications I manually extract information from the journal articles. The informa-
tion I include is the start and end date of the trial, publication date, first date presented at
a conference, the treatment and control outcomes for three measures of efficacy (overall survival,
progression free survival, overall response rate) and whether or not the difference was statistically
significant, and finally whether or not the author’s report the trial result as positive, negative or
neutral. Some of the conference dates were not reported in the articles, and these were found
by looking up the lead author’s presentation history in the historical conference records for the
American Society of Clinical Oncologists Annual meeting.

3.5 Market & Drug Characteristics

Data on drug and market (by cancer type) are used to predict the firm’s decision to invest in a
trial. The drug characteristics are the price at the launch of the drug, length of patent remaining
and the efficacy reported in the trial for the first use. The market characteristics are the incidence
of the cancer, the probability of surviving one year with the disease, the relationship between the
cancers in terms of the number of approved drugs shared, off-label prescribing per person and the
number of approved drugs in the market.

Data on the launch prices is taken from Bach (n.d.). The length of patent remaining was
computed using data from the FDA orange book (Food & Drug Administration, 2017). The
efficacy data for the trials for the first use were taken from drug labels, which are available from
drugs@FDA (Food & Drug Administration, 2015b). The probability of surviving one year and
the incidence of the cancer were obtained from the SEER program (Howlader et al., 2017). The
number of drugs approved on the market and the number of drugs shared between two cancers
were calculated using the data described in Section 3.1. The amount of off-label prescribing per
person was computed in the Medicare data.
4 Empirical Facts

4.1 Off-label Prescribing

This section examines how widespread off-label prescribing is amongst Medicare beneficiaries, and whether or not those off-label prescriptions are for tested or untested uses. Both the whole market and only drugs released from 1996-2010 ("new drugs"), which are the focus of much of the paper, are considered. The data reveals that 23\% of prescriptions of cancer drugs in Medicare are for an off-label use. Amongst “new drugs”, the share of off-label prescriptions is 18\%. Off-label prescriptions account for 13\% of Medicare spending on cancer drugs. In the sample, average real spending per year on these drugs is $1.26 billion ($2013).

The vast majority of prescriptions are for drugs that have an RCT. Drugs that are never tested, or drugs that have not yet been tested, are used very little. On average, untested “new drugs” only account for 4.4\% of the total market. However, they account for 10.9\% of the market for new drugs. As a share of Medicare spending on cancer drugs, new untested uses account for 6\%.

Figure 1: Market Share: New Drugs

Notes: This graph shows the market share of the 27 cancer drugs that were released between 1996 and 2010.

Figure 1 shows the market share of “new drugs” over the sample period. The market is defined as all new treatment episodes in each period for each drug.\textsuperscript{6} It increases from 20\% to 40\% over

\textsuperscript{6}A treatment episode begins when a patient is prescribed a drug for the first time, and continues until there is a break in treatment of more than two cycles. If a person is already being prescribed one drug and begins a new
time, this is partly because the drugs are released over time. Figure 2 shows the average take-up of first uses of cancer drugs released to the market between 1996 and 2010 compared with secondary uses. First uses of drugs, which always have an RCT available by the time of their release to the market, are taken-up at a faster rate and by more people. Initial demand for the new drugs is higher on average by 200-300 people for first uses relative to secondary uses. This difference could be due to the drugs having an RCT or underlying differences in the quality of the drug, which is the reason for the estimation approach proposed in this paper.

Figure 2: Take-up and Expenditure on new drugs: First vs Secondary Uses

Notes: Computed using the 27 drugs that were released between 1996 and 2010. First uses are prescriptions of the drug for the cancer type that was the first to be approved by the FDA for each drug. Secondary uses are prescriptions of the drug for all other cancer sites. Secondary uses that have less than five claims per year are excluded.

4.2 Clinical Trials

Table 2 shows characteristics of the clinical trials used in the event study in Section 5 and the firm entry problem in Section 6. The event study includes trials where the results were released between 1999 and 2012 that were sponsored by the drug company. There are 32 trials that meet this criteria. These trials cover 12 different drugs and 15 different cancer types. The average drug, there is a new treatment episode for the new drug, while the treatment episode for the second drug is dated based on when treatment with that drug began.

Secondary uses are defined as all uses of the drug that are not the first approved use. Some of these uses may start out as off-label and later become on-label and some have RCTs and some do not.
Table 2: Descriptive Statistics: Clinical Trials

<table>
<thead>
<tr>
<th>Event Study: Solid &amp; Blood</th>
<th>Positive result</th>
<th>Non-positive result</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td># Trials</td>
<td>10</td>
<td>22</td>
<td>32</td>
</tr>
<tr>
<td>Efficacy effect size (months)</td>
<td>2.84</td>
<td>0.41</td>
<td>1.17</td>
</tr>
<tr>
<td>Addition to combination (vs alternative)</td>
<td>50%</td>
<td>54.6%</td>
<td>53%</td>
</tr>
<tr>
<td>#Unique Drugs</td>
<td>7</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>#Unique Disease</td>
<td>7</td>
<td>13</td>
<td>15</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Firm entry: Solid Cancers Only</th>
<th>Positive result</th>
<th>Non-positive result</th>
<th>All</th>
</tr>
</thead>
<tbody>
<tr>
<td># Trials</td>
<td>13</td>
<td>26</td>
<td>39</td>
</tr>
<tr>
<td>Efficacy effect size (months)</td>
<td>3.64</td>
<td>0.42</td>
<td>1.5</td>
</tr>
<tr>
<td>Addition to combination (vs alternative)</td>
<td>53.9%</td>
<td>57.7%</td>
<td>56.4%</td>
</tr>
<tr>
<td>#Unique Drugs</td>
<td>8</td>
<td>11</td>
<td>13</td>
</tr>
<tr>
<td>#Unique Disease</td>
<td>7</td>
<td>13</td>
<td>14</td>
</tr>
</tbody>
</table>

Notes: This table shows the descriptive statistics for the trials included in the demand problem (event study) and the firm investment problem. The trials differ because the demand problem only includes trials with results released between 1999-2012. The firm problem focuses on solid cancers, and drugs first released for solid cancers, but includes trial released between 1996-2016 (as it does not rely on the prescription data).

An increase in life expectancy of the tested drugs over the control group reported in the trials is 1.17 months. The firm problem only includes solid cancers and drugs first approved for solid cancers. However, since this analysis does not make use of the prescription data, trials that are first released between 1996 and 2016 are included.

The trials have been classified as positive or non-positive based on whether or not overall survival is statistically significant. Improvement in overall survival (which is how long a person lives from beginning treatment) or improvement in quality of life are the goal of cancer treatment. Most clinical trials included in this study did not include a direct measure of quality of life. They tend to measure changes in overall survival, progression free survival and the overall response rate. Progression free survival is the amount of time that the cancer does not advance and the overall response rates is the percentage of patients that experience a response to the drug, where the response is defined in advance of the trial. There is limited evidence linking improvements in progression free survival and the overall response rate with increased survival in solid cancers (Booth and Eisenhauer, 2012; Amir et al., 2012).

---

8The reason for this is that there are greater associations within solid cancers and within blood cancers than their are across the two types. Blood cancers also measure efficacy using different outcomes.

9Progression free survival is sometimes considered a measure of quality of life because progression would tend to be associated with worsening quality of life. However, it is not a direct measure because the treatment and control
It is uncommon for a trial to show a negative result. In terms of statistical significance, there is only one trial included in this study where a statistically significant reduction in overall survival was reported. The lack of negative results has two potential explanations. First, firms may be good at selecting which drugs to trial. Second, firms may be able to hide negative results. I examine empirically whether or not it appears that firms are good at selecting which drugs to trial in Section 6. There is a large body of evidence from the literature suggesting that firms are able to suppress negative trial result. For example, Stensland et al. (2014) find that 20% of cancer trials are canceled before completion. Pak, Rodriguez and Roth (2015) find that 30% of canceled trials report that the reason for cancellation is a failure of the treatment group to meet efficacy expectations. Both of these studies are based on an analysis of the US clinical trial registry (clinicaltrials.gov). There is nothing preventing firms from terminating a trial and potentially even beginning a new trial as long as the existence of the trials is reported. The implications of the lack of negative trial results is shown in Appendix B.

An analysis of the abstracts of the trial publications shows that out of non-positive trials only 14% frame the trial results as negative. Forty-five percent refrain from making a statement and simply report the efficacy and safety outcomes and the remaining 41% are framed as positive. Again, the literature has documented an excess in positivity in interpreting trial results (see for example Vinkers, Tijdink and Otte, 2015; Abola and Prasad, 2016; Boutron et al., 2010).

5 How do RCT’s affect Demand?

In this section I estimate the effect of RCTs on cancer drug demand. First, I present a model of cancer drug demand that has a role for clinical trials. Then I estimate the effect of the release of the results of an RCT on demand for off-label uses of cancer drugs that have a trial during the sample period. I find that demand increases in response to positive information about drug efficacy and does not respond to non-positive information.

5.1 Model

Doctors make decisions about cancer treatments on behalf of their patients. The doctor acts to maximize the patient’s utility; there is no agency.\textsuperscript{10} At each time period $t$ the doctor $i$ treats the drugs might have differing effects on quality of life – cancer drugs tend to have severe side effects that diminish quality of life.

\textsuperscript{10}A core principle in medicine is that the physician should act in the patients best interest. However, it is well known that physicians have financial conflicts of interest. In the case of cancer drugs, physicians receive reimbursement based on a percentage of the drug price, which means that drugs that cost more provide the doctor with greater income (Polite, Conti and Ward, 2015). There is limited empirical evidence on the extent to which
patients that arrive at their medical practice. These patients have a type of cancer $k$, which is accurately diagnosed by the doctor. Since cancer drugs are often consumed in combinations, the doctor selects any number of the drugs $d$ out of the set of existing drugs $D_t$.\footnote{Combining drugs is a core principle in cancer treatment (Chu and DeVita, 2016). I observe in the data that most patients are treated with combination therapy.} Combining drugs is guided by scientific principles – doctors do not combine every drug because combinations can be toxic and some drugs will have reduced efficacy when used in combination (Chu and DeVita, 2016).

I model the decision to choose each individual drug, rather than the choice of the combination. The implication of this assumption is that the drugs are being modeled as if they complements.\footnote{The reason for this choice is driven by the data and several facts about this market. First, many drugs are complements with each other. More than half of the new drugs considered in the event study are tested as additions to existing drug combinations (see Table 2). Second, the relationship between trial results and demand is complicated by the fact that trials are framed around testing the efficacy of a single drug. In reality, what is being tested is the efficacy of a combination of drugs, one of which is the new drug of interest. In the data I observe a very large number of combinations being prescribed almost all of which are untested as combinations. This means there is no data on the efficacy of the combination relative to other combinations. When a trial is released, demand for the drug both as part of the tested and other untested combinations appears to respond to the trial results. Third, consumers do not face a financial trade-off when choosing a drug, they choose the drug with the most preferable safety, efficacy and evidence profile. Without data on the relative efficacy of every combination, or a relationship between prices and quantity, it is very difficult to set up a model that can be estimated that allows for more complicated patterns of substitution and complementarities. Modeling the consumption decision of each drug individually provides an interpretable and direct measure of the effect of the trial results on demand for the drug.}

The doctor/patient’s utility, shown in Equation 1, is a linear function of the benefit of the expected efficacy ($b_{dkt}$) of the drug $d$ for the cancer $k$ less the cost of the expected side effects ($s_{dkt}$). These can change over time as new information is released. Patients are risk averse and dislike uncertainty around the benefits and side effects of the drugs ($\sigma^2_{dkt}$). This creates a role for running a trial. I assume no monetary cost to the patient because patients are fully insured. There is heterogeneity in doctors beliefs about the benefits and side effects of the drugs ($\epsilon_{idkt}$).

\begin{equation}
U_{idkt} = \alpha \mathbb{E}(b_{dkt}) - \beta \mathbb{E}(s_{dkt}) - \frac{\rho}{2} \sigma^2_{dkt} + \epsilon_{idkt} \tag{1}
\end{equation}

The doctor will choose the drug $d$ for the disease $k$ if $U_{idkt} \geq 0$. Under the assumption that $\mathbb{E}(\epsilon_{idkt}=0)$ and that the demand for each drug is independent of the demand for every other drug, the aggregate demand ($q_{dkt}$) for each drug for each disease is given by Equation 2.
\[ q_{dkt} = \alpha \mathbb{E}(b_{dkt}) - \beta \mathbb{E}(s_{dkt}) - \frac{\rho}{2} \sigma^2_{dkt} \]  

When the results of a trial are released there are two potential effects. First it can change the belief about the expected benefit (\( \mathbb{E}(b_{dkt}) \)) of the drug for the disease that was tested. Second, it reduces the uncertainty around the expected benefit (\( \sigma^2_{dkt} \) decreases). I assume that the trial does not affect the expected side effects because these were revealed in the trial for the first use. Since I aim to estimate the effect of a trial on demand, I recast the demand equation as a function of factors that vary depending on whether or not there is a trial (\( trial_{dkt} \)) and factors that do not (\( \delta_{dk} \)).

\[ q_{dkt} = \gamma_{trial_{dkt}} + \delta_{dk} \]  

5.2 Estimation

I estimate the demand function in Equation 3 using an event study. The event is the date the results of an RCT were first presented at a conference - most of the trials were presented at the American Society of Clinical Oncologists conference. The quantity demanded is defined as a new treatment episode. An episode is defined as a patient’s use of a drug without a break of the length of more than two treatment cycles.\(^{13}\) I allow the effect of the trial to vary depending on the result by splitting the responses into positive and non-positive results.\(^{14}\) A positive result is a statistically significant increase in overall survival. A non-positive result has a statistically insignificant difference in overall survival.\(^{15}\)

The event study, Equation 4, is estimated using data on the 32 drug-disease pairs that have a trial released between 1999-2012 that was financed by the drug company. I regress the quantity

\(^{13}\)A cycle refers to the frequency at which the drug is administered. For example, if drug is administered every three weeks with a one week break then the cycle length is four weeks.

\(^{14}\)The reason for using a binary variable to control for the effect of a trial, rather than the level efficacy measured in the trial, is that the size of the efficacy gain is not comparable across drug markets and may not be meaningful at all. For example, if consumers are only interested in choosing the drug with the highest efficacy then it will not matter whether or not the increase is one month or one year. Moreover, the markets that are defined in this paper are based on the IC9-CM definition of a cancer site. The IC9-CM codes are not as detailed as the target markets defined for the drugs in trials. For example, breast cancer drugs are tested for subtypes of cancer on the basis of their hormone receptiveness. The implication of this is that a three month improvement in life expectancy may lead to a small change in a drugs demand in breast cancer because the trial result is only targeted at 25% of the market. While in another disease market the drug is targeted at the whole market (as defined in this paper) but only provides a one month improvement.

\(^{15}\)There is only one trial that reports a statistically significant negative result. This trial is excluded from the regression.
demanded of a drug $d$ for a cancer $k$ on event dummies, which are allowed to vary based on whether or not the trial result is eventually positive ($pos_{dk}$). I include time fixed effects (year-quarter), drug-disease fixed effects and a fixed effect for how long the drug has been available on the market ($elapsed_{dt}$). The latter aims to capture take-up of the drug that would have occurred over time in the absence of the trial. The drug-disease fixed effects control for time invariant characteristics of the drug-disease pair, such as beliefs about the drug efficacy, which come from the non-randomized phase 2 trials as well as differences in physician reimbursement. The standard errors are clustered by drug-disease pair.

$$q_{dkt} = \sum_{l=-12}^{l=12} \alpha_l \mathbb{I}(t - t_{event}^{dk} = l) \times pos_{dk} + \sum_{l=-12}^{l=12} \beta_l \mathbb{I}(t - t_{event}^{dk} = l) + elapsed_{dt} + \delta_t + \gamma_{dk} + \epsilon_{dkt} \quad (4)$$

The coefficients of interest ($\alpha_l, \beta_l$) trace out changes in drug demand relative to demand for drugs that have a trial result that is ultimately non-positive at the time of the release of the trial. These coefficients are identified using variation in the timing of the release of the trial results relative to when the drug was first released (both across drugs and within drugs for different diseases). Variation in the result provides a placebo test. The key identifying assumption is that the time at which the trial is released relative to when the drug was first released to the market is uncorrelated with unobserved time-varying beliefs about drug quality. An example of a threat to identification is if the drug company has hidden information about drug quality and is tracking the off-label use of the drug. When they observe demand is lower than would be expected given their belief about the quality, they run a trial. This would lead the treatment effect to be upwardly biased. This threat is mitigated by the difficulty historically of tracking off-label use in real time. While the drug company likely tracks overall sales, attributing demand to individual off-label uses requires detailed claims data linked to an off-label characterization, which is only now becoming feasible to track.

The estimation result is interpreted as the effect of having an RCT over not having an RCT on efficacy due to institutional features, explained in detail in Section 2. Briefly, the off-label uses of drugs that are observed being prescribed in the claims data have some evidence favoring efficacy from a non-randomized trial (otherwise, Medicare will not reimburse them and the drugs are not very accessible without insurance due to the high prices). These trials are very small and measure efficacy using outcomes that do not necessarily translate into an improvement in overall survival. It is also difficult to compare the efficacy outcomes from the non-randomized trials across treatment options as the trials are conducted in different sub groups of people. It is assumed that

\footnote{As a robustness check, I include the physician reimbursement amount as an explanatory variable. It is statistically insignificant and small, and does not affect the coefficients of interest.}
the safety profile of the drug is known from the RCTs that were conducted for the first use of the drug. Hence, the release of the RCT in the event study is the first randomized evidence on the efficacy of the drug for the off-label use.

5.3 Results

Figure 3 shows the event study coefficients from Equation 4. The first panel shows the additional effect on demand of a positive trial result relative to a non-positive result in each quarter ($\beta_l$). In the twelve quarters before the trial result is released there does not appear to be any difference in demand for the drug. After the trial is released, there is a statistically significant and large increase in demand. By the second year following the trial result, demand has increased by 140 patients beginning the treatment per quarter. On average there are 90 people per quarter taking up these off-label uses of drugs. Hence, this increases represents a 150% increase in demand. The second panel shows the effect of the trial for non-positive results ($\alpha_l$). Demand before and after the trial is not statistically different from demand at the time of the release of the trial results.

It is important to note that demand is measured in terms of new take-up of the drug. While it is possible that doctors respond to trials by increasing the duration of treatment for patients already receiving the treatment, this likely a small effect. Cancer drugs are typically prescribed until treatment failure is observed. There is no reason to adjust treatment duration in response to a trial result once there is information on how the patient’s cancer is responding to the treatment.

6 Why do Firms Run Trials?

Firms have the choice to conduct an RCT for an off-label use of a drug. In this section I provide an empirical characterization of the firm’s investment decision and examine whether or not market and drug characteristics predict the success of the trials that firms choose to conduct. There are three components of the firms choice to run a trial that are of interest. First are firms testing drugs that they know have higher efficacy? Second, are they testing drugs in larger markets? Third, can the outcome of the trial be predicted based on observables? The first two questions are of interest because if firms are selecting the drug based on efficacy or market size, we would expect that as additional trials are run, due to some policy change that encourages trials, the marginal effect of the trial will be diminishing. Selection on efficacy would mean that firms are already testing the most promising drugs and policies that encourage greater testing will lead to less efficacious drugs being tested on the margin. Selection on market size means that firms are testing in the markets where the impact of the trial will be largest. Moving into smaller markets as a result of a policy change will likely mean the trials affect fewer people. The third question relates to the efficiency of
Figure 3: Event study results

Notes: These graphs show the effect of the release of an RCT on take-up of off-label uses of a drugs. The coefficients are estimated relative to the demand before the release of the trial result for drugs that eventually have a non-positive trial result. The top panel shows the effect of a positive trial result relative to a non-positive trial result in each quarter. The bottom panel shows the effect of a non-positive trial result. The sample only includes off-label uses of drugs that are tested in an RCT between 1999-2012.
running these trials. If the results of the trial are predictable based on observables, the trials are not necessarily generating new information. Firms may only be running them due to restrictions on advertising.

6.1 Framework

The firm has the choice of whether or not to invest in an RCT for each of its drugs for each type of cancer. The trial is a fixed cost that is paid upfront, while the revenue resulting from the trial is accrued in the future. This is similar to the standard entry problems described in Berry (1992) and Scott Morton (1999). Firms have a flow of new product ideas, some of which they will choose to research and eventually bring to market. I take this process of producing the first use of a drug as exogenous. Once the drug is approved for its first use, the firm has a one time choice to invest in an RCT for each of the potential off-label uses. At the time of the entry decision, they are the only firm making an entry choice. The model is static, which means that the firm takes the current market conditions as given and does not consider the possibility of future entrants. There is no exit because once the trial investment has occurred and the information is revealed it will exist regardless of the firm’s actions.

The firm’s profit from investing in a trial is given in Equation 5. The revenue gained from a trial for drug \( d \) in cancer market \( k \) is the change quantity sold as a result of having the trial relative to not having the trial \( \Delta q_{dk} \) multiplied by the price of the drug \( p_k \). There is a fixed cost \( F_{dk} \) of conducting the trial. Prices are fixed across markets \( k \), which is a feature of the market. I assume that prices do not adjust to the trial results, which is supported by the data (see Appendix A for discussion and evidence). The marginal cost of producing the drug is assumed to be constant across drugs, which means that the relative prices of the drug reflect the relative profit per unit sold (for simplicity I set it to zero).

\[
\mathbb{E}(\pi_{d,\text{trial}}) = p_d \mathbb{E}(\Delta q_{dk}) - F_{dk}
\]  

(5)

In equilibrium, the firm will choose to run the trial (“enter”) if the expected revenue from the trial exceeds the cost of the trial investment (Equation 6).

\[
p_d \mathbb{E}(\Delta q_{dk}) - F_{dk} \geq 0
\]

(6)

6.2 Estimation

The probability that the firm invests in an RCT in a cancer market \( k \) with a drug \( d \) is estimated using a Probit model. The data for this section is described in detail in Section 3.5. I only consider solid cancers and drugs first approved for solid cancers. The reason is twofold. First, solid cancers
are more likely to have association with other solid cancers. Second, blood cancers use different trial endpoints, which makes them more difficult to compare with solid cancers.

A data point is a drug-cancer pair that could potentially be tested. There are 16 drugs and 14 cancer types. First uses have a 1 by definition as I am only considering approved drugs: 16 first uses are therefore dropped. This leaves 208 potential off-label entry opportunities. The entry decision is estimated using a reduced form of the expected profit function in Equation 5. This is shown in Equations 7. The expected profit is a function of observable characteristics of the drug and the cancer market ($Z_{dk}$) and characteristics that are unobservable to the econometrician but are known by the firm ($\nu_{dk}$). This unobservable component is the firm’s belief about how successful the drug will be for the cancer type. The trial is observed occurring if the expected profit from the trial is greater than zero.

$$
E(\pi_{dk}^{trial}) = Z_{dk}\beta + \nu_{dk} \tag{7}
$$

$\text{Trial}_{dk} = 1(Z_{dk}\beta + \nu_{dk} > 0)$

The characteristics included in $Z_{dk}$ are: the market size (measured using the prevalence of the the cancer in 1996); the percentage of people with the cancer surviving one year; the number of drugs already approved for the disease at the time of the entry decision; the amount of off-label prescribing per person being treated with drugs in the cancer market; a measure of the association between the cancer being considered and the cancer that the drug was first approved for (measured as the number of treatments approved for both the cancer $k$ and the cancer of the drug was first approved for); the percentage increase in efficacy the drug provided in its first use; the launch price of the drug; and the number of years remaining on the patent at the date the drug is first released to the market. The identifying assumption is that the firms unobservable belief about the payoff of conducting the trial is uncorrelated with the observable characteristics of profit. Typically the characteristic of profit that is endogenous with unobserved quality is the price. However, in this case the prices are set in the market for the first use before the entry decision into additional markets.

The results of the Probit model shed light on the observable market and drug characteristics that make investing in an RCT more profitable for the firm. However, it is plausible that the firm also has a belief about the potential efficacy of the drug. Prior to conducting an RCT the firm will have already conducted one or more uncontrolled trials from which they will have an indication of the drug’s potential efficacy. Moreover, they may have information on the potential efficacy of the drug for an off-label use from doctors who are using the drug off-label. I estimate whether or not

---

17 The cancer types thyroid, male genitourinary, gynecologic and urinary are excluded as control variables computed using the Medicare data are based on insufficient sample.
the outcome of the trials that are conducted can be predicted by the observable characteristics of profit. Since the trial outcome can only be observed if the trial occurs, I use a Heckman two-step estimator to account for the selection.

Equation 8 is the relationship between the outcome of the trial and observables that may predict the outcome of the trial. This outcome is only observed if the firm has selected to do the trial (Equation 9). The selection equation is the same as the Probit model estimated in Equation 7.

\[
y_{dk} = X_{dk}\beta + \varepsilon_{dk}
\]  

The trial outcome \(y_{dk}\) is only observed if

\[
Z_{dk}\beta + \nu_{dk} \geq 0
\]

The \(X_{dk}\) characteristics that affect the trial outcome are: the percentage increase in efficacy the drug provided in its first use; a measure of the association between the cancer being considered and the cancer that the drug was first approved for (measured as the number of treatments approved for both the cancer \(k\) and the cancer that the the drug was first approved for); the launch price of the drug; and the number of years remaining on the patent at the date the drug is first released to the market. The \(Z_{dk}\) characteristics include the \(X\) characteristics as well as measures of the market size and competition: the market size (measured using the incidence of the the cancer in 1996); the percentage of people with the cancer surviving one year; the number of drugs already approved for the disease at the time of the entry decision and the amount of off-label prescribing per person being treated with drugs in the cancer market.

I consider four measures of the trial outcome: the gain in months of life from the new drug relative to the control group; the percentage gain in life from the new drug relative to the control group; and these two measures interacted with the statistical significance of the result (which make the gain zero if the result was statistically insignificant). The characteristics include in the \(X_{dk}\) vector were chosen because they relate to the success of the drug in its first use. For example, the patent remaining or the price of the drug is expected to be greater if the drug was more successful in the first use. If there is a relationship between the success in the first use and other tested uses then these would also be predictive of the outcome of the trials for off-label uses. The additional characteristics in the \(Z_{dk}\) vector are characteristics that affect the profit of the firm but should not affect the outcome of the trial. The characteristics I exclude from the trial outcome regression relate to the potential market size for the drug: the number of people with the disease and how much competition there is in the market. An argument against the exclusion of these characteristics is that in market with more drugs/competition, it might be more difficult to find an increase in efficacy over the existing drugs - perhaps increasing life expectancy from a month to a year is not that difficult, but obtaining a second year is substantially harder. Moreover, in
larger markets, it could be less costly to run a larger trial (or multiple trials), which would increase the probability of finding a positive result. If there is no characteristic in the selection equation that can plausibly excluded from the outcome equation, it is not possible to determine the nature of the selection because there is no exogenous variation around the unobservable component of the decision. I proceed with the assumption that the market size and competition measures are not related to the outcome of the trial. As a robustness check, I include a version with only the market size excluded from the outcome equation, as this is should at least not be related to the outcomes that do not incorporate the statistical significance of the trial outcome (although again an important caveat is that the market size could be correlated with the competition measures).

6.3 Results

Table 3 shows the results of the Probit estimation for the firm’s decision to conduct an RCT for off-label uses of a drug. I report the marginal effects, computed using the delta method. The first column shows the effect of the incidence of the disease (the market size) only and the second column includes all of the observable characteristics of profit described above. The effect of the market size on the probability of entry is small and statistically insignificant: an increase of one thousand people being diagnosed with the disease per year increases the probability of a trial by 0.05 percentage points. Once other covariates are included the effect becomes negative but remains statistically insignificant.

Many of the other covariates are statistically significant. I find that each additional approved drug in the market increases the probability of a trial by 1.4 percentage points. This effect could have been positive or negative. On the one hand, more approved drugs mean greater competition in the market. On the other, with more approved drugs available it may be difficult to compete without a trial. A one unit increase in the number of off-label prescriptions per patient in the market decreases the probability of a trial by 1.06 percentage points. Again, the anticipated sign of this coefficient was ambiguous. While more off label prescribing might mean that doctors are more amenable to prescribing off-label, more off-label prescribing might occur when there are fewer approved drugs, increasing the potential market gain from a successful RCT. A one percentage point increase in the efficacy gain reported in the trial for the first use increases the probability of a trial for an off-label use by 0.47 percentage points. Greater success in the first use could increase the probability of success for a second use because there are relationships between solid cancers. However, it could also reduce the need to do a trial if doctors are more optimistic about the drugs potential and willing to prescribe it off-label without the trial. Finally, both the launch price and the years remaining on the patent have negative effects on the probability of the trial, which is unexpected. However, the data used to compute the patent length is based on the ex-post patent
length. At the time of approval the firms expectation of the patent length may have been longer or shorter, depending on plans to pursue additional protection strategies.

Table 4 shows the Heckman selection-corrected results described by Equation 8. I show only the coefficients from the outcome regression, the selection coefficient estimates are identical to those shown in column (2) of Table 4. The first and second columns show respectively the effect of observables on the months of life gained from using the new drug relative to the control group as it is reported in the trial, and the months of life gained interacted with the statistical significance. The third column is the months of life gained from the new drug relative to the control group as a percentage of the overall survival in the control group. The fourth column is this percentage interacted with the statistical significance of the the overall survival difference. None of the covariates are statistically significant predictors of the outcome of the trial at the 5% level. At the 10% level the percentage gain in the first use and the number of approved drugs shared between the cancer of the first use and off-label uses are positive predictors of the percentage gain found in a trial for an off-label use. The coefficient on the inverse mills ratios are not statistically significant at the 5% level. However, they are large particularly for the percentage change outcome variables, one of which is statistically significant at the 10% level. They are positively signed, which means that the unobserved factors that make a trial more likely tend to be associated with higher gains in life expectancy in the trials.

Since the results are inconclusive on the question of whether or not firms are able to select trials that will have greater efficacy, I plot the outcome of the trials against the market size and the predicted probability of entry (based on the fitted values from the Probit model). This is shown in Figure 4. If there was selection, we would expect to see that as the market size and the number of approved drugs increases, the trial outcome decreases. Again, it is difficult to determine if this is the case.

Overall, I find that firms choose to invest in RCTs for off-label uses in markets where there are already more approved drugs, less usage of off-label drugs and where there is a stronger association between the market under consideration and the market of the drug’s first use. It does not appear that the outcome of the trial is predictable from the observables that I include in my analysis. This suggests that there is information to be gained from conducting an RCT for an off-label use of a drug. If, for example, the success in the first use was a strong predictor of success in additional uses, we might question whether or not the trial is producing new information or simply being conducted for advertising purposes. While I find that firms do choose to test off-label uses where they will be able to capture a greater market share with the trial than without, it is unclear whether or not the firms are able to select drugs on the basis of the efficacy revealed by the trial. An important caveat to this result is that the market size and competition variables that were used to identify the selection may not be exogenous.
Table 3: Probit Results: Probability of an RCT for a potential off-label use

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Incidence ('000s)</td>
<td>0.000541</td>
<td>-0.00174</td>
</tr>
<tr>
<td></td>
<td>(0.000745)</td>
<td>(0.000982)</td>
</tr>
<tr>
<td>Launch price (monthly treatment cost $2014)</td>
<td>-0.00242*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00104)</td>
<td></td>
</tr>
<tr>
<td>Patent remaining (years)</td>
<td>-0.0268***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00809)</td>
<td></td>
</tr>
<tr>
<td>Change in efficacy for first use (%)</td>
<td>0.00477***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00142)</td>
<td></td>
</tr>
<tr>
<td>One year survival rate (%)</td>
<td>-0.163</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.105)</td>
<td></td>
</tr>
<tr>
<td>Off-label prescriptions per patient</td>
<td>-0.0106*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00463)</td>
<td></td>
</tr>
<tr>
<td># Approved drugs</td>
<td>0.0145*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.00711)</td>
<td></td>
</tr>
<tr>
<td># Approvals shared with first use</td>
<td>-0.0191</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.130)</td>
<td></td>
</tr>
<tr>
<td>r2-p</td>
<td>0.0023</td>
<td>0.1193</td>
</tr>
<tr>
<td>N</td>
<td>208</td>
<td>208</td>
</tr>
</tbody>
</table>

Standard errors in parentheses
Coefficients are marginal effects, computed using the delta method.
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Table 4: Heckman Selection-Corrected Regression Results

<table>
<thead>
<tr>
<th></th>
<th>(1) # Months gained</th>
<th>(2) # Months gained(^{(a)})</th>
<th>(3) % Δ Months</th>
<th>(4) % Δ Months(^{(a)})</th>
</tr>
</thead>
<tbody>
<tr>
<td>Launch price</td>
<td>0.0101</td>
<td>0.0129</td>
<td>-0.0451</td>
<td>-0.00674</td>
</tr>
<tr>
<td>(monthly treatment cost $2014)</td>
<td>(0.0275)</td>
<td>(0.0266)</td>
<td>(0.163)</td>
<td>(0.133)</td>
</tr>
<tr>
<td>Patent remaining (years)</td>
<td>-0.307</td>
<td>-0.272</td>
<td>-0.932</td>
<td>-0.244</td>
</tr>
<tr>
<td></td>
<td>(0.271)</td>
<td>(0.261)</td>
<td>(1.603)</td>
<td>(1.309)</td>
</tr>
<tr>
<td>Change in efficacy for first use (%)</td>
<td>0.0702</td>
<td>0.0609</td>
<td>0.517(^+)</td>
<td>0.413(^+)</td>
</tr>
<tr>
<td></td>
<td>(0.0501)</td>
<td>(0.0484)</td>
<td>(0.292)</td>
<td>(0.242)</td>
</tr>
<tr>
<td># Approvals shared with first use</td>
<td>3.852</td>
<td>1.104</td>
<td>28.66(^+)</td>
<td>19.99</td>
</tr>
<tr>
<td></td>
<td>(2.836)</td>
<td>(2.734)</td>
<td>(16.80)</td>
<td>(13.69)</td>
</tr>
<tr>
<td>Constant</td>
<td>-6.944</td>
<td>-1.880</td>
<td>-81.03(^*)</td>
<td>-60.45(^+)</td>
</tr>
<tr>
<td></td>
<td>(6.701)</td>
<td>(6.440)</td>
<td>(40.55)</td>
<td>(32.35)</td>
</tr>
<tr>
<td>mills</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lambda</td>
<td>1.904</td>
<td>1.466</td>
<td>21.72(^+)</td>
<td>9.113</td>
</tr>
<tr>
<td></td>
<td>(2.147)</td>
<td>(2.066)</td>
<td>(12.89)</td>
<td>(10.37)</td>
</tr>
<tr>
<td>ρ</td>
<td>0.540</td>
<td>0.444</td>
<td>0.871</td>
<td>0.536</td>
</tr>
<tr>
<td>σ</td>
<td>3.525</td>
<td>3.300</td>
<td>24.93</td>
<td>17.00</td>
</tr>
<tr>
<td>N</td>
<td>208</td>
<td>208</td>
<td>208</td>
<td>208</td>
</tr>
</tbody>
</table>

Notes: (a) Months of life gained are set to zero if the change in overall survival reported in the trial was statistically insignificant.
Standard errors in parentheses
\(^+\) p < .1, \(^*\) p < .05, \(^{**}\) p < .01
Figure 4: Efficacy reported in RCT's for off-label drug uses by market size & trial probability
It is surprising that the market size itself is not a significant predictor of investment into trials, however, an important caveat is that it is difficult to measure the market size. I have used the incidence of the disease, which is the number of people newly diagnosed each year. This is a measure of the flow into the market. However, there is also a stock of people at any one time with the disease. In the official statistics that are collected on cancer, the stock of people, or the “prevalence” is measured based on anyone who has ever had the disease, including people who are in remission or have been cured and no longer require treatment. The market size that is relevant for this paper would be the stock of people with the disease who are currently being treated. No such measure exists.\footnote{It is not possible to measure the potential market size, at least amongst the elderly, in this data because not all drugs are covered by Medicare Part B.}

7 Welfare Implications

The requirement that firms produce evidence from randomized controlled trials is based on the premise that firms would otherwise under provide the information relative to what a social planner would choose. From a theoretical standpoint this isn’t necessarily the case. Clinical trials can be characterized as part of the fixed cost of a drug entering the market. It is a well-known result that in markets with a fixed cost of entry, free entry can lead to socially inefficient high or low entry (Mankiw and Whinston, 1986). In this context, under testing will occur when the firms’ private incentive to run a trial is less than the social benefit of the trial. From the consumer perspective, trials should be conducted when the expected consumer surplus exceeds the cost of the trials. However, the firm will only conduct trials where the expected revenue exceeds the fixed cost. The revenue and surplus will only be the same size if the firm can perfectly price discriminate. There will be under testing if the firms failure to capture the surplus is the dominating effect. Over-testing will occur when testing a drug leads to the firm capturing a rivals surplus (“business stealing”), rather than generating new value. Take the most extreme example, where a firm invests in demonstrating that a drug is exactly the same as an existing drug. Overall, the trial of the second drug is a waste of total surplus because there is an increase in cost but no change in overall surplus, though there is potentially a large transfer from firms to consumers due to subsequent competition.

In the case of off-label uses of drugs there is an additional distortion created because prices are not specific to the indication. Currently, drugs are priced by the unit and it is not possible to price discriminate on the basis of the usage of the drug (unless there is some difference in how the drug is packaged, for example, because it needs to be administered differently). The implication of this is that the price per unit of efficacy will vary within a drug across cancer types. Figure 5
Figure 5: Difference in months of life gained/cost between secondary and first use of drugs

Notes: Each dot is the difference between the months of life gained per $10,000 for a secondary use of the drug and the first use of the drug. Drugs without a tested secondary uses are excluded from the graph. Some drugs have multiple secondary uses on the same point.

shows the difference between the months of life gained per $10,000 in drug cost for the secondary and the first uses of each drug. The majority of secondary uses of drugs that have been tested in an RCT have less effectiveness per $10,000 spent than the first uses (equality would be if the secondary use is positioned at zero). However, the secondary uses are not necessarily less cost effective within the cancer type that they are being tested in. Figure 6 shows the months of life gained per $10,000 of expenditure on treatment for first and secondary uses by cancer type. Many secondary uses are tested in cancer types that have not had a new cancer drug approved between 1999-2010. For those that have, for example lung cancer, the effectiveness relative to the cost is on par with the the first uses.

The price relative to the efficacy affects the firm’s incentive to invest in an RCT for an off-label use. This can be illustrated using a simple model. The value to the firm of testing a new off-label use is the present value of the revenue generated from the trial: price (p) multiplied by the expected change in quantity from the trial (E(Δq)), relative to the cost of running the trial (F).

19 The months of life gained per $10,000 is the inverse of the cost per month of life gained, which I compute because many of the secondary uses of the drugs have zero effectiveness, making the cost per month of life gained infinity.
Figure 6: Months of life gained/cost, secondary and first use of drugs

Notes: Each dot is the number of months gain in life relative to the cost of a treatment episode (in $10,000). These are the actual months of life gained relative to the cost for first and secondary uses by disease. The previous figure shows the secondary uses relative to the first uses by drug.
The value to the consumer of testing the new product is the present value of the expected change in consumption of the drug multiplied by the value of this change to the consumer \((v)\) less the fixed cost of running the trial.

\[ p\mathbb{E}(\Delta q) - F \]  

(10)

Under the assumption that the number of years that the trial gain persists in the same for firms and consumers, and that the discount rate is the same, the inefficiency depends the relative magnitudes of the price and the value to the consumer of the increased efficacy.

The firm over-tests if:

\[ p > \frac{F}{\mathbb{E}(\Delta q)} > v \]

The firm under-tests if

\[ v > \frac{F}{\mathbb{E}(\Delta q)} > p \]

Figure 7 shows the cost relative to months of life gained. The orange line is the cost per year of life of $129,000, which is the standard value of a statistical life. It can be seen that around half of the first uses are cost effective but most of the second uses are not. That the price is so much higher than the effectiveness would suggest that if this is the true VSLY, on the margin the firm has an incentive to over invest in trials relative to the social optimum.

An alternative strategy to examine the the optimality of the firm’s investment decision is to compare the social value of the trials that have been conducted to an estimate of the cost. This is a comparison of the average benefit of a trial to the average cost. If the marginal cost of producing the drug is zero, the total surplus collapses to the consumer surplus, as the firms profit becomes a transfer of surplus between the consumer and the firm. When consumers do not face prices it is not feasible compute the consumer surplus using the area under the demand curve. Kwerel (1980) proposes that the value of information from clinical trials depends on how many consumers gain from trials when the trial changes their choice. The value of this gain depends on how much they would be willing to pay to have the information. To approximate this I take the estimate of the average change in quantity from the event study and apply the life years gained assuming that people gain the amount of life expectancy reported in the trial. This is converted into a dollar value using the value of a statistical life year (VSLY). The per period value of the trial is then summed over the life of the drug and discounted to the present value.

Equation 12 shows the formula for the consumer valuation of the trial and Table 5 shows the values used for each of the inputs. Since I separately estimate the average change in quantity
Figure 7: Cost per year of life month gained: first and secondary uses

Notes: The cost is the cost of a treatment episode. This is defined as the monthly cost of the drug multiplied by the average number of months patients were treated in the trial. The efficacy is the difference in overall survival between the treatment and the control group.
from a trial based on whether or not the trial result was positive, I rewrite the expected change in quantity multiplied by efficacy as the probability weighted sum of these two effects (Equation 13). The event study finds that on average the change in quantity for a non-positive trial ($\Delta q^{np}$) is zero. Hence the second term in Equation 13 drops out. The remaining effect is scaled by a factor $S$ to account for the sample’s share of the total cancer market.

$$v = \sum_{t=\text{trial}+1}^{T} \beta^t (E(\Delta q \times e) \times VSLY)$$

where

$$E(\Delta q \times e) = (P(pos)\Delta q^{pos}e^{pos} + (1 - P(pos))\Delta q^{np}e^{np}) \times S$$

$$= (P(pos)\Delta q^{pos}e^{pos}) \times S$$

Table 5: Social value of a trial

<table>
<thead>
<tr>
<th>Variable</th>
<th>Value</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\beta$</td>
<td>0.97</td>
<td>Standard discount rate in cost effectiveness analysis: see Sanders et al. (2016)</td>
</tr>
<tr>
<td>$T$</td>
<td>11</td>
<td>Average time remaining on patent for drugs released 1996-2010.</td>
</tr>
<tr>
<td>$P(pos)$</td>
<td>0.32</td>
<td>Share of positive trials in the data.</td>
</tr>
<tr>
<td>$E(\Delta q^{pos})$</td>
<td>580</td>
<td>Estimated in the event study in Section 4.</td>
</tr>
<tr>
<td>$E(e)$</td>
<td>2.84</td>
<td>Mean efficacy change reported in the positive trials that were included in event study.</td>
</tr>
<tr>
<td>$VSLY$</td>
<td>$173,000$</td>
<td>Dialysis benchmark from Lee, Chertow and Zenios (2009) inflated from $2003 to $2017 using the CPI</td>
</tr>
<tr>
<td>$S$</td>
<td>0.18</td>
<td>Estimate of the share that the sample represents. See Section 3.1 for discussion.</td>
</tr>
</tbody>
</table>

The parameters in the calculation are set as follows: the discount rate is 3% (following Sanders et al. (2016)); $T$ is measured as the average number of years remaining on the drug patent from the date of the release of the trial; the probability of a positive trial $P(pos)$ is 32%, which is the share of positive trials in the event study, the expected change in quantity for a positive trial
result per year is set to the average change found in the event study; the expected efficacy gain from a positive trial is set to the mean in the data (2.84 months); the VSLY used is the dialysis benchmark, computed by Lee, Chertow and Zenios (2009) inflated from $2003 to $2017 using the CPI; and I assume that my data is an 18% share of the US cancer market. Based on these parameters, each year after the trial consumers gain $42.9 million from increased life expectancy. If this effect persists until the end of the patent period the payoff to consumers is on average $390.9 million dollars.

Table 6: Cost of RCTs

<table>
<thead>
<tr>
<th>Source</th>
<th>Cost estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sertkaya et al. (2016)</td>
<td>$22 million</td>
</tr>
<tr>
<td>DiMasi, Grabowski and Hansen (2016)</td>
<td>$130.1 million</td>
</tr>
<tr>
<td>Prasad and Mailankody (2017)</td>
<td>$129.6 million</td>
</tr>
</tbody>
</table>

There are a range of cost estimates available in the literature. I consider the two most prominent as well as a recent third estimate that is specifically about newer cancer drugs. Table 6 shows the estimates of the cost of a phase 3 trial in cancer based on these research. All three suggest that the firm is on average under investing in trials.

There are two important caveats to this analysis, in addition to the assumptions already mentioned. First, the computation has valued the trial based on the average outcome, allowing the effect of the trial result to vary only by whether or not it is positive. This is done because the quantity change is the average across the drugs that have been tested. The actual value of information depends on the absolute value of the change in demand. If the variance in the change in demand is larger, the gain to consumers will be larger. However, in the computation used here, there is no role for the variance. Second, the welfare analysis has not been adjust for selection. The average gain from the trials that are conducted is compared with a measure of costs. If there is selection, the cost and benefit estimates would not be directly comparable, as the cost estimates are not specifically for the off-label trials included in this analysis.

8 Conclusion: Lessons from off-label prescribing

This paper is the first to show that RCTs have a causal effect on demand for new cancer drugs. I find that the release of an RCT showing a statistically significant increase in overall survival on

---

20 I use the coefficient estimates for the first 4 quarters. I then take the average over the remaining quarters and use that as the per quarter gain.

21 See discussion in Section 3.1.
average increases demand for the drug for that cancer type by 150%. A non-positive result does not affect demand. I do not find evidence that firms are able to select the trials on the basis of efficacy. However, I also do not find clear evidence that firms select the off-label uses to trial based on profitability, suggesting that the results are better interpreted as inconclusive. The welfare conclusions are mixed. I find that on average the trials that firms are conducting have a greater social value than the likely cost, however, on the margin firms have an incentive to invest in trials that are not socially valuable.

The off-label setting differs from first uses in several important ways, however, there is still much that can be learned about drug regulation more broadly. First, and most importantly, it appears that prices are set at the time of the launch of the first use, which is often before RCTs for additional uses have been concluded (or even begun). A change in the evidentiary requirement for entry would likely affect the initial pricing strategy. The anticipated effect on prices from reducing the evidentiary requirement would be downwards. Lower prices are likely (if anything) to lead to less prescribing due to physician agency because patients are typically fully insured. The second difference is that off-label uses of drugs arguably have information on safety and side effects from the RCTs of the first use. If the evidentiary requirements were relaxed there might be drugs available that have less information on safety and side effects. However, there are policy mechanisms that could address this including, for example, post market registration of adverse events or setting a standard for safety evidence. Moreover, in the case of cancer and other terminal illness with short life expectancy precise comparisons of safety and adverse events between treatment choices may be of less interest than in other types of disease. Finally, on-label uses can be advertised while off-label cannot, which may provide an important incentive for firms to invest in RCTs. As long as advertising continued to reflect evidence, this difference should not have a material effect on the interpretation of the findings from this paper.

In favor of reduced entry restrictions on new drugs, I show that evidence has an important role in demand for drugs. The majority of prescribing is for uses of drugs that have been tested, with untested drugs being a small proportion of the market. This suggests that new untested drugs would not be widely adopted. An important caveat to this is that I only consider a specific type of off-label prescribing. I also find that firms have an incentive to run trials voluntarily, possible exceeding the socially optimum, because demand increases in response to RCTs that show an improvement in overall survival and because prices are so high. Against reduced entry restriction, or in favor of greater regulation around off-label prescribing, I find that demand for drugs is persistent in the face of evidence that could be considered negative. I find that doctors do not respond to a failure to find an increase in overall survival as if it were a negative finding and demand persists among drugs with a trial even if the trial result was not favorable. The policy implication of this is unclear. On the one hand, spending occurs on drugs that would not have
been available if an entry rule were in place because they have failed to demonstrate efficacy. On the other hand, allowing the drug to be prescribed provides an increased incentive for firms to run the trial in the first place.

References


Boutron, Isabelle, Susan Dutton, Philippe Ravaud, and Douglas G Altman. 2010. “Reporting and Interpretation of Randomized Controlled Trials with Statistically Nonsignificant Results for Primary Outcomes.” *JAMA*, 303(20): 2058–64.


**Food & Drug Administration.** 2006. “Promoting Safe and Effective Drugs for 100 Years.” Available at [https://www.fda.gov/AboutFDA/WhatWeDo/History/ProductRegulation/PromotingSafeandEffectiveDrugsfor100Years/](https://www.fda.gov/AboutFDA/WhatWeDo/History/ProductRegulation/PromotingSafeandEffectiveDrugsfor100Years/). Last Access 10/10/2017.


To determine whether or not prices paid by Medicare rise following the release of an RCT for an off-label use of the drug, I redo the event study shown in Equation 4 with real prices as the outcome variable. Prices are computed as the median reimbursement price of the drug per unit per quarter. They are converted to 2013 dollars using the CPI.

The results are shown in Figure 8. The top panel shows the effect of the release of a non-positive RCT on prices. There does not appear to be any effect. The second panel shows the effect of the release of a positive result from an RCT, again there does not appear to be any effect. The bottom panel shows the difference in the response between a positive and negative result, again there is no response.

The lack of price response could be explained by two different things. First, it could be that the off-label use of the drug is a small proportion of the drugs overall market. Since it is not currently

A Prices
Figure 8: Event study result: Real Prices
possible to price drugs based on indication, the firm must price the drug taking into account its usage in each market. Off-label uses tend to be a small proportion of a drug’s sales for new drugs (see Figure 9). Second, although the prices in the Medicare data reflect Medicare’s expenditure on the drug, they may not be representative of the prices the firm receives from the rest of the market. Most cancer drugs covered under Part B use a reimbursement or buy-and-bill model, meaning that providers purchase the drug and then claim reimbursement from Medicare. Since 2005, Medicare has reimbursed physicians at the Average Sales Price (ASP) plus 6%. Prior to this the reimbursement rate was 95% of the drug’s average wholesale price (AWP). The average sale price is the net sales of a drug by a manufacturer in a calendar quarter divided by the number of units sold. The ASP is net of any price concessions, such as volume discounts, prompt pay discounts, cash discounts, free goods contingent on purchase requirements, chargebacks, and rebates other than those obtained through the Medicaid drug rebate program. Also, only manufacturers participating in the Medicaid rebate program are required to report this information to Medicare. The average wholesale price is supposed to be the average price at which drugs are purchased at the wholesale level. However, it is a list price and does not include discounts, it is also not based on a formalized reporting system and can be manipulated.
B Effect of No Negative Results

This section shows the implication of trials rarely showing a negative result (as shown in Section 2) or being interpreted as being negative (as seen in the event study, doctors do not respond to non-positive trials as if this is a negative finding). I propose a simple model of the firm investment decision where the firm invests in a trial if the net present value of the revenue stream \( pE(\Delta q) \) from having the trial, versus not having the trial, is greater than the cost of the trial \( F \).

\[
pE(\Delta q) \geq F \tag{15}
\]

Equation 15 can be rewritten in terms of the change in revenue from a successful trial (a statistically significant positive result) and a non-positive trial (a statistically insignificant change in overall survival).

\[
p[\text{prob}(success)\Delta q^{\text{success}} + (1 - \text{prob}(success))\Delta q^{\text{fail}}] \geq F
\]

The results of the event study in Section 4 show that \( \Delta q^{\text{fail}} = 0 \), which means that the firms invest if:

\[
p[\text{prob}(success) \ast \Delta q^{\text{success}}] \geq F
\]

If the doctors had responded to the non-positive trial results as being negative, or if there were negative trial results then \( \Delta q^{\text{fail}} < 0 \). This means that the lack of negative trial results, or the lack of response to negative trial results, provides firms with an additional incentive to invest in trials.