NBER WORKING PAPER SERIES

EFFECTS OF EXPANDING HEALTH SCREENING ON TREATMENT - WHAT SHOULD WE EXPECT? WHAT CAN WE LEARN?

Rebecca Mary Myerson
Darius Lakdawalla
Lisandro D. Colantonio
Monika Safford
David Meltzer

Working Paper 24347
http://www.nber.org/papers/w24347

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

We are grateful for valuable suggestions from Daniel Bennett, Marguerite Burns, Lenisa Chang, James Heckman, Richard Hirth, Bryant Kim, and Edward Norton, as well as participants of the NBER summer institute, ASHEcon, and APPAM conferences. This work was supported by the National Institute of Neurological Disorders and Stroke [grant number U01 NS041588], the Agency for Healthcare Research and Quality [grant number R36HS023964-01], and the National Heart, Lung, and Blood Institute [grant numbers R01 HL080477 and K24 HL111154]. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed a financial relationship of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w24347.ack

Effects of Expanding Health Screening on Treatment - What Should We Expect? What Can We Learn?

Rebecca Mary Myerson, Darius Lakdawalla, Lisandro D. Colantonio, Monika Safford, and David Meltzer

NBER Working Paper No. 24347
February 2018
JEL No. D0,D8,I1

ABSTRACT

Screening interventions can produce very different treatment and health outcomes, depending on the reasons why patients went unscreened in the first place. Economists have paid scant attention to these complexities and their implications for evaluating screening programs. In this paper, we propose a simple economic framework to guide policy-makers and analysts in designing and evaluating the impact of screening on treatment uptake. We apply these insights to several salient empirical examples that illustrate the different kinds of effects screening programs might produce. Our empirical examples focus on contexts relevant to the top two causes of death in the United States, heart disease and cancer, and match three predictions from the framework. First, currently unscreened patients differ from currently screened patients in important ways, leading to lower predicted uptake of recommended treatment if these patients were diagnosed. Second, there are diminishing clinical returns to screening, which can be reversed if patients with low access to care are targeted with a bundled intervention. Third, changes in the composition of diagnosed patients can produce misleading conclusions during policy analysis, such as spurious reductions in measured health system performance as screening expands.

Rebecca Mary Myerson
University of Southern California
635 Downey Way, VPD Office 414M
Schaeffer Center for Health Policy
and Economics
Los Angeles, CA 90089
rmyerson@usc.edu

Darius Lakdawalla
University of Southern California
635 Downey Way, VPD 414-K
Schaeffer Center for Health Policy
and Economics
Los Angeles, CA 90089-7273
and NBER
dlakdawa@usc.edu

Lisandro D. Colantonio
Department of Epidemiology,
School of Public Health
Lister Hill Library, Suite 446,
1700 University Bl.
Birmingham, AL 35233
lcolantonio@uab.edu

Monika Safford
Joan and Sanford I. Weill Department
of Medicine
Weill Cornell Medical College 1300
York Ave
New York, NY 10065
mms9024@med.cornell.edu

David Meltzer
Section of Hospital Medicine
University of Chicago
5841 S. Maryland, MC 5000
Chicago, IL 60637
and NBER
dmeltzer@medicine.bsd.uchicago.edu

1 Introduction

Many people – particularly those in vulnerable groups – suffer from undiagnosed conditions that result in missed opportunities to improve health. Diabetes, high cholesterol, hypertension, and cancer are chief contributors to avoidable premature mortality in the United States: despite available treatment, many patients with these conditions are undiagnosed or untreated (Cowie et al., 2009; Global Burden of Disease Collaborators, 2013; McDonald et al., 2009; Olives et al., 2013; Patel et al., 2015; Zweifler et al., 2011). Increasing access to screening for these chronic conditions has been a focus of both public-sector and private-sector efforts in recent years.[1]

Improving access to screening will at least weakly increase treatment of chronic conditions, because detection leads to treatment. However, the magnitude of this effect varies, because additional screening might diagnose patients with lower uptake of medical treatment after diagnosis. These gaps in treatment would shape the total costs and health benefits of policies and programs that subsidize screening, and affect physician practice after access to screening is expanded.

There are two main reasons that lower-cost screenings could diagnose patients with lower treatment rates after diagnosis. First, there is a clinical channel: additional screening may "over-diagnose" and result in treatment that does little to improve long-term health outcomes. For instance, epidemiologists have long observed that increases in cancer screening may fail to improve health outcomes at all, if it instead picks up small, slow-growing tumors that would never have harmed a patient's health (Ahn et al., 2014; Bleyer and Welch, 2012; Loeb et al., 2014). Thus, screening programs may be reaching patients for whom the benefit of treatment is lower than average or even zero.

Second, there is an economic channel. Patients who went unscreened may have done so because they perceive higher barriers to medical treatment.[2] These barriers to care could include out-of-pocket costs, or non-pecuniary costs such as distance to a physician, language barriers, or psychological costs (Carpenter, 2010; Hyman et al., 1994; Lange, 2011; Kenkel, 1994; Manning et al., 1987; Musa et al., 2009). These same barriers could then translate to lower treatment rates for conditions detected after lowering screening costs, despite clinical

[1]For example, Medicare added a free "Welcome to Medicare" visit for new enrollees in which screening needs are discussed and addressed, and the Affordable Care Act required health insurance plans to offer preventive care, including free screening for diabetes, high cholesterol, hypertension and cancer to people at high risk. In the private sector, pharmacy chains such as CVS, Walgreens, and stores such as Ralph's and Sam's Club now offer screening for diabetes, high cholesterol, and hypertension in convenient retail locations.

[2]For evidence on screening, see Hyman et al. (1994); Lostao et al. (2001); Oster et al. (2013); Wilson (2011). For a discussion of self-selection into treatment and related econometric approaches, see, e.g., Carneiro et al. (2010); Eisenhauer et al. (2010); Heckman (2010).

benefit.

These examples suggest a broader lesson: screening interventions can produce very different treatment and health outcomes, depending on the reasons why patients went unscreened in the first place. Economists have not studied the implications for evaluating screening interventions. To fill this gap, we propose a simple economic framework to help illustrate these points and to guide policy-makers and analysts in designing and evaluating the impact of screening interventions on uptake of relevant treatment. We then apply these insights to several salient empirical examples that illustrate the different kinds of effects screening programs might produce.

The paper begins by presenting an economic framework of screening and treatment. Our framework is very parsimonious and requires two key assumptions: demand for screening is downward sloping, and screening provides the option of accessing treatment. As screening expands, newly diagnosed patients are expected to have higher ex ante net cost of treatment, i.e., higher cost or lower benefit to treatment, than previously diagnosed patients. Accordingly, these patients are predicted to have lower treatment uptake. Yet, this low treatment uptake could mask potentially high potential health benefits to treatment in certain cases. If many sick patients were unscreened due to high barriers to care, then the benefits of screening and treating these patients could be high. In this case, a bundled screening and access to care intervention - targeted to patients with low access to care - could have significant health impact. In contrast, if access to care was already high so that patients remained unscreened chiefly due to low predicted benefit, diminishing returns to screening would be unavoidable.

Our empirical exercises are tightly related to this theoretical framework, and focus on the top two causes of death in the United States: heart disease and cancer. We first focus on heart disease risk factors which are commonly undiagnosed and untreated: diabetes, hypertension, and high cholesterol. Data on these disease risk factors are used to assess our prediction that as screening expands, the additional diagnosed patients are expected to have higher cost or lower benefit to treatment than previously diagnosed patients. Using data from the National Health and Nutrition Examination Survey (NHANES), we show that people who were not recently screened for undiagnosed diabetes or high cholesterol tend to show larger barriers to treatment and/or appear ex ante to be healthier, with lower overall heart disease risk. Subsequently, these same factors are associated with lower propensity to treat conditions after diagnosis. In a simulation analysis, we find that these factors could be responsible for about a one-half to one percentage-point decline in treatment of diagnosed conditions for every 10 percentage point increases in screening.

This analysis provides support for our hypothesis that currently unscreened patients differ from currently screened patients in important ways, which predict lower uptake of

recommended treatment if these patients were diagnosed. Analyses of the potential treatment uptake of currently undiagnosed patients using other datasets find supportive results. We first study participants in the REasons for Geographic and Racial Differences in Stroke study (hereafter, REGARDS). The REGARDS study randomly contacted older adults across the continental United States, conducted biomarker assessments in the participants' homes, compensated participants for their time, and informed participants of their biomarker results (Howard et al., 2005).

Using merged individual-level Medicare claims for the REGARDS study participants, we find that conditions diagnosed as part of the biomarker study are less likely than previously diagnosed conditions to receive annual doctor visits for evaluation and management. Additional analyses using the Oregon Health Insurance Experiment and NHANES use different methods but find supportive results: time periods and patients with lower screening uptake show lower treatment rates after diagnosis as well as poorer biomarker control after diagnosis (Centers for Disease Control and Prevention, 2014; Finkelstein, 2013). Together, our findings indicate that screening currently unscreened patients could increase the fraction of diagnosed patients who don't receive relevant doctor visits for their condition, don't use recommended treatment, and possibly remain uncontrolled.

Although these findings seem to indicate diminishing returns to screening, we demonstrate that this need not be the case with a targeted, bundled intervention. Targeting is important because screening patients with high access to care could produce diminishing health effects. We hypothesize that screening patients with low access to care could yield significant health benefits if these patients' barriers to care are simultaneously addressed. These hypotheses are supported by data from the largest cancer registry in the United States. We exploit an exogenous increase in cancer detection and treatment as people age into Medicare, which not only provides access to screening but also provides access to care. Results indicate that health benefits arise only for racial and ethnic minorities, a group that previously faced higher barriers to care. In contrast, non-minority patients showed no significant improvements in post-diagnosis survival. This evidence is consistent with our prediction that diminishing health returns to screening could be reversed if patients with barriers to care are targeted and their barriers to care are addressed.

Our final empirical analyses demonstrate the importance of these findings for policy analysis and health system performance evaluation. First, our findings imply that changes in patient composition could mask the benefits of expanding access to screening, as captured by commonly used measures of health system performance. This problem arises because the true prevalence of conditions is not observed, whereas diagnosis status is observed. As a result, commonly used health system performance metrics focus on treatment and control

of conditions that are diagnosed (Center for Medicare and Medicaid Services, 2011, 2016a; National Committee on Quality Assurance, 2016; Song et al., 2011, 2014). However, use of these metrics can produce misleading conclusions - for example, that completeness of care for chronic conditions declines rather than improves as more patients become diagnosed. We demonstrate this possibility using the REGARDS data. Similarly, national Medicare data aggregated by hospital referral region show that hospital referral regions with higher diagnostic intensity, as calculated by Finkelstein et al. (2017), show lower use of maintenance care such as eye exams and hba1c checks for patients with diabetes as calculated in the Dartmouth atlas of health care quality (The Center for the Evaluative Clinical Sciences, Dartmouth Medical School, 2017; Fisher et al., 2008).

This paper provides several novel insights. First, based on the patterns we uncovered, expanded screening as a stand-alone program is likely to be less cost-effective than previously anticipated due to low treatment uptake among marginally screened patients. To our knowledge, these effects are not currently accounted for in cost-effectiveness analyses that simulate the impact of screening expansions (The CDC Diabetes Cost-Effectiveness Study Group, 1998; Glümer et al., 2006; Hoerger et al., 2004; Kahn et al., 2010; Nathan and Herman, 2004; Wang et al., 2011). Accounting for these effects could change the coverage policies selected in health systems that make decisions based on cost-effectiveness analysis.

Second, screening program design must account for the reasons why patients are unscreened. Numerous screening expansions have shown lower effects than anticipated, including programs in national health systems and screening programs for underinsured women (Ahn et al., 2014; Kim et al., 2017; Lantz et al., 1997). By showing how potential health effects are linked with the reasons why patients went unscreened in the first place, our framework clarifies when low health effects are avoidable and how they can be avoided.

Third, our research contributes to the small but growing literature on the unintended effects of quality reporting (Casalino et al., 2007; Dranove et al., 2003; Harris et al., 2016; Karve et al., 2008). In multiple pay-for-performance systems such as Accountable Care Organizations, providers have financial incentives to maintain high treatment rates for diagnosed conditions as well as high screening rates (Center for Medicare and Medicaid Services, 2011, 2016a; National Committee on Quality Assurance, 2016; Song et al., 2011, 2014). However, our research suggests that expanding access to screening could carry a penalty by reducing other, treatment-related quality metrics. This would suggest reconsideration or reweighting of the metrics used in pay-for-performance systems, to avoid penalizing health systems that expand screening in diverse patient populations.

Finally, in addition to providing a venue to test our theoretical predictions, our empirical analysis on cancer expands the literature on the health effects of Medicare (Card et al., 2008;

McWilliams et al., 2009). Little was previously known about the impacts of Medicare on cancer diagnosis and survival. Our findings related to timely cancer detection and post-diagnosis survival are important for population health because post-diagnosis survival is a commonly used quality metric for cancer care and racial-ethnic disparities in this metric are substantial (Du et al., 2007; Jatoi et al., 2003; Ward et al., 2004).

The paper proceeds as follows. Section 2 compares this study with previous literature and articulates our contributions. Section 3 presents our theoretical framework. Section 4 presents our main empirical analysis. Section 5 demonstrates the implications of our findings for policy analysis and health system performance measurement. Section 6 concludes.

2 Comparison with the literature

Anticipated costs and benefits of health care can differ across individuals, influencing individuals' willingness to seek care (Egan and Philipson, 2014; Eisenhauer et al., 2010; Heckman, 2010). This premise underlies commonly used public health models such as the health belief model.[3] It follows that anticipated net benefits of particular health services can vary across individuals (Vanness and Mullahy, 2012). In certain cases, distributions of these individual-level net benefits can be estimated (Basu and Heckman, 2007; Carneiro et al., 2010; Eisenhauer et al., 2010). These distributions are useful because changes to out-of-pocket costs of health care will attract different patients to treatment, depending on their anticipated cost and benefit (Basu and Meltzer, 2007; Goldman and Philipson, 2007; Pauly and Blavin, 2008).

A number of recent papers use new econometric methods to estimate distributions of net benefits of specific health services. These papers typically focus on how patients choose between treatments for their conditions (i.e., the intensive margin) (Basu and Heckman, 2007; Basu and Manning, 2009; Basu, 2011, 2013; Huang et al., 2006; Meltzer and Huang, 2007; Sculpher, 2008). In contrast, our theoretical model considers how the anticipated net benefits of screening are distributed in the population. These net benefits govern which conditions are *not* treated (i.e., the extensive margin). We present a simple and intuitive framework that describes cost, benefit, and the demand for care. It would be straightforward to extend these insights to dynamic models of health investment, e.g., the Grossman (1972) health capital model, as we show in Appendix B.[4]

[3]See Glanz and Bishop (2010) for a review of commonly used health behavior models in the public health field. The health belief model includes perceived benefits and perceived barriers as a key construct, and these are the constructs that are most strongly predictive of behavior in empirical tests (Rosenstock et al., 1988; Carpenter, 2010).

[4]To this point, we also offer a comparison of the model in the appendix with prior versions of the Grossman

Our study can also be situated in the literature on screening and the demand for information. In the economics literature, the demand for screening has been empirically related to patient perceptions of disease risk and treatment effectiveness. Demand for screening is low if no effective treatment yet exists, and becomes higher once treatment is available (Oster et al., 2013; Wilson, 2011). Research examining conditions with available treatment has found that people who know about their elevated risk for a condition have higher demand for screening (Lange, 2011). In the public health literature, research on the health behavior model shows that participants who anticipate higher risk of the condition are more motivated to take action, whereas those who face logistical barriers are less likely to take action; these same variables are found to be predictive of screening (Carpenter, 2010; Hyman et al., 1994; Lostao et al., 2001).[5] Research on cancer screening in the medical and public policy literature has shown that screening expansions attract patients with less severe conditions, resulting in concern about overdiagnosis (Ahn et al., 2014; Kadiyala and Strumpf, 2016; Loeb et al., 2014). Yet at the same time, patients not reached at all by screening interventions are sicker (Kim and Lee, 2017). This underscores the point that factors other than clinical need can play an important role in determining who is screened. In summary, the literature has shown that costs or benefits of treatment empirically predict patients' uptake of screening.

This self-selection process implies that the composition of diagnosed patients should change as access to screening changes. Our paper pushes this literature forward by providing a framework to synthesize points about screening and preventive care made across disparate fields, and by showing how these determinants of screening can interact to determine the impact of policy changes. We also contribute by demonstrating the implications of patient composition effects after screening expansions for policy analysis and health care quality measurement.

Accordingly, our paper also contributes to the literature on health care quality measurement. As strategies to improve population health and promote health equity, the success of public reporting and pay-for-performance programs hinges on selection of appropriate metrics. Previous research has shown that some metrics used in existing public reporting schemes create incentives to select certain types of patients for care, because providers' scores

(1972). Broadly, in the Grossman (1972) model, agents make decisions about how much time and money to invest in health to maximize their utility given practical constraints. The original model had no uncertainty: agents had perfect knowledge about their health and about the health production process. Previous research has incorporated uncertainty about how health investments translate to future health and productivity into the model using random shocks (Liljas, 1998; Grossman, 1982, 2000). Many of these papers, such as those of Chang (1996), Dardanoni and Wagstaff (1987) and Selden (1993), focus on health investment motivated by labor market returns. We allow agents' source of uncertainty about their health to be a lack of screening rather than exogenous shocks, as in Oster et al. (2013); Boozer and Philipson (2000) and others.

[5]Stigma related to testing can also play a role, although this is less relevant in the disease contexts from which we draw our empirical examples (Godlonton and Thornton, 2012).

decrease if they treat vulnerable or sick patients (Dranove et al., 2003; Harris et al., 2016; Konetzka et al., 2013). These findings have raised concerns that public reporting could create a less inclusive health system depending on the metrics chosen (Casalino et al., 2007; Karve et al., 2008).

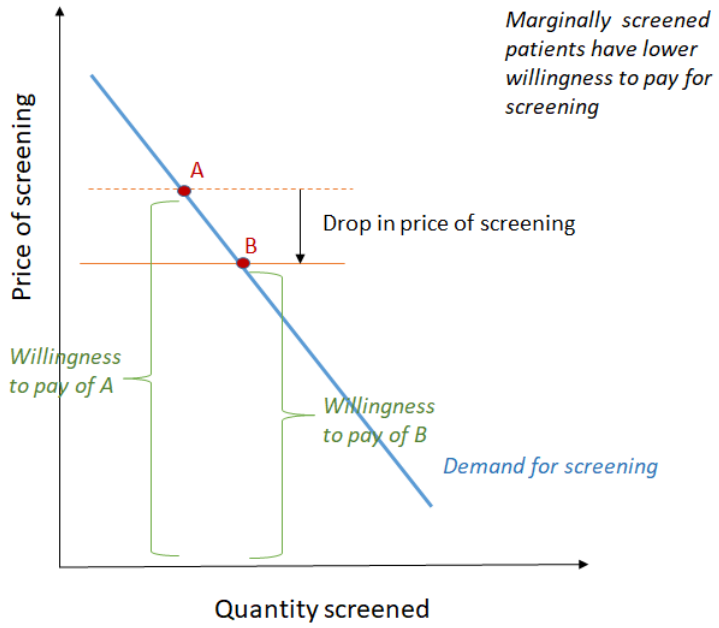
Our study contributes to this literature by generalizing previous findings for the case of screening. We find that expanding the set of diagnosed patients makes a health system more inclusive but carries a "quality penalty," in the form of decreased treatment rates for diagnosed conditions. This is important because treatment rates for diagnosed conditions are commonly used as health care quality metrics (CDC, 2012; Center for Medicare and Medicaid Services, 2011; Dale et al., 2016; National Committee on Quality Assurance, 2016; Agency for Healthcare Research and Quality, 2013).

Finally, in addition to providing a venue to test our theory, our empirical analysis on cancer contributes to the literature on the health effects of Medicare. Previous studies have exploited the fact that Medicare eligibility abruptly changes at age 65 to explore the effects of the program. Card et al. (2008) use a regression-discontinuity framework to analyze survey data from the 1999-2003 National Health Interview Survey (NHIS), and find that reaching age 65 is associated with an increase in overall insurance coverage. Although Card et al. (2008) find that increases in the use of medical care services vary across groups and the type of service, it is most relevant to our analysis that routine doctor visits increase more for racial and ethnic minority groups. This matches our finding that early cancer detection increases disproportionately for racial and ethnic minority patients. Additionally, our finding that racial and ethnic disparities in post-diagnosis cancer mortality decline at age 65 builds upon the findings of McWilliams et al. (2009), who find that disparities in systolic blood pressure, hemoglobin A1c levels and total cholesterol levels decline upon aging into Medicare using 1999-2006 NHANES data. Our findings related to timely cancer detection and post-diagnosis survival are important for population health because post-diagnosis survival is a commonly used quality metric for cancer care and racial-ethnic disparities in this metric are substantial (Du et al., 2007; Jatoi et al., 2003; Ward et al., 2004).

3 Theoretical model

In this section, we present a simple economic framework of self-selection into screening to inform policy analysis and policy design. We employ two key assumptions: demand for screening is downward sloping, and screening provides the option of accessing medical treatment. This framework is then used to discuss the different kinds of effects screening programs might produce. Although our treatment of the theoretical model is intuitive to

Figure 1: Screening demand curve



keep assumptions to a minimum, similar points could be made using a formal model; see Appendix B.

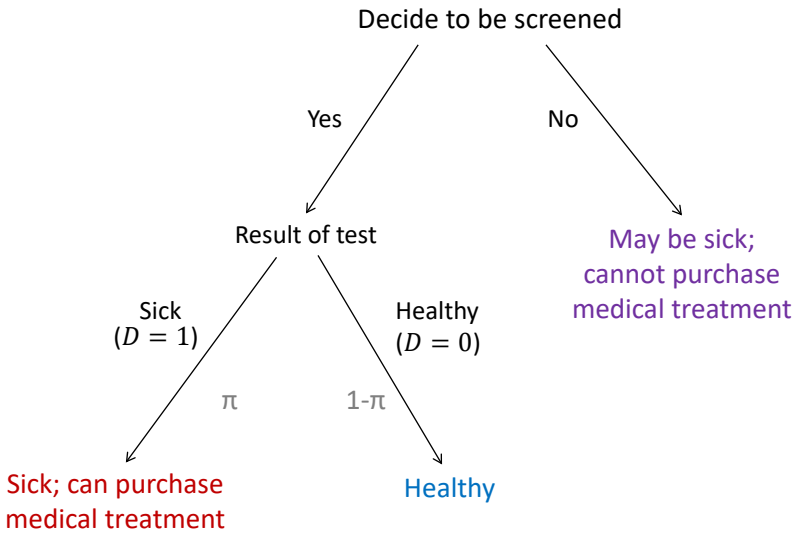
3.1 Which patients are screened?

Suppose the value of screening differs across patients. In some cases, this variation occurs for reasons that are unobservable to patients or their physicians. For instance, some patients might be genetically predisposed to developing complications from a disease if untreated while others might be relatively immune to complications. Variation that is unobservable to the patient is less problematic for policy analysis because it will not systematically affect a patient's willingness to be screened.

Greater complexity arises when this variation is observable to the patient or her physician. In this case, unscreened patients may anticipate lower value from screening. In economic terms this is a downward sloping demand curve, as in Figure 1. The vertical axis here indicates the price of screening and horizontal axis denotes quantity of people screened. Patients above the horizontal price line are screened, and patients below are not screened. This setup implies that at high prices of screening, few patients are screened but those that are screened have high willingness to pay due to high anticipated value. At lower prices of screening, many more patients are screened, including those who anticipate low value.

Naturally, not all patients face the same price of screening. Some patients face high out-of-pocket costs, live farther from a facility, face language barriers, or live in an area with

Figure 2: Screening decision tree



low health care supply. These facts are not at odds with a downward sloping demand curve. Higher costs of screening could be reflected by allowing different patients to face different supply curves, as we will show below.

3.2 Where do patients sit on the demand curve?

We next consider the determinants of demand for screening. Why might some patients (in consultation with their doctors) anticipate more value from screening than other patients? This topic has been analyzed in the public health and economics literature (Boozer and Philipson, 2000; Hyman et al., 1994; Lostao et al., 2001; Oster et al., 2013; Wilson, 2011). We provide a brief discussion here with the aid of a diagram, Figure 2.

Screening is valuable to patients as an input to improved health and well-being (Glanz and Bishop, 2010; Carpenter, 2010). Figure 2 depicts one way screening can improve health, namely, by providing patients with the option to pursue medical treatments that are only available after diagnosis. (Hereafter, we will use the term "prescription-only medical treatment" or just "treatment.") This is represented by the left-most branch of Figure 2. Ex post, this option is not used by patients with a negative screening result, as shown in the middle branch of Figure 2. Patients who have not been recently screened do not know their disease state and cannot access treatment, as shown in the right branch of Figure 2. The

key take-away from this figure is that the value of screening is entwined with the value of treatment, because screening is a gateway to treatment.

Which patients are likely to value the option to pursue treatment? Two general predictions are proposed. First, patients who know they would have difficulties accessing or affording treatment – including pecuniary costs as well as non-pecuniary costs such as language barriers or distance to a facility – should find the option to pursue treatment less valuable, all else equal. Second, patients who think they are unlikely to benefit from treatment – either because they are unlikely to have the condition, or because the treatment is unlikely to work – should find screening less valuable, all else equal (Kim and Lee, 2017). Accordingly, demand for screening is low for conditions which lack effective treatments, including HIV before the development of highly active anti-retroviral therapy and Huntington's disease (Oster et al., 2013; Wilson, 2011). In summary, the option to be treated is less valuable to patients who anticipate high pecuniary or non-pecuniary costs of treatment and/or little clinical benefit.

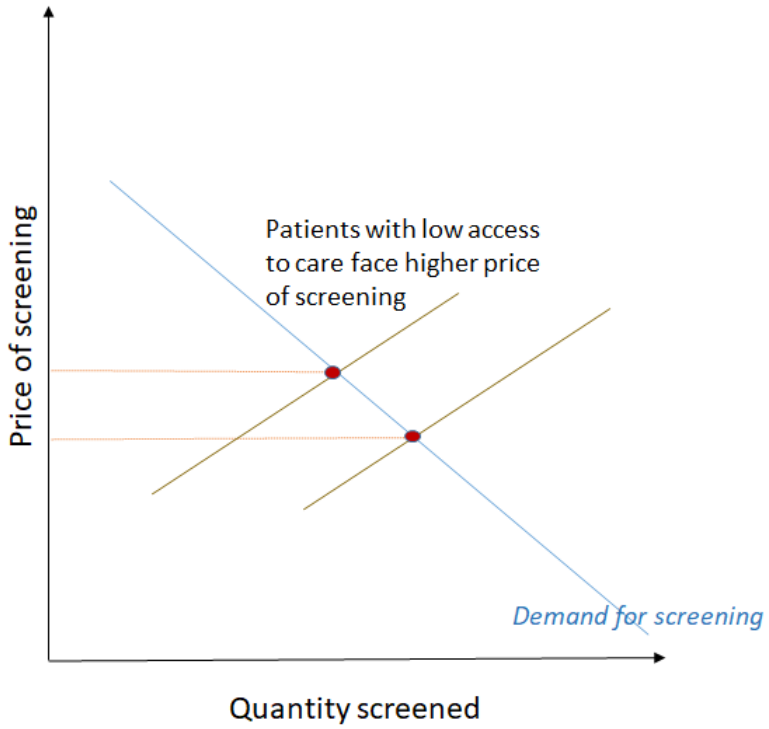
3.3 What outcomes can we expect after screening expands?

By the logic above, a patient's expected costs and benefits of treatment help to determine where that patient sits on the screening demand curve. People facing higher costs of treatment will not only have lower uptake of treatment if diagnosed, but also have lower net expected value from screening. As a result, these patients may be among the last to be screened. Similar logic applies for patients with lower anticipated benefit of treatment.

For these reasons we expect marginally screened patients, as screening expands, to have lower uptake of treatment after diagnosis. This point can be demonstrated by revisiting Figure 1. Suppose that the price of screening starts very high and is slowly lowered due to a national screening program, screening mandates for insurance plans, or other interventions. When the price of screening is high, patients who anticipate high clinical benefits and low treatment costs may be willing to pay. As the price of screening declines, more patients are screened, including those who anticipate little clinical benefit but think treatment would be affordable and manageable. After prices decline sufficiently, most patients are screened. This includes patients for whom screening presents less value, perhaps because they are unlikely to need treatment or unlikely to complete treatment if diagnosed. As a result, the last patients screened would have low uptake of treatment – and the impact of screening on treatment would not be as large as one might otherwise expect.

This prediction remains if different patients face different prices for screening. Figure 3 shows the downward sloping demand curve for screening again but with two upward sloping

Figure 3: Screening demand curve and multiple supply curves



supply curves added, reflecting the fact that different patients can face different prices for screening. Based on the points of intersection between the supply and demand curves, fewer patients are screened in the group facing higher prices for screening. Could high treatment rates could be obtained to by targeting a screening intervention to the less-screened group, since they are higher on the demand curve? Yes, if barriers to screening are uncorrelated with barriers to treatment or benefits to care. Yet, such a situation would be highly unusual - for example, a new screening technology is introduced in a national health system, and some regions have lower access to the technology than other regions for idiosyncratic reasons. In a far more empirically likely case, patients facing high barriers to screening will also face high barriers to treatment. For example, patients who live in rural areas, who are uninsured, or who face language barriers effectively have high barriers to screening and high barriers to treatment. In an insurance scheme where patient co-payments are set based on value, patients facing high prices should also show lower clinical benefits. Both factors reinforce our original prediction that the marginally screened patients, as screening expands, should have lower uptake of treatment after diagnosis on average.

Our framework can also be applied to predict the potential health impacts of screening expansions. We noted above that as screening expands, the additional diagnosed patients are expected to have higher ex ante net cost than previously diagnosed patients. In defining

health impacts, it matters whether patients were mainly selected into screening based on barriers vs. benefits. If sick patients were unscreened chiefly due to high barriers to care, then the benefits of screening and treating these patients could be high. In this case, a bundled screening and access to care intervention - targeted to patients with low access to care - could have significant health impact. In contrast, if access to care is already high and patients remain unscreened due to low predicted benefit, diminishing health returns to screening should be unavoidable.

We assess these predictions empirically in section 4, focusing on contexts relevant to the top two causes of death in the United States, cancer and heart disease. Evidence from heart disease risk factors indicates that currently unscreened patients differ from currently screened patients in important ways, leading to lower predicted uptake of recommended treatment if these patients were diagnosed. Furthermore, evidence from cancer demonstrates diminishing clinical returns to screening, which can be reversed if patients with low access to care are targeted with a bundled intervention. Both findings match the predictions above.

In section 5, we explore the practical implications of these changes in the composition of diagnosed patients for policy analysis. We use three datasets to demonstrate how changes in the composition of diagnosed patients can produce misleading conclusions during policy analysis, such as false reductions in measured health system performance after barriers to screening are removed.

4 Empirical analysis

4.1 Treatment outcomes: Evidence from heart disease risk factors

We now assess our theoretical prediction that newly screened patients are less likely to seek treatment after diagnosis, holding other factors constant. To this end, the main goal of this analysis is to establish an association between patients' propensity to be screened and their uptake of physician-recommended treatment after diagnosis. This analysis focuses on three conditions which are important risk factors for heart disease and are commonly undiagnosed: diabetes, high cholesterol, and hypertension.

We use two empirical approaches to provide policy-relevant evidence on the factors underlying links between screening and treatment and the implications for screening interventions. First, we provide practical evidence on the potential impact of screening interventions by conducting a comparison across patients. In particular, we use claims data to compare uptake of relevant doctor visits among patients who were diagnosed during a biomarker assessment with patients who were already diagnosed prior to the intervention. The results demonstrate

a shortfall in treatment among newly diagnosed patients, consistent with our hypothesis. These findings are helpful for understanding policies that increase screening access across a broad population.

Second, we aggregate data to the patient-level and analyze common patterns across patients' multiple prevalent conditions. This analysis highlights the role of over-arching contextual and patient-level factors as determinants of screening and treatment uptake, underscoring the point that factors other than clinical severity can produce a link between screening and treatment. We also conduct a simulation analysis to assess the contribution of specific individual characteristics to the effects of screening interventions. The results indicate that currently unscreened patients differ from currently screened patients in important ways, leading to lower predicted uptake of recommended treatment if these patients were diagnosed. These findings are important for understanding policies that specifically target patients who are rarely screened.

4.1.1 Analysis of claims data after a biomarker assessment

This analysis provides a direct test of whether biomarker assessment interventions can close gaps in treatment uptake. We exploit an exogenous increase in biomarker assessment among participants in the REGARDS study.

The REGARDS study provided participants with biomarker assessments and then informed participants about their previously undiagnosed conditions (Howard et al., 2005). The REGARDS baseline biomarker data have been merged with individual-level Medicare claims data, allowing us to track doctor visits to treat specific conditions before and after their biomarker assessment via REGARDS (Muntner et al., 2014).

We assume that lack of diagnosis prior to biomarker assessment via the REGARDS study is related to prior uptake of screening, and therefore place prior diagnosis status on the right-hand side of the model. To use the example of diabetes, we run:

$$Pr\left(\text{Diabetes is Treated}\right) = f(\text{Diabetes was Undiagnosed Prior to Biomarker Assessment Via REGARDS})$$

and likewise for the other two conditions. More precisely, we use models of the following form to compare annual treatment of previously diagnosed vs. previously undiagnosed conditions data the year after biomarker assessment via REGARDS:

$$M_{ijt} = \tilde{\alpha} + U_{ij,t-2}\tilde{\gamma} + X_i\tilde{\beta} + \epsilon_{ijt} \tag{1}$$

We do not parse out Hawthorne effects (i.e., the effect of biomarker assessment via REGARDS on treatment of already-diagnosed conditions) because we find no evidence of such effects for our outcomes of interest in a companion paper (Myerson et al., 2017).

Our predictor of interest is $U_{ij,t-2}$. This variable takes the value 1 if individual i's prevalent condition j was undiagnosed prior to biomarker assessment via REGARDS, and 0 if condition j was diagnosed prior to biomarker assessment via REGARDS. Time t denotes our period of observation: this model includes data from the 12-24 months after each participant had his or her biomarkers assessed via the REGARDS study. We selected the year after REGARDS participation for ease of interpretation. (This also provides a conservative estimate, as our findings are larger in magnitude if we include additional data from the first 12 months after biomarker assessment REGARDS.)

In addition to bivariate models, we also present models which adjust for health measures related to condition severity, denoted X_i, with the purpose of assessing whether condition severity fully accounts for the gaps in treatment that we find. These health measures include BMI, glucose measures (fasting plasma glucose), cholesterol measures (HDL and LDL cholesterol, total cholesterol, triglycerides), the average of two systolic and diastolic blood pressure measures, and an indicator for presence of multiple co-morbid conditions. All continuous variables are binned into four categories of equal size based on quartiles of the sample distribution to allow non-linearity in the relationship between these variables and doctor visits; missing values are given their own indicator. Because we estimate linear probability models and the M_{ij} outcomes are binary, we account for heteroskedasticity in ϵ_{ij} by using robust standard errors.

4.1.2 Analysis of cross-sectional survey data

We also aggregate data to the patient-level and analyze common patterns across patients' multiple prevalent conditions. This analysis exploits the fact that a person may have multiple chronic conditions, and uses lack of screening for other conditions as an indicator of screening uptake on the patient-level. We use two cross-sectional datasets to run models of the following form:

$$Pr\left(\text{Condition is Treated After Diagnosis}\right) = f(\text{Recent Blood Test to Screen for } \textit{Other} \text{ Undiagnosed Conditions})$$

While it may seem unusual to analyze multiple conditions in a single model, the purpose of this model is to highlight the role of over-arching contextual and patient-level factors as determinants of screening and treatment uptake. Furthermore, our conditions of interest

(diabetes, high cholesterol, and hypertension) are frequently discussed together in the medical literature. They are often co-morbid and contribute to a cluster of risk factors called metabolic syndrome (Grundy, 2004; Sowers et al., 2001).

We focus on screening via blood tests rather than blood pressure measurements for two main reasons. First, none of our data sets include data on recent blood pressure tests conducted in a clinical setting. Second, blood pressure is considered a vital sign and is therefore measured at most clinic visits; in contrast, blood tests are conducted among undiagnosed patients for the express purpose of screening.

These models can be represented by the following general notation:

$$M_{ij} = \alpha + U_{i,-j}\gamma + X_i\beta + \epsilon_{ij} \tag{2}$$

M_{ij} indicates medical care received by person i for prevalent, diagnosed condition j. $U_{i,-j}$ takes the value 0 if person i recently had a blood test to screen for conditions $-j$ for which they were not already diagnosed, and 1 otherwise.

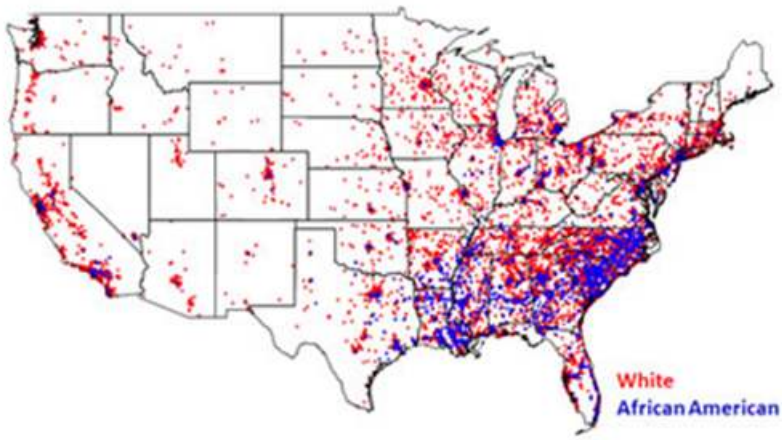
In addition to bivariate models, we present models which adjust for health measures related to condition severity, denoted X_i, as in the previous analysis. We use the same set of health related covariates as in the previous analysis, with two exceptions. First, in NHANES data analysis we add controls for diabetic retinopathy, which is the only consistently measured symptom for our conditions of interest and not measured in REGARDS. Second, due to the availability of hba1c, a more stable measure of glucose, we use this rather than fasting plasma glucose. As in the previous analysis, we account for heteroskedasticity in ϵ_{ijt} by using robust standard errors and bin all covariates into quartiles included as dummy variables, giving missing values their own indicator.

4.1.3 Data

We use data from three studies: the REasons for Geographic and Racial Differences in Stroke study (REGARDS), the National Health and Nutrition Examination Survey (NHANES), and the Oregon Health Insurance Experiment (OHIE). In all three studies, participants reported their diagnosed conditions in a survey, had their biomarkers taken, and were paid for their time. Table 1 summarizes the sample selection and characteristics of included participants from these three studies.

The REasons for Geographic and Racial Differences in Stroke (REGARDS) study recruited community-dwelling participants into an epidemiological longitudinal cohort study designed to answer questions about racial differences in stroke mortality. Recruitment was conducted from 2003-2007 and was accomplished by randomly selecting numbers from com-

Figure 4: Location of REGARDS participants (Source: Howard et al., 2011)



mercially available lists of residential phone numbers in the 48 contiguous United States (i.e., excluding Alaska and Hawaii). Sampling was stratified across African Americans and whites and three regions: the stroke belt (Alabama, Arkansas, Mississippi, and Tennessee), stroke buckle (North Carolina, South Carolina and Georgia) and all other states in the continental United States. Individuals who were under 45 years of age, did not identify as either African American or white, were non-English speaking, undergoing cancer treatment, or who resided in or were on a waiting list to enter a nursing home were excluded from the REGARDS study (Howard et al., 2005). Figure 4 shows the geographic distribution of African American and white participants.

Participants were first interviewed, including questions about whether they had been diagnosed with high blood pressure, diabetes or high cholesterol by a doctor or nurse. For the in-home visit, participants were instructed to fast for 8-10 hours,[6] and had their blood glucose, blood pressure and lipid panel plus other biomarkers assessed in their home on a morning of their choosing. Participants were compensated $30 for their time, and were notified of their results and advised to seek medical care for abnormal results using three levels of notification: (1) by telephone if any value is in the critical range, with instructions to immediately seek care; (2) by mail when a value is in the alert range with instructions to promptly seek care, and (3) general mail notification otherwise. The text of the mail notification for notification of high cholesterol or blood glucose and cards used for notification

[6]About 80% of participants met the fasting requirement at the time that their labs were taken. We use fasting- or non-fasting specific cutoffs where applicable when judging participants' disease status based on their biomarkers.

of high blood pressure are shown in Figure A.1 in the Appendix.

The REGARDS data have been linked with administrative records of doctor visits for participants enrolled in traditional Medicare (Muntner et al., 2014). We use the ICD-9 codes in the claims data to identify which of the patient's prevalent conditions were addressed in any given evaluation and management visit with a doctor; a single visit could address multiple conditions. (Myerson et al. (2017) provides additional discussion.)

NHANES is a nationally representative biomarker survey run by the Center for Disease Control and Prevention. Comparable data have been collected on a rolling basis from 1999-2014, and these are the data most commonly used to track awareness of chronic conditions over time on the national level (Centers for Disease Control and Prevention, 2014). Data on recent screening for diabetes are only available starting in 2005; we therefore use data from 2005-2014.

Finally, we use publicly available data from the OHIE in-person biomarker data collection and 12-month mail-in survey, which were conducted during 2009-2010. Participants in these surveys had entered a lottery to apply for Medicaid in Oregon in 2008 (Allen et al., 2010; Baicker et al., 2013). In the OHIE data, both self-reported and validated measures of current medications are available, although the data are collected at slightly different times. To ensure that participants' treatment, screening, and diagnosis status are measured at the same time, we measure all of these variables using data from the 12 month follow-up survey in the main analysis; out of necessity, we measure the biomarkers using data from the in-person survey.[7] (All findings we report using the OHIE data remain qualitatively similar if we use the medication measures collected during the in-person survey rather than the 12-month follow-up survey.)

These three data sources have different advantages and disadvantages for our analysis. The merged REGARDS-Medicare data present the unique advantage of using administrative data to track participants' relevant doctor visits after their biomarkers were assessed.[8] We therefore estimate Model (1) using the REGARDS data, comparing doctor visits for newly diagnosed conditions vs. previously diagnosed conditions after the REGARDS study assessed patients' biomarkers. In contrast, we can only use NHANES and OHIE for estimating Model (2), because the NHANES and OHIE ask about recent screening whereas REGARDS does

[7]Although both surveys were implemented over 2009-2010, the median gap in time between the in-person survey and 12-month follow-up was just over 6 months; the in-person survey was completed later than the 12-month survey for most respondents. In the OHIE data, codebooks of the publicly available data indicate that questions about screening for high cholesterol were collected as part of the in-person survey but questions about screening for diabetes were not. In contrast, screening of both conditions was asked about in the 12-month follow-up survey.

[8]The timing of biomarker assessment in OHIE precludes us from examining the impact of biomarker assessment on self-reported doctor visits or use of medications using the publicly available data.

Table 1: Characteristics of included participants from the three biomarker surveys

	REGARDS	**NHANES**	**OHIE**
Survey Inclusion Criteria	In traditional Medicare past 2 years; black or white; English speaking	Nationally representative	Applicants to expanded Medicaid (in both the in-person survey and 12 month follow-up)
Geography of Sample	National	National	Oregon
Year of Biomarker Collection	2003-2007	2005-2014	2009-2010
Age Range in Analysis	67+	All	19+
Participants with Any Condition(s) of Interest	5,721	18,735	3,482
Participants with Undiagnosed Condition(s) of Interest	1,077	6,281	1,546
Among Participants with Condition(s) of Interest:			
Average Age	74	55	45
Had Health Insurance	100%	81%	48%
African American	30%	22%	9%
Participants with Diabetes	1,309	4,282	705
Aware of Diabetes	1,192	3,482	666
Treating with Medication	1,161	2,991	500
Participants with Hypertension	4,502	11,576	1,917
Aware of Hypertension	4,170	10,193	1,657
Treating with Medication	3,846	7,680	1,056
Participants with High Cholesterol	4,268	13,716	2,663
Aware of High Cholesterol	3,542	9,030	1,301
Treating with Medication	2,457	4,860	672

not.

The NHANES and OHIE data also have different advantages and disadvantages for estimating Model (2). The NHANES data include information on whether a doctor had ever recommended managing hypertension and high cholesterol using a prescription, whereas the OHIE (and REGARDS) data do not. This is important because national guidelines recommend treating less severe cases of these conditions with diet and exercise before prescribing medication (James et al., 2014; Stone et al., 2014). By tracking medication use only among participants who report that their doctor recommended medication, we can ensure that our results are not driven by medication non-use among patients whose doctors recommended controlling the condition through diet and exercise alone. As a nationally representative survey, the NHANES also samples the most diverse group of participants. In contrast, the OHIE data have a different advantage for the present analysis. Adding these data allows us to pursue a focused analysis of a group of importance given recent health policy changes: applicants to expanded Medicaid. In Medicaid expansion states, many patients who become diagnosed due to the Affordable Care Act could come from this group (Kaufman et al., 2015; Myerson and Laiteerapong, 2016; Simon et al., 2016; Wherry and Miller, 2016).

Tracking diagnosed and undiagnosed conditions, for Model (1) We code participants as having a particular chronic condition (diabetes, hypertension, and/or high cholesterol) if they report prior diagnosis for the condition at the time of participation, with the appropriate exclusions for diagnosis during pregnancy, or if their biomarkers meet standard definitions for the condition after taking their fasting status into account (American Diabetes Association, 2014; Stone et al., 2014; James et al., 2014). Table A.1 in the Appendix includes details of each definition. Individuals are classified as undiagnosed for the condition if they meet the biomarker definitions for a condition, but report no prior diagnosis for that condition.

Using the merged REGARDS-Medicare data, we are able to correct for patients' underreporting of diagnosis using Medicare claims. We accomplish this by also classifying participants as diagnosed if they meet biomarkers criteria of the condition and also meet Chronic Conditions Warehouse definitions for the condition based on their recent Medicare claims. This process increases the number of diagnosed cases of high cholesterol by 148 (4%), the number of diagnosed cases of diabetes by 26 (2%), and the number of diagnosed cases of hypertension by 119 (2%).

Tracking screening of undiagnosed conditions, for Model (2) The questions about recent blood tests to screen for diabetes and high cholesterol in the OHIE and NHANES

have slightly different look-back periods. The questions about screening for diabetes and high cholesterol in the OHIE data in the 12-month follow up survey focus on screening within the past 12 months (Finkelstein, 2013).[9] In contrast, the look-back period for diabetes screening in the NHANES data is 3 years. We combined data from multiple variables in the NHANES to construct measures of high cholesterol screening within the past year and within the past two years.[10] We present results using the two-year look-back period, but findings are similar using the one-year look-back period.

4.1.4 Results

Doctor visits A key question is whether patients who are diagnosed after outreach to encourage screening would be less likely to treat conditions that become diagnosed as a result of this intervention. To address this question, we use Medicare claims data from individuals whose biomarkers were assessed by the REGARDS study to estimate Model (1). This model estimates the gap in annual doctor visits for evaluation and management of previously diagnosed vs. previously undiagnosed conditions after all participants learned about their biomarkers due to participation in REGARDS.

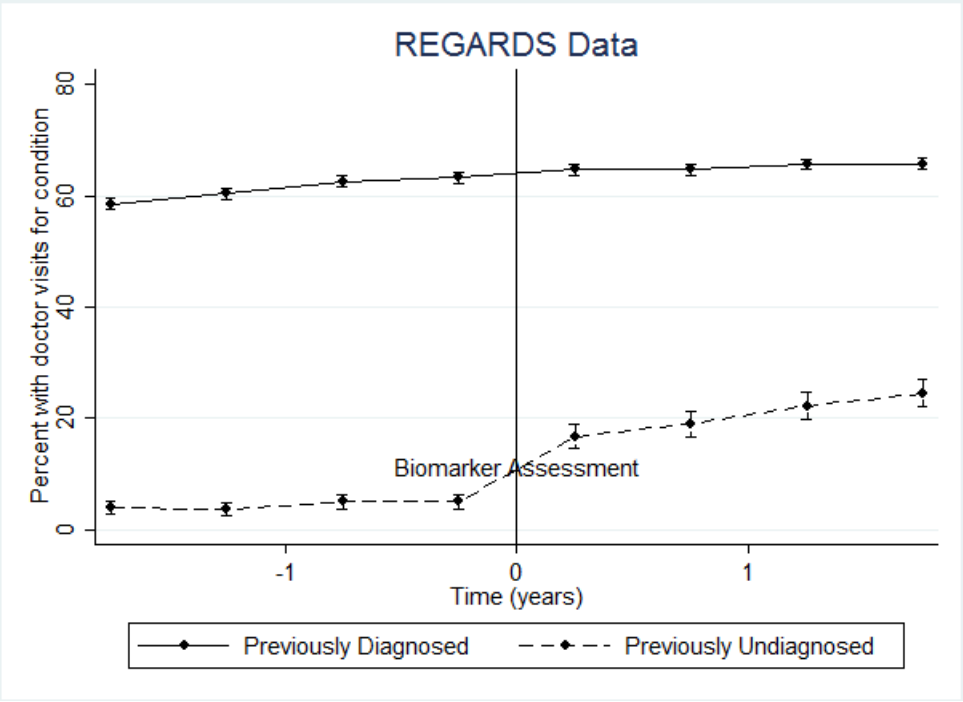
As shown in Table 2, we find that previously undiagnosed conditions were less likely to receive an annual evaluation and management doctor visit than previously diagnosed conditions. We present unadjusted gaps in doctor visits as well as the gap after adjustment for a biomarker-measured condition severity. These controls adjust for the possibility that less severe cases of hypertension or high cholesterol could be evaluated by a physician on less than an annual basis. (In the case of diabetes, foot exams, eye exams, and multiple hba1c measurements by a physician are recommended on an annual basis for all diabetes patients regardless of severity (American Diabetes Association, 2014).) The results indicate that condition severity does not account for the gaps in treatment we find.

The findings are qualitatively similar if we analyze number of visits per year or if we include data from the first year after biomarker assessment via REGARDS. Notification by mail for the diabetes and high cholesterol results is unlikely to account for the observed

[9]The relevant questions in 12-month follow-up survey are as follows: "Have you ever had your blood cholesterol checked?" and "Have you ever had a blood test for high blood sugar or diabetes?" The response options include "Yes, within the last year," "Yes, but it's been more than a year," and "Never" . In this case, we code both the second and third response options as a negative response and determine "recent" screening to be screening within the past year.

[10]Timing of blood cholesterol screening is assessed using two questions: "Have you ever had your blood cholesterol checked?" and "About how long has it been since you last had your blood cholesterol checked? Has it been..." with the options "Less than a year ago," "1 year but less than 2 years ago," "2 years but less than 5 years ago," or "5 years or more." Timing of diabetes screening is assessed using the question: "Have you had a blood test for high blood sugar or diabetes within the past three years?" with responses either "Yes" or "No." (We code "Refused" or "Don't Know" as missing.)

Figure 5: Previously undiagnosed conditions remain less likely than previously diagnosed conditions to receive doctor visits (REGARDS raw data)



This figure compares semi-annual doctor visits after biomarker assessment via REGARDS for evaluation and management of previously diagnosed vs. previously undiagnosed diabetes, hypertension, and high cholesterol. The previously undiagnosed conditions were conditions REGARDS participants became aware of through participation in the study; all patients with abnormal biomarkers were advised to see a doctor. The x-axis indicates years since biomarker assessment via REGARDS; the 0-point indicates the month of biomarker assessment, which maps to a different calendar time for different participants due to rolling recruitment. The y-axis indicates the percent of conditions with any doctor visits on a semi-annual basis, as measured using Medicare claims data and categorized as relevant to each prevalent condition using ICD-9 codes.

shortfall in doctor visits for newly diagnosed conditions, because the gap in doctor visits exists for all three conditions, including high blood pressure.[11] Finally, Figure 5 shows a similar relationship in the raw data: doctor visits for previously undiagnosed conditions increased after biomarker assessment, but only to about half the level of previously diagnosed conditions.

Uptake of recommended treatment Using the NHANES data and OHIE, we show that the use of recommended treatment for diagnosed conditions is lower among individuals not recently screened for other, undiagnosed conditions. By combining data from multiple

[11]Participants received their blood pressure results immediately, in person.

Table 2: Previously undiagnosed conditions remain less likely than previously diagnosed conditions to receive doctor visits (REGARDS data, regression-based comparison)

		Any Relevant Doctor Visits Per Year			
	Obs.	Previously undiagnosed condition	Previously diagnosed condition	Unadjusted difference	Adjusted difference
(1) Relevant doctor visits for prevalent diabetes	1,309	0.46	0.92	-0.46***	-0.46***
				(0.01)	(0.05)
(2) Relevant doctor visits for prevalent high cholesterol	4,268	0.68	0.29	-0.39***	-0.344***
				(0.02)	(0.02)
(3) Relevant doctor visits for prevalent hypertension	4,502	0.85	0.38	-0.47***	-0.47***
				(0.03)	(0.03)
Adjust for bio-marker measured severity and co-morbid conditions				No	Yes

This table compares annual doctor visits after biomarker assessment via REGARDS for evaluation and management of previously diagnosed vs. previously undiagnosed diabetes, hypertension, and high cholesterol. The previously undiagnosed conditions were conditions REGARDS participants became aware of through participation in the study; all patients with abnormal biomarkers were advised to see a doctor. The outcomes are annual doctor visits from the 12-24 months after biomarker assessment via REGARDS, measured using Medicare claims data and categorized as relevant to each prevalent condition using ICD-9 codes. Robust standard errors are in parentheses.

prevalent conditions, this analysis captures the role of over-arching contextual and patient-level factors as determinants of screening and treatment uptake.

First, bivariate regressions in the NHANES data indicate that participants who were not recently screened for undiagnosed conditions were less likely to have a foot exam or eye exam over the past year for their diagnosed diabetes, or to report taking the medications their doctor prescribed for their diagnosed hypertension or high cholesterol.[12] See Table A.2 in the Appendix. Adjusting for patient health does not eliminate these findings. This pattern of results indicates that condition severity does not fully account for the lower uptake of treatment we find among rarely screened patients.

Our findings are similar when we replicate these analyses in the OHIE data, subject to some caveats related to data restrictions. Because questions about recommended treatment are only asked in the NHANES, we cannot restrict the sample to only patients whose doctors recommended treatment using medications (rather than dietary modification alone) when analyzing the OHIE data (Finkelstein, 2013). The OHIE data on use of medication for diabetes, hypertension, and high cholesterol are therefore presented with the caveat that our treatment metric is an imperfect measure of compliance with recommended treatment. Nonetheless, our findings using OHIE data on treatment, shown in Table A.3 in the Appendix, are qualitatively similar to the NHANES results. As was found in the NHANES data, the gaps in treatment that we observe in the OHIE data diminish but do not disappear when we control for biomarker based measures of health.

One might argue that if people who are rarely screened tend to have less severe conditions, lower treatment of diagnosed conditions in this group would represent an appropriate allocation of resources. However, our findings are to the contrary. In the OHIE and NHANES data, we find people who are rarely screened had *more* severe biomarkers for their diagnosed conditions. See Tables A.4 and A.5 in the Appendix. Findings are similar in the REGARDS data, with the caveat that these data do not measure recent screening. Participants in REGARDS with undiagnosed conditions show more severe biomarkers for their other, diagnosed conditions. See Table A.6 in the Appendix. The low adherence to treatment we observe could contribute to this finding, because treatment helps patients to control their

[12]Doctors' recommendations to control hypertension and high cholesterol using medication are asked about in the NHANES, enabling us to track medication use only among diagnosed patients for whom medication was recommended. However, there is no comparable question for diabetes. However, annual foot exams and eye exams are recommended for all people with diabetes as standard care (American Diabetes Association, 2014).

conditions.[13]

4.1.5 Evidence of possible mechanisms

We next investigate the possible mechanisms that could account for these findings. When outlining the theoretical framework in the previous section, we had hypothesized that barriers to care and condition severity would predict which patients were screened, and also predict which patients were treated after diagnosis. In this sub-section we use individual-level data from the nationally representative NHANES survey to show that these factors could attenuate the impact of screening expansions on treatment.

We construct two key variables to measure screening and treatment. The screening variable takes the value 1 if the patient has been recently screened for a particular never-diagnosed condition and 0 otherwise, with missing values for patients already diagnosed for the condition of interest. Likewise, the treatment variable takes the value 1 if the patient is complying with doctor-recommended treatment for their diagnosed condition and 0 otherwise, with missing values for patients never diagnosed for the condition of interest.

Our findings support the hypothesis that factors related to barriers to care and potential clinical benefit predict screening and treatment in the directions hypothesized. Table 3 summarizes the sign of the statistically significant correlation coefficients found in the data. These correlations show that patients who were not recently screened (for a condition for which they have never been diagnosed) show larger barriers to care including lack of health insurance, low income, and lack of a usual source of care. These same factors are, in turn, associated with lower likelihood treating diagnosed conditions. We also find that patients who appear healthier ex ante - patients who were younger, had better self-reported health or lower Framingham risk score, fewer co-morbid conditions, or no recent hospitalizations - were less likely to be screened for undiagnosed conditions and also less likely to treat their diagnosed conditions.

We next examine whether these factors could account for lower treatment uptake among patients reached by expanded screening, by conducting a policy simulation. To minimize assumptions in our policy simulation, we use patients' real treatment outcomes. Our exercise therefore involves selectively dropping the treatment data of patients who had in fact been screened, rather than imputing the treatment data of patients who were never screened. The policy simulation exercise proceeded as follows. In one counterfactual, we only allowed patients with propensity scores over 70% to be "screened" and treatment data for all others

[13]Severity at diagnosis is not measured in these data for already-diagnosed conditions. Therefore, we are unable to parse out the extent to which gaps in biomarkers were pre-existing at diagnosis - because patients selected into care due to non-clinical factors - versus exacerbated by low uptake of treatment. We analyze data on severity at diagnosis in another dataset in the subsequent sub-section.

Table 3: Patient-level characteristics associated with increased screening are also associated with increased treatment after diagnosis (NHANES data)

	Screened if Undiagnosed	Treated if Diagnosed
Barriers to care		
No health insurance	↓	↓
No usual place for health care	↓	↓
Lower income	↓	↓
Benefits to care		
Poor self-reported health	↑	↑
Recently hospitalized overnight	↑	↑
Co-morbid condition	↑	↑
Framingham risk score	↑	↑
Age	↑	↑

This table shows that patient-level characteristics associated with increased screening are also associated with increased treatment after diagnosis in the NHANES data. The arrows represent the sign of statistically significant correlations between the variable listed in the column title and the variable listed in the row title. The sign and significance of the findings are the same for both diabetes and high cholesterol, two conditions whose screening and treatment is measured in the NHANES data.

were dropped, even if these patients truly had been screened. In another counterfactual, all patients with propensity scores over 60% were allowed to be "screened" and others were dropped. In each case, we calculated the average treatment rate using patients' real data, including only patients who would have been diagnosed under such a scheme. We repeated this exercise for policy counterfactuals with propensity score cutoffs ranging from 0.7 to 0, in 0.1 unit intervals. (The cutoff of 0 represents the case where no patients were dropped.)

Findings from this exercise indicate that treatment rates for diagnosed patients would decline as more patients are screened, with approximately 0.5 percentage point decline in treatment for each 10 percent increase in screening. When we focus on patients without health insurance, a group that includes the patients in our sample with the lowest access to care, we found that treatment rates declined more rapidly, up to 1 percentage point decline in treatment for each 10 percentage point increase in screening. This is consistent with our prediction that targeting patients with high barriers to screening cannot overcome the pattern of marginally diagnosed patients having lower treatment uptake.

Because pharmaceutical treatment is not recommended for some diagnosed patients, we also repeated these analyses using compliance with recommended treatment as the outcome of interest. For this exercise, we restricted the high cholesterol sample to only include patients who reported that their doctor had recommended taking medication for their high cholesterol. For diabetes, we focused on eye exams and foot exams, which are recommended

for all patients with diabetes even when prescription medications are not. Our findings were qualitatively similar with this change. See Table A.7 in the Appendix.

4.1.6 Summary

The empirical findings in this sub-section supported our key point that the composition of diagnosed patients changes as screening expands, in ways that can depress treatment uptake. First, we found that conditions diagnosed as part of a biomarker study are less likely than previously diagnosed conditions to receive any doctor visits over a one-year period. Second, we found that people who have not recently had blood tests to screen for undiagnosed conditions were less likely to adhere to recommended treatment for their other, diagnosed conditions. Correspondingly, these patients showed worse biomarkers for their diagnosed conditions. Together, these findings indicate that screening currently unscreened patients could increase the fraction of diagnosed patients who don't receive relevant doctor visits for their condition, don't use recommended treatment, and possibly remain uncontrolled.

Finally, we showed that these associations between screening and treatment could be produced by the intermediate factors discussed in section 3. Namely, patients who appear less sick or face higher barriers to care are less likely to be screened and also less likely to be treated after diagnosis. Because the gaps in treatment we observe remain after adjusting for condition severity, we hypothesize that barriers in access to care could play an important role. These issues are explored further in the next analysis.

4.2 Health outcomes: Evidence from cancer

Although the findings in the previous sub-section seem to indicate diminishing returns to screening, we demonstrate that this need not be the case with a targeted, bundled intervention. This analysis assesses two hypotheses arising from our theoretical framework. First, we hypothesized that screening patients with high access to care could produce diminishing health benefits as screening expands. Second, we hypothesize that screening patients with low access to care could yield significant health benefits if these patients' barriers to care are simultaneously addressed.

We exploit an exogenous increase in health insurance, which provides access to both screening and treatment. In particular, we exploit the previously established increases in health insurance coverage at age 65, the age of near-universal eligibility for Medicare (Card et al., 2008; McWilliams et al., 2013). We track changes in cancer detection, severity of detected conditions, and patient health outcomes at age 65 using cancer registry data. Results from a regression discontinuity analysis indicate that at age 65, one-year survival after can-

cer diagnosis increases for racial and ethnic minorities, a group that previously faced higher barriers to care. In contrast, non-minority patients showed no significant improvements in post-diagnosis survival. This evidence is consistent with our prediction that diminishing health returns to screening could be reversed if patients with barriers to care are targeted and their barriers to care are addressed.

4.2.1 Data and analysis

Our empirical analysis requires data on age and some welfare-relevant outcomes, such as severity of cancer at detection and one-year survival after cancer detection. Our analysis uses the Surveillance, Epidemiology and End Results (SEER) 2000–2014 program database, which combines data from 18 cancer registries across the United States. Together, the registries capture data on the majority of newly detected tumors in 11 states (Alaska, California, Connecticut, Hawaii, Iowa, Kentucky, Louisiana, Georgia, New Jersey, New Mexico, and Utah) and two additional cities (Detroit and Seattle). These regions cover about one-quarter of the US population, according to 2010 population numbers. We exclude years prior to 2000 to maintain a balanced panel of SEER registries.

The SEER data include information on patient age at the time of cancer detection, the patient's sex, race and ethnicity, and the cancer's stage at the time of detection (i.e., the extent to which the cancer has spread throughout the body). The SEER data also include information on subsequent mortality based on linked mortality records. We classified tumors as detected early if the tumor had not yet spread beyond the organ in which it originated (i.e., metastasized) at the time of diagnosis; this maps to the in situ or localized stage in the SEER classification.

Sample selection issues are minimal. We include data from all cancers in the SEER registries except those which lacked information on whether the cancer had metastasized prior to diagnosis.[14] The final data set includes over 1.4 million cancer cases which fall within a bandwidth of 6 years from age 65, including over 250,000 cancer cases for patients who are racial or ethnic minorities.

To estimate the size of the discontinuity in cancer detection and survival, we follow standard methods for analysis of a regression discontinuity analysis (Lee and Lemieux, 2009; Imbens and Lemieux, 2008). As such, our analytic strategy resembles previous research on age discontinuities in Medicare eligibility and age discontinuities in cancer screening (Card et al., 2008; Srikanth and Strumpf, 2016). First, we restrict the data to a small window

[14]This exclusion criterion eliminates prostate cancer, which SEER codes differently than other cancers by combining with local and regional stage in a single category. This exclusion criterion also eliminates cancers of unknown stage.

around the Medicare eligibility threshold (age 65) and estimate a local-linear regression using the *rdrobust* Stata command (Calonico et al., 2014a). We use a triangle kernel which places higher weight on observations with closer distance to the threshold, and report the robust bias-corrected standard errors recommended in Calonico et al. (2014b). Second, we use data within our bandwidth and estimate the following model for patient i of age a_i:

$$\begin{aligned} Y_i =& \beta_0 + \beta_1 \left(Age \geq 65\right)_i + \beta_2 \left(Age \geq 65\right)_i \times (a_i - 65) \\ &+ \beta_3 \left(Age < 65\right)_i \times (a_i - 65) + \beta_4 g_i + \epsilon_i \end{aligned}$$

Y_i indicates outcomes of interest such as severity of the cancer at detection or one-year survival after cancer detection, and $\left(Age \geq 65\right)_i$ indicates that the patient is age-eligible for Medicare (that is, strictly above age 64). We allow age trend terms to vary above vs. below the cutoff, and adjust for patient gender g_i. This model is estimated by OLS with heteroskedasticity robust standard errors. For both analyses, we select an optimal bandwidth using the *rdbwselect* Stata command, which suggests a bandwidth of 6 years for our data (Calonico et al., 2014a).

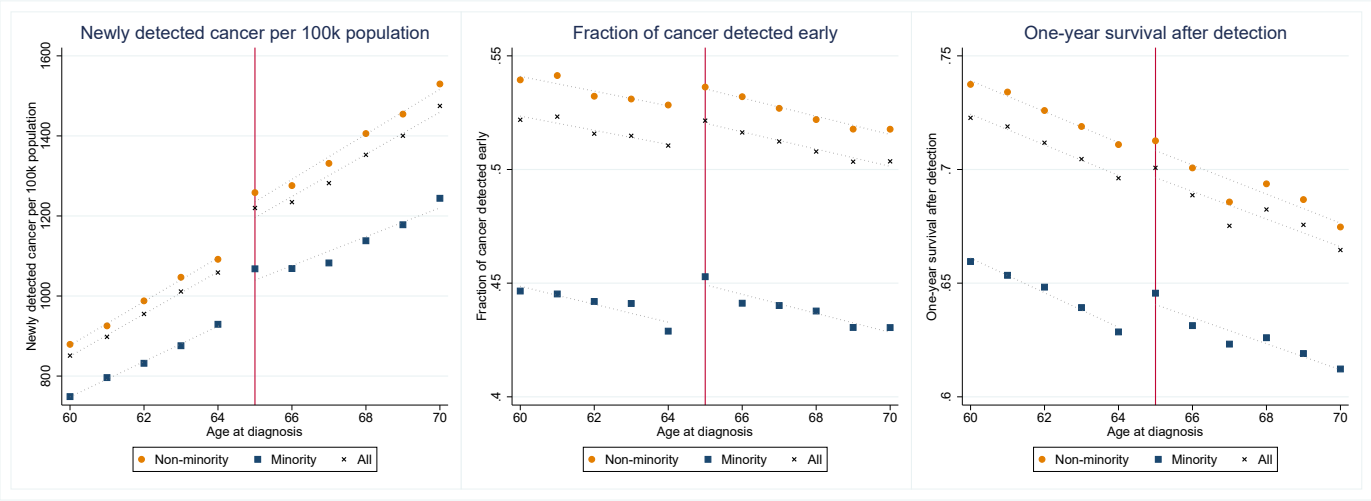
This regression discontinuity design permits us to assess potential heterogeneity in outcomes and early detection across racial and ethnic groups by stratifying the models. As such, we estimate each model separately for patients who are not racial or ethnic minorities (i.e., non-Hispanic white patients) vs. patients who are racial or ethnic minorities. Finally, disparities between racial/ethnic minority patients and non-minority patients in cancer survival are also policy-relevant and widely tracked in the cancer literature. Therefore, we compare the disparities found at ages 63-64 with the disparities at age 65 implied by the measured treatment effect.

4.2.2 Results

Figure 6 depicts cancer detection rates per 100,000 population, the fraction of cancers detected early (while still contained within a single organ), and one-year survival after cancer detection. These data have multiple notable characteristics. First, detection of cancer increases and survival decreases in general as age increases, reflecting the overall aging process. Second, late detection of cancer is relatively common and one-year survival after cancer diagnosis is relatively low in this age range, particularly among racial and ethnic minority patients. Finally, there are visible discontinuities in each of the graphs at age 65, the age at which patients become age-eligible for Medicare.

Table 4 reports our findings from the regression-discontinuity analyses, which account for secular trends in aging as well as patient gender. The first row of each table displays our

Figure 6: Detection of cancer, fraction of cancers detected early, and one-year survival after cancer detection (SEER registry data)



These graphs show an increase in cancer detection at age 65, as well as corresponding changes in the fraction of cancers detected early and changes in one-year survival after cancer diagnosis at age 65. In each graph, the x-axis is age at diagnosis; a vertical line is drawn at age 65, the age at which patients become age-eligible for Medicare. The dotted lines are linear regression lines, estimated separately below vs. above age 65.

findings when all patients are pooled together. We find that cancer detection increases by 66 per 100,000 population at age 65. This implies about a 6% increase in cancer detection compared to the mean detection rates at ages 63-64 (the "untreated" group in the regression discontinuity design). Furthermore, cancers detected at age 65 are 1 percentage point more likely to be detected prior to metastasis than cancers detected at ages 63-64. This is consistent with the epidemiological literature on cancer, which has shown that screening expansions tend to diagnose less severe cancers (Ahn et al., 2014; Loeb et al., 2014; Srikanth and Strumpf, 2016).

However, these average effects mask the importance of prior access to care in driving the impact of becoming age-eligible for Medicare. For example, previous research has shown that compared to non-minority patients, racial and ethnic minority patients have lower levels of health insurance coverage and fewer annual doctor visits at ages 63-64 (Card et al., 2008). Likewise, the SEER data indicate that at ages 63-64, racial and ethnic minority patients are 19% more likely to have metastatic cancer at the time of detection.[15] Subsequently, at age 65, racial and ethnic minority patients show larger gains in health insurance coverage and uptake of annual doctor visits than non-minority patients at age 65, according to previous

[15] 56% of cancers detected in racial and ethnic minority were detected after metastasis, compared to 47% of cancers detected among non-minority patients. See Table 4. A detailed histogram is provided in Figure A.2 in the Appendix.

Table 4: Cancer diagnosis and outcomes just before age 65 and estimated discontinuities at age 65 (SEER registry data)

A. Estimates from local linear regression

	Cancers diagnosed per 100k population per year		Fraction of diagnosed cancer detected prior to metastasis		One-year survival after cancer diagnosis	
	Age 63-4	RD at 65	Age 63-4	RD at 65	Age 63-4	RD at 65
All	1002	68.22***	0.51	0.011**	0.70	0.005
		(12.17)		(0.005)		(0.005)
White, non-Hispanic	1013	70.57***	0.53	0.009	0.71	0.002
		(14)		(0.006)		(0.004)
Racial or Ethnic Minority	990	65.66***	0.44	0.022***	0.63	0.019***
		(20.20)		(0.006)		(0.006)

B. Estimates from ordinary least squares regression

	Cancers diagnosed per 100k population per year		Fraction of diagnosed cancer detected prior to metastasis		One-year survival after cancer diagnosis	
	Age 63-4	RD at 65	Age 63-4	RD at 65	Age 63-4	RD at 65
All	1002	66.08***	0.51	0.011***	0.70	0.005***
		(10.61)		(0.002)		(0.002)
White, non-Hispanic	1013	66.46***	0.53	0.009***	0.71	0.002
		(9.27)		(0.002)		(0.001)
Racial or Ethnic Minority	990	66.08***	0.44	0.021***	0.63	0.018***
		(18.80)		(0.004)		(0.004)

*** p<0.01, ** p<0.05, * p<0.1.

Entries in odd-numbered columns are percentages of age 63-64 year-olds with cancer detection characteristics shown in the column heading. Entries in even-numbered columns are estimated regression discontinuities at age 65 after adjusting for patient gender. In Table A, the models are estimated using local linear regression with a triangular kernel. In Table B, the models are estimated using ordinary least squares. Robust standard errors are in parentheses.

research (Card et al., 2008).

These disparities in access to care and delays in detection prior to age 65 map precisely to the changes in cancer detection and outcomes we find at age 65. In particular, we find that gains in early detection and post-diagnosis survival at age 65 are concentrated among racial-ethnic minorities. Racial and ethnic minority patients experience a statistically significant 2 percentage point increase in detection of cancer prior to metastasis, and a 2 percentage point increase in one-year survival after cancer diagnosis. In contrast, non-minority patients show no change in early detection or one-year survival after diagnosis according to local linear models, despite accounting for most of the additional cancers detected at age 65.[16] These disparate effects create a decline in health disparities upon aging into Medicare: the gap in one-year survival between racial and ethnic minority cancer patients vs. other cancer patients shrinks by one-quarter at age 65, from 8 percentage points to 6 percentage points.

In summary, these data show that the low treatment uptake shown in the previous section could mask different potential health outcomes, depending on the reasons patients went unscreened in the first place. Patients with high access to care were likely unscreened due to low benefit rather than high cost. Expansions of screening in this group rapidly hit diminishing clinical returns, as evidenced by our data on non-minority patients and previous literature from national health systems. In contrast, patients with low access to care likely may be unscreened because of high barriers to care rather than low benefits. In this group, diminishing clinical returns to screening could be reversed by providing these patients with access to care, as evidenced by our data on racial and ethnic minority patients.

5 Implications for policy analysis and measurement of health system performance

In the previous section, we provided evidence that screening expansions may disproportionately diagnose patients who are less likely to treat their diagnosed conditions. In doing so, we built on existing literature demonstrating that patients self-select into health care based on their anticipated costs and benefits. If the impact of screening expansions can be undermined by patient composition effects, this can help to explain the numerous previous studies showing low impact of cancer screening expansions on treatment and health (Ahn et al., 2014; Kim and Lee, 2017; Lantz et al., 1997; Loeb et al., 2014).

This section discusses the consequences of these patient composition effects for policy analysis and measurement of health system performance. We first argue that changes in

[16] OLS models show similar results with one exception: non-minority patients show a significant increase in early detection, but the point estimate is half as large as that observed for minority patients.

the composition of diagnosed patients can produce misleading conclusions during policy analysis, such as false reductions in measured health system performance after screening expands. We demonstrate this point empirically using the REGARDS data. We also show suggestive evidence using repeated cross-section data on the national level from NHANES and Dartmouth Atlas (Centers for Disease Control and Prevention, 2014; The Center for the Evaluative Clinical Sciences, Dartmouth Medical School, 2017). In the NHANES data, a rise in diagnosis of diabetes, hypertension and high cholesterol in recent years coincided with a fall in treatment of these conditions if diagnosed. Likewise, in the Dartmouth Atlas data, regions with higher diagnostic intensity show lower completeness of care for diagnosed diabetes.

5.1 Potential for common quality metrics to yield misleading conclusions about health system performance

Changes in the composition of diagnosed patients could mask the benefits of expanding access to screening, as captured by commonly used measures of health system performance. This problem arises because true prevalence of conditions is not observed, whereas diagnosis status is observed. As a result, a number of health system performance metrics focus on treatment and control of diagnosed conditions (Center for Medicare and Medicaid Services, 2011, 2016b; National Committee on Quality Assurance, 2016). However, tracking the rate of treatment given diagnosis could lead to the incorrect conclusion that quality of care for chronic conditions declines as more patients become diagnosed.

We are best able to demonstrate this point using the REGARDS data. Biomarker assessment via REGARDS increased doctor visits for undiagnosed conditions without changing doctor visits for diagnosed conditions: in total, the impact on doctor visits was positive (Myerson et al., 2017). When data from previously diagnosed and previously undiagnosed conditions are graphed separately (the solid and dashed lines in Figure 5), these data show an increase in doctor visits after biomarker assessment. In contrast, if the running average of doctor visits for currently diagnosed conditions were graphed, these data would show a decrease in doctor visits after biomarker assessment.[17] This reversal of sign is driven by changes to the group of diagnosed conditions before versus after biomarker assessment.

We can also illustrate the link between higher diagnostic intensity and lower performance on quality metrics using the Dartmouth atlas of health care data on quality and effective

[17]The running average from before biomarker assessment via REGARDS only includes data on conditions that were diagnosed prior to recruitment into REGARDS. In contrast, the running average from after biomarker assessment via REGARDS includes these conditions, plus any previously undiagnosed conditions detected via REGARDS biomarker assessment.

Table 5: Bivariate relationship between diagnostic intensity and treatment of diagnosed diabetes

Outcome	Correlation with diagnostic intensity score
Average annual percent of Medicare enrollees with diabetes age 65-75 having eye examination	-0.22***
Average annual percent of Medicare enrollees with diabetes age 65-75 having hba1c test	-0.31***

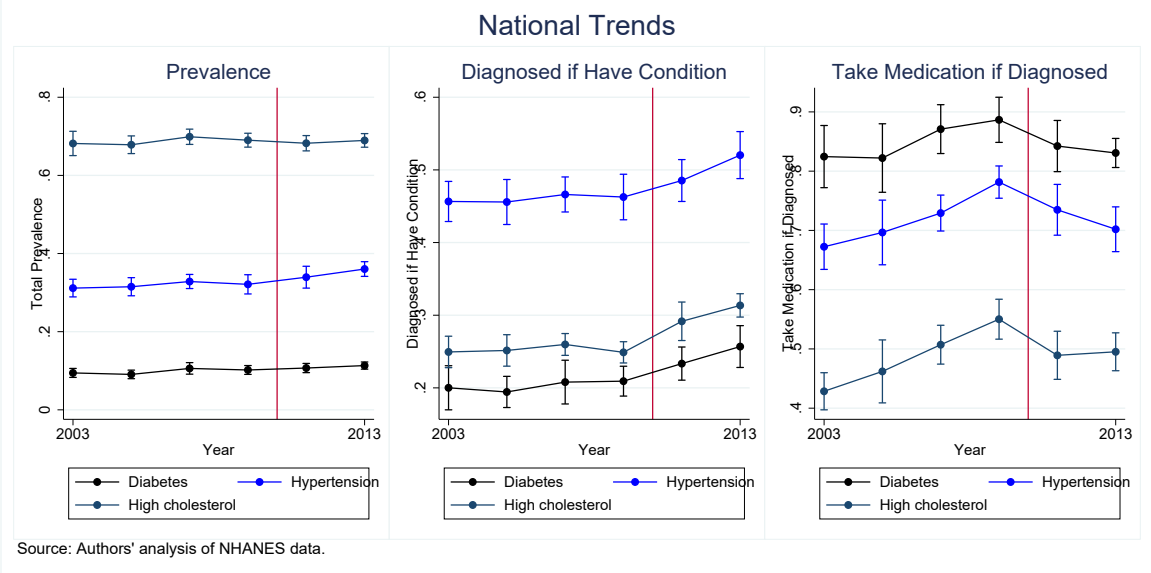
* p <0.1, ** p <0.05, *** p <0.01

This table shows that Hospital Referral Regions that have higher diagnostic intensity for reasons unrelated to the health of their population, as measured by Finkelstein et al. (2017), also have lower uptake of required treatment for diabetes as measured by the Dartmouth Atlas (The Center for the Evaluative Clinical Sciences, Dartmouth Medical School, 2017). This finding is important because uptake of required treatment for diabetes is a popularly used metric of health care quality (CDC, 2012; Center for Medicare and Medicaid Services, 2011; Dale et al., 2016; National Committee on Quality Assurance, 2016; Agency for Healthcare Research and Quality, 2013) .

care (The Center for the Evaluative Clinical Sciences, Dartmouth Medical School, 2017). These data provide a summary of quality of care metrics among Medicare beneficiaries by Hospital Referral Region, including uptake of recommended care for patients with diagnosed diabetes. We combine these data with data from Finkelstein et al. (2017) which summarize doctors' propensity to diagnose conditions ("diagnostic intensity") vary across Hospital Referral Regions after accounting for patients' underlying health needs. Table 5 shows the correlations between the diagnostic intensity variable from the Finkelstein et al paper and quality metrics from the Dartmouth Atlas from the same time period. The results indicate that that Hospital Referral Regions with higher diagnostic intensity have lower rates of treatment of diabetes after diagnosis.

Trends in diagnosis and care for chronic conditions in the national data also support this point. Figure 7 depicts nationally representative estimates of the following three quantities of interest: (a) total prevalence on the population level, including undiagnosed conditions; (b) the fraction of people who truly have the condition who report being diagnosed, (c) and the fraction of people who are diagnosed for the condition who report taking medication to treat the condition. The national estimates are calculated using the NHANES repeated cross-section data, using the survey analysis commands to take into account the complex sampling scheme. These data demonstrate that an increase in diagnosis of diabetes, hypertension, and high cholesterol in recent years coincided with a fall in treatment of diagnosed conditions.

Figure 7: Increased diagnosis of chronic conditions is associated with decreased rates of medical care for diagnosed conditions (NHANES data)



This figure uses repeated cross-sectional data from the NHANES study to demonstrate that an increase in diagnosis of diabetes, hypertension, and high cholesterol on the national level in recent years coincided with a fall in treatment of diagnosed conditions. The left panel depicts total prevalence, including undiagnosed conditions; the middle panel depicts the fraction of prevalent conditions that are diagnosed; and the right panel depicts the fraction of diagnosed conditions that are treated with medications.

6 Conclusion

The key message of this paper is that if anticipated costs and benefits of treatment determine which patients went unscreened before screening interventions, they should also shape the impact of screening interventions. In locations with unequal access to care, such as the United States, barriers sometimes can play a more important role than clinical severity in determining which patients are screened and treated. Regardless, clinical benefits and barriers to care can both create a situation where treatment uptake diminishes as screening expands. This finding has a number of implications for policy analysis, policy design, and health services research.

First, our framework and our findings have implications for policy analysis. Based on the patterns we uncovered, expanded screening as a stand-alone program is likely to be less cost-effective than previously anticipated due to low treatment uptake among marginally screened patients. To our knowledge, these effects are not currently accounted for in cost-effectiveness analyses that simulate the impact of screening expansions, and accounting for these effects

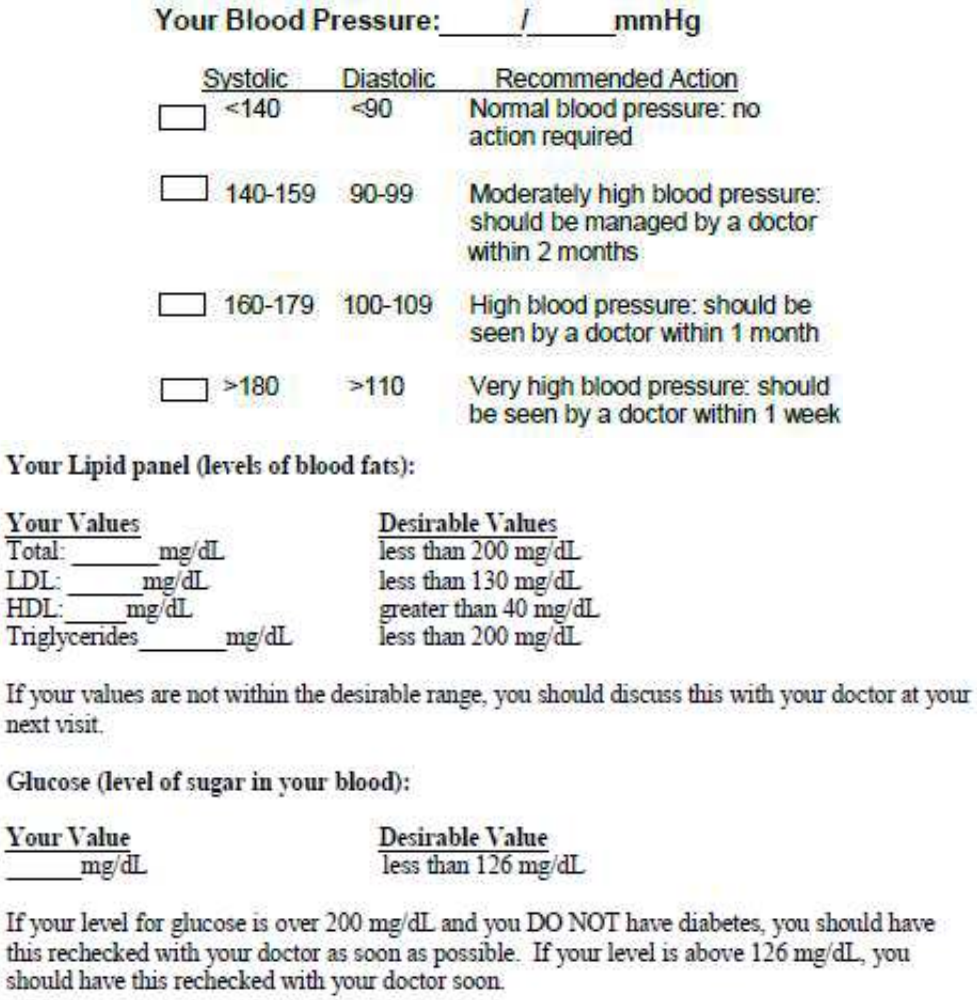
could change the coverage policies selected in health systems that make decisions based on cost-effectiveness analysis. Additionally, when analyzing public policies that expand access to screening, we caution against using treatment of diagnosed conditions as an outcome metric. Analysts who use treatment of diagnosed conditions as an outcome metric in policy analysis could risk conflating changes in the composition of diagnosed patients with declines in health system performance.

Our findings also inform policy design in multiple ways. First, in pay-for-performance systems where providers have financial incentives to maintain high treatment rates for diagnosed conditions, such as Accountable Care Organizations, expanding access to screening could carry a penalty by reducing other quality metrics. This would suggest reconsideration or reweighting of the metrics used in pay-for-performance systems, to avoid penalizing health systems that expand screening in diverse patient populations. Second, our analysis raises questions about how to target patients with low access to care. A program that screens patients with high barriers to screening could have minimal impact on treatment if these same patients also face high barriers to treatment, as shown in our policy simulation. However, our cancer analysis demonstrates that expanding access to screening could yield high impact on health if patients' barriers to treatment are also addressed.

Appendix

A Supplemental tables and figures

Figure A.1: Text from the card and letter given to REGARDS participants informing them about their blood pressure and the results of their lab tests



Your Blood Pressure:_____/_____mmHg

	Systolic	Diastolic	Recommended Action
☐	<140	<90	Normal blood pressure: no action required
☐	140-159	90-99	Moderately high blood pressure: should be managed by a doctor within 2 months
☐	160-179	100-109	High blood pressure: should be seen by a doctor within 1 month
☐	>180	>110	Very high blood pressure: should be seen by a doctor within 1 week

Your Lipid panel (levels of blood fats):

Your Values	Desirable Values
Total: _____mg/dL	less than 200 mg/dL
LDL: _____mg/dL	less than 130 mg/dL
HDL: ____mg/dL	greater than 40 mg/dL
Triglycerides_____mg/dL	less than 200 mg/dL

If your values are not within the desirable range, you should discuss this with your doctor at your next visit.

Glucose (level of sugar in your blood):

Your Value	Desirable Value
_____mg/dL	less than 126 mg/dL

If your level for glucose is over 200 mg/dL and you DO NOT have diabetes, you should have this rechecked with your doctor as soon as possible. If your level is above 126 mg/dL, you should have this rechecked with your doctor soon.

Table A.1: Definitions used for diabetes, hypertension, and high cholesterol

Condition	Status	Definition
Diabetes	No condition	No self-reported diagnosis of diabetes and FPG<126 mg/dl or NFPG<200mg/dl
	Undiagnosed	No self-reported diagnosis of diabetes, but FPG>126 mg/dl or NFPG>200mg/dl
	Diagnosed	Self-reported diagnosis of diabetes (when non-pregnant for women)
Hypertension	No condition	No self-reported diagnosis, SBP<140mmHg, and DBP<90mmHg
	Undiagnosed	No self-reported diagnosis of hypertension, but SBP>140mmHg or DBP>90mmHg
	Diagnosed	Self-reported diagnosis of hypertension (when non-pregnant for women)
High cholesterol	No condition	No self-reported diagnosis, total cholesterol <200 mg/dl, LDL cholesterol<160 mg/dl, and HDL cholesterol>40 mg/dl
	Undiagnosed	No self-reported diagnosis, but total cholesterol >200 mg/dl, LDL cholesterol>160 mg/dl, or HDL cholesterol<40 mg/dl
	Diagnosed	Self-reported diagnosis

Note: FPG=fasting plasma glucose; NFPG=non-fasting plasma glucose; SBP=systolic blood pressure; DBP=diastolic blood pressure; HDL=high-density lipoprotein, LDL= low-density lipoprotein. In the REGARDS data, we calculated LDL cholesterol using the Friedewald equation (Friedewald et al., 1972). Because neither LDL cholesterol nor triglycerides were available in the OHIE data, we could not calculate LDL cholesterol and therefore defined high cholesterol using HDL and total cholesterol only.

Table A.2: Individuals not recently screened for undiagnosed conditions are less likely to use recommended treatment for their other, diagnosed conditions (NHANES data)

		Treatment of Diagnosed Conditions			
	Obs.	Patients not recently screened for other, undiagnosed conditions	Patients recently screened for other, undiagnosed conditions	Unadjusted difference	Adjusted difference
(1) Annual eye exam for diagnosed diabetes	1,373	0.36	0.66	-0.30***	-0.25***
				(0.04)	(0.04)
[Population: patients with diagnosed diabetes]					
(2) Annual foot exam for diagnosed diabetes	792	0.42	0.73	-0.31***	-0.26***
				(0.04)	(0.04)
[Population: patients with diagnosed diabetes]					
(3) Taking medication for diagnosed high cholesterol	6,960	0.68	0.78	-0.10***	-0.06***
				(0.02)	(0.02)
[Population: patients with diagnosed high cholesterol whose doctor recommended medication]					
(4) Taking medication for diagnosed hypertension	4,038	0.76	0.89	-0.13***	-0.12***
				(0.02)	(0.02)
[Population: patients with diagnosed hypertension whose doctor recommended medication]					
Adjust for bio-marker measured severity and co-morbid conditions				No	Yes

*** p<0.01, ** p<0.05, * p<0.1. This table shows that individuals not recently screened for undiagnosed conditions are less likely to use recommended care for their other, diagnosed conditions; the last column shows that this relationship is not driven by biomarker-measured condition severity. "Screened patients" are patients who were recently screened for undiagnosed conditions, and "unscreened patients" are patients who were not recently screened for undiagnosed conditions. Patients who have been diagnosed for all conditions of interest are dropped from the sample because they can no longer be screened. Look-back periods for screening for undiagnosed diabetes and high cholesterol are two and three years, respectively. Robust standard errors are in parentheses.

Table A.3: Individuals not recently screened for undiagnosed conditions are less likely to use medication for their other, diagnosed conditions (OHIE data)

		Treatment of Diagnosed Conditions			
	Obs.	Patients not recently screened for other, undiagnosed conditions	Patients recently screened for other, undiagnosed conditions	Unadjusted difference	Adjusted difference
(1) Taking medication for diagnosed diabetes	254	0.51	0.83	-0.32***	-0.28***
				(0.06)	(0.06)
[Population: all patients with diagnosed diabetes]					
(2) Taking medication for diagnosed high cholesterol	641	0.17	0.59	-0.42***	-0.35***
				(0.03)	(0.0357)
[Population: patients with diagnosed high cholesterol]					
(3) Taking medication for diagnosed hypertension	1,140	0.41	0.77	-0.37***	-0.32***
				(0.03)	(0.03)
[Population: patients with diagnosed hypertension]					
Adjust for bio-marker measured severity and co-morbid conditions				No	Yes

*** p<0.01, ** p<0.05, * p<0.1.

This table shows the relationship between screening for undiagnosed conditions within the past year and use of medication to treat other, diagnosed conditions. The last two columns include coefficients and standard errors obtained from linear probability models after adjusting for the listed control variables. The outcomes include (1) use of diabetes medication among participants reporting prior diagnosis of diabetes; (2) use of cholesterol-lowering medication among participants reporting prior diagnosis of high cholesterol; and (3) use of anti-hypertensive medication among participants reporting prior diagnosis of hypertension.

Table A.4: Individuals not recently screened for undiagnosed conditions have worse biomarker-based health measures for their other, diagnosed conditions (OHIE data)

		Severity of Diagnosed Conditions		
	Obs.	Unscreened patients	Screened patients	Unadjusted difference
(1) HbA1c among patients with diagnosed diabetes	276	6.0	6.3	-0.22 (0.14)
(2) Total cholesterol among patients with diagnosed high cholesterol	723	222.3	213.8	8.52*** (2.83)
(3) Systolic blood pressure among patients with diagnosed hypertension	1,252	130.7	128	2.72** (1.13)
(4) Diastolic blood pressure among patients with diagnosed hypertension	1,252	82.1	85.0	2.87*** (0.78)

*** p<0.01, ** p<0.05, * p<0.1.

This table shows that individuals not recently screened for undiagnosed conditions have worse biomarkers for their other, diagnosed conditions. "Screened patients" are patients who were recently screened for currently undiagnosed conditions, and "unscreened patients" are patients who were not recently screened for currently undiagnosed conditions. Look-back periods for screening for undiagnosed diabetes and high cholesterol are two and three years, respectively. The outcomes include (1) hba1c among participants reporting prior diagnosis of diabetes; (2) total cholesterol among participants reporting prior diagnosis of high cholesterol; and (3) systolic blood pressure and (4) diastolic blood pressure among participants reporting prior diagnosis of high blood pressure. LDL cholesterol is not reported in the OHIE data and there is insufficient information to calculate it using the Friedewald et al. (1972) equation. Robust standard errors are in parentheses.

Table A.5: Individuals not recently screened for undiagnosed conditions have worse biomarker-based health measures for their other, diagnosed conditions (NHANES data)

		Severity of Diagnosed Conditions		
	Obs.	Patients not recently screened for other, undiagnosed conditions	Patients recently screened for other, undiagnosed conditions	Unadjusted difference
(1) HbA1c among patients with diagnosed diabetes	821	7.9	7.4	0.6*** (0.2)
(2) LDL among patients with diagnosed high cholesterol	2,013	132.4	125.1	7.3*** (1.9)
(3) Total cholesterol among patients with diagnosed high cholesterol	4,269	215.6	207.6	8.0*** (1.4)
(4) Systolic blood pressure among patients with diagnosed hypertension	4,774	133.2	130.5	2.7*** (0.7)
(5) Diastolic blood pressure among patients with diagnosed hypertension	4,774	75.3	73.0	2.3*** (0.5)

*** p<0.01, ** p<0.05, * p<0.1. Note: HbA1c=glycated hemoglobin; LDL= low-density lipoprotein. This table shows that individuals not recently screened for undiagnosed conditions have worse biomarkers for their other, diagnosed conditions. "Screened patients" are patients who were recently screened for currently undiagnosed conditions, and "unscreened patients" are patients who were not recently screened for currently undiagnosed conditions. Look-back periods for screening for undiagnosed diabetes and high cholesterol are two and three years, respectively. Robust standard errors are in parentheses.

Table A.6: Individuals with undiagnosed conditions have worse biomarker-based health measures for their other, diagnosed conditions (REGARDS data)

		Severity of Diagnosed Conditions		
	Obs.	Patients with undiagnosed conditions	Patients with only diagnosed conditions	Unadjusted difference
(1) FPG among patients with diagnosed diabetes	1,132	148	132.3	15.7*** (4.4)
(2) LDL among patients with diagnosed high cholesterol	3,016	117	109.6	7.4*** (2.8)
(3) Total cholesterol among patients with diagnosed high cholesterol	3,542	197.1	187.2	9.9*** (3.0)
(4) Systolic blood pressure among patients with diagnosed hypertension	4,169	136.9	133	3.9*** (0.8)
(5) Diastolic blood pressure among patients with diagnosed hypertension	4,169	78	75.6	2.4*** (0.4)

*** p<0.01, ** p<0.05, * p<0.1.

Note: FPG=fasting plasma glucose; LDL= low-density lipoprotein. This table shows that individuals not recently screened for undiagnosed conditions have worse biomarkers for their other, diagnosed conditions. We use FPG rather than glycated hemoglobin (HbA1c) in the REGARDS data because HbA1c is not measured. In addition, we calculated LDL cholesterol in the REGARDS data using the Friedewald equation (Friedewald et al., 1972). Robust standard errors are in parentheses.

Table A.7: Simulation analysis: Reduced screening and subsequent treatment of diagnosed conditions

A. All patients

	Diabetes		
Who is screened?	**Cases detected**	**Average treatment if diagnosed**	**Foot and eye exams if diagnosed**
Propensity 0.7 and higher	1931	89%	42%
Propensity 0.6 and higher	2636	88%	41%
Propensity 0.5 and higher	3146	88%	39%
Propensity 0.4 and higher	3373	87%	39%
Propensity 0.3 and higher	3499	87%	38%
Propensity 0.2 and higher	3573	86%	38%
Propensity 0.1 and higher	3621	86%	38%
Everyone	3621	86%	38%

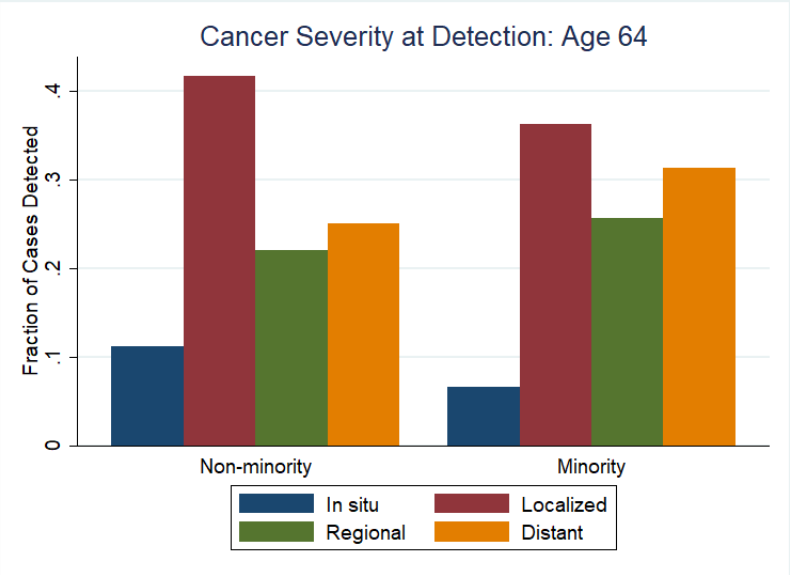
	High Cholesterol		
Who is screened?	**Cases detected**	**Average treatment if diagnosed**	**Average treatment if doctor recommended treatment**
Propensity 0.7 and higher	6793	62%	85%
Propensity 0.6 and higher	9238	58%	84%
Propensity 0.5 and higher	11174	55%	82%
Propensity 0.4 and higher	12595	53%	81%
Propensity 0.3 and higher	13577	51%	80%
Propensity 0.2 and higher	14501	50%	80%
Propensity 0.1 and higher	15106	50%	79%
Everyone	15608	50%	79%

B. Patients without health insurance

	Diabetes	
Who is screened?	**Average treatment if diagnosed**	**Foot and eye exams if diagnosed**
Propensity 0.7 and higher	86%	24%
Propensity 0.6 and higher	87%	25%
Propensity 0.5 and higher	85%	24%
Propensity 0.4 and higher	82%	22%
Propensity 0.3 and higher	80%	20%
Propensity 0.2 and higher	78%	20%
Propensity 0.1 and higher	77%	19%
Everyone	77%	19%

	High Cholesterol	
Who is screened?	**Average treatment if diagnosed**	**Average treatment if doctor recommended treatment**
Propensity 0.7 and higher	45%	72%
Propensity 0.6 and higher	50%	74%
Propensity 0.5 and higher	48%	73%
Propensity 0.4 and higher	44%	70%
Propensity 0.3 and higher	40%	67%
Propensity 0.2 and higher	36%	65%
Propensity 0.1 and higher	34%	64%
Everyone	33%	63%

Figure A.2: Severity of cancer at detection and racial/ethnic minority status at age 64



This graph shows that racial and ethnic minority patients (Hispanic and/or non-white patients) had cancers detected at a later stage that non-minority patients at age 64, just prior to aging into Medicare. In situ is the earliest stage of cancer, followed by localized, regional, and distant. Regional and distant cancers have already infiltrated multiple organs. These patterns are consistent with lower access to care among minority patients prior to aging into Medicare eligibility (Card et al., 2008).

B Formal theoretical model

To supplement the intuitive discussion in the main text, we analyze a model of demand for screening and demand for medical treatment after diagnosis to show one reason why these two outcomes could be correlated on the patient-level. The purpose of this exercise is to demonstrate that although our empirical findings may be consistent with any number of theoretical models, no departure from a rational choice model of patients' health care consumption is required to account for our findings. To allow our model to apply to agents not in the labor force, we follow Picone et al. (1998) and others by directly including health in the utility function.

In the model, agents use medical treatment to ameliorate the negative effects of chronic health conditions. Agents who have been recently screened know whether they have a chronic condition, whereas agents who have not been recently screened hold beliefs about the probability they have a chronic condition. Agents differ only in their costs of medical treatment; we separately model pecuniary and non-pecuniary costs. We analyze this model to derive predictions about which agents are willing to pay more for screening.

Agents maximize a continuously differentiable function of health (H) and consumption (C), net of disutility of medical treatment. Disutility of medical treatment due to non-pecuniary costs is linear in units of medical treatment M and the magnitude of disutility from non-pecuniary costs is captured by θ, which varies across agents.[18] The utility function is therefore:

$$u\left(C,\, H\left(M,\, D\right)\right) - \theta M$$

$u\left(\cdot\right)$ is concave in C and H, and agents have weakly higher marginal utility from consumption when they are healthier.

Health does not affect income, as in the pure consumption version of the Grossman model (see Grossman (2000)). To keep notation simple, we assume that agents have assets A and receive no further income. If an agent has a chronic condition, then $D = 1$; otherwise, $D = 0$. If $D = 1$ and the agent has been diagnosed, then he must decide how to divide his funds between medical treatment ($M \geq 0$ units, purchased at a price P per unit where P can vary across agents), and other consumption (C). This yields the budget constraint:

$$C + PM = A$$

If the agent does not have a diagnosed condition, he is not eligible to receive medical treat-

[18] Non-pecuniary costs could be related to factors such as language barriers, distance to a provider, depression symptoms or other psychological factors which provide barriers to accessing care.

ment. In this case, therefore, the entire budget is spent on other consumption: $C = A$.

Health H is a function of medical treatment M and chronic condition status D, as follows. When agents have a chronic condition, health becomes worse: $H(M, 0) > H(M, 1) \forall M$. However, medical treatment improves health for agents with chronic conditions: $\frac{\partial H(M,1)}{\partial M} > 0 \forall M$.

Because doctors only provide medical treatment to patients who are diagnosed for a condition, an agent's utility and decision variables vary based on whether he has been screened and the results of the screening. There are three possible cases:

1. The agent has not been recently screened and does not know whether he has a chronic condition, but has (correct) beliefs about π, the probability that he has a chronic condition. Because the agent is not diagnosed, he cannot receive medical treatment ($M = 0$) and therefore uses all funds for consumption. His expected utility is therefore:

$$\pi u(A, H(0, 1)) + (1 - \pi) u(A, H(0, 0)) \tag{3}$$

2. The agent has been recently screened and knows he does not have a chronic condition ($D = 0$).[19] He is not eligible for medical treatment and therefore uses all funds for consumption. His utility is:

$$u(A, H(0, 0)) \tag{4}$$

3. The agent has been recently screened and knows he has a chronic condition ($D = 1$). Therefore, the agent can choose to use medical treatment. As such, the agent selects M and C to maximize his utility:

$$\max_{C, M} u(C, H(M, 1)) - \theta M \tag{5}$$

   subject to $C + PM = A$.

Screening moves agents from case (1) to case (2) or (3) depending on the results of the test.

Equations (3), (4), and (5) can be combined to describe agents' willingness to pay for screening. In particular, agents are indifferent between being screened and not being screened at out-of-pocket price of screening κ if:

$$\begin{aligned} \pi \left(\max_{M} u(A - PM - \kappa, H(M, 1)) - \theta M \right) + (1 - \pi) u(A - \kappa, H(0, 0)) \\ - (\pi u(A, H(0, 1)) + (1 - \pi) u(A, H(0, 0))) = 0 \end{aligned} \tag{6}$$

[19] For simplicity, we present the case where the test is perfectly informative. This assumption can be relaxed without altering the main results.

We can then define κ^* as the price of screening that makes any given agent just indifferent between being screened and not being screened. As such, κ^* captures the agent's willingness to pay for screening.

Optimal decisions if screened and $D = 1$

In this case, the agent is eligible for medical treatment and can choose his consumption of medical treatment and other goods. The optimal solutions, denoted M^* and C^*, are defined by the first order condition:

$$\frac{\partial u\left(C,\, H\left(M^*,\, 1\right)\right)}{\partial H}\frac{\partial H\left(M^*,\, 1\right)}{\partial M} - \theta = P\frac{\partial u\left(C^*,\, H\left(M^*,\, 1\right)\right)}{\partial C} \tag{7}$$

The left-hand side of Equation (7) indicates the utility gains from consuming a unit of medical treatment. $\frac{\partial u(C, H(M^*,1))}{\partial H}\frac{\partial H(M^*,1)}{\partial M}$ is the utility benefit from improved health and $-\theta$ is the disutility of consuming a unit of medical treatment due to non-pecuniary costs. The right-hand side of Equation (7) indicates the utility gains from spending P additional dollars on consumption rather than on medical treatment. Therefore Equation (7) indicates that at the optimal point, the marginal benefits of purchasing a unit of medical treatment equal the marginal benefits of using the same funds for consumption.

Analysis of marginally screened individuals and empirical predictions

We now show that agents who become willing to be screened after a decrease in the out-of-pocket price of screening use less medical treatment after diagnosis than already screened individuals. This follows from two propositions.

Proposition B.1 *Willingness to pay for screening is decreasing in agents' costs of medical treatment:* $\frac{\partial \kappa^*}{\partial \theta} < 0$ *and* $\frac{\partial \kappa^*}{\partial P} < 0$*, respectively.*

The proofs are based on the envelope theorem. See Appendix B.

Proposition B.2 *Demand for medical treatment after diagnosis is also decreasing in agents' costs of medical treatment:* $\frac{\partial M^*}{\partial \theta} < 0$ *and* $\frac{\partial M^*}{\partial P} < 0$.

See Appendix B for the proofs.

Based on these propositions, higher costs of medical treatment decrease agents' demand for medical treatment after diagnosis, and also decrease agents' willingness to pay for screening. The implications for a policy that decreases the out-of-pocket price of screening when costs of medical treatment vary across agents are as follows. First, decreasing the out-of-pocket price of screening will attract agents with marginally lower willingness-to-pay for screening (κ^*) to become screened. Agents with marginally lower κ^* will also face marginally higher costs (θ and/or P) by Proposition 2.1. In turn, higher costs for medical treatment imply that these agents will use less medical treatment for their diagnosed conditions than previously screened agents by Proposition 2.2. This produces the empirical prediction that patients whose conditions become diagnosed because of a decline in the out-of-pocket price of screening use less medical treatment for their conditions after diagnosis.

Based on these propositions, higher costs of medical treatment decrease agents' demand for medical treatment after diagnosis, and also decrease agents' willingness to pay for screening. The implications for a policy that decreases the out-of-pocket price of screening when costs of medical treatment vary across agents are as follows. First, decreasing the out-of-pocket price of screening will attract agents with marginally lower willingness-to-pay for screening (κ^*) to become screened. Agents with marginally lower κ^* will also face marginally higher costs (θ and/or P) by Proposition 2.1. In turn, higher costs for medical treatment imply that these agents will use less medical treatment for their diagnosed conditions than previously screened agents by Proposition 2.2. This produces the empirical prediction that patients whose conditions become diagnosed because of a decline in the out-of-pocket price of screening use less medical treatment for their conditions after diagnosis.

Proofs: Demand for screening is (weakly) decreasing in θ and P

Demand for screening is weakly decreasing in θ

When the price of screening equals willingness to pay for screening κ^*, agents are just indifferent between being screened and not being screened as follows:

$$\pi\left(\max_{M} u\left(A - PM - \kappa^*,\, H\left(M,\, 1\right)\right) - \theta M\right) + \left(1-\pi\right) u\left(A - \kappa^*,\, H\left(0,\, 0\right)\right) \tag{8}$$
$$- \left(\pi u\left(A,\, H\left(0,\, 1\right)\right) + \left(1-\pi\right) u\left(A,\, H\left(0,\, 0\right)\right)\right) = 0$$

Differentiating (8) with respect to θ yields the following expression (by the envelope theorem, we can ignore the fact that the optimal M varies with θ):

$$\pi\left(-\frac{\partial u\left(A-PM^{*}-\kappa^{*},\, H\left(M^{*},\, 1\right)\right)}{\partial C}\frac{\partial \kappa^{*}}{\partial \theta}-M^{*}\right)-(1-\pi)\,\frac{\partial u\left(A-\kappa^{*},\, H\left(0,\, 0\right)\right)}{\partial C}\frac{\partial \kappa^{*}}{\partial \theta}=0 \tag{9}$$

Then rearranging to solve for $\frac{\partial \kappa^{*}}{\partial \theta}$ yields:

$$-\left(\pi\frac{\partial u\left(A-PM^{*}-\kappa^{*},\, H\left(M^{*},\, 1\right)\right)}{\partial C}\frac{\partial \kappa^{*}}{\partial \theta}\right)-(1-\pi)\,\frac{\partial u\left(A-\kappa^{*},\, H\left(0,\, 0\right)\right)}{\partial C}\frac{\partial \kappa^{*}}{\partial \theta}=\pi M^{*}$$

$$\frac{\partial \kappa^{*}}{\partial \theta}\left(-\pi\frac{\partial u\left(A-PM^{*}-\kappa^{*},\, H\left(M^{*},\, 1\right)\right)}{\partial C}-(1-\pi)\,\frac{\partial u\left(A-\kappa^{*},\, H\left(0,\, 0\right)\right)}{\partial C}\right)=\pi M^{*}$$

$$\implies \frac{\partial \kappa^{*}}{\partial \theta}=-\frac{\pi M^{*}}{\pi\frac{\partial u(A-PM^{*}-\kappa^{*},\, H(M^{*},1))}{\partial C}+(1-\pi)\,\frac{\partial u(A-\kappa^{*},\, H(0,0))}{\partial C}}\leq 0$$

We conclude $\frac{\partial \kappa^{*}}{\partial \theta}\leq 0$ because $\frac{\partial u}{\partial C}>0$, $\pi>0$ and $M^{*}\geq 0$.

Demand for screening is decreasing in P

When the price of screening equals willingness to pay for screening κ^{*}, agents are just indifferent between being screened and not being screened as follows:

$$\begin{aligned}\pi\left(\max_{M} u\left(A-PM-\kappa^{*},\, H\left(M,\, 1\right)\right)-\theta M\right)+(1-\pi)\,u\left(A-\kappa^{*},\, H\left(0,\, 0\right)\right) \\ -\left(\pi u\left(A,\, H\left(0,\, 1\right)\right)+(1-\pi)\,u\left(A,\, H\left(0,\, 0\right)\right)\right)=0\end{aligned} \tag{10}$$

Differentiating (10) with respect to P yields the following expression (by the envelope theorem, we can ignore the fact that the optimal M varies with P):

$$\begin{aligned}\pi\left(-\frac{\partial u\left(A-PM^{*}-\kappa^{*},\, H\left(M^{*},\, 1\right)\right)}{\partial C}\frac{\partial \kappa^{*}}{\partial P}+\frac{\partial u\left(A-PM^{*}-\kappa^{*},\, H\left(M^{*},\, 1\right)\right)}{\partial C}\right) \\ -(1-\pi)\,\frac{\partial u\left(A-\kappa^{*},\, H\left(0,\, 0\right)\right)}{\partial C}\frac{\partial \kappa^{*}}{\partial P}=0\end{aligned} \tag{11}$$

Then rearranging to solve for $\frac{\partial \kappa^{*}}{\partial P}$ yields:

$$-\left(\pi\frac{\partial u\left(A-PM^{*}-\kappa^{*},\,H\left(M^{*},\,1\right)\right)}{\partial C}\frac{\partial\kappa^{*}}{\partial P}\right)-\left(1-\pi\right)\frac{\partial u\left(A-\kappa^{*},\,H\left(0,\,0\right)\right)}{\partial C}\frac{\partial\kappa^{*}}{\partial P}$$
$$=\pi\frac{\partial u\left(A-PM^{*}-\kappa^{*},\,H\left(M^{*},\,1\right)\right)}{\partial C}$$
$$\frac{\partial\kappa^{*}}{\partial P}\left(-\pi\frac{\partial u\left(A-PM^{*}-\kappa^{*},\,H\left(M^{*},\,1\right)\right)}{\partial C}-\left(1-\pi\right)\frac{\partial u\left(A-\kappa^{*},\,H\left(0,\,0\right)\right)}{\partial C}\right)$$
$$=\pi\frac{\partial u\left(A-PM^{*}-\kappa^{*},\,H\left(M^{*},\,1\right)\right)}{\partial C}$$

$$\implies\frac{\partial\kappa^{*}}{\partial P}=-\frac{\pi\frac{\partial u(A-PM^{*}-\kappa^{*},H(M^{*},1))}{\partial C}}{\pi\frac{\partial u(A-PM^{*}-\kappa^{*},H(M^{*},1))}{\partial C}+\left(1-\pi\right)\frac{\partial u(A-\kappa^{*},H(0,0))}{\partial C}}<0$$

We conclude $\frac{\partial\kappa^{*}}{\partial P}<0$ because $\frac{\partial u}{\partial C}>0$ and $\pi>0$.

Proofs: Demand for medical treatment is decreasing in θ and P

Demand for medical treatment is decreasing in θ

We show that agents must demand less medical treatment when they have higher non-pecuniary costs of treatment (captured by θ), because to do otherwise would violate the first-order conditions.

Consider the optimal decisions when agents know that $D=1$. (This is the only case where purchase of medical treatment is an option, because medical treatment is not available without a prescription.) Now consider that θ decreases from $\overline{\theta}$ to $\underline{\theta}$. Let $M_{\overline{\theta}}$ and $C_{\overline{\theta}}$ denote the optimal decisions before the change and $M_{\underline{\theta}}$ and $C_{\underline{\theta}}$ denote the optimal decisions after the change.

$M_{\overline{\theta}}$ and $C_{\overline{\theta}}$ must fulfill the first-order conditions summarized in equation (7), as follows:

$$\frac{\partial u\left(C_{\overline{\theta}},\,H\left(M_{\overline{\theta}},\,1\right)\right)}{\partial H}\frac{\partial H\left(M_{\overline{\theta}},\,1\right)}{\partial M}-\overline{\theta}=P\frac{\partial u\left(C_{\overline{\theta}},\,H\left(M_{\overline{\theta}},\,1\right)\right)}{\partial C}\tag{12}$$

After non-pecuniary cost decreases from $\overline{\theta}$ to $\underline{\theta}$, previously optimal decisions $M_{\overline{\theta}}$ and $C_{\overline{\theta}}$ would violate equation (7) as follows:

$$\frac{\partial u\left(C_{\overline{\theta}},\,H\left(M_{\overline{\theta}},\,1\right)\right)}{\partial H}\frac{\partial H\left(M_{\overline{\theta}},\,1\right)}{\partial M}-\underline{\theta}>P\frac{\partial u\left(C_{\overline{\theta}},\,H\left(M_{\overline{\theta}},\,1\right)\right)}{\partial C}\tag{13}$$

To make inequality (13) an equality, M and C must change so that the left-hand side decreases and/or the right-hand side increases. By concavity of the utility function in H and C, the weakly positive cross-partial $\frac{\partial^{2}u(C,H)}{\partial C\partial H}$, and weakly decreasing marginal returns

to medical care, increasing M and decreasing C achieves both. Therefore $M_{\overline{\theta}} < M_{\underline{\theta}}$ and $C_{\overline{\theta}} > C_{\underline{\theta}}$ resolves the contradiction in the first-order conditions. We conclude that $\frac{\partial M^*}{\partial \theta} < 0$.

Demand for medical treatment is decreasing in P

We show that agents must demand less treatment when they have higher cost of medical treatment P, because to do otherwise would violate the first-order conditions.

Consider the optimal decisions when agents know that $D = 1$. (This is the only case where purchase of medical treatment is an option, because medical treatment is not available without a prescription.) Now consider that P decreases from $\overline{P}$ to $\underline{P}$. Let $M_{\overline{P}}$ and $C_{\overline{P}}$ denote the optimal decisions before the change and $M_{\underline{P}}$ and $C_{\underline{P}}$ denote the optimal decisions after the change.

$M_{\overline{P}}$ and $C_{\overline{P}}$ must fulfill the first-order conditions summarized in equation (7), as follows:

$$\frac{\partial u\left(C_{\overline{P}},\, H\left(M_{\overline{P}},\, 1\right)\right)}{\partial H}\frac{\partial H\left(M_{\overline{P}},\, 1\right)}{\partial M} - \theta = \overline{P}\frac{\partial u\left(C_{\overline{P}},\, H\left(M_{\overline{P}},\, 1\right)\right)}{\partial C} \tag{14}$$

After cost of care P decreases from $\overline{P}$ to $\underline{P}$, previously optimal decisions $M_{\overline{P}}$ and $C_{\overline{P}}$ would violate equation (7) as follows:

$$\frac{\partial u\left(C_{\overline{P}},\, H\left(M_{\overline{P}},\, 1\right)\right)}{\partial H}\frac{\partial H\left(M_{\overline{P}},\, 1\right)}{\partial M} - \theta > \underline{P}\frac{\partial u\left(C_{\overline{P}},\, H\left(M_{\overline{P}},\, 1\right)\right)}{\partial C} \tag{15}$$

To make inequality (15) an equality, M and C must change so that the left-hand side decreases and/or the right-hand side increases. As before, increasing M and decreasing C achieves both. Therefore $M_{\overline{P}} < M_{\underline{P}}$ and $C_{\overline{P}} > C_{\underline{P}}$ resolves the contradiction in the first-order conditions. We conclude that $\frac{\partial M^*}{\partial P} < 0$.

References

Agency for Healthcare Research and Quality (2013). 2012 National Healthcare Quality Report. Technical report, US Department of Health and Human Services, Rockville, MD.

Ahn, H. S., H. J. Kim, and H. G. Welch (2014). Korea's thyroid-cancer "epidemic" - screening and overdiagnosis. *New England Journal of Medicine 371*(19), 1765–1767. PMID: 25372084.

Allen, H., K. Baicker, A. Finkelstein, S. Taubman, and B. J. Wright (2010, August). What the Oregon health study can tell us about expanding Medicaid. *Health Affairs (Project Hope) 29*(8), 1498–506.

American Diabetes Association (2014, January). Standards of Medical Care in Diabetes–2014. *Diabetes Care 37 Suppl 1*(October 2013), S14–80.

Baicker, K., S. L. Taubman, H. L. Allen, M. Bernstein, J. H. Gruber, J. P. Newhouse, E. C. Schneider, B. J. Wright, A. Zaslavsky, A. N. Finkelstein, M. Carlson, T. Edlund, C. Gallia, and J. Smith (2013, May). The Oregon experiment–effects of Medicaid on clinical outcomes. *The New England journal of medicine 368*(18), 1713–22.

Basu, A. (2011). Economics of individualization in comparative effectiveness research and a basis for a patient-centered health care. *Journal of Health Economics*.

Basu, A. (2013). Estimating person-centered treatment effects using instrumental variables: an application to evaluating prostate cancer. *Journal of Applied Econometrics*.

Basu, A. and J. Heckman (2007). Use of instrumental variables in the presence of heterogeneity and self selection: an application to treatments of breast cancer patients. *Health Economics 1157*(October), 1133–1157.

Basu, A. and W. Manning (2009). Issues for the next generation of health care cost analyses. *Medical Care 47*(7).

Basu, A. and D. Meltzer (2007). Value of information on preference heterogeneity and individualized care. *Medical Decision Making 27*(2), 112–27.

Bleyer, A. and H. G. Welch (2012). Effect of three decades of screening mammography on breast-cancer incidence. *New England Journal of Medicine 367*(21), 1998–2005. PMID: 23171096.

Boozer, M. A. and T. J. Philipson (2000). The impact of public testing for human immunodeficiency virus. *The Journal of Human Resources 35*(3), 419–446.

Calonico, S., M. Cattaneo, and R. Titiunik (2014a). Robust data-driven inference in the regression-discontinuity design. *Stata Journal 14*(4), 909–946.

Calonico, S., M. D. Cattaneo, and R. Titiunik (2014b). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica 82*(6), 2295–2326.

Card, D., C. Dobkin, and N. Maestas (2008). The impact of nearly universal insurance coverage on health care utilization and health: evidence from Medicare. *American Economic Review 98*(5), 2242–2258.

Carneiro, P., J. Heckman, and E. Vytlacil (2010). Estimating marginal returns to education. *101*(October), 2754–2781.

Carpenter, C. J. (2010, December). A meta-analysis of the effectiveness of health belief model variables in predicting behavior. *Health Communication 25*(8), 661–9.

Casalino, L. P., A. Elster, A. Eisenberg, E. Lewis, J. Montgomery, and D. Ramos (2007, May). Will Pay-For-Performance And Quality Reporting Affect Health Care Disparities? *Health Affairs 26*(3), w405–w414.

CDC (2012). CDC - A Diabetes Report Card for the United States Quality of care in the 1990s - Publications - Diabetes DDT. *136*(8), 565–574.

Center for Medicare and Medicaid Services (2011). Guide to Quality Performance Scoring Methods for Accountable Care Organizations.

Center for Medicare and Medicaid Services (2016a). Publicly available ACO data and ACO performance data sources maintained by CMS. (March).

Center for Medicare and Medicaid Services (2016b). Quality Measures, Reporting and Performance Standards.

Centers for Disease Control and Prevention (1999-2014). National Health and Nutrition Examination Survey.

Chang, F. R. (1996). Uncertainty and investment in health. *Journal of Health Economics 15*(3), 369–376.

Cowie, C., K. Rust, and E. Ford (2009). Full accounting of diabetes and pre-diabetes in the US population in 1988-1994 and 2005-2006. *Diabetes ... 32*(2), 0–7.

Dale, S. B., A. Ghosh, D. N. Peikes, T. J. Day, F. B. Yoon, E. F. Taylor, K. Swankoski, A. S. O'Malley, P. H. Conway, R. Rajkumar, M. J. Press, L. Sessums, and R. Brown (2016). Two-year costs and quality in the comprehensive primary care initiative. *New England Journal of Medicine 374*(24), 2345–2356. PMID: 27074035.

Dardanoni, V. and A. Wagstaff (1987). Uncertainty, inequalities in health and the demand for health. *Journal of Health Economics 6*, 283–290.

Dranove, D., D. Kessler, M. McClellan, and M. Satterthwaite (2003, June). Is More Information Better? The Effects of and quot;Report Cards" on Health Care Providers. *Journal of Political Economy 111*(3), 555–588.

Du, X. L., T. E. Meyer, and L. Franzini (2007). Meta-analysis of racial disparities in survival in association with socioeconomic status among men and women with colon cancer. *Cancer 109*(11), 2161–2170.

Egan, M. and T. Philipson (2014, July). Health Care Adherence and Personalized Medicine. NBER Working Papers 20330, National Bureau of Economic Research, Inc.

Eisenhauer, P., J. Heckman, and E. Vytlacil (2010). Generalized Roy Model and the Cost-Benefit Analysis of Social Programs. *Policy Analysis*.

Finkelstein, A. (2013). Oregon Health Insurance Experiment Public Use Data.

Finkelstein, A., M. Gentzkow, P. Hull, and H. Williams (2017). Adjusting risk adjustment - accounting for variation in diagnostic intensity. *New England Journal of Medicine 376*(7), 608–610. PMID: 28199802.

Fisher, A. E. S., D. C. Goodman, and A. Chandra (2008). Regional and Racial Variation in Health Care among Medicare Beneficiaries: A Brief Report of the Dartmouth Atlas Project. Technical report, The Dartmouth Institute for Health Policy and Clinical Practice.

Friedewald, W., R. Levy, and D. Fredrickson (1972). Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clinical chemistry 18*(6), 499–502.

Glanz, K. and D. Bishop (2010, January). The role of behavioral science theory in development and implementation of public health interventions. *Annual Review of Public Health 31*, 399–418.

Global Burden of Disease Collaborators (2013, August). The state of US health, 1990-2010: burden of diseases, injuries, and risk factors. *JAMA 310*(6), 591–608.

Glümer, C., M. Yuyun, S. Griffin, D. Farewell, D. Spiegelhalter, A. L. Kinmonth, and N. J. Wareham (2006). What determines the cost-effectiveness of diabetes screening? *Diabetologia 49*, 1536–1544.

Godlonton, S. and R. Thornton (2012). Peer effects in learning hiv results. *Journal of Development Economics 97*(1), 118 – 129.

Goldman, D. and T. Philipson (2007). Insurance Design in the Presence of Integrated Medical Multiple Technologies. *American Economic Review 97*(2), 427–432.

Grossman, M. (1972). On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy 80*(2), 223.

Grossman, M. (1982). The demand for health after a decade. *Journal of Health Economics 1*(1), 1–3.

Grossman, M. (2000). The human capital model. *Handbook of Health Economics 1*(Part A), 347–408.

Grundy, S. M. (2004, January). Definition of Metabolic Syndrome: Report of the National Heart, Lung, and Blood Institute/American Heart Association Conference on Scientific Issues Related to Definition. *Circulation 109*(3), 433–438.

Harris, A. H. S., C. Chen, A. D. Rubinsky, K. J. Hoggatt, M. Neuman, and M. E. Vanneman (2016, April). Are Improvements in Measured Performance Driven by Better Treatment or "Denominator Management"? *Journal of General Internal Medicine 31*(S1), 21–27.

Heckman, J. (2010, June). Building Bridges Between Structural and Program Evaluation Approaches to Evaluating Policy. *Journal of Economic Literature 48*(2), 356–398.

Hoerger, T., R. Harris, K. Hicks, K. Donahue, and S. Sorensen (2004). Screening for type 2 diabetes mellitus: a cost-effectiveness analysis. *Annals of Internal Medicine 140*(9), 689–699.

Howard, V., M. Cushman, L. Pulley, C. Gomez, R. Go, R. Prineas, A. Graham, C. Moy, and G. Howard (2005). The reasons for geographic and racial differences in stroke study: objectives and design. *Neuroepidemiology 25*, 135–143.

Howard, V. J., D. O. Kleindorfer, S. E. Judd, L. A. Mcclure, M. M. Safford, J. D. Rhodes, M. Cushman, C. S. Moy, E. Z. Soliman, B. M. Kissela, and G. Howard (2011). Disparities in Stroke Incidence Contributing to Disparities in Stroke Mortality. *Annals of Neurology 69*(4), 619–627.

Huang, E. S., M. Shook, L. Jin, M. H. Chin, and D. O. Meltzer (2006, February). The impact of patient preferences on the cost-effectiveness of intensive glucose control in older patients with new-onset diabetes. *Diabetes Care 29*(2), 259–64.

Hyman, R. B., S. Baker, R. Ephraim, A. Moadel, and J. Philip (1994, August). Health Belief Model variables as predictors of screening mammography utilization. *Journal of Behavioral Medicine 17*(4), 391–406.

Imbens, G. W. and T. Lemieux (2008, February). Regression discontinuity designs: A guide to practice. *Journal of Econometrics 142*(2), 615–635.

James, P., S. Oparil, B. Carter, W. Cushman, C. Dennison-Himmelfarb, J. Handler, D. Lackland, M. LeFevre, T. MacKenzie, O. Ogedegbe, S. Smith, L. Svetkey, S. Taler, R. Townsend, J. Wright, A. Narva, and E. Ortiz (2014, February). 2014 evidence-based guideline for the management of high blood pressure in adults: report from the panel members appointed to the Eighth Joint National Committee (JNC 8). *JAMA 311*(5), 507–20.

Jatoi, I., H. Becher, and C. R. Leake (2003). Widening disparity in survival between white and african-american patients with breast carcinoma treated in the u. s. department of defense healthcare system. *Cancer 98*(5), 894–899.

Kadiyala, S. and E. Strumpf (2016). How effective is population-based cancer screening? regression discontinuity estimates from the us guideline screening initiation ages. *Forum for Health Economics and Policy 19*(1), 87–139.

Kahn, R., P. Alperin, D. Eddy, K. Borch-Johnsen, J. Buse, J. Feigelman, E. Gregg, R. R.

Holman, M. S. Kirkman, M. Stern, J. Tuomilehto, and N. J. Wareham (2010, April). Age at initiation and frequency of screening to detect type 2 diabetes: a cost-effectiveness analysis. *Lancet (London, England) 375*(9723), 1365–74.

Karve, A. M., F.-S. Ou, B. L. Lytle, and E. D. Peterson (2008, March). Potential unintended financial consequences of pay-for-performance on the quality of care for minority patients. *American Heart Journal 155*(3), 571–576.

Kaufman, H. W., Z. Chen, V. A. Fonseca, and M. J. McPhaul (2015). Surge in newly identified diabetes among medicaid patients in 2014 within medicaid expansion states under the affordable care act. *Diabetes Care*.

Kenkel, D. (1994). The demand for preventive medical care. *Applied Economics 26*(4), 313–325.

Kim, H. B., S. A. Lee, and W. Lim (2017). Knowing is Not Half the Battle: Impacts of the National Health Screening Program in Korea. *IZA diecussion paper series*.

Kim, H. B. and S.-m. Lee (2017). When public health intervention is not successful: Cost sharing, crowd-out, and selection in korea's national cancer screening program. *Journal of Health Economics 53*(Supplement C), 100 – 116.

Konetzka, R. T., D. Polsky, and R. M. Werner (2013, March). Shipping out instead of shaping up: Rehospitalization from nursing homes as an unintended effect of public reporting. *Journal of Health Economics 32*(2), 341–352.

Lange, F. (2011). The role of education in complex health decisions: Evidence from cancer screening. *Journal of Health Economics 30*(1), 43 – 54.

Lantz, P., D. Macklem, M. Hare, L. Richardson, L. Sever, and C. Orians (1997). Follow-up and treatment issues in the National Breast and Cervical Cancer Early Detection Program: results from a multiple-site case study - final report. *Baltimore Battelle, Centers for Public Health Research and Evaluation*.

Lee, D. S. and T. Lemieux (2009, February). Regression Discontinuity Designs in Economics. NBER Working Papers 14723, National Bureau of Economic Research, Inc.

Liljas, B. (1998). The demand for health with uncertainty and insurance. *Journal of Health Economics 17*(2), 153–170.

Loeb, S., M. Bjurlin, J. Nicholson, T. L. Tammela, D. Penson, H. B. Carter, and R. Etzioni (2014). Overdiagnosis and overtreatment of prostate cancer. *European Urology 65*(6), 1046–1055.

Lostao, L., T. E. Joiner, J. W. Pettit, P. Chorot, and B. Sandín (2001). Health beliefs and illness attitudes as predictors of breast cancer screening attendance. *European Journal of Public Health 11*(3), 274–279.

Manning, W., J. P. Newhouse, N. Duan, E. B. Keeler, a. Leibowitz, and M. S. Marquis (1987,

June). Health insurance and the demand for Medical Care: evidence from a randomized experiment. *The American economic review 77*(3), 251–77.

McDonald, M., R. Hertz, A. Unger, and M. Lustik (2009). Prevalence, awareness, and management of hypertension, dyslipidemia, and diabetes among United States adults aged 65 and older. *Journals of Gerontology - Series A Biological Sciences and Medical Sciences 64*(2), 256–263.

McWilliams, J. M., E. Meara, A. M. Zaslavsky, and J. Z. Ayanian (2009). Differences in Control of Cardiovascular Disease and Diabetes by Race, Ethnicity, and Education: U.S. Trends From 1999 to 2006 and Effects of Medicare Coverage. *Annals of Internal Medicine*, 505–516.

McWilliams, J. M., E. Meara, A. M. Zaslavsky, and J. Z. Ayanian (2013). Health of Previously Uninsured Adults After Acquiring Medicare Coverage. *JAMA : the journal of the American Medical Association 298*(24).

Meltzer, D. and E. Huang (2007). Effects of patient self-selection on cost-effectiveness of diabetes therapy by age. *iHEA 2007 6th World ...*, 2–3.

Muntner, P., L. Colantonio, M. Cushman, D. Goff, G. Howard, V. Howard, B. Kissela, E. Levitan, D. Lloyd-Jones, and M. Safford (2014, April). Validation of the atherosclerotic cardiovascular disease Pooled Cohort risk equations. *JAMA 311*(14), 1406–15.

Musa, D., R. Schulz, R. Harris, M. Silverman, and S. B. Thomas (2009, Jul). Trust in the health care system and the use of preventive health services by older black and white adults. *American Journal of Public Health 99*(7), 1293–9.

Myerson, R. and N. Laiteerapong (2016). The affordable care act and diabetes diagnosis and care: Exploring the potential impacts. *Current Diabetes Reports 16*(4), 27.

Myerson, R. M., L. D. Colantonio, M. M. Safford, and E. S. Huang (2017). Does identification of previously undiagnosed conditions change care-seeking behavior? *Health Services Research*, n/a–n/a.

Nathan, D. and W. Herman (2004). Screening for Diabetes: Can We Afford Not to Screen? *Annals of Internal Medicine 140*(9), 756–758+I6.

National Committee on Quality Assurance (2016). HEDIS Measures and Performance Measurement.

Olives, C., R. Myerson, A. H. Mokdad, C. J. L. Murray, and S. S. Lim (2013, January). Prevalence, awareness, treatment, and control of hypertension in United States counties, 2001-2009. *PloS one 8*(4), e60308.

Oster, E., I. Shoulson, and E. R. Dorsey (2013, April). Optimal Expectations and Limited Medical Testing: Evidence from Huntington Disease. *American Economic Review 103*(2), 804–830.

Patel, S., M. Winkel, M. Ali, K. M. V. Narayan, and N. Mehta (2015, June). Cardiovascular Mortality Associated With 5 Leading Risk Factors: National and State Preventable Fractions Estimated From Survey Data. *Annals of Internal Medicine 163*(4), 245–253.

Pauly, M. and F. Blavin (2008, December). Moral hazard in insurance, value-based cost sharing, and the benefits of blissful ignorance. *Journal of Health Economics 27*(6), 1407–17.

Picone, G., M. Uribe, and R. Mark Wilson (1998). The effect of uncertainty on the demand for Medical Care, health capital and wealth. *Journal of Health Economics 17*(2), 171–185.

Rosenstock, I., V. Strecher, and M. Becker (1988, January). Social Learning Theory and the Health Belief Model. *Health Education and Behavior 15*(2), 175–183.

Sculpher, M. (2008, January). Subgroups and heterogeneity in cost-effectiveness analysis. *PharmacoEconomics 26*(9), 799–806.

Selden, T. (1993). Uncertainty and health care spending by the poor: The health capital model revisited. *Journal of Health Economics 12*(1), 109–115.

Simon, K., A. Soni, and J. Cawley (2016, May). The impact of health insurance on preventive care and health behaviors: Evidence from the 2014 aca medicaid expansions. Working Paper 22265, National Bureau of Economic Research.

Song, Z., S. Rose, D. G. Safran, B. E. Landon, M. P. Day, and M. E. Chernew (2014). Changes in health care spending and quality 4 years into global payment. *New England Journal of Medicine 371*(18), 1704–1714. PMID: 25354104.

Song, Z., D. G. Safran, B. E. Landon, Y. He, R. P. Ellis, R. E. Mechanic, M. P. Day, and M. E. Chernew (2011). Health care spending and quality in year 1 of the alternative quality contract. *New England Journal of Medicine 365*(10), 909–918. PMID: 21751900.

Sowers, J. R., M. Epstein, and E. D. Frohlich (2001, April). Diabetes, Hypertension, and Cardiovascular Disease : An Update. *Hypertension 37*(4), 1053–1059.

Srikanth, K. and E. Strumpf (2016). How effective is population-based cancer screening? regression discontinuity estimates from the us guideline screening initiation ages. *Forum for Health Economics and Policy 19*(1), 87–139.

Stone, N., J. Robinson, A. Lichtenstein, C. N. Bairey Mertz, C. Blum, R. Eckel, A. Goldberg, D. Gordon, D. Levy, D. Lloyd-Jones, P. McBride, J. S. Schwartz, S. Shero, S. Smith, K. Watson, and P. Wilson (2014, July). 2013 ACC/AHA Guideline on the Treatment of Blood Cholesterol to Reduce Atherosclerotic Cardiovascular Risk in Adults: A Report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *Journal of the American College of Cardiology 63*(25 Pt B), 2889–934.

The CDC Diabetes Cost-Effectiveness Study Group (1998). The Cost-effectiveness of Screening for Type 2 Diabetes. *JAMA 280*(20), 1757–1763.

The Center for the Evaluative Clinical Sciences, Dartmouth Medical School (1996-2017). The Dartmouth Atlas of Health Care: Data on Quality/Effective Care.

Vanness, D. and J. Mullahy (2012, March). Moving Beyond Mean-based Evaluation of Health Care. In *The Elgar Companion to Health Economics, Second Edition*, Chapters, Chapter 52. Edward Elgar.

Wang, C., A. Cheung, K. Bibbins-Domingo, L. Prosser, N. Cook, and L. Goldman (2011). Effectiveness and cost-effectiveness of blood pressure screening in adolescents in the United States. *Journal of pediatrics 158*, 257–64.

Ward, E., A. Jemal, V. Cokkinides, G. K. Singh, C. Cardinez, A. Ghafoor, and M. Thun (2004). Cancer disparities by race/ethnicity and socioeconomic status. *CA: A Cancer Journal for Clinicians 54*(2), 78–93.

Wherry, L. R. and S. Miller (2016, June). Early Coverage, Access, Utilization, and Health Effects Associated With the Affordable Care Act Medicaid Expansions: A Quasi-experimental Study. *Annals of Internal Medicine 164*(12), 795.

Wilson, N. (2011, August). Antiretroviral Therapy and Demand for HIV Testing: Evidence from Zambia. *SSRN Electronic Journal*.

Zweifler, R., L. McClure, V. Howard, M. Cushman, M. Hovater, M. Safford, G. Howard, and D. Goff (2011, January). Racial and geographic differences in prevalence, awareness, treatment and control of dyslipidemia: the reasons for geographic and racial differences in stroke (REGARDS) study. *Neuroepidemiology 37*(1), 39–44.