HCEO WORKING PAPER SERIES

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

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

Working Paper

2018-014

03/2018

HUMAN CAPITAL AND
ECONOMIC OPPORTUNITY
GLOBAL WORKING GROUP

The University of Chicago
1126 E. 59th Street Box 107
Chicago IL 60637

www.hceconomics.org

Effects of expanding health screening on treatment – What should we expect? What can we learn?

Rebecca Myerson*, Darius Lakdawalla, Lisandro D. Colantonio,

Monika Safford, David Meltzer

February 19, 2018

Abstract

Screening interventions can produce very different treatment and health outcomes, depending on the reasons why patients had been 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 undiagnosed patients differ from currently diagnosed 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.

*Corresponding author. Mailing address: Leonard D. Schaeffer Center for Health Policy and Economics and Department of Pharmaceutical and Health Economics, University of Southern California, Verna & Peter Dauterive Hall, Office 414M, 635 Downey Way, Los Angeles, CA 90089. Email: rmyerson@healthpolicy.usc.edu. 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.

Keywords: diagnosis, screening, self-selection, policy analysis, quality measurement. JEL Codes: I10, I12, I14, I18.

Note: 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 content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Neurological Disorders and Stroke or the National Institutes of Health. Representatives of the funding agency have been involved in the review of the manuscript but not directly involved in the collection, management, analysis or interpretation of the data. The authors thank the other investigators, the staff, and the participants of the REGARDS study for their valuable contributions. A full list of participating REGARDS investigators and institutions can be found at http://www.regardsstudy.org.

I 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.

The key message of this paper is that screening interventions can produce very different treatment and health outcomes, depending on the reasons why patients had remained unscreened in the first place. There are two main reasons that patients diagnosed after a screening expansion may have different treatment rates and outcomes from previously diagnosed patients. 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).

[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.

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. Some patients' biomarkers are rarely assessed because they face high barriers to health care.[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 these patients after diagnosis, despite clinical benefit.

Economists have not studied the implications of the differences between currently diagnosed and undiagnosed patients for evaluating screening interventions. To fill this gap, we propose a simple economic framework to guide policy-makers and analysts in designing and evaluating the impact of screening interventions on uptake of relevant treatment. We then apply the insights from this framework 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 high potential health benefits to treatment in certain cases. If many sick patients remain undiagnosed 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

[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).

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 more patients are diagnosed, 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 whose blood glucose and blood cholesterol were not recently assessed 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 high cholesterol and diabetes 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 rates of diagnosed conditions for every 10 percentage point increases in biomarker assessment.

This analysis provides support for our hypothesis that currently undiagnosed patients differ from currently diagnosed 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 NHANES and Dartmouth atlas of health

care quality use different methods but find supportive results: time periods and locations with higher diagnosis rates show lower treatment rates for diagnosed conditions (Centers for Disease Control and Prevention, 2014; Finkelstein, 2013).

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 diagnosis 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.

We also demonstrate the importance of these findings for policy analysis and health system performance evaluation. First, our findings imply that 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 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 treatment rates for chronic conditions decline rather than improve 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 the health effects of diagnosing additional patients depend on the reasons why patients were undiagnosed 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 II compares this study with previous literature and articulates our contributions. Section III presents our conceptual framework. We assess predictions from the framework empirically in sections IV and V, focusing on contexts relevant to the top two causes of death in the United States, cancer and heart disease. Section VI demonstrates the implications of our findings for policy analysis and health system performance measurement. Section VII concludes.

II 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.,

[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).

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.

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).[4] Research on cancer screening in the medical and public policy literature has shown that screening expansions attract patients with less severe conditions,

[4]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).

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. Yet, economists have not studied the implications for evaluating screening interventions. We fill this gap in the literature by exploring how the impact of screening interventions relates to the reason patients were unscreened and undiagnosed in the first place.

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 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).

III Conceptual Framework

In this section, we present a simple economic framework of selection into screening. 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 theory is intuitive for clarity of presentation, similar points could be made using a formal model.

Figure 1: Demand for screening with a single supply curve

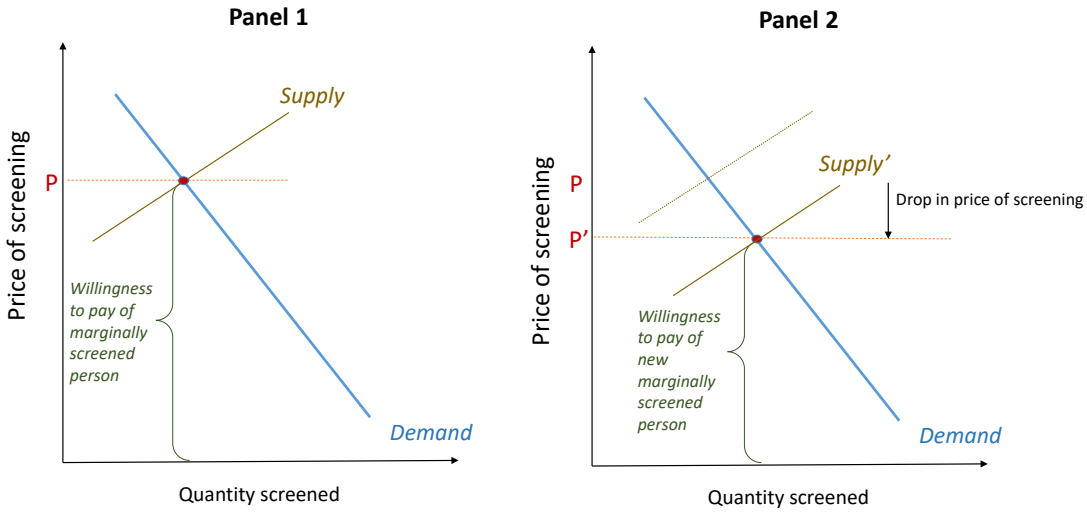
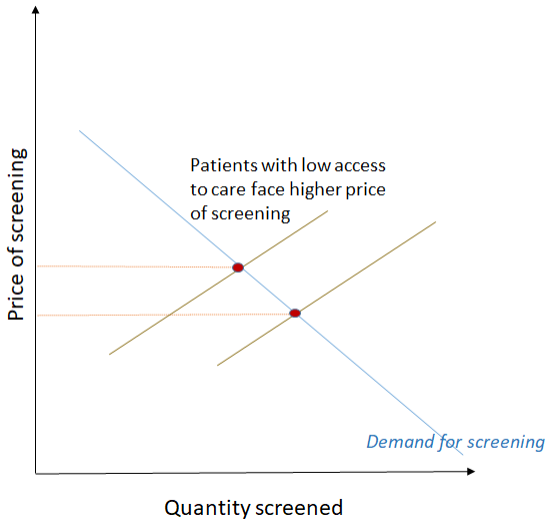


Figure 2: Demand for screening with multiple supply curves



A WHICH PATIENTS ARE SCREENED?

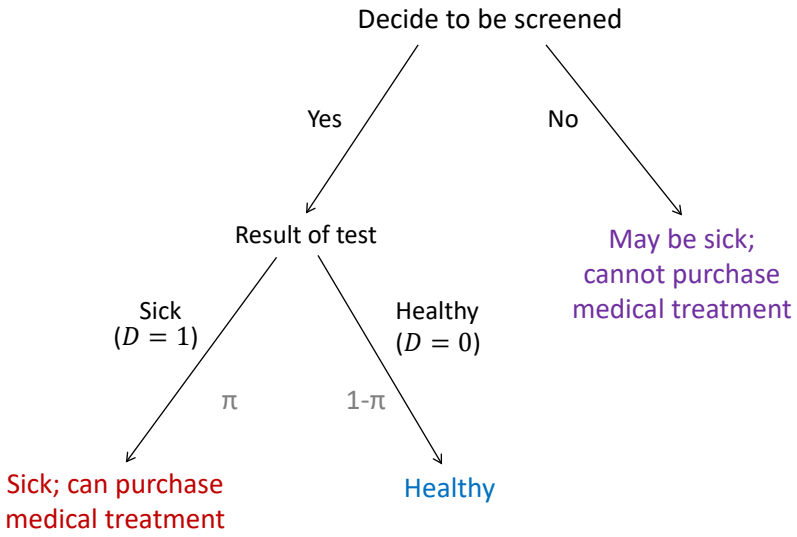
Suppose the value of screening differs across patients. In some cases, this variation occurs for reasons that are unobservable to patients or their physicians. For instance, some patients might be genetically predisposed to developing complications from a disease if untreated while others might be relatively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

Greater complexity arises when this variation is observable to the patient or her physician. In this case, patients with high expected value from screening have higher willingness to pay to be screened. 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.

If all patients face the same price of screening, say price P in the first panel of Figure 1, then the patients with willingness to pay greater than or equal to P are screened, and patients with a lower willingness to pay are not screened. In the second panel of Figure 1, the price of screening shifts down to P'. In this case, many more patients are screened including those with lower willingness to pay, reflecting lower expected value.

In practice, 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 low health care supply. Figure 2 shows the demand curve for screening again but with two supply curves added, reflecting the fact that different patients can face different pecuniary or non-pecuniary costs of 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. For each group, the logic of Figure 1 still applies: if one supply curve were to shift so as to lower the price of screening in one group, the marginally screened patients in that group would expect less value from screening than the previously screened patients.

Figure 3: Screening decision tree



B 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) expect 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 3.

Screening is valuable to patients as an input to improved health and well-being (Glanz and Bishop, 2010; Carpenter, 2010). Figure 3 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 3. Ex post, this option is not used by patients with a negative screening result, as shown in the middle branch of Figure 3. 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 3. 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 expect little value from 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, such as or HIV before the development of highly active anti-retroviral therapy, or 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.

C WHAT CAN WE EXPECT AFTER SCREENING EXPANDS?

Section III.B argued that a because a patient's expected costs and benefits of treatment contribute to the patient's expected value of screening, these factors also affect where that patient sits on the screening demand curve. Below, we argue that these same factors can produce declining treatment rates for newly diagnosed conditions as screening expands. We further argue that the health benefits from treatment of newly diagnosed conditions may decline or increase as screening expands, depending on why patients were previously un-screened.

For simplicity, first consider the case where all patients face the same price of screening. In this case, 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 less clinical benefit or who would have difficulty affording treatment if diagnosed. Because of these higher costs or lower benefits to treatment, the newly diagnosed patients would treat their conditions at lower rates than previously diagnosed patients, on average – and the impact of screening on treatment would not be as large as one might otherwise expect.

Let's now consider the case where different patients face different prices for screening. Could high treatment rates be obtained 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 treatment. This might occur, for example, if 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. Yet, such a situation would be highly unusual. In a far more empirically likely case, patients facing high barriers to screening (for example, patients who are uninsured or live in rural areas) will also face 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 currently undiagnosed patients differ from currently diagnosed patients in ways that predict lower uptake of treatment after diagnosis.

The health impact of treating newly diagnosed conditions could increase or decrease as diagnosis rates rise, depending on why patients were previously unscreened and therefore undiagnosed. If many sick patients were unscreened due to high barriers to care, then the benefits of screening and treating these patients could be high.[5] In this case, a bundled

[5]To clarify, consider the case where access to health services is unequal, so that some patients have low access to both screening and treatment for idiosyncratic reasons. Figure 2 shows that the marginally screened patient in the low-access group sits relatively high on the demand curve. Patients with a higher position on the demand curve can have comparatively high expected health benefits to treatment or comparatively low

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 benefits, diminishing returns to screening would be unavoidable.

We assess these predictions empirically in sections IV and V, 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 undiagnosed patients differ from currently diagnosed 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 diagnosing additional patients, 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 VI, 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 as diagnosis rates rise.

IV Treatment Outcomes: Evidence from Heart Disease Risk Factors

We now assess our theoretical prediction that as screening expands, the additional diagnosed patients will be less likely to receive 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 empiri-

barriers to treatment. However, we assumed that this group of patients face comparatively high barriers to treatment. Therefore, it follows that the marginally screened patients from this group have comparatively high expected health benefits from treatment.

cal approaches to provide policy-relevant evidence on the links between diagnosis rate and treatment rate, and the implications for screening interventions.

First, we provide practical evidence on the potential impact of biomarker assessment 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 important for understanding policies that specifically target patients who are rarely screened.

Next, we examine the over-arching contextual and patient-level factors associated with biomarker assessment, diagnosis, and treatment. We build our findings into a simulation analysis to assess how predicted treatment uptake would change under different policies. The results indicate that patients with infrequent biomarker assessments differ from frequently assessed patients in important ways, corresponding to lower treatment rates in the former group after diagnosis. These findings are helpful for understanding policies that change access to biomarker assessment across a broad population.

A DATA

We use data from two studies: the REasons for Geographic and Racial Differences in Stroke study (REGARDS) and the National Health and Nutrition Examination Survey (NHANES). In both 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.

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 biomarker assessments relevant for diabetes are only available starting in 2005; we therefore use data from 2005-2014.

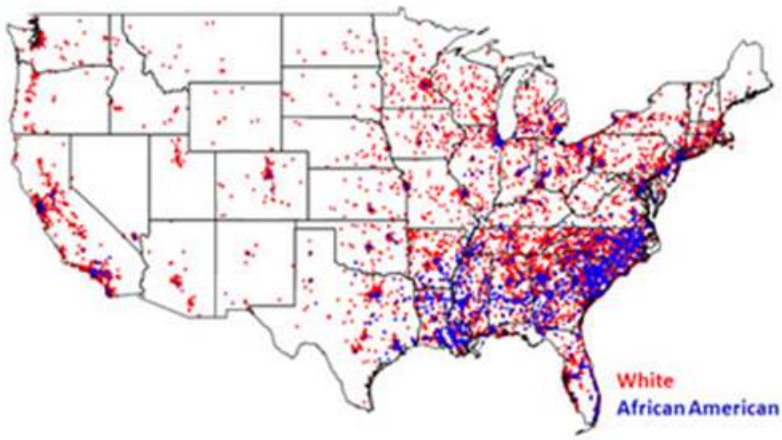
The REasons for Geographic and Racial Differences in Stroke (REGARDS) study recruited community-dwelling participants into an epidemiological longitudinal cohort study designed to answer questions about racial differences in stroke mortality. Recruitment was conducted from 2003-2007 and was accomplished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

In a companion paper, we compare the REGARDS sample with the NHANES sample year by year to show that the REGARDS sample is a close match with the nationally representative NHANES sample when the same sample restrictions are applied (Myerson et al., 2017).

REGARDS 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

[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.

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



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 of high blood pressure are shown in Figure 8 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.)

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 6 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

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

	REGARDS	**NHANES**
Survey Inclusion Criteria	In traditional Medicare past 2 years; black or white; English speaking	Nationally representative
Geography of Sample	National	National
Year of Biomarker Collection	2003-2007	2005-2014
Age Range in Analysis	67+	All
Participants with Any Condition(s) of Interest	5,721	18,735
Participants with Undiagnosed Condition(s) of Interest	1,077	6,281
Among Participants with Condition(s) of Interest:		
Average Age	74	55
Had Health Insurance	100%	81%
African American	30%	22%
Participants with Diabetes	1,309	4,282
Aware of Diabetes	1,192	3,482
Treating with Medication	1,161	2,991
Participants with Hypertension	4,502	11,576
Aware of Hypertension	4,170	10,193
Treating with Medication	3,846	7,680
Participants with High Cholesterol	4,268	13,716
Aware of High Cholesterol	3,542	9,030
Treating with Medication	2,457	4,860

diagnosis for that condition.

Using the merged REGARDS-Medicare data, we are able to make additional corrections for patients' under-reporting of diagnosis. 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%).

Although claims data are not available to validate self-reported diagnosis in the NHANES data, the relatively colloquial wording of the NHANES questions helps to mitigate concerns that patients may not report a prior diagnosis because they are unfamiliar with medical terms. Participants in the NHANES data were considered to be diagnosed for diabetes if they responded positively to the question, "Have you ever been told by a doctor that you have diabetes or sugar diabetes?" Similarly, participants were considered to be diagnosed for high cholesterol if they responded positively to the question "Have you ever been told by a doctor or other health professional that your blood cholesterol level was high?"

The NHANES data additionally include information on recent biomarker tests relevant to high cholesterol and diabetes, also obtained by participants' self-report. Biomarker assessments are tightly related to our topic of study, as the key step through which undiagnosed conditions become diagnosed. On the one hand, patients whose biomarkers are never assessed will never become diagnosed. On the other hand, physicians who observe biomarker assessment results indicative of previously undiagnosed hyperlipidemia or diabetes must ethically proceed according to national guidelines for diagnosis and treatment.[7]

[7]This point can be illustrated with multiple examples. For example, a physician may run a blood test to assess whether a patient who presents with symptoms of diabetes, such as polydipsia or polyuria, in fact has undiagnosed diabetes. Although this test may not have been scheduled in advance, the purpose of this test is to assess whether the patient has undiagnosed diabetes. As another example, a physician may conduct blood tests to track the effectiveness of treatment of diagnosed pre-diabetes that may or may not ultimately develop into diagnosed diabetes. Although one purpose of this test is to monitor the impacts of the treatment for pre-diabetes, ethically the physician must diagnose and treat diabetes if test results are indicative of diabetes.

Recent blood cholesterol assessment is captured using two questions in the NHANES: "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." We present results using the two-year look-back period in this text, but findings are similar if other look-back periods are used. Likewise, NHANES participants were considered to have had a recent biomarker assessment related to diabetes if they responded positively to the question, "Have you had a blood test for high blood sugar or diabetes within the past three years?" (In the NHANES instructions, interviewers coded patients with only urine tests as not having been tested for high blood sugar or diabetes.) The colloquial wording of this question helps to prevent false negative reports among patients unfamiliar with the word "diabetes". Nonetheless, some misclassification of biomarker assessment status or diagnosis status may remain, which would bias our findings toward the null (Black et al., 2000).

B ANALYSIS

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, and assess whether treatment rates for the newly diagnosed conditions reach the level of previously diagnosed conditions by placing prior diagnosis status on the right-hand side of the model. To take the example of diabetes, we run:

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

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

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

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. To account for the fact that patients with multiple co-morbid conditions may treat both in the same doctor visit, we also adjust for an indicator variable that takes the value 1 if the patient has more than one of our conditions of interest (diabetes, hypertension, and high cholesterol) and 0 otherwise. 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. All findings are qualitatively similar if logit or probit models are used.

C RESULTS

As hypothesized, we find that conditions diagnosed after outreach to encourage screening are less likely than previously diagnosed conditions to receive relevant medical care. Table 2 shows that previously undiagnosed conditions were less likely to receive an annual evaluation and management doctor visit than previously diagnosed conditions. The table presents 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.

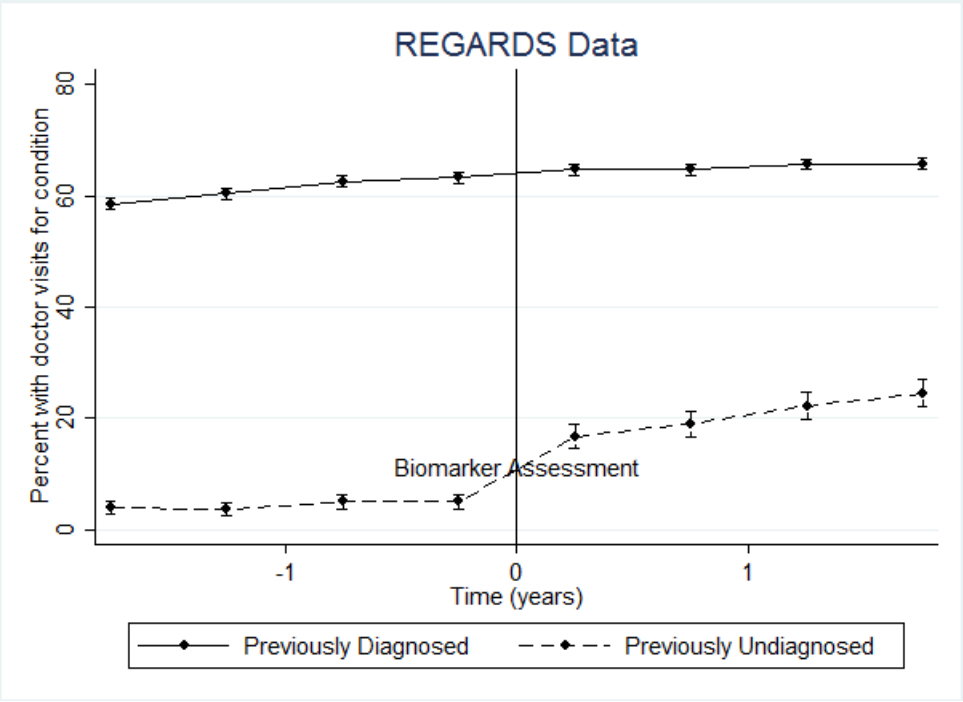
The findings are qualitatively similar if we analyze number of visits per year or if we include data from the first year after biomarker assessment via REGARDS. Additionally, Figure 5 shows a similar relationship in the raw data: doctor visits for previously undiagnosed conditions increased after biomarker assessment, but only to about half the level of previously diagnosed conditions.

Notification by mail for the diabetes and high cholesterol results is unlikely to account for the observed shortfall in doctor visits for newly diagnosed conditions, because the gap in doctor visits exists for all three conditions, including high blood pressure.[8]

Doctor visits to treat chronic conditions are not recommended for patients who are very elderly or near death. Yet, patterns of extreme age or illness are also unlikely to account

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

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.

for our findings for two reasons. First, patients who are very infirm or near death are unlikely to meet inclusion criteria for the REGARDS study. The REGARDS study excluded patients who were not community-dwelling, who lived in a nursing home or were on a waitlist for a nursing home, were in cancer care, or were unable/unwilling to answer a lengthy questionnaire by phone. Second, patients with vs. without undiagnosed conditions do not differ substantially on age, and our results are similar if we adjust in a flexible way by including indicators for five-year age groups.

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***	-0.34***
				(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 health 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. In the last column, we adjust for patients' biomarkers and an indicator variable that captures whether or not the patients have more than one of our conditions of interest. Robust standard errors are in parentheses.

D SIMULATION ANALYSIS TO GENERALIZE THE FINDINGS

We now generalize the previous findings using the NHANES sample which is representative of all adults in the United States. When outlining the theoretical framework in the previous section, we had hypothesized that patients with undiagnosed conditions would differ from patients whose conditions were diagnosed in important ways, which would predict lower treatment uptake after diagnosis. We now compare the factors associated with patients being checked for undiagnosed conditions and the factors associated with treatment after diagnosis, to quantify how our hypothesized intermediate factors could attenuate the impact of diagnosing additional patients.

We construct two key variables to measure (a) which patients' biomarkers were recently assessed, providing the opportunity for undiagnosed conditions to become diagnosed; and (b) which patients' diagnosed conditions were treated. Taking the example of high cholesterol, the biomarker assessment variable would take the value 1 for patients who have never been diagnosed for high cholesterol and have recently had a cholesterol check, and would take the value 0 for patients who have never been diagnosed for high cholesterol and have not recently had a cholesterol check. This variable would be missing for patients already diagnosed for high cholesterol. Likewise, the treatment after diagnosis variable would take the value 1 for patients whose diagnosed high cholesterol is treated, and 0 for patients whose diagnosed high cholesterol is untreated. Similar variables are constructed for diabetes, and we conduct separate analyses for each condition.

Consistent with our hypotheses, we find that patients who have higher barriers to care and who (ex ante) seem healthier are less likely to be checked for undiagnosed conditions, and are also less likely to be treated after diagnosis. Table 3 summarizes the sign of the statistically significant correlation coefficients found in the data. These correlations show that on average, patients who lack recent blood tests relevant to undiagnosed conditions have larger barriers to care: these patients are more likely to lack health insurance, lack a usual

source of care, and have low income. These same factors are, in turn, associated with lower likelihood of treating diagnosed conditions. We also find that patients who appear healthier - 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 receive blood tests relevant to any undiagnosed conditions and also less likely to treat their diagnosed conditions.

Using a simulation, we assess whether these factors could account for lower treatment rates among patients diagnosed after a small policy change. 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 diagnosed, rather than imputing the potential treatment rates of undiagnosed patients. The policy simulation exercise proceeds as follows. In one counterfactual, we only allow patients with biomarker assessment propensity scores over 70% to become diagnosed; all other diagnosed patients are dropped from the sample. In another counterfactual, all patients with propensity scores over 60% are allowed to be diagnosed and others are dropped. In each case, we calculate the average treatment rate of diagnosed conditions using patients' real data, including only patients who would have been diagnosed under such a scheme. We repeat this exercise for policy counterfactuals with propensity score cutoffs ranging from 70% to 0%, in intervals of 10%. (The cutoff of 0% represents the case where no diagnosed patients were dropped.)

Findings from this exercise indicate that treatment rates for diagnosed patients decline as more patients' biomarkers are assessed, with approximately a 0.5 percentage point decline in treatment rates for each 10 percentage point increase in regular biomarker assessment. When we focus on patients without health insurance, a group that includes the patients in our sample with the lowest access to care, we find that treatment rates declined more rapidly, up to 1 percentage point decline in the treatment rate for each 10 percentage point increase in regular biomarker assessment. This is consistent with the notion that patients with low access to screening may also have low access to treatment.

Table 3: Patient-level characteristics associated with biomarker checks for undiagnosed conditions are also associated with increased treatment of diagnosed conditions (NHANES data)

	Biomarkers Checked If Not Diagnosed	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 biomarker checks 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 with the requisite variables measured in the NHANES data.

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. Table 7 in the Appendix includes a comparison of actual treatment rates with counterfactual treatment rates under various policies.

E SUMMARY

The empirical findings in this section support our key point that the composition of diagnosed patients changes as diagnosis rates rise, in a way that decreases treatment rates for diagnosed conditions. We showed 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. We also explored the intermediate factors that could underlie such a finding. Consistent with our hypotheses in section III, patients who appear less sick or face higher barriers to care are less likely to be checked for undiagnosed conditions and also less likely to be treated if diagnosed. The role of access to care in shaping the health effects of screening expansions is explored further in the next analysis.

V Health Outcomes: Evidence from Cancer

Although the findings in the previous section seem to indicate diminishing returns to diagnosing additional patients, 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 diagnosing additional patients with high access to care could produce diminishing health benefits. Second, we hypothesized that diagnosing additional 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 cancer 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 the diminishing health returns to diagnosing additional patients could be reversed if patients with barriers to care are targeted and their barriers to care are addressed.

A 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.[9] 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

[9]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.

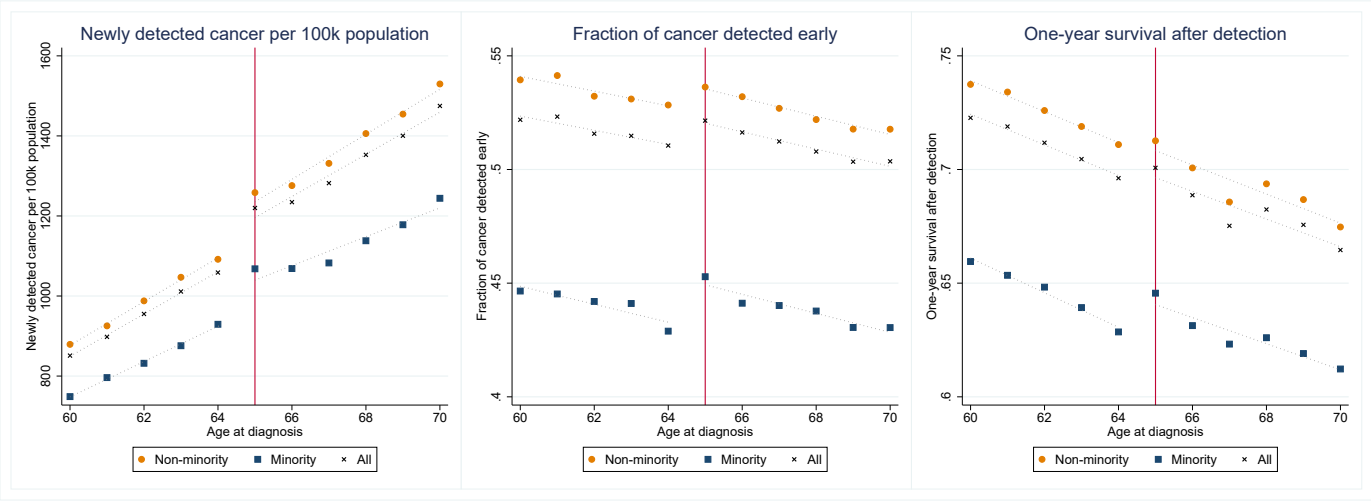
age discontinuities in Medicare eligibility and age discontinuities in cancer screening (Card et al., 2008; Srikanth and Strumpf, 2016). First, we restrict the data to a small window around the Medicare eligibility threshold (age 65) and estimate a local-linear regression using the *rdrobust* Stata command (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.

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.

B 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 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 than non-minority patients to have metastatic cancer at the time of detection.[10] 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, as shown in previous 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.[11] These disparate effects create a decline in health disparities upon aging into Medicare: the

[10] 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 9 in the Appendix.

[11] 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.

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 and secular trends in aging. 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.

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 were undiagnosed in the first place. Patients with high access to care may not have their biomarkers assessed due to low expected benefit rather than high cost. Expanding biomarker assessment rates in this group show 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 skip biomarker assessments 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.

VI Implications for Policy Analysis, and for 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 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.

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 relevant 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. If instead we had graphed the "treatment rate" indicating the fraction of diagnosed conditions receiving doctor visits, this rate would decrease after biomarker assessment.[12] This reversal of sign is driven by changes to the group of diagnosed conditions before versus after biomarker assessment.

[12]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***

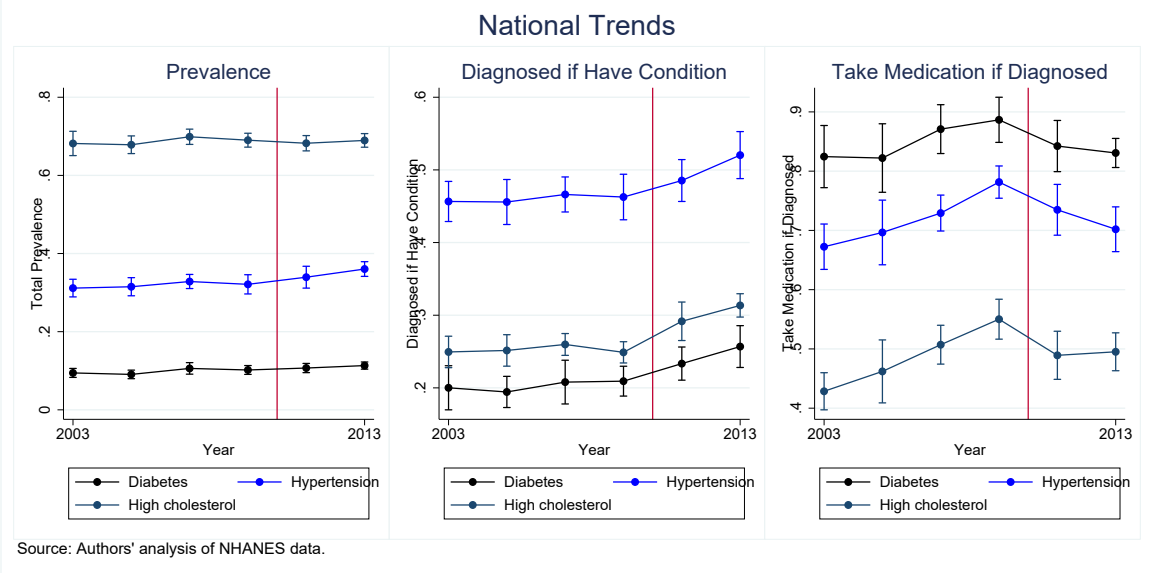
* p <0.1, ** p <0.05, *** p <0.01

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) .

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 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, accounting for patients' underlying health needs. Table 5 shows the correlations between diagnostic intensity 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 appropriate diabetes treatment.

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

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.

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 decrease in rates of treatment among diagnosed conditions.

VII Conclusion

The key message of this paper is that if anticipated costs and benefits of treatment determine which patients had remained undiagnosed prior to 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.

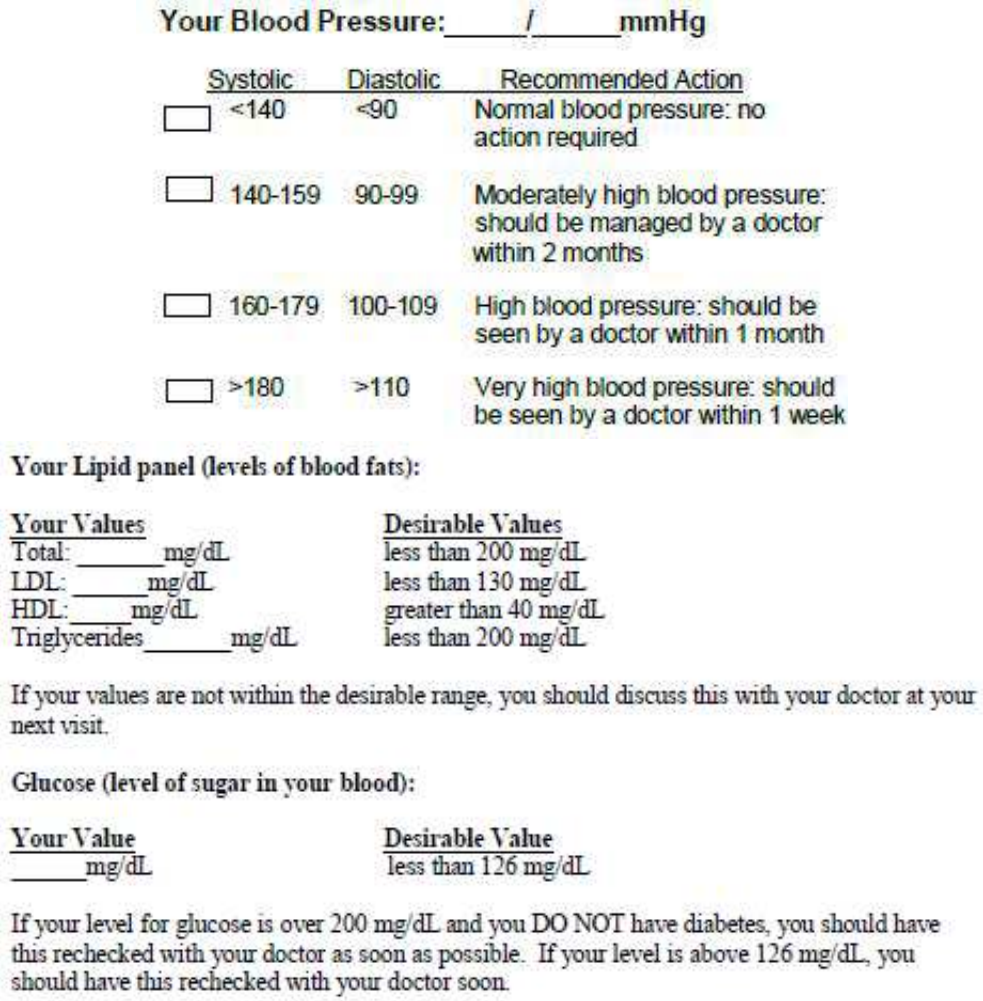
First, our framework and our findings have implications for policy analysis. Based on the patterns we uncovered, expanded screening as a stand-alone program is likely to be less cost-effective than previously anticipated due to low treatment uptake among marginally screened patients. To our knowledge, 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 8: 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 6: 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).

Table 7: Simulation analysis: Biomarker checks for undiagnosed conditions, and treatment rates for diagnosed conditions

A. All patients

	Diabetes		
Who is checked?	**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%

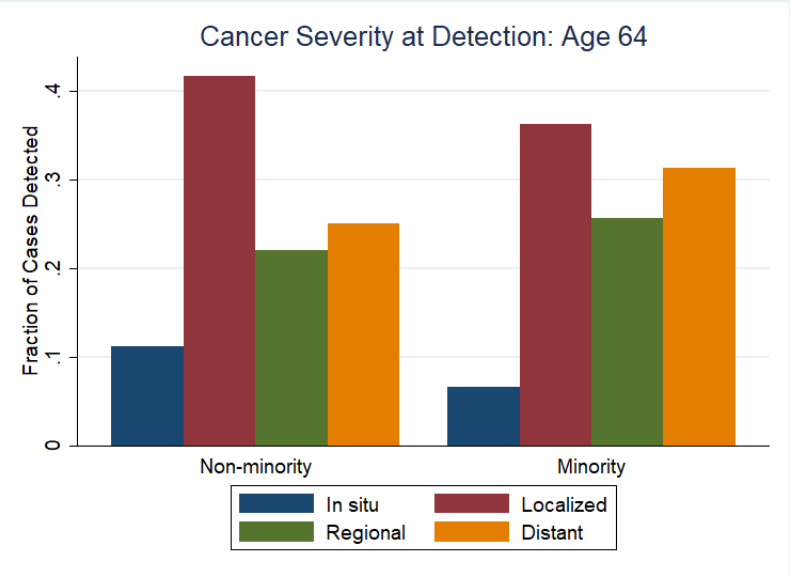
	High Cholesterol		
Who is checked?	**Cases detected**	**Average treatment if diagnosed**	**Average treatment if doctor recommended treatment**
Propensity 0.7 and higher	6793	62%	85%
Propensity 0.6 and higher	9238	58%	84%
Propensity 0.5 and higher	11174	55%	82%
Propensity 0.4 and higher	12595	53%	81%
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 checked?	**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 checked?	**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 9: 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).

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.

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

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.

Black, D. A., M. C. Berger, and F. A. Scott (2000). Bounding parameter estimates with nonclassical measurement error. *Journal of the American Statistical Association 95*(451), 739–748.

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 immun-

odeficiency 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.

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.

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.

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.

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.

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. 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.

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.

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.

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.

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.