Chasing the Missing Diagnoses: Exploring the Unintended Consequences of Low-Cost Health Screenings

Rebecca Myerson, Lisandro Colantonio, Monika Safford, and David Meltzer

October 27, 2016

Abstract
This paper investigates the consequences of removing barriers to information by studying screenings that enable medical treatment of asymptomatic health conditions. We consider two unintended consequences of reducing out-of-pocket costs of screening, as is done in the Affordable Care Act. First, lower-cost screenings attract patients with lower demand for information, who may also have lower demand for treatment. Second, expanding screening could increase adverse selection and reduce the stability of health insurance markets. Using data from three biomarker studies reflecting different populations affected by the Affordable Care Act, we find evidence for the former prediction but not the latter.

*Myerson (corresponding author): University of Southern California. Email: rmyerson@usc.edu. Colantonio, Safford, and Meltzer: University of Alabama at Birmingham, Weill Cornell Medical College, and University of Chicago. Emails: lcolantonio@uab.edu, mms9024@med.cornell.edu, and dmelzter@medicinebsd.uchicago.edu. We are grateful for valuable feedback and suggestions from Dan Black, Marguerite Burns, Lenisa Chang, James Heckman, Elbert Huang, and Roberto Robatto as well as participants of numerous conferences.

†This research project is supported by a cooperative agreement U01 NS041588 from the National Institute of Neurological Disorders and Stroke, National Institutes of Health, Department of Health and Human Service, as well as grants R01 HL080477 and K24 HL111154 from the National Heart, Lung, and Blood Institute. 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.
Many people, including some with health insurance, are not screened for conditions that can be asymptomatic and can be treated to prevent illness. For instance, diabetes, high cholesterol, and hypertension are top contributors to cardiovascular disease and end-stage renal disease in the United States and about one-fifth of cases are undiagnosed (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). To reduce the number of “missing” diagnoses as an entry point for treatment, the Affordable Care Act (hereafter, ACA) requires health insurance plans to offer free screening for diabetes, high cholesterol, and hypertension to people at high risk.\(^1\) This policy has already affected the health plans of an estimated 76 million people (Burke and Simmons, 2014). Due to the effectiveness of available treatments,\(^2\) the policy could have a substantial health impact if this policy increases treatment of chronic conditions.

Although it may seem intuitive that subsidizing screening should increase treatment of chronic conditions, in fact the size of the effect is unclear for two reasons that we explore in this paper. The first reason is that patients with undiagnosed conditions may differ from other patients in their demand for health care after diagnosis. Our logic is as follows. Patients who are unaware of some of their conditions are less likely than patients who are aware of all their conditions to have been screened recently. This lack of screening may have resulted from their lower perceived benefits of medical treatment or higher perceived barriers to medical treatment.\(^3\) 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; Kenkel, 1994; Manning et al., 1987; Musa et al., 2009). These same barriers to medical treatment could then translate to lower treatment rates for newly detected conditions. These changes in patient composition could

---

\(^1\)Section 2713, “Coverage of Preventive Services,” requires that health insurance issuers “not impose any cost sharing requirements for evidenced-based items or services that have in effect a rating of A or B in the current recommendations of the United States Preventive Services Task Force.” Accordingly, no cost sharing is required for blood pressure screening for all adults, diabetes screening for adults aged 40-70 who are overweight or obese, and cholesterol screening for all men aged 35+ (or at lower age with elevated risk) and all women at elevated risk aged 20+. Sections 4103 and 4104 of the Affordable Care Act also eliminate cost sharing for annual wellness visits for Medicare beneficiaries.

\(^2\)See American Diabetes Association (2014); Collaborators (2005); Blood Pressure Lowering Treatment Trialists’ Collaboration (2000); Bindman et al. (1995); Bressler et al. (2014); D’Agostino et al. (2008); Farley et al. (2010); James et al. (2014); Stone et al. (2014); Sytkowski et al. (1990) for evidence on the effectiveness of treatments.

\(^3\)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).
undermine the apparent effects of expanded access to low-cost screening; if analysts ignore these composition effects, they may incorrectly conclude that treatment and control of chronic conditions declines rather than improves as more patients become diagnosed.

Second, increased screening could make medical treatment less affordable due to indirect effects if patients demand more medical care after learning of an undiagnosed condition. In particular, if patients switch into more generous health plans after learning that they are sick and health plans cannot use pre-existing conditions in deciding the premium price, then generous health plans must increase premium prices for all patients to cover the additional costs. This would be an example of adverse selection, the process by which patients sort into more generous insurance based on anticipated health needs that are known to them but are not incorporated into premium prices (Pauly, 1974; Rothschild and Stiglitz, 1976; Wilson, 1977). When adverse selection compounds over time, there can be major implications for how well a health insurance market can function (Culter and Reber, 2011; Cutler and Zeckhauser, 1998). Previous studies have shown that increases in adverse selection can unravel the benefits to providing patients with information (Handel, 2011).

This paper contributes to the literature by providing the first empirical assessment, to our knowledge, of these two concerns. First, we examine whether people who had undiagnosed conditions prior to low-cost screenings differ from others in their uptake of physician-recommended treatment for diagnosed conditions. This is important for providers to know whether additional outreach could be necessary to engage patients whose chronic conditions become diagnosed as a result of a change in the price of screening. We examine this issue using data from three biomarker studies that paid participants to collect their biomarkers. These include two national biomarker studies, the National Health and Nutrition Examination Survey (hereafter, NHANES) and the REasons for Geographic and Racial Differences in Stroke study (hereafter, REGARDS), and one regional biomarker study, the Oregon Health Insurance Experiment (hereafter, OHIE) (Centers for Disease Control and Prevention, 2014; Finkelstein, 2013; Howard et al., 2005). These three studies reflect different groups affected by the ACA policy: all people with health insurance, people with Medicare insurance, and people who would apply for expanded Medicaid, respectively. We find that people who were not recently screened for undiagnosed conditions are less likely to receive recommended medical treatment for their previously diagnosed conditions; likewise, people with some undiagnosed conditions are less likely to treat their previously diagnosed conditions. We therefore project that marginally screened people may also be less likely to receive recommended medical treatment after their currently un-
diagnosed conditions are diagnosed. This could reduce the health benefits of subsidizing screening and affect physician practice.

Second, we use the REGARDS study as a natural experiment to examine whether people select health plans with lower cost-sharing after learning of a previously undiagnosed condition. Understanding how low-cost screening affects health care markets is important to predict the impacts of changes to screening policy in health insurance markets with many undiagnosed patients, such as the health insurance marketplaces established under the ACA. We implement the test using panel data on Medicare plan selection that have been merged with the survey and biomarker data of 9,990 REGARDS participants, 20% of whom had an undiagnosed case of diabetes, high cholesterol and/or hypertension. Because these REGARDS participants were recruited and screened on a rolling basis over 2003-2007, we can compare not-yet-recruited participants to recently-recruited participants to determine the impact of learning of an undiagnosed condition via REGARDS on enrollment in Medicare Advantage while adjusting for secular trends. \(^4\) We find no evidence that participants with undiagnosed conditions became more or less likely to enroll in Medicare Advantage after learning of these conditions. (These results hold both before and after changes to the Medicare market in 2006.) The lack of a link between learning of an undiagnosed condition and plan selection may indicate that participants did not anticipate spending much money on their newly detected conditions. This result should be interpreted with caution, however, as we do not observe switches across Medicare Advantage plans. If generalizable, the results would imply that subsidizing screening will not increase sorting of sicker patients into more generous health plans, i.e., adverse selection. This is important because adverse selection is a prominent concern when health status cannot be incorporated into premium prices, as is the case under the ACA (Handel et al., 2015).

An association between demand for medical treatment and demand for screening could explain both findings. In particular, patients with lower demand for medical treatment would have lower treatment of existing conditions and also less reason to adversely select after learning of a previously undiagnosed condition. To show that such an association is plausible, we analyze a theoretical model based on the Picone et al. (1998) interpretation

\(^4\)We test selection into or out of Medicare Advantage without assuming ex ante that individuals with chronic conditions find Medicare Advantage to be the more-generous plan. (That is, we allow for the possibility that individuals may wish to adversely select out of Medicare Advantage after learning of a previously undiagnosed condition.) Although Medicare Advantage plans had lower cost-sharing than the default fee-for-service Medicare plan by requirement, there is evidence that Medicare Advantage formularies were selected to produce advantageous selection into Medicare Advantage during our sample period (Newhouse et al., 2014; McWilliams et al., 2012; Morgan et al., 1997).
of the Grossman (1972) model. In the model, agents can consume medical treatment to ameliorate the negative health effects of a chronic health condition, but only if they have been screened and diagnosed for the condition. Agents who have been recently screened know whether they have a chronic health condition, whereas agents who have not been recently screened hold beliefs about the probability that they have a condition. Agents differ only in the costs they face for medical treatment. We use the model to show that, as intuited above, costs both reduce patients’ demand for medical treatment after diagnosis and reduce their demand for screening.

The paper proceeds as follows. Section 1 compares our study with previous literature. In section 2, we analyze a theoretical model of demand for health screening and demand for medical treatment to show why the two could be linked. Section 3 shows empirical evidence of a relationship between use of medical treatment for previously diagnosed conditions and use of screening (directly measured, or proxied by presence of undiagnosed conditions). Section 4 shows no evidence of adverse selection after patients learn of a previously undiagnosed condition, and describes how we use the REGARDS epidemiological study as a subsidy to screening with random roll-out to identify this relationship. Section 5 concludes.

1 Comparison with the literature

Anticipated costs and benefits of health care can differ across individuals, influencing individuals’ willingness to pursue 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. It follows that anticipated net benefits of particular health services can vary across individuals (Vanness and Mullahy, 2012). In certain cases, distributions of these individual-level net benefits can be estimated (Basu and Heckman, 2007; Carneiro et al., 2010; Eisenhauer et al., 2010). These distributions are useful because changes to out-of-pocket costs of health care will attract different patients to use the treatment based 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

---

5See 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).
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 distributions of anticipated net benefits of screening, a determinant of which conditions are not treated (i.e., the extensive margin).

When analyzing adverse selection, we use panel data to avoid conflating adverse selection with other factors (Abbring et al., 2003; Cohen, 2005; Hackmann et al., 2012; Spenkuch, 2012). In particular, the observed correlation between health care claims and insurance generosity in cross-sectional data reflects two effects: (a) sicker individuals select more generous insurance based on their private knowledge of their healthcare needs (adverse selection), and (b) generous insurance reduces incentives to avoid claims due to the lower out-of-pocket price of curative care (moral hazard). Although both forces can cause a positive correlation between claims and insurance generosity, the two problems have different implications for public policy (see Einav and Finkelstein, 2011 for a discussion). We sidestep this issue by using a strategy originally proposed by Abbring et al. (2003), namely, using panel data combined with shocks to the cost or benefit of claims to identify the two forces separately. Abbring et al. (2003) exploit shocks in the cost of claims in a bonus-malus car insurance payment scheme to identify moral hazard separately from selection; subsequently, Spenkuch (2012) exploits shocks in the cost of health insurance due to randomized roll-out of Mexico’s Seguro Popular. Our identification strategy exploits changes to the perceived benefits of generous insurance as in Cohen (2005).

With respect to the model, our approach is based on the most commonly used economic framework for health investment, the Grossman (1972) health capital model. In this model, agents make decisions about how much time and money to invest in health to maximize their utility given practical constraints. In Grossman’s original health capital model, there was no uncertainty: agents had perfect knowledge about their health and about the health production process. Previous research has incorporated uncertainty about how health investments translate to future health and productivity into the model using random shocks (Liljas, 1998; Grossman, 1982, 2000). Many of these papers, such as those of Chang (1996), Dardanoni and Wagstaff (1987) and Selden (1993), incorporate uncertainty into a Grossman model where health investment is motivated chiefly by labor market returns. Be-

---

6 As a result, it is difficult to use the model to discuss preventive care as distinct from curative care (Kenkel, 2000). Perhaps because individuals demand more curative care when they are sick, the model’s predictions of a positive relationship between health and demand for health care was not upheld in empirical tests (Wagstaff, 1986); this is still a critique of the Grossman model (Zweifel, 2012).
cause we want our results to generalize to agents who are not in the labor force, we follow
an approach closer to that of Picone et al. (1998), who simulate the effects of uncertainty
about health on demand for preventive medical care among retirees. We expand on this
literature by focusing on endogenous lack of screening, rather than exogenous shocks, as
agents’ key source of uncertainty about their health.

2 Theoretical model

In this section, we analyze a model of demand for screening and demand for medical treat-
ment after diagnosis to show one reason why the two could be mechanically related. In the
model, agents use medical treatment to ameliorate the negative effects of chronic health
conditions. Agents who have been recently screened know whether they have a chronic
condition, whereas agents who have not been recently screened hold beliefs about the prob-
ability they have a chronic condition. Agents differ only in their costs of medical treatment;
we separately model pecuniary and non-pecuniary costs. We analyze this model to derive
predictions about which agents are willing to pay more for screening.

Agents maximize a continuously differentiable function of health \( H \) and consumption
\( C \), net of disutility of medical treatment. Disutility of medical treatment due to non-
pecuniary costs is linear in units of medical treatment \( M \) and the magnitude of disutility
from non-pecuniary costs is captured by \( \theta \), which varies across agents.\(^7\) The utility function
is therefore:

\[
u (C, H(M, D)) - \theta M
\]

\( u (\cdot) \) is concave in \( C \) and \( H \), and agents have weakly higher marginal utility from consump-
tion when they are healthier.

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

\[
C + PM = A
\]

\(^7\)Non-pecuniary costs could be related to factors such as language barriers, distance to a provider, depres-
sion symptoms or other psychological factors which provide barriers to accessing care.
If the agent does not have a diagnosed condition, he is not eligible to receive medical treatment. In this case, therefore, the entire budget is spent on other consumption: \( C = A \).

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

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

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

\[
\pi u(A, H(0, 1)) + (1 - \pi) u(A, H(0, 0)) \tag{1}
\]

2. The agent has been recently screened and knows he does not have a chronic condition \( (D = 0) \).\(^8\) He is not eligible for medical treatment and therefore uses all funds for consumption. His utility is:

\[
u(A, H(0, 0)) \tag{2}
\]

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

\[
\max_{C, M} u(C, H(M, 1)) - \theta M \tag{3}
\]

subject to \( C + PM = A \).

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

Equations (1), (2), and (3) can be combined to describe agents’ willingness to pay for screening. In particular, agents are indifferent between being screened and not being

\(^8\)For simplicity, we present the case where the test is perfectly informative. This assumption can be relaxed without altering the main results.
screened at out-of-pocket price of screening $\kappa$ if:

$$\pi \left( \max_M u (A - PM - \kappa, H (M, 1)) - \theta M \right) + (1 - \pi) u (A - \kappa, H (0, 0))$$
$$- \left( \pi u (A, H (0, 1)) + (1 - \pi) u (A, H (0, 0)) \right) = 0 \quad (4)$$

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

### 2.1 Optimal decisions if screened and $D = 1$

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

$$\frac{\partial u (C, H (M^*, 1))}{\partial H} \frac{\partial H (M^*, 1)}{\partial M} - \theta = P \frac{\partial u (C^*, H (M^*, 1))}{\partial C} \quad (5)$$

The left-hand side of Equation (5) indicates the utility gains from consuming a unit of medical treatment. $\frac{\partial u (C, H (M^*, 1))}{\partial H} \frac{\partial H (M^*, 1)}{\partial M}$ is the utility benefit from improved health and $-\theta$ is the disutility of consuming a unit of medical treatment due to non-pecuniary costs. The right-hand side of Equation (5) indicates the utility gains from spending $P$ additional dollars on consumption rather than on medical treatment. Therefore Equation (5) indicates that at the optimal point, the marginal benefits of purchasing a unit of medical treatment equal the marginal benefits of using the same funds for consumption.

### 2.2 Analysis of marginally screened individuals and empirical predictions

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

**Proposition 2.1** Willingness to pay for screening is decreasing in agents’ costs of medical treatment: $\frac{\partial \kappa^*}{\partial \theta} < 0$ and $\frac{\partial \kappa^*}{\partial P} < 0$, respectively.
The proofs are based on the envelope theorem. See Appendix D.

**Proposition 2.2** Demand for medical treatment after diagnosis is also decreasing in agents’ costs of medical treatment: \( \frac{\partial M^*}{\partial \theta} < 0 \) and \( \frac{\partial M^*}{\partial P} < 0 \).

See Appendix E for the proofs.

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

If we allowed the benefits of medical treatment rather than the costs of medical treatment to vary across agents, we could produce the same empirical predictions in some cases. However, this exercise would be complicated by the fact that patients learn about the benefits of medical treatment through screening. In practice, patients are unaware of the presence or severity of their asymptomatic conditions prior to screening, and these factors determine the benefits of medical treatment. This presents the problem that patients’ ex ante beliefs about the benefits of medical treatment determine their willingness to pay for screening, whereas their (different) ex post beliefs determine demand for medical treatment. If all patients hold correct beliefs about the benefits of medical treatment after screening, then gaps in care of diagnosed conditions among newly-screened and previously-screened patients would only appear if the true benefits to treatment were lower for newly diagnosed patients. This could be the case if newly diagnosed patients have less severe conditions overall. We explore this possibility in the empirical analysis to follow, and conclude that condition severity is unlikely to account for our findings.
3 Links between screening and treatment of previously diagnosed conditions on the individual-level

In this section we show that individuals who are not regularly screened, or whose undiagnosed conditions are detected only after an epidemiological study pays them to have their biomarkers assessed, differ from other individuals in their use of treatment for their previously diagnosed conditions. This pattern remains after adjusting for insurance status, condition severity, and prevalence of diagnosed and undiagnosed comorbid conditions. We also replicate the results using data from three groups that are likely to be affected by ACA provisions limiting cost-sharing for screening. If treatment of previously diagnosed conditions predicts care for newly diagnosed conditions, these results imply that the impact of low-cost screenings on treatment of diabetes, hypertension, and high cholesterol could be blunted by the lower use of medical treatment among patients with previously undiagnosed conditions.

3.1 Data

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

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 1999-2013, 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). REGARDS is an epidemiological study of older adults that recruited participants across the continental United States over 2003-2007 using a commercial list of residential phone numbers (Howard et al., 2005). The REGARDS data have been linked with administrative records of doctor visits for participants enrolled in traditional Medicare (Muntner et al., 2014). Finally, the Oregon Health Insurance Experiment (OHIE) baseline biomarker study was conducted during 2009-2010 and sampled adults who entered a lottery to apply for Medicaid in Oregon in 2008 (Allen et al., 2010; Baicker et al., 2013).
These three datasets have different advantages and disadvantages for our analysis. First, the NHANES data provide self-reported information on whether a doctor had ever recommended managing hypertension and high cholesterol using a prescription, whereas the REGARDS and OHIE data do not. This is important because national guidelines recommend treating less severe cases of these conditions with diet and exercise before prescribing medication (James et al., 2014; Stone et al., 2014). By tracking medication use only among participants who report that their doctor recommended medication, we can ensure that our results are not driven by medication non-use among patients whose doctors recommended controlling the condition through diet and exercise alone. As a nationally representative survey, the NHANES also samples the most diverse group of participants.

In contrast, the OHIE baseline data have a different advantage for the present analysis. Adding these data allows us to pursue a focused analysis of a group of importance for the ACA: applicants to expanded Medicaid. In Medicaid expansion states, many patients who become diagnosed due to the ACA could come from this group.

Finally, the OHIE and NHANES data have the disadvantage of relying exclusively on participant self-report to measure diagnosis and treatment of conditions, which could be an important source of measurement error (Meyer et al., 2015). To address this issue, we also analyze data from participants in REGARDS with merged Medicare claims from the the two years prior to participation. For this group, we can analyze claims data on doctor visits for evaluation and management of diagnosed diabetes, high cholesterol, and/or hypertension rather than relying only on self-reported data.

**Identifying diagnosed and undiagnosed conditions** In each of the three datasets, 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, if their biomarkers meet standard definitions for the condition after taking their fasting status into account, or if they are taking medications that indicate the condition in a medication review (if applicable for that study) (American Diabetes Association, 2014; Stone et al., 2014; James et al., 2014). Table 7 in the Appendix includes details of each definition. Individuals are classified as undiagnosed for the condition if they meet the biomarker definitions for a condition, but report no prior diagnosis for that condition.

In the REGARDS data, we correct for under-reporting by also classifying participants as diagnosed if they show biomarkers relevant to the the condition and their doctors have been regularly evaluating and managing their condition based on recent claims. In par-
<table>
<thead>
<tr>
<th>Survey Inclusion Criteria</th>
<th>NHANES</th>
<th>OHIE Baseline</th>
<th>REGARDS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Geography of Sample</td>
<td>National</td>
<td>Oregon</td>
<td>National</td>
</tr>
<tr>
<td>Age Range in Analysis</td>
<td>All</td>
<td>19+</td>
<td>67+</td>
</tr>
<tr>
<td>Participants with Any Condition(s) of Interest</td>
<td>25,332</td>
<td>7,108</td>
<td>5,884</td>
</tr>
<tr>
<td>Participants with Undiagnosed Condition(s) of Interest</td>
<td>8,676</td>
<td>4,845</td>
<td>1,322</td>
</tr>
</tbody>
</table>

**Among Participants with Condition(s) of Interest:**

- **Average Age**
  - 55
  - 43
  - 75
- **Had Health Insurance**
  - 82%
  - 44%
  - 100%
- **African American**
  - 21%
  - 11%
  - 30%
- **Participants with Diabetes**
  - 5,627
  - 1,007
  - 1,309
  - Aware of Diabetes
    - 4,569
    - 847
    - 1,166
  - Treating with Medication
    - 3,911
    - 819
    - 1,161
- **Participants with Hypertension**
  - 15,598
  - 3,117
  - 4,502
  - Aware of Hypertension
    - 13,675
    - 2,169
    - 4,051
  - Treating with Medication
    - 10,154
    - 1,530
    - 3,846
- **Participants with High Cholesterol**
  - 18,384
  - 5,714
  - 4,268
  - Aware of High Cholesterol
    - 11,876
    - 1,511
    - 3,394
  - Treating with Medication
    - 6,139
    - 976
    - 2,457
ticular, we also classify participants as diagnosed if they meet the biomarker definitions for a condition and had two evaluation and management visits coded as relevant to that condition in the 24 months prior to REGARDS participation, i.e., meet Chronic Conditions Warehouse definitions for the condition based on their claims data at the time of REGARDS participation. 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%).

3.2 Analysis

We first use NHANES data to show that use of recommended treatment for diagnosed conditions is lower among individuals with undiagnosed conditions. Bivariate regressions indicate that participants with undiagnosed conditions are less likely to report taking their prescribed medications for diagnosed hypertension or high cholesterol, or having a foot exam or eye exam over the past year for their diagnosed diabetes.9 (Doctors’ recommendations to control hypertension and high cholesterol using medication are asked about in the NHANES, enabling us to track medication use only among diagnosed patients for whom medication was recommended. However, there is no comparable question for diabetes.) See row 1 of Table 2.

One might argue that if people with undiagnosed conditions have less severe conditions overall, lower treatment rates and screening rates in this group would represent an appropriate allocation of resources. As noted previously, the tractability of this argument is limited by the fact that people cannot know the severity of asymptomatic conditions without screening and most cases of hypertension and high cholesterol and many cases of diabetes are asymptomatic. Furthermore, we find in all three biomarker datasets that participants with undiagnosed conditions show more severe, not less severe, biomarkers for their diagnosed conditions. See Appendix Table 8.

We address this argument in the analysis by adding controls for patients’ biomarkers including LDL and HDL cholesterol, HbA1c, and systolic and diastolic blood pressure to the bivariate analysis above. We also adjust for self-reported retinopathy, a diabetes symptom that is consistently measured across different waves of the NHANES survey, to account for the possibility that onset of diabetes symptoms could spur demand for treatment of diabetes and screening for other conditions. As shown in row 2 of Table 2, findings are

---

9 Annual foot exams and eye exams are recommended as part of standard diabetes care (American Diabetes Association, 2014).
similar. Findings are also similar when we adjust for prevalence and comorbidity of other conditions rather than using continuous biomarkers, as shown in row 3 of Table 2.

Another possibility is that high out-of-pocket costs for treating diagnosed conditions could reduce patients’ willingness to be screened and also reduce their use of treatment of diagnosed conditions, as proposed in Section 2. We cannot quantify the importance of this channel without knowing more about participants’ out-of-pocket costs of care. Using the available data, however, we find that results remain similar when we control for predictors of out-of-pocket costs and access to care such as current health insurance status, lack of coverage at any point during the past year, current Medicaid coverage, current Medicare coverage, year of age, current calendar year, black race, and current prescription coverage; we allow the effect of health insurance coverage on use of health care to vary by year by adding interactions. See rows 4 and 5 of Table 2. We also find similar results when the analysis is repeated including only participants with health insurance; see row 6 of Table 2.

The analysis above documents the relationships between prevalence of undiagnosed conditions and treatment of diagnosed conditions. However, prevalence of undiagnosed conditions is determined by two rates: prevalence of conditions, and frequency of screening. Therefore, we conduct an additional test to isolate the relationships between frequency of screening and treatment. In this test, we compare use of recommended treatment for diagnosed conditions among patients who do vs. do not self-report taking a blood test for diabetes or high cholesterol if not already diagnosed for these conditions in the past 3 years.\(^\text{10}\) In other words, we compare participants who have vs. have not had a recent blood test to check for diabetes if they have not been diagnosed for diabetes, or blood test to check for high cholesterol if they have not been diagnosed for high cholesterol. (NHANES does not include data on time since last blood pressure test.) Table 3 shows the results. In harmony with the previous results, we find that use of recommended treatment for diagnosed conditions is lower for patients who have not taken a blood test to check for diabetes or high cholesterol in the past 3 years.\(^\text{11}\) As in the previous analysis, we find that this relationship holds after adjusting for severity of diagnosed conditions, prevalence of comorbid conditions, and factors related to access to care such as race, age, and health insurance status.

\(^\text{10}\)Similar questions are not available in the REGARDS data, and only available for cholesterol in the OHIE data. However, because the NHANES are nationally representative, results obtained using only the NHANES are still helpful for policy purposes.

\(^\text{11}\)Because we do not know the timing of diagnosis, we code the variable so that participants who have diagnosed diabetes need not be screened for diabetes, and participants with diagnosed high cholesterol need not be screened for high cholesterol.
Table 2: Prevalence of undiagnosed conditions and use of recommended care for diagnosed diabetes, hypertension, or high cholesterol (NHANES data)

<table>
<thead>
<tr>
<th></th>
<th>(1) Diabetic</th>
<th>(2) Diabetic</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Eye Exam</td>
<td>Foot Exam</td>
<td>Hypertension Meds</td>
<td>Cholesterol Meds</td>
</tr>
<tr>
<td>(1) No controls</td>
<td>-0.138***</td>
<td>-0.0813***</td>
<td>-0.0340***</td>
<td>-0.0276</td>
</tr>
<tr>
<td></td>
<td>(0.0238)</td>
<td>(0.0229)</td>
<td>(0.00838)</td>
<td>(0.0169)</td>
</tr>
<tr>
<td>Control for:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Biomarkers</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0975***</td>
<td>-0.0902***</td>
<td>-0.0162</td>
<td>-0.0282</td>
</tr>
<tr>
<td></td>
<td>(0.0351)</td>
<td>(0.0335)</td>
<td>(0.0127)</td>
<td>(0.0231)</td>
</tr>
<tr>
<td>(3) Comorbidity</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.148***</td>
<td>-0.101***</td>
<td>-0.0660***</td>
<td>-0.0865***</td>
</tr>
<tr>
<td></td>
<td>(0.0251)</td>
<td>(0.0239)</td>
<td>(0.00864)</td>
<td>(0.0173)</td>
</tr>
<tr>
<td>(4) Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.104***</td>
<td>-0.0585**</td>
<td>-0.0105</td>
<td>-0.0475***</td>
</tr>
<tr>
<td></td>
<td>(0.0235)</td>
<td>(0.0228)</td>
<td>(0.00794)</td>
<td>(0.0164)</td>
</tr>
<tr>
<td>(5) All controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0771**</td>
<td>-0.109***</td>
<td>-0.0129</td>
<td>-0.0657***</td>
</tr>
<tr>
<td></td>
<td>(0.0361)</td>
<td>(0.0345)</td>
<td>(0.0127)</td>
<td>(0.0237)</td>
</tr>
<tr>
<td>(6) Insured only</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>-0.0674*</td>
<td>-0.108***</td>
<td>-0.0100</td>
<td>-0.0586**</td>
</tr>
<tr>
<td></td>
<td>(0.0396)</td>
<td>(0.0371)</td>
<td>(0.0126)</td>
<td>(0.0237)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

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

This table shows the relationship between use of recommended treatment for diagnosed conditions and prevalence of undiagnosed diabetes, hypertension, and/or high cholesterol. The rows include coefficients and standard errors obtained from linear probability models after adjusting for the listed control variables. The outcomes in columns 1-2 are self-reported foot exams or eye exams in the past year among participants who reported prior diagnosis of diabetes. The outcome in column 3 is self-reported use of medication for hypertension among participants among participants who reported prior diagnosis of hypertension and reported that a doctor recommended anti-hypertensive medication. The outcome in column 4 is self-reported use of medication for high cholesterol among participants who self-reported prior diagnosis of high cholesterol and reported that a doctor recommended cholesterol-lowering medication.
Table 3: Recent use of a blood test to screen for asymptomatic undiagnosed conditions and use of recommended care for diagnosed diabetes, hypertension, or high cholesterol (NHANES data)

<table>
<thead>
<tr>
<th>(1) No controls</th>
<th>(2) Biomarkers</th>
<th>(3) Comorbidity</th>
<th>(4) Demographics</th>
<th>(5) All controls</th>
<th>(6) Insured only</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Diabetic Eye Exam</td>
<td>Diabetic Foot Exam</td>
<td>Hypertension Meds</td>
<td>Cholesterol Meds</td>
<td></td>
</tr>
<tr>
<td>Not screened last 3 years</td>
<td>-0.339***</td>
<td>-0.320***</td>
<td>-0.0372***</td>
<td>-0.0799***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0617)</td>
<td>(0.0604)</td>
<td>(0.0110)</td>
<td>(0.0142)</td>
<td></td>
</tr>
<tr>
<td>Control for:</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) Biomarkers</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not screened last 3 years</td>
<td>-0.353***</td>
<td>-0.229***</td>
<td>-0.0395**</td>
<td>-0.0537***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0882)</td>
<td>(0.0871)</td>
<td>(0.0159)</td>
<td>(0.0194)</td>
<td></td>
</tr>
<tr>
<td>(3) Comorbidity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not screened last 3 years</td>
<td>-0.339***</td>
<td>-0.322***</td>
<td>-0.0460***</td>
<td>-0.0569***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0778)</td>
<td>(0.0767)</td>
<td>(0.0123)</td>
<td>(0.0152)</td>
<td></td>
</tr>
<tr>
<td>(4) Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not screened last 3 years</td>
<td>-0.263***</td>
<td>-0.260***</td>
<td>-0.0391***</td>
<td>-0.0607***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0619)</td>
<td>(0.0612)</td>
<td>(0.0105)</td>
<td>(0.0137)</td>
<td></td>
</tr>
<tr>
<td>(5) All controls</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not screened last 3 years</td>
<td>-0.322***</td>
<td>-0.256**</td>
<td>-0.0436***</td>
<td>-0.0385*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0982)</td>
<td>(0.0997)</td>
<td>(0.0158)</td>
<td>(0.0198)</td>
<td></td>
</tr>
<tr>
<td>(6) Insured only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Not screened last 3 years</td>
<td>-0.310**</td>
<td>-0.300**</td>
<td>-0.0344**</td>
<td>-0.0391*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.128)</td>
<td>(0.125)</td>
<td>(0.0158)</td>
<td>(0.0203)</td>
<td></td>
</tr>
</tbody>
</table>

Standard errors in parentheses

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

This table shows the relationship between use of recommended treatment for diagnosed conditions and use of a blood test to check for asymptomatic diabetes or high cholesterol in the past 3 years. The rows include coefficients and standard errors obtained from linear probability models after adjusting for the listed control variables. The outcomes in columns 1-2 are self-reported foot exams or eye exams in the past year among participants who reported prior diagnosis of diabetes. The outcome in column 3 is self-reported use of medication for hypertension among participants who reported prior diagnosis of hypertension and reported that a doctor recommended anti-hypertensive medication. The outcome in column 4 is self-reported use of medication for high cholesterol among participants who self-reported prior diagnosis of high cholesterol and reported that a doctor recommended cholesterol-lowering medication.
tus, and that the relationship holds when we restrict the sample to only include participants with health insurance. Furthermore, the results are more consistently significant in Table 3 than Table 2, which would support a hypothesis that the relationship between undiagnosed conditions and treatment is driven by the underlying relationship between screening and treatment.

The analysis thus far has the shortcoming that diagnosis and treatment of chronic conditions are only measured using self-reported data. To address this shortcoming, we conduct additional checks using claims data available in the merged Medicare-REGARDS data. In this analysis, our main outcome of interest is doctor visits for evaluation and management of diagnosed conditions in the previous year, measured using Medicare claims assigned to conditions based on the Chronic Conditions Warehouse classifications. As shown in Table 4, we find that participants with previously undiagnosed conditions had fewer doctor visits for their previously diagnosed conditions. As before, this relationship holds after adjusting for severity of diagnosed conditions, prevalence of comorbid conditions, and factors related to access to care such as race, age, and Medicaid dual eligibility.

The role of health insurance is worth exploring further, because the ACA regulations that prevent health insurance plans from imposing cost-sharing for screening (for high risk patients) directly change out-of-pocket costs for screening among individuals with health insurance. Individuals with health insurance after the ACA can be broken into two groups: individuals who were insured prior to the ACA and remained insured, and individuals who become insured after the ACA. Some individuals in the latter group may have become insured as a result of elements of ACA implementation such as health insurance mandates, state-level exchanges, and Medicaid expansions. Table 5 replicates the analysis in Table 2 for each of these groups of interest, adjusting for prevalence of comorbid conditions and biomarker measures of condition severity. In row 1 of Table 5, we include data from individuals who were insured before the ACA provisions came into effect; this analysis uses the NHANES data. In rows 2 and 3 of Table 5, we include data from individuals who applied for expanded Medicaid in Oregon; this analysis uses the OHIE data. We separately analyze all applicants to the Oregon Medicaid program vs. applicants who were uninsured at the time of application. Finally, because additional ACA provisions eliminate cost-sharing for annual wellness visits among Medicare enrollees, row 4 of Table 5 includes

---

12 An exception is the so-called “grandfathered” health plans, which covered enrollees before the ACA became law (March 23, 2010) and have not made substantial changes since that time. The 2015 Kaiser Family Foundation’s Employer Health Benefits Survey found that 25% of covered workers were enrolled in a grandfathered health plan in 2015 (Claxton et al., 2015).
### Table 4: Prevalence of undiagnosed conditions and number of annual doctor visits for diagnosed diabetes, hypertension, or high cholesterol (REGARDS data)

<table>
<thead>
<tr>
<th></th>
<th>(1) Diabetes Claims</th>
<th>(2) Hypertension Claims</th>
<th>(3) High Cholesterol Claims</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(1) No controls</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>-1.089***</td>
<td>-0.617***</td>
<td>-0.446***</td>
</tr>
<tr>
<td></td>
<td>(0.380)</td>
<td>(0.159)</td>
<td>(0.141)</td>
</tr>
<tr>
<td><strong>(2) Control for biomarkers</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>-1.122**</td>
<td>-0.743***</td>
<td>-0.377**</td>
</tr>
<tr>
<td></td>
<td>(0.435)</td>
<td>(0.178)</td>
<td>(0.152)</td>
</tr>
<tr>
<td><strong>(3) Control for comorbidity</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>-1.260***</td>
<td>-0.864***</td>
<td>-0.543***</td>
</tr>
<tr>
<td></td>
<td>(0.387)</td>
<td>(0.164)</td>
<td>(0.144)</td>
</tr>
<tr>
<td><strong>(4) Control for demographics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>-1.396***</td>
<td>-0.685***</td>
<td>-0.408***</td>
</tr>
<tr>
<td></td>
<td>(0.378)</td>
<td>(0.155)</td>
<td>(0.141)</td>
</tr>
<tr>
<td><strong>(5) All controls</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>-1.140***</td>
<td>-0.895***</td>
<td>-0.353**</td>
</tr>
<tr>
<td></td>
<td>(0.435)</td>
<td>(0.182)</td>
<td>(0.155)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

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

This table shows the relationship between doctor visits for diagnosed conditions and prevalence of undiagnosed diabetes, hypertension, and/or high cholesterol. Due to the use of Medicare claims data, we only include individuals who had fee-for-service Medicare insurance the two years prior to REGARDS participation. The rows indicate coefficients of a linear regression model after adjusting for the listed control variables. The outcome in column 1 is the number of evaluation and management visits from the prior year coded as relevant to diabetes among participants with prior diagnosis of diabetes. Likewise, the outcomes in columns 2 and 3 are the number of evaluation and management visits from the prior year coded as relevant to hypertension or high cholesterol for participants with prior diagnosis of hypertension or high cholesterol, respectively.
Table 5: Replicating the previous analysis for groups likely impacted by ACA provisions that reduce cost-sharing for screening (NHANES, OHIE and REGARDS data)

<table>
<thead>
<tr>
<th></th>
<th>(1) Diabetes Meds</th>
<th>(2) Hypertension Meds</th>
<th>(3) Cholesterol Meds</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NHANES data:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(1) All insured prior to ACA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>0.0004</td>
<td>-0.0841***</td>
<td>-0.0767**</td>
</tr>
<tr>
<td></td>
<td>(0.0337)</td>
<td>(0.0212)</td>
<td>(0.0317)</td>
</tr>
<tr>
<td><strong>OHIE data:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(2) All applicants for Medicaid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>-0.0634*</td>
<td>-0.0975***</td>
<td>-0.125***</td>
</tr>
<tr>
<td></td>
<td>(0.0338)</td>
<td>(0.0262)</td>
<td>(0.0461)</td>
</tr>
<tr>
<td>(3) Uninsured applicants for Medicaid</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>-0.0644</td>
<td>-0.104***</td>
<td>-0.138**</td>
</tr>
<tr>
<td></td>
<td>(0.0516)</td>
<td>(0.0358)</td>
<td>(0.0601)</td>
</tr>
<tr>
<td><strong>REGARDS data:</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(4) Medicare insured prior to ACA</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>0.0103</td>
<td>-0.0165</td>
<td>-0.111***</td>
</tr>
<tr>
<td></td>
<td>(0.0175)</td>
<td>(0.0149)</td>
<td>(0.0321)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

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

This table uses data from the NHANES, OHIE, and REGARDS studies to investigate the relationship between undiagnosed conditions and treatment of diagnosed conditions. These three data sources are used to study individuals who had any health insurance prior to the ACA, who wished to apply for expanded Medicaid insurance prior to the ACA, and who had Medicare insurance prior to the ACA, respectively. All models adjust for severity of conditions measured using biomarkers and prevalence of comorbid conditions.
data from Medicare enrollees prior to the ACA using the REGARDS data. The findings in Table 5 indicate that, among groups likely to be affected by the ACA provisions related to cost-sharing for screening, the patterns found in Table 2 largely persist.

The findings in this section indicate that demand for screening is related to demand for medical treatment of previously diagnosed conditions in a way not fully explained by prevalence of comorbid conditions, condition severity, or health insurance status. In the next section, we show additional evidence of low demand for medical treatment among patients with undiagnosed conditions: a lack of adverse selection after participants are informed of a previously undiagnosed condition.

4 Diagnosis of a chronic condition and health plan selection

When patients sort into health insurance based on health risk factors that are known to them but cannot be incorporated into premium prices, this pattern - adverse selection - can cause health insurance markets to unravel. If patients demand more generous health insurance after learning about previously undiagnosed conditions, this would comprise adverse selection because diagnoses can no longer be taken into account in premium prices after the ACA (Handel et al., 2015). Therefore, it is important to investigate whether patients switch to more generous health insurance after learning about previously undiagnosed conditions.

4.1 Data

To determine the impact of learning about previously undiagnosed conditions on health insurance plan selections, we use the REGARDS study as a natural experiment. REGARDS recruited participants from across the continental United States over a period of four years (2003-2007) by making random selections using a commercial list of residential phone numbers (Howard et al., 2005). Participants first reported their diagnosis status in a survey, and then had their fasting blood glucose, blood pressure, and lipid panel assessed in their home on a morning of their choosing. All participants were notified of their results using standard text and paid $30 as compensation for their participation. It is because of the notifications and compensation that REGARDS can be considered an intervention that provides free screenings in addition to an epidemiological study. Additional details on REGARDS data collection procedures are included in Appendix B.
Outcome of interest

Linked Medicare data are available for REGARDS participants who were Medicare beneficiaries. These data track participants’ health insurance plan selections on a monthly basis, including our outcome of interest: enrollment in Medicare Advantage, the voluntary HMO alternative to Medicare’s default fee-for-service health plan (hereafter, traditional Medicare).

Enrollment in Medicare Advantage is not the only relevant margin for adverse selection in the Medicare market. Adverse selection could also involve purchasing supplemental coverage (Medigap plans), enrolling in physician coverage (Medicare Part B), or switching to a more generous Medicare Advantage plan. However, our data do not include information on Medigap plans and do not detail the generosity of Medicare Advantage plans. In addition, 95% of participants who were age-eligible for Medicare were already enrolled in Medicare Part B prior to participation in REGARDS; see Figure 1. Therefore, participants who were not enrolled in Medicare Part B prior to REGARDS are likely a highly selected sample which would raise concerns about including Part B enrollment as an outcome of interest.

Sample selection

This analysis uses data from REGARDS participants who were 63 years old or older at the time of REGARDS participation. As in the previous section, we only include participants who became Medicare beneficiaries as a result of turning 65 (rather than due to illness or disability). Including people who participated in REGARDS before age 65 enables us to identify the effect of REGARDS participation on Medicare Advantage enrollment separately from age effects. This is important because age is a predictor of Medicare Advantage enrollment, as shown in Figure 1.

To ensure that any observed health plan switches could plausibly be attributed to REGARDS, we narrow our window of observation to the 24 months around the time of REGARDS participation. We therefore exclude participants with no health plan selection data available from the 24 months after their participation in REGARDS. These exclusion criteria yield a dataset with 9,990 REGARDS participants. See Appendix Table 9 for additional details.
Figure 1: Insurance decisions by age, measured in the month prior to REGARDS participation

This figure shows two trends in the Medicare plan selection data that inform our analytic strategy. First, almost all age-eligible Medicare enrollees were enrolled in Medicare Part B prior to REGARDS participation. (All participants in Medicare Advantage are also enrolled in Part B.) Second, Medicare Advantage enrollment varies by age: enrollment almost doubles from age 65 to age 67 and declines again after age 77.
4.2 Context of the Medicare Advantage market and predicted effects

To provide context for the analysis to follow, it is useful to discuss the structure of the Medicare market, the role of Medicare Advantage, and how Medicare Advantage policy changed during our sampling period. Upon aging into Medicare at age 65, beneficiaries are automatically enrolled in traditional Medicare. Beneficiaries may then choose to enroll in Medicare Advantage plans in any subsequent year, including their first year of Medicare; beneficiaries may switch between Medicare Advantage and traditional Medicare as many times as they wish. Medicare Advantage plans receive capitated, risk-adjusted payments from the government and are required by law to be as generous as traditional Medicare in terms of covering a minimal suite of services, with total out-of-pocket payments (premiums and cost-sharing) not higher than traditional Medicare (Newhouse et al., 2014). Plans attract customers by offering supplemental benefits such as vision or dental coverage, or by reducing patients’ cost-sharing, made possible by negotiating physician payments in restricted provider networks (Newhouse, 2002).

Given the capitated-payment reimbursement system, Medicare Advantage plans can earn more money by attracting the patients with lowest health care costs within each capitated payment risk-adjustment bin. There is evidence that plans accomplished this task during the period of our data, despite requirements to accept all applicants (Newhouse et al., 2014). Indeed, analysis of patients who switched plans indicate advantageous selection into Medicare Advantage: that is, patients with increasing health care costs were likely to switch to traditional Medicare, and vice versa (Newhouse et al., 2014; McWilliams et al., 2012; Morgan et al., 1997). Plans reduced their exposure to high-cost patients by not entering higher-cost counties and by structuring provider networks and drug formularies so as to influence which patients wished to enroll (Cao and McGuire, 2003; Frank et al., 2000; Rothschild and Stiglitz, 1976).

A number of changes occurred in the Medicare Advantage market in 2006. First, Medicare Part D was introduced in 2006, allowing traditional Medicare to become more comparable to HMO coverage. Second, Medicare fought against advantageous selection into Medicare Advantage by phasing in an improved model for calculating individual-level risk-adjustment payments, the hierarchical condition categories model (Newhouse et al., 2014). This model was over six times as predictive of expenditures as the previous model (Pope et al., 2004). Third, Medicare sought to lower prices and increase generosity of Medicare Advantage plans to consumers by providing “rebates” as supplemental benefits to all enrollees in a given plan using a bid-based system (Newhouse and McGuire, 2014;
Fourth, open enrollment and lock-in periods were enforced beginning in January 2006. Before 2006, beneficiaries could change plans once per month. In 2006, a lock-in period was introduced covering the latter six months of the year; in 2007 it was extended to 9 months (McGuire et al., 2011). This lock-in period resembles the open enrollment periods in the health insurance marketplaces established under the ACA.

Based on the particular features of the Medicare Advantage market, we hypothesize that a Medicare beneficiary who is currently enrolled in traditional Medicare will become more likely to enroll in Medicare Advantage after screening detected a previously undiagnosed chronic condition if several conditions hold: (a) she anticipates an increase in doctor visits as a result of the diagnosis; (b) there is a Medicare Advantage plan available to her that provides her anticipated bundle of services for a lower out of pocket price than that of traditional Medicare; and (c) the benefits to enrolling in Medicare Advantage are not outweighed by switching costs or hassles of a restricted provider network. If any of these conditions do not hold, the beneficiary will not wish to switch to Medicare Advantage after learning of a screen-detected condition. Furthermore, if (b) does not hold, beneficiaries who are enrolled in Medicare Advantage may wish to switch back to traditional Medicare after learning of a screen-detected condition.

4.3 Analytic plan

Our empirical strategy compares health plan selections of recently screened vs. not-yet-screened REGARDS participants with vs. without undiagnosed conditions. We conduct these comparisons before and after policy changes that occurred in 2006 using the following specification of a linear probability model:

\[ MA_{it} = \sum_{r \in \text{pre 2006, post 2006}} (\mu_{0r}^T + T_{it}^r \phi_{1r}^T + T_{it}^r U_i^r \phi_{2r}^T) + X_{it} B + \alpha_i + \omega_{it} \]  

\( MA_{it} \) indicates whether individual \( i \) was enrolled in Medicare Advantage \( t \) months before or since REGARDS, \( r \) indicates the policy regime at time \( t \) (before vs. after changes in January 2006), and \( U_i \) denotes whether the participant has any undiagnosed conditions. \( X_{it} \) denotes the control variables listed below and lower-level interaction terms, and \( \alpha_i \) denotes an individual-level random effect. We use heteroskedasticity-robust standard errors clustered by individual.
The $\phi_2^r$ coefficients are the coefficients of interest in the model. If the $\phi_2^r$ are significantly different from zero, this would support a hypothesis that screening for diabetes, hypertension, and high cholesterol has an additional effect on plan selection for patients with previously undiagnosed conditions.

In additional analyses, we examine the influence of switching costs by controlling for polynomials of participants’ log duration of enrollment in their current Medicare plan type (Eberwein et al., 2002). This strategy leveraged the finding from previous literature that duration of plan enrollment can proxy for switching costs or “status quo bias” in the Medicare market (Samuelson and Zeckhauser, 1988; Sinaiko et al., 2014).

As a robustness check, we also model switches into Medicare Advantage rather than modeling current enrollment. In this analysis, the specification of the right-hand side of the regression is the same as equation (6), but the outcome is a switch into Medicare Advantage. In this analysis, we restrict the data to only include person-months in traditional Medicare (i.e., time when participants were at "risk" for switching into Medicare Advantage).

To examine whether particular groups are driving the results, we also interact the coefficients of interest by participant characteristics relevant to plan selection, such as prior healthcare use or duration of enrollment in the current plan type. As before, all relevant lower-order interaction terms are included in the model.

**Selection of control variables**

Control variables are selected to address two possible biases. First, we expect that secular trends contribute to observed changes in doctor visits after REGARDS participation. For example, all participants were older after REGARDS participation than before REGARDS participation, and policy changes were implemented during our period of observation. These secular trends could bias our estimates if not controlled for in the model. The rolling nature of the implementation of screenings in the REGARDS study makes controlling for secular trends possible: there is random variation in age and calendar time of REGARDS participation because participants received their offers to participate using random phone calls over a number of years. To this end, we include a number of time-varying control variables including year dummies, interactions between region and year, and individual age, divided into 8 bins of equal size to allow for a non-linear relationship between age and doctor visits. We also include an indicator variable for open enrollment season and an interaction with this variable with an indicator for 2006 or later.

Second, our results might be biased if the type of individual willing to participate in RE-
GARDS changed over time. This would be problematic because the models compare not-yet-recruited participants with recently-recruited individuals to control for secular trends. Therefore, we also include a number of time-invariant control variables such as physical health measures taken at the time of REGARDS participation and a number of demographic and health-related characteristics from the REGARDS survey. In particular, we control for waist size in centimeters, BMI, glucose, lipid panel, the average of two blood pressure measures (both systolic and diastolic) and reported physical health from the SF-12; type of condition (high cholesterol, hypertension, or diabetes), and whether the condition was previously undiagnosed; race (African American or white), sex (male or female), income (less than $20,000, $20,000-$35,000, $35-$75,000, and over $75,000), education (less than high school education, high school, some college education, or graduated from college), fair or poor self-reported health, usual health provider at the time of the interview (self-reported having or not-having a usual health provider), self-reported smoking status (current smoker, past smoker, or non-smoker), number of alcoholic drinks per week, fasting status at the time of the interview (fasting or not), cognitive status according to a short memory test (impaired or not), Medicaid dual eligibility in 2008 (eligible or not), status of county as a primary care health professional shortage area (all, part, or none of the county is a designated health professional shortage area), and the fraction of residents in poverty in the participant’s county of residence. All continuous variables are binned into four categories of equal size to allow non-linearity in the relationships between these variables and health plan selections.

4.4 Results

We have data on Medicare Advantage enrollment status in the month prior to REGARDS participation for 9,239 included participants. 15% of these participants were enrolled in a Medicare Advantage plan and 95% of participants were enrolled in Medicare Part B. Participants who were enrolled in Medicare Advantage prior to REGARDS participation are more likely to be African American or low income than those who were not enrolled in Medicare Advantage, but have similar prevalence of undiagnosed conditions. See Appendix Table 10. These findings match previous published findings based on linked survey and Medicare records data from similar years (McWilliams et al., 2011).

The regression models show no evidence that participants switched from traditional Medicare to Medicare Advantage or vice versa after learning of a previously undiagnosed
Table 6: Learning of previously undiagnosed diabetes, hypertension, or high cholesterol via REGARDS and subsequent enrollment in Medicare Advantage

<table>
<thead>
<tr>
<th>Prior to January 1, 2006</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Currently in MA</td>
<td>-0.019*</td>
<td>-0.021</td>
<td>0.001</td>
<td>5.15e-05</td>
</tr>
<tr>
<td>Currently in MA</td>
<td>(0.010)</td>
<td>(0.015)</td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>After January 1, 2006</td>
<td>-0.013</td>
<td>-0.023</td>
<td>-0.000</td>
<td>0.003</td>
</tr>
<tr>
<td>After January 1, 2006</td>
<td>(0.0169)</td>
<td>(0.0199)</td>
<td>(0.003)</td>
<td>(0.004)</td>
</tr>
</tbody>
</table>

| Observations | 165,642 | 75,276 | 137,295 | 58,769 |
| Number of participants | 3,405 | 1,682 | 3,010 | 1,368 |

Control for duration dependence | N | Y | N | Y

Robust standard errors in parentheses

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

MA: Medicare Advantage. This table shows that we fail to reject the null hypothesis of no significant association between learning of a previously undiagnosed condition and enrollment in, or switches into, Medicare Advantage. The rows of the table include coefficients of interest from the regression models described in equation (6). “Undiagnosed” denotes the participant had at least one undiagnosed condition (hypertension, high cholesterol, and/or diabetes) prior to participation in REGARDS. The results in columns 2 and 4 show that controlling for polynomials of log duration of enrollment in the current plan type, our proxy for inertia or switching costs, does not affect the results.

These patterns hold both before and after the policy changes in 2006, and hold regardless of whether we analyze current enrollment in Medicare Advantage or analyze switches directly. The findings also do not change after we control for duration of enrollment in the current plan type, our proxy for inertia or switching costs. The coefficients of interest from these models are included in Table 6. The raw data show a similar pattern; see Appendix Figure 2.

The findings are also unchanged when we interact the coefficients of interest by participant characteristics relevant to plan selection such as prior claims or self-reported health. See Appendix Tables 11 and 12.

In summary, we find no evidence that free screenings provided by the REGARDS study resulted in adverse selection. This finding follows intuitively from our prediction in sec
tion 2 that marginally screened patients will have lower demand for medical treatment. If marginally screened patients have lower demand for medical treatment, they would also have less reason to adversely select after learning of a previously undiagnosed condition.

However, there are multiple reasons to interpret our empirical results with caution or doubt the external validity. Our data do not include details on cost sharing for specific health plans and we do not observe switches across Medicare Advantage plans. In addition, the results may be shaped by the particular environment of the Medicare Advantage market, such as the use of risk-adjusted per capita payment schemes. We conclude that although these estimates contribute to the literature by examining the relationship between screening and adverse selection in the Medicare market, this test bears repeating using richer data and data from other health insurance markets.

5 Conclusion

To incentivize early detection and treatment, the Affordable Care Act removes cost sharing for diabetes, high cholesterol, and hypertension screening for people at high risk. This paper explores two possible unintended consequences related to patient composition and adverse selection. In addition to suggesting new directions for health economics research, our two main findings have implications for policy-makers, health providers, and health insurance markets.

First, we find that patients whose conditions were not detected prior to free screenings seek less medical treatment than patients who were aware of all their conditions. We replicate this result using biomarker data from three datasets, with a focus on groups likely to be affected by the ACA provisions related to cost-sharing for screening. This analysis bridges the health policy literature on expanding access to care with models from economics and public health in which individuals act based on their anticipated costs and benefits. In addition, our findings are important for health providers and policy-makers. Given that health insurance expansions and changes in the price of screening should increase access to screening for a variety of patients, lower use of treatment among patients with newly diagnosed conditions could be an important trend in the care of chronic conditions nationwide. Furthermore, if analysts ignore these composition effects, they may incorrectly conclude that treatment and control of chronic conditions declines rather than improves as more patients become diagnosed. This point would be important to consider when designing pay for performance schemes.
Second, we find no evidence that patients switch into more generous health plans after learning of a previously undiagnosed condition. This result builds logically from the previous finding in the sense that patients who use less medical care have less reason to switch health plans after learning of a previously undiagnosed condition. However, because of the limitations of our data and the particular regulatory environment of the Medicare Advantage market, we urge that this result should be interpreted with caution. We encourage replication of the analysis using data from other health insurance markets and from other, more common screening contexts such as a doctor’s office or pharmacy. Regardless, the analysis raises important questions for health insurance markets about the links between access to screening, patients’ plan selections, and possible reactions to these selections by health insurance providers. This topic is important because sorting of sicker people into generous health plans is a prominent concern when health status cannot be incorporated into health insurance premium prices, as is the case in the ACA marketplaces.

Our findings suggest three additional directions for future research. First, as noted above, our findings have implications for pay for performance schemes. Accountable Care Organizations, established under the ACA, are health provider organizations that are allocated financial rewards based in part on their performance on quality metrics including screening and control of chronic conditions. Our results suggest that increasing performance on the screening metrics could reduce performance on the control metrics. Future research can examine whether this is the case and whether the quality metrics should be redesigned so that practices are not penalized for expanding screening. In addition, future research can build on our findings to investigate which additional engagement strategies are most successful for increasing treatment and control among patients with previously undiagnosed conditions and consider methods to bring these strategies to scale in locations with an influx of new patients. Effective intervention design would require more detailed data on the key barriers faced by patients with newly diagnosed conditions; based on previous work on uninsured individuals, factors such as health literacy may play an important role. Finally, our findings suggest new directions for research on the economics of health care demand. In particular, classic health capital models should be revisited to see if conclusions drawn about the economics of health care demand change when agents can determine their own level of uncertainty about their health by choosing to be screened; this is a topic we plan to pursue in future work.
References


States. *JAMA ophthalmology* 132(2), 168–73.


A), 347–408.


Muntner, P., L. Colantonio, M. Cushman, D. Goff, G. Howard, V. Howard, B. Kissela,


Appendix

A Additional tables and figures

| Table 7: 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 and 2013 NHANES data, we calculated LDL cholesterol using the Friedewald equation (Friedewald et al., 1972). Because neither LDL cholesterol nor triglycerides were available in the OHIE data, we could not calculate LDL cholesterol and therefore defined high cholesterol using HDL and total cholesterol only.
Table 8: Participants with undiagnosed conditions show more severe biomarkers for their other, previously diagnosed conditions than do patients who are aware of all their conditions

<table>
<thead>
<tr>
<th></th>
<th>(1) HbA1c or FPG</th>
<th>(2) SBP</th>
<th>(3) DBP</th>
<th>(4) LDL</th>
<th>(5) TChol</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>NHANES data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>0.323***</td>
<td>2.073***</td>
<td>1.309***</td>
<td>18.00***</td>
<td>18.90***</td>
</tr>
<tr>
<td></td>
<td>(0.0824)</td>
<td>(0.616)</td>
<td>(0.431)</td>
<td>(2.011)</td>
<td>(1.795)</td>
</tr>
<tr>
<td><strong>OHIE data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>0.400***</td>
<td>2.181</td>
<td>3.077***</td>
<td></td>
<td>15.15***</td>
</tr>
<tr>
<td></td>
<td>(0.129)</td>
<td>(1.512)</td>
<td>(1.007)</td>
<td></td>
<td>(5.431)</td>
</tr>
<tr>
<td><strong>REGARDS data</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>17.85***</td>
<td>3.473***</td>
<td>2.961***</td>
<td>13.65***</td>
<td>16.26***</td>
</tr>
<tr>
<td></td>
<td>(4.534)</td>
<td>(0.827)</td>
<td>(0.464)</td>
<td>(2.766)</td>
<td>(3.030)</td>
</tr>
</tbody>
</table>

Standard errors in parentheses

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

This table shows that participants with some undiagnosed conditions show more severe biomarkers for their other, previously diagnosed conditions. The models are adjusted for demographic factors and prevalence of comorbid conditions. Glycated hemoglobin (HbA1c) or fasting plasma glucose (FPG) are only included for participants with diagnosed diabetes; systolic blood pressure (SBP) and diastolic blood pressure (DBP) are included only for participants with diagnosed hypertension; and low-density lipoprotein cholesterol (LDL) and total cholesterol (TChol) are included only for participants with diagnosed high cholesterol. LDL cholesterol is not measured in the OHIE data and cannot be calculated using the Friedewald equation because data on triglycerides are also not available. We use FPG rather than HbA1c in the REGARDS data because HbA1c is not measured in these data.
Table 9: Participants cascade for analyzing the impact of screening on health insurance selections

<table>
<thead>
<tr>
<th>Inclusion criterion</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>All REGARDS participants</td>
<td>30,239</td>
</tr>
<tr>
<td>Removed for data anomalies</td>
<td>-56</td>
</tr>
<tr>
<td>&lt;63 years old at REGARDS participation</td>
<td>-12,990</td>
</tr>
<tr>
<td>No Medicare-linked data</td>
<td>-3,368</td>
</tr>
<tr>
<td>Not observed for 24 months after participation</td>
<td>-2,391</td>
</tr>
<tr>
<td>Included</td>
<td>9,990</td>
</tr>
</tbody>
</table>
Table 10: Characteristics of participants in traditional Medicare vs. Medicare Advantage in the month prior to screening via REGARDS

<table>
<thead>
<tr>
<th></th>
<th>Traditional Medicare</th>
<th>Medicare Advantage</th>
<th>MA-TM</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>(%)</td>
<td>N</td>
</tr>
<tr>
<td>Total</td>
<td>7806</td>
<td>(85%)</td>
<td>1345</td>
</tr>
<tr>
<td>Male</td>
<td>3893</td>
<td>(50%)</td>
<td>615</td>
</tr>
<tr>
<td>African American</td>
<td>2490</td>
<td>(32%)</td>
<td>674</td>
</tr>
<tr>
<td>Low income</td>
<td>1354</td>
<td>(17%)</td>
<td>311</td>
</tr>
<tr>
<td>Undiagnosed condition(s)</td>
<td>1562</td>
<td>(20%)</td>
<td>232</td>
</tr>
</tbody>
</table>

"MA-TM" denotes the traditional Medicare estimate is subtracted from the Medicare Advantage estimate. * denotes p<0.1, ** denotes p<0.05, and *** denotes p<0.01.
Figure 2: Enrollment in Medicare Advantage before and after REGARDS participation

This figure shows the fraction of participants enrolled in Medicare Advantage rather than traditional Medicare before and after participation in REGARDS, for participants who did vs. did not learn of a previously undiagnosed condition by participating in REGARDS.
B REGARDS data collection procedures

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 through the use of commercially available lists of residential phone numbers and included 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 3 shows the geographic distribution of African American and white participants.

Figure 3: Location of REGARDS participants (Source: Howard et al., 2011)

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,\textsuperscript{13} and had their blood

\textsuperscript{13}About 80\% of participants met the fasting requirement at the time that their labs were taken. We control for fasting status in our panel data analysis and use fasting- or non-fasting specific cutoffs where applicable.
glucose, blood pressure and lipid panel plus other biomarkers assessed in their home on a morning of their choosing. Blood pressure was measured as an average of two measurements taken by a trained technician using a regularly tested aneroid sphygmomanometer, after the participant was seated with both feet on the floor for 5 minutes. Glucose and the lipid panel were measured using colorimetric reflectance spectrophotometry with the Ortho Vitros 950 IRC Clinical Analyzer (Johnson and Johnson Clinical Diagnostics) after being shipped on ice packs overnight to a central laboratory. Participants were compensated $30 for their time, and were notified of their results and advised to seek medical care for abnormal results using three levels of notification: (1) by telephone if any value is in the critical range, with instructions to immediately seek care; (2) by mail when a value is in the alert range with instructions to promptly seek care, and (3) general mail notification otherwise. The text of the mail notification and cards used for notification of high blood pressure are shown in Appendix Figure 4 below.

when judging participants’ disease status based on their biomarkers.
Your Blood Pressure: ____/____ mmHg

<table>
<thead>
<tr>
<th>Systolic</th>
<th>Diastolic</th>
<th>Recommended Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt;140</td>
<td>&lt;90</td>
<td>Normal blood pressure: no action required</td>
</tr>
<tr>
<td>140-159</td>
<td>90-99</td>
<td>Moderately high blood pressure: should be managed by a doctor within 2 months</td>
</tr>
<tr>
<td>160-179</td>
<td>100-109</td>
<td>High blood pressure: should be seen by a doctor within 1 month</td>
</tr>
<tr>
<td>&gt;180</td>
<td>&gt;110</td>
<td>Very high blood pressure: should be seen by a doctor within 1 week</td>
</tr>
</tbody>
</table>

Your Lipid panel (levels of blood fats):

<table>
<thead>
<tr>
<th>Your Values</th>
<th>Desirable Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total: _____ mg/dL</td>
<td>less than 200 mg/dL</td>
</tr>
<tr>
<td>LDL: _____ mg/dL</td>
<td>less than 130 mg/dL</td>
</tr>
<tr>
<td>HDL: _____ mg/dL</td>
<td>greater than 40 mg/dL</td>
</tr>
<tr>
<td>Triglycerides: _____ mg/dL</td>
<td>less than 200 mg/dL</td>
</tr>
</tbody>
</table>

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):

<table>
<thead>
<tr>
<th>Your Value</th>
<th>Desirable Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>_____ mg/dL</td>
<td>less than 126 mg/dL</td>
</tr>
</tbody>
</table>

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.
C Additional robustness checks

In the main text, we report finding no significant relationship between learning of an undiagnosed condition and health plan selection. One might be concerned that null findings when all participants are pooled together might mask significant findings among certain patient sub-groups who face higher health risk.

Therefore, we also report results from an additional analysis wherein we add interactions to allow effects to vary based on participants’ health care use in the past and self-rated health. As shown in Tables 11 and 12 our findings remain unchanged.
Table 11: No association between learning of a previously undiagnosed condition and current enrollment in Medicare Advantage

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prior to January 1, 2006</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After participation × Undiagnosed</td>
<td>-0.011</td>
<td>-0.019*</td>
<td>-0.003</td>
<td>-0.020</td>
</tr>
<tr>
<td></td>
<td>(0.012)</td>
<td>(0.011)</td>
<td>(0.016)</td>
<td>(0.016)</td>
</tr>
<tr>
<td>After × Undiag × High claims</td>
<td>-0.013</td>
<td>-0.0410</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.015)</td>
<td>(0.027)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After × Undiag × Low health</td>
<td>-0.002</td>
<td>-0.015</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.020)</td>
<td>(0.037)</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>After January 1, 2006</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After participation × Undiagnosed</td>
<td>-0.010</td>
<td>-0.004</td>
<td>-0.002</td>
<td>-0.011</td>
</tr>
<tr>
<td></td>
<td>(0.0256)</td>
<td>(0.0166)</td>
<td>(0.028)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>After × Undiag × High claims</td>
<td>-0.008</td>
<td>-0.041</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.037)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After × Undiag × Low health</td>
<td>-0.091</td>
<td>-0.091</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.081)</td>
<td>(0.090)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>165,642</td>
<td>165,642</td>
<td>75,276</td>
<td>75,276</td>
</tr>
<tr>
<td>Number of participants</td>
<td>3,405</td>
<td>3,405</td>
<td>1,682</td>
<td>1,682</td>
</tr>
<tr>
<td>Duration dependence</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

This table shows that we fail to reject the null hypothesis of no association between learning of a previously undiagnosed condition and current enrollment in Medicare Advantage for a variety of groups. “High claims” denotes above-median outpatient claims two calendar years prior to participation in REGARDS. The rows of the table list the coefficients and standard errors from panel data models. The columns indicate four regression specifications. “Low health” denotes fair or poor self-reported health in the REGARDS participant survey. “Undiagnosed” denotes the participant was undiagnosed for hypertension, high cholesterol, and/or diabetes prior to participation. Controlling for polynomials of log duration of enrollment in the current plan type (“Duration dependence”), our proxy for inertia or switching costs, does not affect the results.
Table 12: No association between learning of a previously undiagnosed condition and switching into Medicare Advantage

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Prior to January 1, 2006</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After participation × Undiagnosed</td>
<td>0.001</td>
<td>0.001</td>
<td>-4.05e-05</td>
<td>0.001</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.00358)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>After × Undiag × High claims</td>
<td>0.000</td>
<td>-0.000</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.004)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>After × Undiag × Low health</td>
<td>0.003</td>
<td></td>
<td>-0.008*</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td></td>
<td>(0.005)</td>
<td></td>
</tr>
<tr>
<td><strong>After January 1, 2006</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>After participation × Undiagnosed</td>
<td>-0.001</td>
<td>-1.87e-05</td>
<td>0.000</td>
<td>0.003</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.002)</td>
<td>(0.007)</td>
<td>(0.004)</td>
</tr>
<tr>
<td>After × Undiag × High claims</td>
<td>0.001</td>
<td></td>
<td>0.006</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.005)</td>
<td></td>
<td>(0.007)</td>
<td></td>
</tr>
<tr>
<td>After × Undiag × Low health</td>
<td>-0.004</td>
<td></td>
<td>-0.008</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.011)</td>
<td></td>
<td>(0.015)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>137,295</td>
<td>137,295</td>
<td>58,769</td>
<td>58,769</td>
</tr>
<tr>
<td>Number of participants</td>
<td>3,010</td>
<td>3,010</td>
<td>1,368</td>
<td>1,368</td>
</tr>
<tr>
<td>Duration dependence</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
</tr>
</tbody>
</table>

Robust standard errors in parentheses

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

This table shows that we fail to reject the null hypothesis of no association between learning of a previously undiagnosed condition and enrolling in Medicare Advantage for a variety of groups. “High claims” denotes above-median outpatient claims two calendar years prior to participation in REGARDS. The rows of the table list the coefficients and standard errors from panel data models. The columns indicate four regression specifications. “High claims” denotes above-median outpatient claims two calendar years prior to participation in REGARDS. “Low health” denotes fair or poor self-reported health in the REGARDS participant survey. “Undiagnosed” denotes the participant was undiagnosed for hypertension, high cholesterol, and/or diabetes prior to participation. Controlling for polynomials of log duration of enrollment in the current plan type (“Duration dependence”), our proxy for inertia or switching costs, does not affect the results.

### C.1 A note on imputing missing claims data

Imputation of missing values is necessary for an analysis of prior health care claims because our data do not capture health care claims for individuals who were in Medicare Advantage or who were younger than 65 prior to screening. In addition, this exercise has
theoretical appeal. Each individual’s history of medical care claims is determined not only by their health needs, but also by the out-of-pocket prices they face in their health plan. Out-of-pocket prices can vary for each type of care across plans and can vary widely in the Medicare Advantage market across time and location. We eliminate the influence of out-of-pocket prices on demand for care by imputing claims using data from a single health plan, traditional Medicare.

To impute the data, we predict the number of out-patient Medicare claims during the two years prior to participation (which were collected for some participants) using demographic and health variables (which were collected for all REGARDS participants). The predictors include blood pressure, lipid panel measurements, and blood glucose; measured waist circumference and BMI; county poverty and Health Professional Shortage Area status; age; year and region; race, sex, education, marital status, income, smoking, drinking, self-reported health, and cognitive functioning. All continuous variables are modeled as four categorical dummies to allow for non-linearity.

Health care utilization data typically have a skewed distribution and many zeroes, which makes selecting an appropriate model challenging. Although a number of modeling approaches have been explored in the literature, the current consensus is that no one approach dominates the others in all circumstances (Buntin and Zaslavsky, 2004; Duan et al., 1983; Manning, 1998; Manning et al., 2005; Manning, 2014). We follow recommendations set forth by Manning and Mullahy (2001) to decide between models. The algorithm involves examining the kurtosis of residuals from a log OLS model, and then conducting a Park (1966) style test to examine the relationship between the mean and variance of residuals from a gamma GLM model if the residuals do not show substantial kurtosis. Our results indicate that the log-residuals have heavy tails, with coefficients of kurtosis in the range of 3.9-4.0, so that the GLM models are likely to be problematic (Manning and Mullahy, 2001). This is why the results shown in Tables 11 and 12 include imputed claims produced by log-transformed OLS models.
D Proofs: Demand for screening is (weakly) decreasing in θ and P

D.1 Demand for screening is weakly decreasing in θ

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

\[ \pi \left( \max_{M} u (A - PM - \kappa^*, H (M, 1)) - \theta M \right) + (1 - \pi) u (A - \kappa^*, H (0, 0)) \]

\[ - (\pi u (A, H (0, 1)) + (1 - \pi) u (A, H (0, 0))) = 0 \]  

(7)

Differentiating (7) with respect to θ yields the following expression (by the envelope theorem, we can ignore the fact that the optimal M varies with θ):

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

(8)

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

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

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

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

We conclude \( \frac{\partial \kappa^*}{\partial \theta} \leq 0 \) because \( \frac{\partial u}{\partial C} > 0, \pi > 0 \) and \( M^* \geq 0 \).
D.2 Demand for screening is decreasing in $P$

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

$$ \pi \left( \max_M u (A - PM - \kappa^*, H(M, 1)) - \theta M \right) + (1 - \pi) u (A - \kappa^*, H(0, 0)) - (\pi u (A, H(0, 1)) + (1 - \pi) u (A, H(0, 0))) = 0 \quad (9) $$

Differentiating (9) with respect to $P$ yields the following expression (by the envelope theorem, we can ignore the fact that the optimal $M$ varies with $P$):

$$ \pi \left( -\frac{\partial u (A - PM^* - \kappa^*, H(M^*, 1))}{\partial C} \frac{\partial \kappa^*}{\partial P} + \frac{\partial u (A - PM^* - \kappa^*, H(M^*, 1))}{\partial C} \right) - (1 - \pi) \frac{\partial u (A - \kappa^*, H(0, 0))}{\partial C} \frac{\partial \kappa^*}{\partial P} = 0 \quad (10) $$

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

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

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

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

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

53
E Proofs: Demand for medical treatment is decreasing in $\theta$ and $P$

E.1 Demand for medical treatment is decreasing in $\theta$

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

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

$M_{\overline{\theta}}$ and $C_{\overline{\theta}}$ must fulfill the first-order conditions summarized in equation (5), as follows:

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

After non-pecuniary cost decreases from $\overline{\theta}$ to $\underline{\theta}$, previously optimal decisions $M_{\overline{\theta}}$ and $C_{\overline{\theta}}$ would violate equation (5) as follows:

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

To make inequality (12) an equality, $M$ and $C$ must change so that the left-hand side decreases and/or the right-hand side increases. By concavity of the utility function in $H$ and $C$, the weakly positive cross-partial $\frac{\partial^2 u(C, H)}{\partial C \partial H}$, and weakly decreasing marginal returns to medical care, increasing $M$ and decreasing $C$ achieves both. Therefore $M_{\overline{\theta}} < M_{\underline{\theta}}$ and $C_{\overline{\theta}} > C_{\underline{\theta}}$ resolves the contradiction in the first-order conditions. We conclude that $\frac{\partial M^*}{\partial \theta} < 0$.

E.2 Demand for medical treatment is decreasing in $P$

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

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

$M_{\bar{P}}$ and $C_{\bar{P}}$ must fulfill the first-order conditions summarized in equation (5), as follows:

$$\frac{\partial u(C_{\bar{P}}, H(M_{\bar{P}}, 1))}{\partial H} \frac{\partial H(M_{\bar{P}}, 1)}{\partial M} - \theta = \frac{\partial u(C_{\bar{P}}, H(M_{\bar{P}}, 1))}{\partial C}$$

(13)

After cost of care $P$ decreases from $\bar{P}$ to $\underline{P}$, previously optimal decisions $M_{\bar{P}}$ and $C_{\bar{P}}$ would violate equation (5) as follows:

$$\frac{\partial u(C_{\bar{P}}, H(M_{\bar{P}}, 1))}{\partial H} \frac{\partial H(M_{\bar{P}}, 1)}{\partial M} - \theta > \frac{\partial u(C_{\bar{P}}, H(M_{\bar{P}}, 1))}{\partial C}$$

(14)

To make inequality (14) an equality, $M$ and $C$ must change so that the left-hand side decreases and/or the right-hand side increases. As before, increasing $M$ and decreasing $C$ achieves both. Therefore $M_{\bar{P}} < M_{\underline{P}}$ and $C_{\bar{P}} > C_{\underline{P}}$ resolves the contradiction in the first-order conditions. We conclude that $\frac{\partial M_c}{\partial P} < 0$. 

55