What Difference Does a Diagnosis Make? Evidence from Marginal Patients*

Mattan Alalouf† Sarah Miller‡ Laura R. Wherry§

November 22, 2022

Abstract

This paper explores the impact of receiving a diagnosis of type 2 diabetes among patients who are close to the diagnostic threshold using a regression discontinuity design. Using data from a large national insurer, we find that a marginally diagnosed patient with diabetes spends $1,097 more on drugs and diabetes-related care annually after diagnosis. This increase in spending persists over the 6-year period we observe the patients, despite many who are not initially diagnosed receiving a later diagnosis during this time frame. These marginally diagnosed patients experience improved blood sugar after the first year of diagnosis. However, this improvement is not statistically significant in subsequent years, and in some post-test years our confidence intervals rule out *any* improvement in this measure. Other clinical measures of health, cholesterol and mortality, do not change significantly at the cutoff, although our confidence intervals include meaningfully-sized effects. The diagnosis rates for preventable disease-related conditions such as diabetic retinopathy, neuropathy, and kidney disease increase following a diagnosis, likely due to more intensive screening.

Keywords: health, diabetes

JEL codes: I10, I12

*This work benefited from helpful comments from seminar participants at the University of Michigan, UCLA, Harvard Kennedy School, UC Santa Barbara, CU Denver, and the NBER Health Economics and Health Care groups. We would also like to thank anonymous referees and Doug Bell, Marianne Bitler, John Mafi, Tannaz Moin, and Yusuke Tsugawa for helpful feedback, as well as Amanda Do, Rob Follett, and Javier Sanz for their help in accessing and interpreting the UCLA medical records data. An additional thank you to Sonia Iyengar for research assistance. Laura Wherry benefited from non-financial support from the California Center for Population Research at UCLA (CCPR), which receives core support (R24-HD041022) from the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD). Research with UCLA medical records data was supported by NIH National Center for Advancing Translational Science (NCATS) UCLA CTSI Grant Number UL1TR001881. Mattan Alalouf was supported by the National Institute on Aging of the National Institutes of Health under Award Number T32AG000221. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

†Mathematica. Email: mattan@umich.edu

‡University of Michigan Ross School of Business and NBER. Email: mille@umich.edu

§NYU Wagner Graduate School of Public Service and NBER. Email: laura.wherry@nyu.edu

Since the late 1990s, the criteria for diagnosing several chronic diseases have been relaxed. For example, diagnosis guidelines have changed for osteoporosis (in 2003), hyperlipidemia (in 1998), and hypertension (in 1997 and again in 2017), generating millions of new patients who were considered healthy under previous versions of these guidelines (Welch et al., 2011). Sociological research demonstrates that, throughout the 20th and 21st centuries, new diagnostic and therapeutic tools have driven trends of medicalization and of expanding diagnostic criteria. Greene (2007) and Jutel (2011) emphasize that there are many stakeholders in the diagnosis decision beyond the patient; providers, researchers, manufacturers of diagnostic equipment and pharmaceuticals all have a financial stake in the decision to diagnose a patient with an illness. The trend towards diagnosing and treating progressively healthier patients has led some in the medical field to express concern that such interventions may be, at best, wasteful, if not actually harmful and counterproductive (Welch et al., 2011). Yet, little is known regarding the implications of diagnosing and treating patients at the margin of being diagnosed on either health care spending or population health.

This paper explores the impact of a diagnosis on such "marginally" sick patients by exploiting a cutoff in the diagnosis criteria for a common chronic illness, type 2 diabetes mellitus.[1] Type 2 diabetes is a condition characterized by the body's inability to regulate blood sugar and its diagnosis presents an ideal setting in which to explore the treatment of marginally diagnosed patients for several reasons. First, it is a common disease, making it an empirically relevant example; over 24 million individuals in the United States are currently diagnosed with diabetes, and annual spending on medical care related to diabetes and complications of diabetes is estimated at $237 billion, with over $50 billion spent on insulin and diabetes-controlling drugs alone (American Diabetes Association, 2018b).

Second, diabetes is diagnosed using a sharp cutoff in blood sugar levels, which are smoothly distributed in the population. The American Diabetes Association (ADA) recommends diagnosing diabetes in patients whose glycated hemoglobin (A1c) levels are at or above 6.5% (American Diabetes Association, 2018b). Clinical trials that determine treatment recommendations, however, typically focus on the average effects of treatment among all individuals with diabetes or those with high blood sugar (e.g. Garber et al., 1997; Ponssen et al., 2000), not individuals close to the diagnostic threshold. Further, there is mixed evidence regarding the validity of the 6.5% diagnostic threshold.[2] Therefore,

[1] We focus on type 2 diabetes given its prevalence among those affected with diabetes (affecting about 90% of individuals with diabetes). Unlike type 1 diabetes, which is a genetic condition that tends to reveal itself early in life, type 2 diabetes develops over time and is more commonly diagnosed among adults. Here and throughout the paper, when we mention diabetes, we are referring to type 2 diabetes, unless it is explicitly indicated otherwise.

[2] The threshold was derived from analyses of cross-sectional epidemiological studies that examined the prevalence of dia-

the question of the threshold's health and utilization impacts on marginally diagnosed patients is of particular interest.

Third, the blood test used to diagnosis diabetes is difficult to manipulate, making both self-selection by patients and strategic coding by providers extremely unlikely. This supports our use of a regression discontinuity design to estimate the effect of a diabetes diagnosis for patients with similar underlying health but with blood sugar levels falling on either side of the 6.5% cutoff.

Marginally diagnosed patients are a particularly relevant group to study in the case of chronic illness, such as diabetes, because they represent a large fraction of the patients diagnosed with a given disease, due to the underlying distribution of health in the population (Welch et al., 2011). Most people have blood glucose levels below the 6.5% cutoff in the normal range (between 4% and 5.6%) and the share of the population decreases as the values become more extreme. As a result, even though spending is comparatively low among the "marginal" patients with diabetes whose A1c falls just above the diagnostic threshold, they have an outsized impact on total costs due to their large representation among patients with diabetes. The distribution of A1c values among patients receiving an initial A1c test in our data makes this clear: over one quarter of all patients whose first A1c test puts them in the diabetes range have an A1c value within 0.1 percentage points of the cutoff, and over 15 percent have an A1c value of exactly 6.5%, the lowest possible value one can receive and still be considered to have diabetes.

Our analysis is motivated by existing work in economics that evaluates the impact of diagnoses and subsequent care on health.[3] Most relevant to this paper, two recent papers use identification

betic retinopathy, a long-term complication of the disease, across a range of A1c levels (The International Expert Committee, 2009). However, a number of studies have documented the better performance of alternative cutoff values in predicting the presence of diabetic retinopathy (Kim and Kim, 2015). In addition, some of the epidemiological studies used to select the diagnostic cutoff included individuals who were already diagnosed with diabetes, which may not be representative of disease development or complications among undiagnosed individuals (Nakagami et al., 2017). Longitudinal studies aiming to validate the use of the diagnostic threshold among undiagnosed patients have found differing results (see discussion in Tsugawa et al., 2012).

[3]One of the first studies in this literature documents that infants born near the diagnosis threshold for "very low birthweight" (birthweights of less than 1500g) receive more intensive medical care and experience a reduction in infant mortality (Almond et al., 2010). Follow up work has found evidence of longer-run effects of this diagnosis on health and human capital outcomes in the U.S., Chile, and Norway (Bharadwaj et al., 2013; Chyn et al., 2019), indicating high cumulative returns to spending on this margin. There has also been some work examining the sensitivity of some of these results when addressing non-random heaping in birthweight (Barreca et al., 2011, 2016). More recent work examining other types of medical conditions also highlights that there may be important effects on patient behavior, or psychological effects of diagnosis. For instance, among female students in New York City who receive BMI report cards with weight classifications, those with values of BMI just above an "overweight" threshold experience small *increases* in BMI and body weight during the subsequent year compared to those just below the threshold (Almond et al., 2016). However, individuals receiving a diagnosis of abdominal aortic aneurysm and told they will receive increased monitoring in a Swedish screening program have improved psychological well-being, compared to those just missing the diagnosis criteria but informed that they are at risk (Dahlberg et al., 2016). Other examples include: Zhao et al. (2013) find that individuals in China with systolic blood pressure above the hypertension diagnosis cutoff reduce their fat intake upon receiving a diagnosis; and, Fukuma et al. (2020) show that

strategies similar to ours to examine the effects of diagnoses of specific chronic health conditions, including diabetes, in Japan and Korea. Iizuka et al. (2021) find no evidence that Japanese patients with test values above the diabetes threshold receive more care or see health improvements; however, among a subset of patients with other health conditions they do detect some improvements in blood sugar and BMI.[4] Next, Kim et al. (2019) test for discontinuities at "medium risk" and "high risk" thresholds for diabetes under a Korean national health screening program. They find no effects for individuals receiving a medium risk diagnosis, but high risk individuals are more likely to take diabetes medication and experience weight loss in the short-run. The authors find no effects on exercise behavior, however, and the effects on weight loss appear to dissipate in the long-run (3-4 years after testing).

Several related papers have also examined the behavioral effects of diagnosis among the broader set of patients receiving a diabetes diagnosis in the U.S. In a novel use of household scanner data, Oster (2018) finds relatively limited behavior change among households in which someone was recently diagnosed with diabetes, documenting small calorie reductions and improvements in diet. Myerson et al. (2018) examine health care utilization among Medicare patients who were advised to seek medical care for abnormal biomarker screening results, including for blood glucose level, in a national epidemiological study. Over the next two years, individuals with previously undiagnosed diabetes were 45 percentage points more likely to receive semi-annual doctor visits for evaluation and management of their condition. In a similar study, Edwards (2018) examines behavioral response among previously undiagnosed older adults in the Health and Retirement Study notified of elevated A1c biomarker readings (greater than 7.0%). He finds a 40 percentage point increase in a diabetes diagnosis and a similarly sized increase in diabetes medication usage. In addition, he observes an 2.2 percent decrease in self-reported weight following notification receipt, as well as evidence of less drinking and more frequent exercise, with the latter finding also observed among the respondent's spouse.

Our work builds on these studies in several ways. First, we focus on the effects of a diabetes diagnosis for marginally diagnosed patients, which as we describe earlier represent a sizable share of patients with diabetes. Second, we contribute by examining the effects of a marginal diabetes diagnosis in the U.S. context, which is unique in its high medical spending and poor health outcomes

individuals in Japan with a cardiovascular risk factor whose waist circumference falls above a cutoff experience weight loss following lifestyle counseling.

[4]They also examine the pre-diabetes threshold and find patients who exceed the threshold have small but statistically significant increases in diabetes-related doctor visits and outpatient medical spending in the following year. They find no evidence, however, that this increased utilization leads to improvements in health over the short- or medium-term.

when compared to other high income countries (Papanicolas et al., 2018).

Similar to the studies in Japan and Korea mentioned above, we aim to examine the effects of a marginal diagnosis on both health care spending and patient health. Examining these outcomes is tricky in the U.S. context, however, since it requires data on lab values, in addition to medical claims, for a large sample of individuals with A1c values close to the diagnosis cutoff. Since the U.S. does not have a national health system, the type of population-level administrative data often used in other settings (e.g. Korea in Kim et al., 2019) does not exist. In addition, national health surveys are not well-suited to study this question due to their lack of clinical data, small sample sizes, and limited information on utilization and spending. Instead, we take advantage of medical claims records from a large, national US health insurer that includes laboratory test results from affiliated labs. With 61 million enrollees over our study period, this data source captures a sizeable share of the U.S. population with private health insurance coverage (nearly 30 percent of the privately insured population in the U.S. in 2018, calculation based on Berchick et al., 2019). Using this data source, we are able to examine the impact of a diabetes diagnosis for the marginally diagnosed patient on different types of spending, use of care, and on clinical measures of health (blood sugar, cholesterol, and mortality). We supplement these data with matched survey records that collect self-reported information on health and healthy behaviors.[5]

Finally, our work contributes to a broader literature in economics devoted to measuring the marginal impact of additional medical spending in the U.S. This work seeks to disentangle patient selection into care from the benefit of the care itself, usually with the goal of determining whether certain types of care reflect wasteful "flat of the curve" spending. In general, the results that have been previously documented on this topic are mixed, with some studies showing large benefits of additional medical spending (e.g. Doyle, 2011; Doyle et al., 2015) and other papers showing limited or no benefit associated with additional spending (e.g. Fisher et al., 2003b,a; Frakes, 2013; Currie and Slusky, 2020). A limitation in the existing literature is that much of it focuses on care for the onset of acute conditions (such as heart attack or hip fracture) and mortality, with very little attention to treatment for chronic illnesses. Understanding the marginal impact of treatments for chronic illness is particularly relevant given that they represent a substantial fraction of total medical spending (e.g. 86 percent of health care spending is for patients with one or more chronic illnesses, see Gerteis et al.,

[5] An earlier version of this paper included analyses of additional clinical health measures (blood pressure and BMI) using electronic health records from a large health system. However, differences in these outcomes measured prior to the lab test were associated with the diabetes diagnosis cutoff, a pattern not observed in our main data sources (see discussion in Section B), and this led us to remove these analyses.

2014), with over over 100 million individuals in the United States diagnosed with diabetes or pre-diabetes (Centers for Disease Control and Prevention, 2017).

We find that patients with A1c values that fall right above the 6.5% threshold are 11 percentage points more likely to have a diabetes diagnosis appear on their health insurance claim in the first year following an initial A1c test.[6] In addition, receiving a diagnosis at the cutoff is associated with approximately $1,097 in additional spending on recommended diabetes-related care and $139 in endocrinology specialty care during that first year. Spending increases related to office-based care and laboratory tests are particularly large (representing most of the observed increase in spending), although we also detect increases in the probability of purchasing diabetes-related drugs and spending on vision screening.

Similar to other authors, we attempt to trace out both the short- and long-term effects of disease diagnosis and follow patients through year 6 after diagnosis. Notably, this elevated rate of spending for patients with A1c values just above the threshold persists for all six years that we observe patients; there is no evidence that patients with slightly lower A1c values catch up, despite the natural progression of disease that leads many who were initially below the cutoff to be diagnosed at later points during our study period. It appears that receiving a "marginal" diagnosis sets patients on a path of persistently higher spending.

The evidence for health-related benefits is less clear. We find that rates of diagnoses for complications related to diabetes (diabetic retinopathy, neuropathy, and diabetic kidney disease) increase for individuals right above the cutoff, likely due to more intensive screening. We see no evidence that diagnosis rates for these complications fall over time among the marginally diagnosed group. We find no evidence of changes in self-reported health or in healthy behaviors such as diet, exercise, drinking alcohol, or smoking following the initial lab test, even though clinical practice recommendations include lifestyle education and management for patients regarding diet, exercise, alcohol intake and smoking. This latter finding could be explained by the receipt of similar lifestyle recommendations by patients with A1c values just below the cutoff, who may be identified as at increased risk of later developing diabetes.

When we examine clinical measures of health, we find short-term improvements in blood sugar among marginally diagnosed patients, but these improvements do not persist past the first year, de-

[6] We require individuals to be in the sample for at least one year prior to the initial A1c test in order to exclude those with previous diagnoses of diabetes. This also allows us to test for and rule out pre-test differences in utilization for patients on either side of the cutoff. As discussed later in Section B, results are similar if we use a two year observation window to exclude those with previous diagnoses of diabetes.

spite the persistence of higher levels of diabetes-related medical spending over the entire period. We also find no evidence of improvements in total cholesterol or mortality. However, our sample size decreases as more time passes from the original A1c test, and as individuals gradually switch insurance carriers. This attrition, which does not appear to be related to diagnosis, decreases the precision of our estimates and limits our ability to make strong conclusions regarding a lack of health benefits based on our estimates over this longer-term period.

Notably, our results indicate that many patients who receive an A1c value just above the diagnostic cutoff do not receive a diagnosis for diabetes. In additional analyses, we document that patients who are marginally diagnosed (compliers) differ on observable characteristics from those whose A1c test result puts them above the diagnosis threshold but who nevertheless do not receive the prescribed interventions (never takers). Notably, the compliers are less likely to face high cost sharing compared to the never takers. We also find that patients of specialists, as opposed to primary care doctors, are more likely to receive a diabetes diagnosis. These differences could be due to the different practice styles of the physicians, e.g. if specialists tend to favor more intensive intervention for diseases than primary care physicians. Alternatively, the differences may be due to unobserved heterogeneity in patient health across these two types of providers. Additional research is needed to fully characterize how plan characteristics, provider behavior, and other factors, affect who receives timely care and under what circumstances.

I Background

Many chronic illnesses are diagnosed by comparing the outcome of a laboratory test to a "cutoff" value specified in guidelines created by health or medical associations.[7] To diagnose type 2 diabetes, physicians primarily rely on a measure of blood sugar called glycated hemoglobin A1c ("A1c"), which is collected via a blood draw. A1c measures the exposure of hemoglobin cells to plasma glucose (blood sugar) over the past three months. Normal A1c levels range from 4.0% to 5.6%. High A1c levels are considered a marker of diabetes, as they indicate poor control of glucose.

Although diabetes may be diagnosed using alternative tests of single or episodic glucose levels, A1c tests are considered the preferred diagnostic tool since they measure chronic glycemic exposure.[8]

[7] Interestingly, recent work suggests that additional discontinuities may also exist in terms of *who is tested* due to rules of thumb used by physicians for administering diagnostic tests. Coussens (2018) finds that, in the emergency department setting, physicians use heuristics based on patient age when determining whether to test for ischemic heart disease, resulting in a discontinuity in diagnosis for this condition.

[8] Other tests that can be used to screen for diabetes are the measurement of fasting plasma glucose (FPG) and plasma glucose during a two-hour oral glucose tolerance test (OGTT). The OGTT is not commonly used other than among pregnant

The International Expert Committee recommended A1c as a superior method to diagnose diabetes in a 2009 consensus report, noting that it is convenient (does not require the patient to fast) and more stable (less fluctuation within and between days).[9] Using data from cross-sectional epidemiological studies, this committee observed that the rate of any retinopathy, a long-term complication of diabetes, "increases substantially at A1c values starting between 6.0 and 7.0%." The committee also noted that when A1c levels were examined in 0.5% increments, the incidence of diabetes-specific "moderate" retinopathy noticeably increases at A1c levels $\geq$6.5%. The analysis included data on approximately 28,000 individuals from nine countries. The IEC also considered specificity and sensitivity and relied on a receiver operating characteristic curve analysis when determining the optimal cutpoint. Based on these analyses, they recommended using an A1c cutoff value of $\geq$6.5% for the diagnosis of diabetes, although they note that "this cut point should not be construed as an absolute dividing line between normal glycemia and diabetes." The committee also emphasized that while there is evidence for a continuum of risk for the development of diabetes based on A1c levels, "there does not appear to be a specific level at which risk for diabetes clearly begins" (The International Expert Committee, 2009).

Concordant with the IEC guidelines, the American Diabetes Association subsequently issued the recommendation of a diagnosis of diabetes if a patient's A1c is greater than or equal to 6.5% (American Diabetes Association, 2010).[10] A1c tests are recommended for patients 45 years old and older, as well as patients under 45 who are overweight and have one other risk factor. Risk factors include family history, sedentary lifestyle, and a history of gestational diabetes (Centers for Disease Control and Prevention, 2018). Once diagnosed, recommended care includes a comprehensive medical evaluation to form a care management plan, annual diagnostic tests (urinary albumin, lipid panel, glomerular filtration rate) to detect comorbid conditions or complications, A1c monitoring at least twice per year, antidiabetic medication, statin therapy (for those age 40 and above),[11] annual assessment for diabetic retinopathy by an ophthalmologist, annual comprehensive foot evaluation, and receipt of flu, pneumonia, and hepatitis B vaccines (American Diabetes Association, 2017). Patients are also advised to engage in regular exercise (2-3 times per week), decrease sedantary behavior, eat a healthful diet, and

women to screen for gestational diabetes.

[9]The IEC was convened in 2008 to consider the diagnosis of diabetes among nonpregnant individuals. Members were appointed by the American Diabetes Association, the European Association for the Study of Diabetes, and the International Diabetes Federation.

[10]Beginning in 2011, the ADA guidelines specify a value of 6.5% or higher on two separate tests be used to diagnose diabetes, unless the patient exhibits "unequivocal" hyperglycemia. However, in practice, we do not find that patients are receiving multiple tests. For example, in our data, less than 0.5% of patients who received an A1c value of 6.5 received a follow-up test within two weeks.

[11]This guideline for statin therapy was recommended in 2013 by the American College of Cardiology and the American Heart Association and confirmed by the American Diabetes Association at the end of 2014.

quit smoking if relevant. In addition, the ADA recommends that patients with diabetes receive more aggressive treatment for comorbid conditions. For example, patients with diabetes with blood pressure above 140/90 mmHg should have prompt initiation of drug therapy at the "maximum tolerated dose" to achieve blood pressure goals, while patients at high risk for kidney disease require additional monitoring and treatment.

In this paper, we use the cutoff in diabetes classification to evaluate the impact of receiving a diabetes diagnosis on the health care utilization and associated costs, health, and behaviors for patients at the margin of being diagnosed. Because we expect that individuals with A1c values falling just below or above the 6.5 threshold have similar underlying health, we rely on the experiences of patients just missing the diagnosis criteria as a credible counterfactual and use a regression discontinuity design that approximates random assignment in meeting the diagnosis criteria. Further contributing to this plausibly exogeneous source of variation in diagnosis is the measurement error present in the A1c test, due to both analytical variation across labs and within-person biological variation (Phillipov and Phillips, 2001).

Patients whose A1c falls below 6.5 are still expected to receive clinical guidance from their physician, but less diabetes-related care then their counterparts with A1c values at or above 6.5. Patients who miss the 6.5 cutoff but have A1c values at or above 5.7 are considered to have "pre-diabetes," with impaired glucose tolerance and at a higher risk of developing diabetes in the future. As such, they also receive some care aimed at lowering their risk for developing diabetes; specifically, the ADA recommends that the patients have their A1c tested annually and that the provider counsel them to make lifestyle changes such as losing weight and increasing exercise. The ADA also recommends that drug therapy should be discussed as an option for certain risk groups (e.g., those with a BMI of at least 35). Our analyses will, therefore, document the average difference in utilization and spending for individuals on the margin of receiving diagnoses of pre-diabetes versus diabetes.[12]

II Data

A MEDICAL CLAIMS AND LAB DATA

To assess the impact of a diagnosis on health care utilization and spending, self-reported health and healthy behaviors, and subsequent lab values (blood sugar and total cholesterol), we use Optum's

[12] While we do not explicitly evaluate the effects of a pre-diabetes diagnosis, our RD graphs include the 5.7 cutoff and do not indicate an observable discontinuity on spending or receipt of diabetes-recommended care at this A1c value, in contrast to the effects that we observe at the 6.5 cutoff.

de-identified Clinformatics ©Data Mart Database (herein after referred to as "CDM") from a large, national US health insurer with data that span 2009 to 2017.[13] CDM provides claims data on over 61 million patients over our sample period. The data include all claims associated with these patients that are paid for by the health insurance company. The claims data include procedure and diagnosis codes, as well as the date the procedure occurred and the cost of the claim (including the amount paid by the patient and by the insurance company); for pharmacy claims, the data include the national drug code, quantity dispensed, and cost. Additionally, the data contain some limited information on demographic characteristics of the patient, such as age and gender. CDM data are primarily composed of private or employer-sponsored health insurance plans, but includes some Medicare Advantage plans as well (see Table 1).

When an A1c test is used for diagnosis, the provider sends a blood sample to a lab for evaluation. In the CDM data, we observe the lab test results for laboratories that are affiliated with CDM. We do not observe test results for unaffiliated laboratories, although we do see spending associated with these tests. In nearly all cases, lab test results are rounded to the nearest 0.1 percent; for the small number of lab results that are reported with greater precision, we round to the nearest 0.1 percent so the data are comparable across all labs reported.

The data include unique patient identifiers that allow us to track patient spending and utilization over time. We are able to track patient spending if patients change insurance plans as long as the insurance plan is offered by an insurance company associated with CDM. If patients drop their insurance coverage, or switch to an insurance plan offered by a different company, we are no longer able to observe them in the data. However, CDM provides records of patients' enrollment, so we can distinguish patients who dropped coverage from patients who had no claims in a given year. We do not find any evidence of differential attrition from the sample based on A1c value relative to the diagnosis threshold (see Section C). Therefore, as described in more detail below, we include patients in analyses for all years when they are enrolled in an CDM plan, regardless of whether they exit at a later point during the sample period. In Section V, we examine the sensitivity of our estimates to restricting the analysis to members enrolled during the entire sample period.

To construct our sample, we pull all adult patients in CDM who have an A1c test at an CDM-affiliated lab during our study period. From this group, we drop individuals who had a diagnosis of diabetes on their claim at any time before their first observed A1c test in order to eliminate patients

[13] Although CDM data are available for earlier years, we use only individuals with the initial A1c test taken in 2009 or later because this was the first year that the A1c test was recommended for the diagnosis of diabetes.

who are receiving the test for monitoring, rather than the diagnosis of, diabetes.[14] We also exclude individuals with a pregnancy-related diagnosis during the year prior to their first A1c test in order to drop cases associated with gestational diabetes, which has different treatment and diagnostic guidelines than type 2 diabetes (Vandorsten et al., 2013).[15] We further restrict to patients who are in the sample at least 12 months prior to and 12 months following the initial lab test in order to observe their annual spending before and after the test. This leaves us with 142,541 patients in year 1 (the calendar year following the initial lab test). Figure A1 in the Appendix provides a diagram of our sample definition with patient counts for each exclusion criterion.

We observe patients for up to 6 years following their initial test. For subsequent years following the initial test, we require that patients are enrolled in a CDM plan for at least half the year to be included in the analysis. We define year relative to the date of the initial test. For example, if an individual was tested on March 1, 2010 we would consider the first year after the test to be March 2, 2010 through March 1, 2011. For patients enrolled in a CDM plan for less than a year, we scale their spending proportionally to their enrollment time in order to annualize the outcome. For example, if the patient is enrolled for 6 months and spends $300, we double spending to $600 to account for the fact that we observe only one half of the year's spending. We do this to make patients observed for a partial year comparable to those observed for the full year; however, our results are essentially unchanged if we instead assume that partial year enrollees had zero spending during the time of the year we do not observe them.

We first estimate the change in the diagnosis of diabetes at the A1c cutoff value of 6.5%. We define the diagnosis of diabetes as the presence of any claim with a diagnosis code for diabetes during the first year following the test. In later analyses, we also examine how the likelihood of diagnosis evolves over time by looking at the cumulative likelihood of diagnosis through year 2, year 3, etc. Our primary outcome measures look at spending for all types of care recommended for patients with diabetes by the American Diabetes Association (ADA). This includes doctor office visits and consultations related to diabetes; spending on diabetes-controlling drugs such as metformin; spending on two other types of drugs that are commonly prescribed for diabetes: statins and Angiotensin-converting enzyme (ACE) inhibitors; lab tests for relevant conditions; and spending related to vision testing for diabetic

[14] Although we are examining A1c tests starting in 2009 only, we have CDM claims dating back to 2001 and the entire period of claims data is used to implement this sample restriction.

[15] We do not explicitly exclude individuals who have a diagnosis of type 1 diabetes, however individuals with type 1 diabetes are considerably less likely to be diagnosed as adults. Roughly three fourths of type 1 diabetes cases present in childhood (Levitsky et al., 2021). Overall, only 5 to 10 percent of adults with diagnosed diabetes have type 1 diabetes (Inzucchi et al., 2021).

retinopathy. We examine total spending on recommended care as well as each category individually. We also examine spending on endocrinology specialty care, which is not specifically recommended, but that may be affected by a diabetes diagnosis. In addition, we evaluate how receiving a diagnosis affects the quantity of care received by looking at the number of diabetes-related visits, the number of A1c or glucose lab tests, indicators that either of these two outcomes are positive, and an indicator for the presence of any spending on diabetes-controlling drugs.

Next, we evaluate changes in diabetes-related conditions and overall health at the cutoff. We group health outcomes into three categories: diabetes-related complications, lab results, and mortality. Diabetes-related complications are arguably the most consequential health outcome for the patients in the study. These complications are relatively common among diabetes patients (Wexler, 2021) and materially affect quality of life. We examine diabetic retinopathy, damage to the retina that can result in blindness; diabetic neuropathy, nerve damage that can result in, among other symptoms, pain, dizziness, and loss of feeling; and, diabetic kidney disease, chronic loss of kidney function that can result in end-stage renal disease.[16] These complications were used by the IEC to determine diagnostic criteria for diabetes (American Diabetes Association, 2010), and we believe that any favorable impacts on the rate of these complications would provide the most direct evidence that treatment for marginal patients is beneficial.

We examine changes in individuals' A1c scores and total cholesterol from their baseline levels taken on the date of the first test (for A1c) or the year prior to the first test (for total cholesterol). Lab results, though they are less direct measures of patients' health and quality of life, are measured frequently, and may be more likely than complications to provide early evidence of the benefits of treatment. We estimate the effects of a marginal diagnosis on patients' A1c scores over time, as well as their cholesterol levels. High A1c scores are a risk factor for diabetes-related complications (Inzucchi and Lupsa, 2020), and high cholesterol is a risk factor for cardiovascular events (e.g. heart attack, stroke) (Cholesterol Treatment Trialists' (CTT) Collaborators, 2012). Clinical guidelines recommend managing levels of A1c to prevent or treat diabetes-related complications (Inzucchi and Lupsa, 2020). These lab results complement our estimates of the effects of a diagnosis on complications since improvements in these scores are likely to manifest earlier than diabetes-related complications.

As we demonstrate in Appendix Table A1, individuals with initial A1c test values just above the

[16] A prior version of this paper also examined the incidence of diabetic coma and diabetic ketoacidosis, which are less common but more severe. Given their low prevalence rates, we found we were underpowered for analyses of these particular outcomes and removed them from the paper.

diagnostic cutoff have a significantly greater number of A1c tests in the year following the initial test but are not more likely to have any A1c test.[17] This is also true in subsequent years (see Appendix Table A2). To address this, we randomly select one A1c test per individual per year. For consistency across measures, we similarly select a random total cholesterol value in each year following the initial A1c test, although there does not appear to be a significant difference in the number of tests for this measure at the diagnosis cutoff.

Finally, we examine differences in mortality at the cutoff. Any effect on mortality would be meaningful, but mortality is uncommon in our sample, and our power to detect effects is low. Since we require individuals to be in the sample for at least one year following the initial test, we examine mortality in years 2 and later. Additional details on the diagnosis and procedure codes used to construct the outcome variables may be found in Appendix Table A3.

Table 1 presents descriptive statistics for the CDM patients in the left panel. Of these patients, 52 percent are female, average age is 51.1 years, and 28 percent are enrolled in a Medicare Advantage plan with the remaining on individual private plans (such as those purchased from the Affordable Care Act health insurance exchanges) or employer-sponsored insurance.

B HEALTH RISK ASSESSMENT DATA

A small subset of patients represented in the CDM data complete a Health Risk Assessment (HRA) survey in any given year.[18] This survey asks a wide range of questions about self-assessed health and behaviors and is typically administered by the insurance company at the behest of an employer. As such, patients who take the HRA are not necessarily representative of patients in the CDM database.

The second column of Table 1 presents statistics for the HRA sample. Because the HRA tends to be administered for employers, the sample is composed almost entirely of individuals enrolled in private (non-Medicare) insurance plans. This sample is also somewhat younger (44.2 years old on average) and more female (55.4 percent). However, the characteristics of individuals who take the HRA appear to be balanced across the A1c threshold (see Section V). Therefore, our estimates provide an internally valid measure of these self-assessed outcomes within the HRA subsample.

A full list of variables considered from the HRA is found in Appendix Section A. In order to avoid conducting an excessive number of statistical tests, which might generate a large number of false pos-

[17] This is consistent with testing guidelines, which recommend screening once per year for patients with prediabetes and two to four times per year for patients with diabetes (Mayo Foundation for Medical Education and Research, 2021).

[18] Since HRA survey responses are typically observed only once per year, we do not impose the 6-month enrollment requirement described above for the CDM claims data; we include any patient with a survey response in any given year for analyses using the HRA data.

itives, and to improve our statistical power, we construct three summary indices for the following question domains: self-assessed health, physical and emotional well-being, and health behaviors. We do this by constructing a standardized z-score (subtracting the mean and dividing by the standard deviation) for each component and taking the average. This approach applies equal weight to each index component topic and is similar to the approach used in Finkelstein et al. (2012) and Kling et al. (2007). Positive values of the index indicate better outcomes (i.e., better health or lower rates of unhealthy behavior).

Self-assessed health is coded as 1 if the respondent reports health that is very good or excellent and zero otherwise. Physical and emotional well-being questions are asked about a variety of settings (e.g. whether in the past 2 weeks, did physical/emotional problems make it difficult to feel capable; see Appendix Section A). Each of these questions are given equal weight. For the health behavior index, we combine multiple measures of tobacco use (currently smoking cigarettes, cigars, or using smokeless tobacco) into a single variable that equals one if the respondent reported using any tobacco product. We also combine reports of engaging in physical activity 3 times per week and reports that the respondent recently made changes to increase physical activity into a single sub-index related to physical activity. The health behavior index therefore contains 5 equally-weighted components: tobacco use, alcohol use, fast food consumption, physical activity, and reported attempts to improve weight.

Of these three indices, we might expect the index related to health behaviors to be most directly affected by the diabetes diagnosis because the ADA recommends that physicians counsel newly diagnosed individuals to alter their behaviors on these dimensions. In contrast, self-reported health and physical/emotional well-being may be indirectly affected as the newly diagnosed engage in different behaviors or learn more about their underlying health status through increased contact with health providers.

III Empirical Strategy

We use a regression discontinuity (RD) design framework to analyze the impact of a diagnosis of diabetes across patients with similar underlying health. This is an example of a "fuzzy" regression discontinuity design since a diabetes diagnosis may depend on patient- and physician-specific factors separate from an A1c value. We estimate two versions of this model. The first version is a "reduced form" model that looks at the change in the outcome at the cutoff. The second version is an instru-

mental variables estimate of the effect of receiving a diagnosis of diabetes, using the diagnostic cutoff as an instrument for being diagnosed. To implement these analyses, we use local linear regression and the Fuji et al. (2009) optimal bandwidth selector, although we also assess the sensitivity of our results to alternative bandwidths and parametric regression methods (see discussion in Section A).

If all patients with A1c values below the cutoff were not diagnosed with diabetes, and all patients above the cutoff did receive a diagnosis, the reduced form and IV estimates would be identical. However, the change in the prevalence of a diagnosis does not change from 0 to 1 at the 6.5% threshold. As will be seen in the analysis that follows, a non-trivial share of patients with an A1c value below the threshold receive a diabetes diagnosis within the first year. This may be due to physician discretion, but it may also be the result of a later positive test. Clinical guidelines recommend that physicians repeat the test in 3 to 6 months for patients with test values close to the threshold (American Diabetes Association, 2018a).

It is also the case that a sizeable share of individuals with values above the cutoff do not receive a diabetes diagnosis. This may result from physician discretion in assigning a diagnosis. Later in Section B, we examine the characteristics of patients who receive a change in diagnosis as a result of crossing the diagnostic cutoff (i.e. the "compliers"), and how they compare to the broader population. We also explore whether we might expect effects of a diabetes diagnosis to look similar for the group of patients whose A1c value puts them above the cutoff, but who nevertheless do not receive a diagnosis ("never takers").

We note that there may also be some error present in our measurement of diagnosis. For instance, if physicians do not record the diagnosis of diabetes on the claim or medical record, or a patient does not seek additional care after learning about the diagnosis (and thus never generates additional claims), we would not be able to observe that the patient is diagnosed. That is, some patients may not appear to have received the diagnosis when, in fact, they did. In this case, our first stage would be attenuated and the IV estimate biased upwards, as it scales the reduced form estimate by the observed (not actual) change in diagnosis prevalence. To explore this further, we conduct a bounding exercise for the IV estimate in Appendix Table A5 that assumes that all patients with A1c values at 6.5 and above receive a diabetes diagnosis, regardless of whether we observe a diabetes diagnosis in their claims or medical record. This approach naturally results in a larger first stage and smaller IV estimates; however, our results still indicate that meaningful increases in spending occur at the diagnosis cutoff.

A MANIPULATION OF THE FORCING VARIABLE

A potential concern about the validity of our design is whether or not patients or physicians can manipulate the test score in some way so as to generate selection across the threshold. However, due to the nature of the test, this is unlikely in our setting. The patient's A1c reflects the plasma glucose levels of the patient during the past 3 months, making the test difficult to manipulate with short term changes to diet or exercise. In addition, the test values are reported to the data provider or recorded in the patient's medical record directly from the lab, making physician or patient manipulation impossible.

The absence of manipulation is also apparent in the first panel of Figure 1, which shows the sample size in our data for each A1c value from the initial lab test. There are three noticeable features. First, the density is smooth across the threshold, with no apparent evidence of bunching on either side. Second, given the shape of the distribution of A1c values within the population of test takers, a substantial fraction of patients whose A1c is at or above the threshold fall very close to the threshold. Indeed, 15 percent of patients with A1c at or above 6.5% in the CDM data have an A1c of 6.5% exactly; over 25 percent have an A1c value of 6.5% or 6.6%.

B TESTING FOR DISCONTINUITIES IN BASELINE CHARACTERISTICS

A second concern is that there are other changes that may occur at the diagnosis cutoff that affect the outcomes being studied. The regression discontinuity model relies on the assumption that patients with A1c values near the threshold differ only in the way they are treated by the medical system; i.e., that there are no systematic underlying differences between patients with A1c values just above and just below the 6.5% cutoff. We can assess this empirically by evaluating whether there are discontinuities in demographic characteristics, spending, and health measures at the cutoff in the year prior to the initial test. These variables are not expected to change at the diagnosis threshold, providing a check on the assumption behind our research design. If we find discontinuities in these characteristics, it suggests that there may be pre-existing differences in individuals above and below the cutoff.

We conduct this test by evaluating whether age, gender, Medicare enrollment, and enrollment in a consumer-driven health plan change discontinuously at the cutoff (results in the top panel of Appendix Table A6 and Appendix Figure A7). In the second panel, we conduct a similar test using spending measures observed in the year prior to the initial lab test. We examine total spending, spending on recommended care, statins, ACE inhibitors, and endocrinology specialty care. We oper-

ationalize this test by using each of these demographic, insurance, and spending characteristics as a dependent variable in our reduced form model for each of the study samples.

We do not find statistically significant discontinuities in the two demographic characteristics we consider (age and gender), in enrollment in a consumer-driven health plan, or in spending within categories. However, we do find a statistically significant reduction in Medicare coverage and baseline total spending at the cutoff. In later analyses presented in Section C, we demonstrate that our estimates are not sensitive to controlling for these covariates.

C TESTING FOR DISCONTINUITIES IN SAMPLE INCLUSION BY YEAR

The sample sizes for our analyses decline with each year of observation for two reasons. First, the length of follow up for each patient depends on the timing of their initial A1c test. For example, if a patient has her initial test in 2016, we are only able to observe one post-test year (2017), so the patient will not be in the sample for regressions that estimate effects in year 2 and beyond. Second, our sample is comprised only of patients who are enrolled in a plan with the health insurance carrier. Some patients leave plans included in CDM before the end of our study period. This effectively reduces the sample sizes and precision of our estimates for each year of follow-up.

This type of attrition should only be a concern for our research design, however, to the extent that patients on either side of the cutoff are differentially more likely to exit the sample over time. For example, if a diabetes diagnosis causes patients to switch out of plans included in CDM, this could affect the interpretation of our outcomes for later years. We test for differential attrition and report the results in Appendix Table A9. Our dependent variable in these models is whether or not the individual appears in the sample in each subsequent year (i.e. years 2-6).

We find no significant discontinuities in the probability of attrition, demonstrating that selective attrition out of the sample based on A1c relative to the diagnosis cutoff is unlikely. The results from these tests indicate that differential selection at the cutoff is not a threat to the internal validity of our estimates.[19]

[19] We should note, however, that it is necessarily the case that the estimates of the long-term effects of diagnosis rely on the sample of patients who were tested during the earlier part of the sample period. To the extent that the treatment of diabetes (or pre-diabetes) has changed since 2009, this limits the generalizability of the results to individuals who are currently screened for diabetes. However, we have little reason to think that physician practice has meaningfully changed since 2009. Practice guidelines for the treatment of diabetes have not changed considerably since then; essentially the only practice guideline that meaningfully changed was the recommendation to initiate statin therapy for all patients with diabetes age 40 and older regardless of cholesterol levels beginning in 2013.

IV Results

A SHORT RUN EFFECTS: YEAR 1

We first analyze the impact of having an A1c value exceeding the 6.5% threshold on diabetes diagnosis and medical care spending in the first 12 months following the initial lab test. The first panel of Figure 2 plots the rates of diagnosis by A1c value for the year following the initial test. We see a large and discontinuous increase in diagnosis rates at the cutoff value of 6.5%. At the 6.5% value, diagnosis rates jump by over 10 percentage points.

While we see a significant jump in diagnoses at the threshold, it is notable that not *all* patients whose A1c puts them above the diagnosis threshold receive an observable diagnosis. As discussed earlier, this may reflect measurement error in our claims data (e.g., the patient is told he or she has diabetes but it is not recorded on the claim) or alternatively it could be caused by the patient not following up for care or the physician not reaching out to the patient following the lab test. Since this lack of follow-up care could have negative implications for patient health, we conduct further analysis to characterize the "compliers" (whose diagnosis status changes as a result of crossing the threshold) and compare them to the "never takers" (who do not receive a diagnosis of diabetes despite having A1c above the threshold) in the next section.

The second panel of Figure 2 shows average spending on all recommended diabetes care. Here, we observe a clear discontinuous increase in spending at the diagnosis threshold. The subsequent panels show the results for spending on specific types of diabetes-related care: office-based care (consultations and evaluation), vision care, lab test spending, and endocrinology specialty care. We observe noticeable jumps in all outcomes, although spending on endocrinology specialty care is particularly noisy.

Figure 3 displays similar graphs for two other measures of utilization: number of diabetes-related office visits and number of labs for glucose or A1c. The figure also displays graphs for spending on drugs that are commonly prescribed to patients with diabetes: antiglycemic medications such as metformin, statins, and ACE inhibitors. We observe noticeable jumps for the first two outcomes, but less evidence of an increase in prescription drug spending in the first year. Because the first-line diabetes controlling drug, metformin, is available as an inexpensive generic (about $4 for a 30-day supply), it may be difficult to observe spending increases even if use of this drug increases. For that reason, we also examine an indicator that the patient has any spending for a diabetes-controlling drug.

This indicator exhibits a small but less obvious jump at the diagnosis threshold. In the next two plots, we see some evidence of an increase in spending for statins and ACE inhibitors, with a larger jump for the latter drug.

Table 2 presents the local linear regression estimates associated with the outcomes in Figure 2. We see a statistically significant increase in the presence of a diabetes diagnosis of 10.8 percentage points (row 1). The next row shows the effect of crossing the diagnosis threshold on spending on recommended care, as defined by the ADA recommendations. At the cutoff of A1c equal to 6.5%, we observe a discontinuous increase in spending of $208. We scale this increase in spending by the fraction of the sample who gain a diagnosis of diabetes at the cutoff with an IV estimate in the next column. This estimate shows that a diagnosis of diabetes is associated with approximately $1,097 of additional spending on recommended care.[20]

When we break this spending down into subcategories, we see that the spending increase in the first year is split primarily between office consultations and evaluations, which exhibit an increase of $508 associated with a marginal diagnosis, and spending related to lab work, which experiences an increase of $505. We also see a significant increase in spending related to screening for diabetic retinopathy (i.e. vision services), of $28 in the first year. In addition, we estimate a marginally significant increase in spending on endocrinology specialty care ($139) among the marginally diagnosed.

Table 3 presents similar results for the additional measures of utilization and drug spending. We find that a diagnosis of diabetes is associated with about 3 additional office visits for diabetes-related care in the 12 months following diagnosis and about 0.4 more A1c or glucose labs. A diagnosis of diabetes also significantly increases the probability that a patient takes a diabetes-controlling drug by 4.3 percentage points. We also find increases in total spending on ACE inhibitors, but no significant effect on spending on statins or the amount spent on diabetes-controlling drugs.

It is notable that, although the average amount spent on diabetes-related care increases at the threshold, relatively few patients receive *all* elements of recommended care. In Appendix Table A10, we examine the probability that a patient has at least some spending in the following categories: diabetes-related office visits, diabetes-related vision services, diabetes-related lab tests, and diabetes-related prescription drugs. A diabetes diagnosis increases the probability a patient receives at least some care from each of these categories by only 5.7 percentage points, suggesting that many patients receive an incomplete bundle of interventions.

[20]Note that because the bandwidth selector chooses a different bandwidth for every outcome based on the variance of that outcome, the "first stage" varies across rows.

We next examine complications related to diabetes and health outcomes in the CDM data. Figure 4 shows the effect of receiving an A1c value above the diabetes diagnosis cutoff on the incidence of preventable conditions associated with diabetes: diabetic retinopathy (eye damage), diabetic neuropathy (nerve damage), and diabetic kidney disease. The incidences of diabetic neuropathy and diabetic kidney disease appear to increase at the A1c cutoff. This is not necessarily surprising given that the ADA guidelines stipulate that much of the recommended care for diabetes involves screening for these conditions. The associated estimates are reported in Table 4. We find statistically significant increases in the diagnoses of all three conditions.

The remaining panels in Figure 4 display the results for the direct health outcomes we measure in the CDM data: change in A1c from the initial test value, change in total cholesterol from pre-test values, and mortality. There appears to be a larger negative value in the difference between current and initial A1c levels, indicating greater improvement relative to the initial test for those above the cutoff relative to those just below. We see no visible evidence of a change in mortality.

The RD estimates are presented in the bottom three rows of Table 4. A1c relative to the prior year falls by about 0.06 percentage points for patients at the cutoff, and is associated with a reduction in A1c of about 0.34 percentage points for the marginally diagnosed. We find no effect on changes in total cholesterol (relative to levels observed prior to the initial test), although the confidence intervals do not rule out clinically meaningful effect sizes. We also do not detect effects on mortality; and, while the confidence intervals include negative estimates, the direction of the point estimates do not suggest mortality benefits. This may be unsurprising since the marginally diagnosed are the healthiest group of patients with diabetes and likely carry a much lower risk of short-term health complications.

Finally, we examine the effect of diabetes diagnosis on measures of self-reported health and healthy behaviors in the Health Risk Assessment (HRA) survey data (Figure 5 and Table 5).[21] As with the main CDM sample, we observe a significant increase in the incidence of diabetes diagnosis in this sample (by 16 percentage points). However, we do not find any statistically significant changes in self-reported health, physical or emotional well-being, or health behaviors associated with the diagnosis of diabetes. In fact, the direction of the coefficient estimates suggests that the marginally diagnosed report worse subjective health. Results for the components of these indices also indicate no effect of diagnosis on self-reported health or behaviors; these are reported in Appendix Table A4.

[21] Note that graph (a) in Figure 5 differs from Figure 2(a) in that it estimates the change in diabetes diagnoses specific to the HRA sample.

B CHARACTERIZING "COMPLIERS"

Not all patients in our data receive an observable diagnosis of diabetes when they receive an A1c value that puts them above the threshold. Our IV estimates show the effect of a diabetes diagnosis among "compliers"–those whose diagnosis status changes as a result of crossing the cutoff. These compliers may differ from the broader population, and may also systematically differ from "never takers" (those whose A1c value puts them above the cutoff, but who nevertheless do not receive a diagnosis) and from "always takers" (those whose A1c value puts them below the cutoff, but who nevertheless do receive a diagnosis). Of particular interest to public health officials and other policy makers are the never takers, individuals who should receive a diagnosis based on clinical guidelines but do not. From the perspective of clinicians, these are patients who are not getting needed treatment.

While it is not possible to individually identify compliers, we can learn how their composition compares to the full population by examining the differences in the first stage across subgroups. We do so following the procedure outlined in Section 4.4.4 of Angrist and Pischke (2009) and report the results in Appendix Table A11. This method examines the ratio of the first stage within a subgroup (e.g., women) to the first stage observed in the entire sample. Groups with a larger first stage are over-represented among compliers, while those with a smaller first stage are under-represented, when compared to the full sample. For example, if women have a first stage that is twice as large as what is estimated in the full sample, we would expect them to be over-represented by a factor of two among compliers. We then multiply this ratio by the average rate of this characteristic in the population near the cutoff (within the optimal bandwidth in the first stage regression). Following the above example, if women were 35 percent of the sample, but over-represented among compliers by a factor of two, we would expect 70 ($35\% \times 2$) percent of compliers to be women and 30 percent to be men.

We also report averages for those whose A1c value falls just above the cutoff but do not receive a diagnosis in the first year ("never takers") and those whose A1c value falls just below the cutoff yet do receive a diagnosis in the first year ("always takers"). For never takers, we use patients with A1c values in the range [6.5,7) who receive no diagnosis; for always takers, we use patients in the range [6.0,6.5) who do receive a diagnosis. This relies on the typical IV assumption that there are no "defiers" (i.e. patients who are only diagnosed if their A1c places them below the diabetic range and not diagnosed if their A1c places them above).

On most dimensions, compliers look similar to the control group (patients with A1c test values below 6.5 and greater than or equal to 6), although they are somewhat younger, less likely to be

female, and less likely to be in a consumer-driven health plan. Compliers are also much more likely to be taking a statin at the time they receive the A1c test. While never takers are similar to compliers in most respects, they are notably more likely to be in a consumer driven health plan. It is possible that facing high cost sharing is a barrier that prevents these patients from receiving a diagnosis or follow up care.

It is not possible to examine the impact of a diagnosis on never takers as their diagnosis status does not change at the threshold. However, we can estimate effects among a subgroup of patients that resembles the never takers based on a number of different pre-test characteristics. When we do this, it indicates that this group experiences larger health improvements and smaller increases in spending than patients who are more dissimilar from the never takers. This suggests that if these "missed" diagnoses could be realized, these patients may experience beneficial effects. See Appendix Section 4 for more details.

C MEDIUM RUN EFFECTS: YEARS 2 THROUGH 6

The initial effects of a diabetes diagnosis may not persist if, for example, those with A1c values directly below the cutoff gain a diagnosis relatively soon after the initial test. Therefore, we first examine the likelihood of later diagnosis in each year. We also examine the persistence of spending and health effects by re-estimating our IV model using outcomes from years 2 to 6 following the initial test. These estimates demonstrate how a diagnosis in year 1 affects spending and health in subsequent years.

Appendix Figure A2 shows how diagnosis and spending evolve over time, in years 2 through 6 following the initial lab test. The first panel shows the reduced form estimates on the probability of ever being diagnosed with diabetes by each subsequent year. After the initial year, the difference in diagnosis rates across the threshold closes somewhat but does not entirely disappear until year 6; we observe significantly higher diagnosis rates 3 years after the initial test and positive point estimates for the first 5 years.

The second panel of Appendix Figure A2 shows the effect of diagnosis in year 1 on spending on subsequent years (i.e. IV estimates). In contrast to the diagnosis estimates, which show convergence between the marginally diagnosed and marginally undiagnosed, we see fairly stable impacts of a year 1 diagnosis on spending on recommended care in each subsequent year, with little evidence that the gap in spending between those above and below the threshold closes. The estimates become less precise due to some individuals leaving the sample over the study period, but we see consistently

positive effects for all years and a statistically significant effect even five years after the initial test. We see similar patterns for spending on other types of diabetes-related care and for measures of utilization (Appendix Figure A3). Indeed, spending on diabetes-related drugs and vision care appears to grow larger over time.

Next, we examine how the impact of a diabetes diagnosis affects the diagnosis of complications related to diabetes and health. These results, displayed in Appendix Figure A4, show significantly higher diagnoses of neuropathy and diabetic kidney disease in years 1 and 2. Subsequent years show neither statistically significant increases nor reductions in the rates of complications related to diabetes. Note that because we have smaller samples in later years, however, we lose precision for many estimates starting in year 4 or year 5 after the initial test. The last three panels of Appendix Figure A4 show the effects on changes in A1c, total cholesterol, and mortality in the years following the test. In contrast to the spending effects, which persisted over the entire period, the improvement in A1c observed in the first year does not continue in subsequent years. We also do not observe any changes in total cholesterol or mortality in subsequent years. However, the confidence intervals on some of our estimates, particularly during the later years of the observation period, do not allow us to rule out meaningful health improvements.

D SUMMARY OF FINDINGS

To summarize, we find that the diagnosis of diabetes for individuals with A1c values just above the diagnosis cutoff leads to greater spending on diabetes-related medical care. This spending increase is not only present in the year following the test but persists over time. Our point estimates suggest that a diagnosis of diabetes is associated with a cumulative $8,107.86 in additional spending on recommended care over 6 years, in present value terms.[22]

In terms of health benefits associated with this spending, we see only minimal improvements in health over our observation period. We find a short-term improvement in A1c during the year following the initial test. This change in A1c appears short-lived with no evidence of additional A1c improvement over the following 5 years. We find no evidence of changes in other clinical health measures (total cholesterol or mortality), nor do we find changes in patient-reported health or health behaviors. The coefficient estimates are not only small in size, but often take the opposite direction of a health benefit.

[22]We arrive at $8,107 by aggregating the IV estimates of the effect of a diagnosis on recommended care over the 6 years for which we have estimates, assuming additional spending related to a diagnosis is $0 for subsequent years.

We interpret these null results with caution. The results may provide evidence that the additional healthcare spending does not correspond to health improvements at the margin. However, recommended care for both patients with prediabetes and patients with diabetes includes changes in health behaviors including diet and exercise. We are unable to precisely track health behaviors with our data, and an alternative explanation for the null results is that changes in health behaviors are symmetrical around the A1c cutoff and translate to similar health outcomes for patients above and below the cutoff.

In addition, we only observe patients for 6 years following the initial diagnosis of diabetes; and, it may be the case that additional improvements in health manifest in year 7 or later. Indeed, we find strong evidence that receiving a diagnosis of diabetes is associated with early diagnosis of several complications associated with diabetes, such as retinopathy, neuropathy, and diabetic kidney disease; since these complications were detected earlier, we might expect patients to experience improved health trajectories over a longer time horizon. There may also be additional value to patients of learning about these conditions or their diabetes diagnosis (see additional discussion in Iizuka et al., 2021), in general, which is not captured in our analysis.

V Specification Checks

In this section, we conduct several analyses to probe the robustness of our results to different sample and model specification choices.

A SENSITIVITY TO SPECIFICATION AND BANDWIDTH CHOICES

First, we assess the robustness of our results to using a quadratic model and alternative bandwidths in the local linear model, reported in Appendix Tables A12-A14. The quadratic regression is given by:

$$y_i = \alpha + \beta A1c \geq 6.5_i + \gamma_0 c + \gamma_1 c^2 + \gamma_2 A1c \geq 6.5 \times c + \gamma_3 A1c \geq 6.5 \times c^2 + \epsilon_i \quad (1)$$

where we include a quadratic function in the A1c value c centered at 6.5 that is allowed to differ on the two sides of the cutoff point. Following recent guidance on implementation of RD designs with discrete running variables, we do not cluster the standard errors by the running variable, but rather estimate robust standard errors (Kolesár and Rothe, 2018).

In the first column, we re-report the results from our main reduced form specification to facilitate comparison. We then report the results from a model that estimates the RD effect using a quadratic in

initial lab value that is allowed to vary at the cutoff. Spending results with this model are similar to our baseline estimates for most outcomes (Appendix Table A12), although often larger in size. Subsequent columns show the results for larger (twice) and smaller (one half) multiples of the optimal bandwidth in the local linear regression estimation. Results are very similar both quantitatively and qualitatively when alternative bandwidths are used, although larger bandwidths tend to result in somewhat larger estimates. The estimates for preventable conditions, clinical health measures, and self-reported health measures (Appendix Tables A13-A14) look, for the most part, similar across specifications. Estimates of the change in A1c from its baseline value, however, shows a larger decline under a smaller bandwidth, suggesting that the difference in values is most apparent for those individuals closest to the threshold.

B SENSITIVITY TO SAMPLE INCLUSION CRITERIA

We also examine the sensitivity of our results to our sample inclusion criteria. In particular, we require that individuals be enrolled in a plan included in CDM data for at least 1 year prior to their initial A1c test so that we can adequately assess whether they have previously been diagnosed with diabetes. It may be the case, however, that some individuals in our sample were misclassified as having no diabetes diagnosis prior to their first test when, in fact, they had already been diagnosed. This might be the case if, for example, they were diagnosed before entering the CDM data. In Appendix Tables A15 and A16, we conduct our analysis but only include those who were enrolled in an included plan for 2 years prior to the initial test. Approximately 70 percent of our main analytic sample met this criterion. Our spending results are very similar, indicating that our initial sample definition likely performed well at excluding those who were already diagnosed with diabetes. Omitted here in the interest of space, results on self-reported health and healthy behaviors are also similar under this sample refinement.

C SENSITIVITY TO CONTROLS FOR DISCONTINUOUS BASELINE CHARACTERISTICS

In section B, we tested whether baseline characteristics changed discontinuously at the cutoff. We found no changes in demographic characteristics, baseline spending by category, or enrollment in a consumer-driven health plan, but we did find evidence of a statistically significant reduction in Medicare coverage and baseline total spending at the cutoff. To evaluate whether these differential characteristics are driving our results, we re-estimate our spending models to include baseline spending and Medicare enrollment. The inclusion of these variables has little effect on our results, suggesting that

the spending effects are not driven by differential baseline characteristics (see Appendix Tables A7 and A8). Furthermore, if we exclude the Medicare sample from our data, we find very similar effects as those observed using the full sample, but no discontinuity in baseline spending. Finally, we do not see any significant effects in other variables measuring spending as presented in the bottom panel of Appendix Table A6. Taken together this information suggests that there are not systematic differences in patient characteristics at the diabetes diagnosis threshold that are confounding our analysis.

D SENSITIVITY TO REMOVING ATTRITERS

Although we find no evidence of sample attrition associated with the diagnosis cutoff, it could be the case that individuals who remain in our sample all six years vary systematically from those we only observe for one or two years. As a result, the effects we estimate in years 2-6 following the initial lab test may reflect changes in the underlying sample rather than time-varying treatment effects. We examine this by comparing the results obtained from our main (unbalanced) sample, from those we obtain when we restrict the sample to only those who remain enrolled all 6 years (balanced). The results of this exercise are presented in Figures A5 and A6, which plot the coefficient in each year for the unbalanced panel (black) against those obtained in the balanced panel (red). We provide results just for the spending and utilization outcomes, given the low rates of mortality and some of the preventable conditions, and the relatively small size of the balanced panel. Although the balanced panel has much larger confidence intervals, the point estimates indicate that diagnosis rates and spending are similar across the two samples. Based on these analyses, there do not appear to be systematic differences in the effect of diabetes diagnosis across the group that remains in the sample continuously and the main sample.

VI Additional Analyses

We next undertake some additional analyses to further explore and understand our main findings.

A CHANGES IN OVERALL UTILIZATION

We explore whether the increase in diabetes-related care occurs during separate encounters or visits, or whether it alternatively changes the focus or style of treatment that occurs during interactions with the health care system that would take place regardless of the diagnosis. For example, diabetes and non-diabetes patients may have the same number of office visits, but diabetes patients receive information and treatment specific to their illness, whereas non-diabetes patients receive more general

information on their health. We assess this in two ways.

First, we look at the total number of office visits for evaluation and management and total spending on this category, including visits unrelated to diabetes. These results are presented in the first panel in Appendix Table A18. We find no significant change in the total number of office visits or total spending on office visits, despite the fact that we see increases in visits and spending for diabetes-related care. This suggests that, at least for office visits, diabetes-related care could be delivered in the context of existing health care interactions, either representing additional care received during the same encounter or crowding out visit time that would have otherwise been focused on other problems or diseases. However, our standard errors are large enough that we cannot rule out annual increases in total office spending smaller than $188 or increases in office visits smaller than 0.6.

Next, we examine diabetes-related care that more clearly represents new utilization by focusing on the elements of recommended care that are *not* identified by a diagnosis code: spending on statins, ACE inhibitors, diabetes-controlling drugs, endocrinology specialist care, and spending on glucose labs. Among this subset of recommended care, we find that diagnosis results in a statistically significant increase in spending of about $250 in the first year, which represents about one-quarter of our main spending estimate. Thus, we find clear evidence of an increase in some types of "new" utilization associated with a diabetes diagnosis.

B HETEROGENEITY BY PROVIDER TYPE

Next, we conduct additional analyses to explore whether patients of different types of providers experience different changes in outcomes upon being diagnosed with diabetes. We identify the provider who ordered lab analysis for the A1c test among patients by finding outpatient visits within one week of the observed lab that contains a procedure code for an A1c test. If no outpatient visit exists, but we do observe an office visit for the patient in the week the lab is processed, we assign the provider at that office visit as the one who ordered the A1c lab.[23] Using these patient-provider linkages, we focus on two broad groups of providers: primary care physicians (those with a specialty of family medicine or internal medicine), and specialists (those with other specialties). We then conduct an RD analysis within each subgroup defined by the provider's specialty.

The results of this heterogeneity analysis are found in Appendix Table A19. As may be seen from the first row of this table, an A1c value above the cutoff is more likely to result in a diabetes diagnosis

[23] We are able to successfully link about 60% of our sample to providers using this method. Additional details on how we link providers to lab orders are found in Appendix Section 1.

for patients of specialists versus primary care doctors. The point estimates also indicate that patients diagnosed by specialists experience somewhat higher spending as a result ($1458 vs $1200), which appears to be accounted for by higher spending on office-related diabetes care ($637 vs $531) and on diabetes-related labs ($637 vs $550). However, our estimates are not precise enough to reject that the changes at the cutoff are similar. Furthermore, it is not clear from this analysis whether differences in the effect at the cutoff by provider specialty reflect provider-specific factors (such as providers' practice style or preferences) or underlying differences in the patient population.[24] More work is needed to understand the interplay between provider guideline adherence and patient outcomes.

VII Conclusion

The distribution of health in the population is such that many individuals with a chronic illness are near the margin of being diagnosed. This paper explores the impact of a diagnosis on such "marginal" patients using longitudinal medical claims records for a large private insurer that includes clinical data. We find that individuals whose blood glucose levels are just above the threshold for a diabetes diagnosis have significantly higher spending on recommended care than similar patients whose blood glucose levels fall just below this cutoff. Receiving such a diagnosis appears to put patients on a path of persistently higher spending and utilization of diabetes-related care. Even after 5 years, patients whose initial test put them above the diagnosis cutoff have spending that is significantly higher than those with lab tests just below this cutoff.

Among these marginally diagnosed patients, this additional spending appears to be associated with only minimal improvements in health over our observation period. We see a reduction in A1c in the first year following diagnosis, but this improvement does not persist through years 2 through 6 and is somewhat sensitive to model specification. We do not detect any improvements in total cholesterol or mortality rates. We also do not detect any changes in self-reported measures of health or healthy behaviors. We do find evidence that receiving a diagnosis of diabetes is associated with diagnosis of several complications associated with diabetes (retinopathy, neuropathy, and diabetic kidney disease), which may have repercussions for quality of life or health over a longer time horizon.

Our results also shed light on the fact that not all patients who receive test values above the diagnosis threshold ultimately receive a diagnosis. The patients who are marginally diagnosed (compliers)

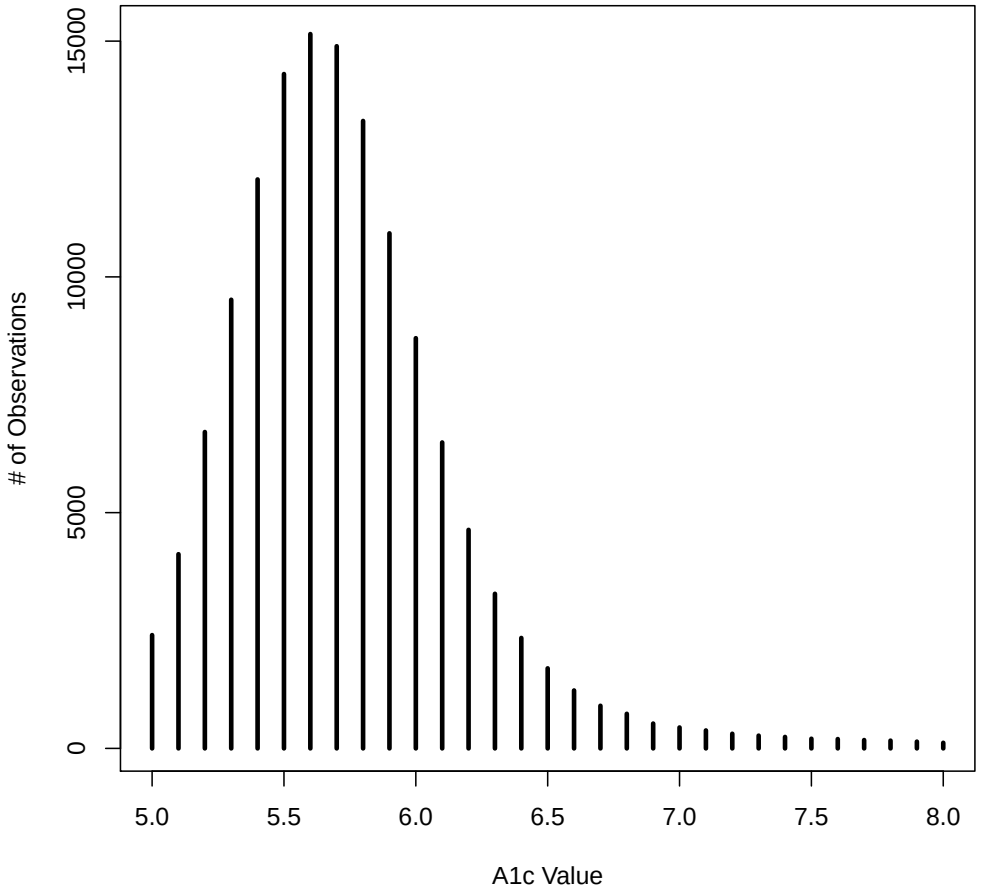
[24] Indeed, we see that spending on recommended care is higher even among patients of specialists who do not cross the diagnosis cutoff threshold, as evidenced by the differences in the "control mean" of those whose A1c falls between 5.9 and 6.4.

differ on observable characteristics from those whose A1c test result puts them above the diagnosis threshold but who nevertheless do not receive the prescribed interventions (never takers). Notably, the compliers are less likely to face high cost sharing compared to the never takers. We also find that patients of specialists, as opposed to primary care doctors, are more likely to receive a diabetes diagnosis. Additional research is needed to fully characterize how plan characteristics, provider behavior, and other factors, affect who receives timely care and under what circumstances.

We note that patients just missing the diagnostic guideline for diabetes are classified as having prediabetes, or being at higher risk for developing diabetes in the future. Clinical guidelines recommend provider counseling on lifestyle changes for these patients as well as continued monitoring, which may contribute to the unobserved differences in healthy behaviors or health outcomes we observe at the diabetes cutoff. In an alternative setting where the diagnosis cutoff generates more of a "bright line" regarding clinical intervention, we might expect the effects of crossing this line on the use of care–and potentially health–to be larger.

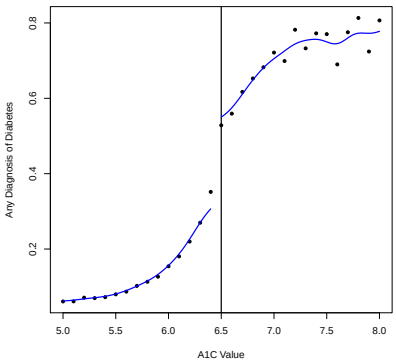
In addition, the length of our analysis is limited by attrition over time in the sample, which is particularly large in year 5 and year 6 following the initial A1c test. During the later years of observation, the confidence intervals on our health measures often do not allow us to rule out clinically meaningful improvements. In the absence of population-level data on clinical health measures in the U.S., it will be difficult to assess any longer-term consequences for patient health. Current efforts underway to longitudinally link electronic health records across health systems, as well as to link all-payer claims databases and EHR records within specific states, may enable researchers to better examine this question at some point in the future.

Figure 1: Sample size by A1c Value

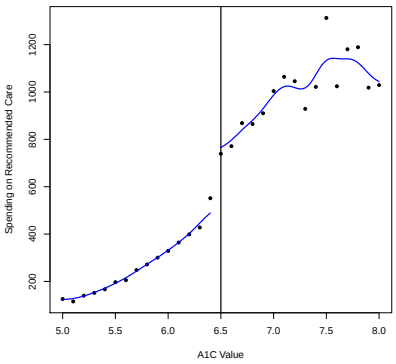


Source: Authors' calculations from the Optum claims data. Horizontal axis indicates each value of A1c. Vertical axis indicates sample size within each A1c bin.

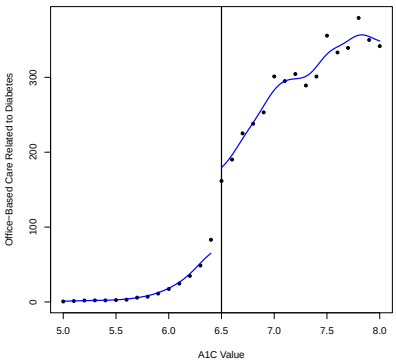
Figure 2: Diabetes Diagnosis and Spending in 12 Months Following Initial Lab Test



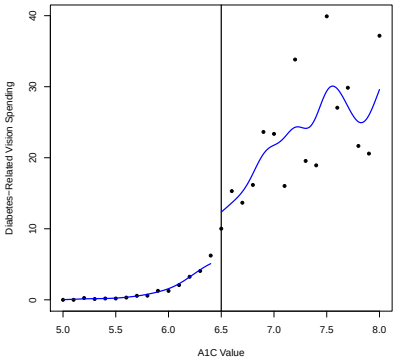
(a) Any Diagnosis



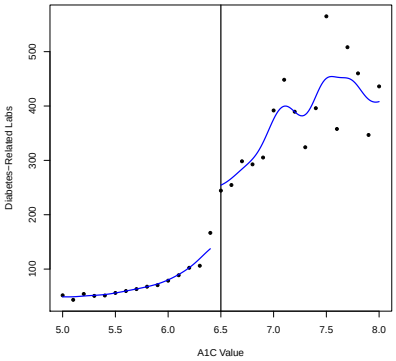
(b) Recommended Care



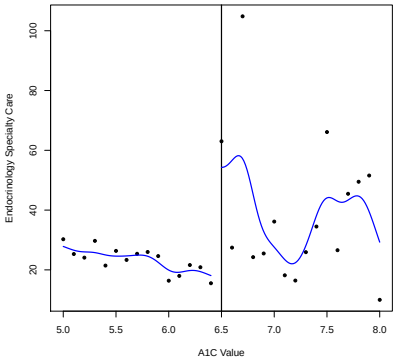
(c) Diabetes-related Office Visits



(d) Diabetes-related Vision Care



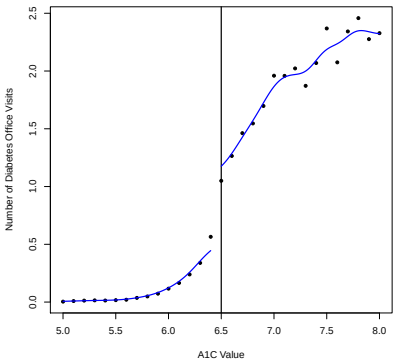
(e) Diabetes-related Labs



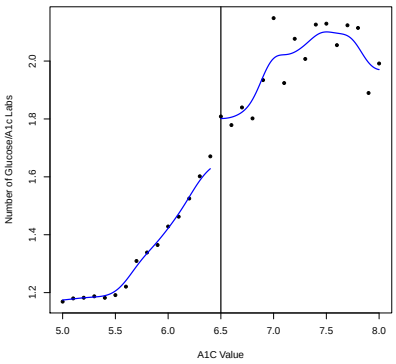
(f) Endocrinology Specialist Care

Source: Authors' calculations from the Optum claims data. Horizontal axis displays the test value. Vertical axis displays within test value averages of outcome variable. Vertical line at 6.5 is the diagnosis cutoff for diabetes.

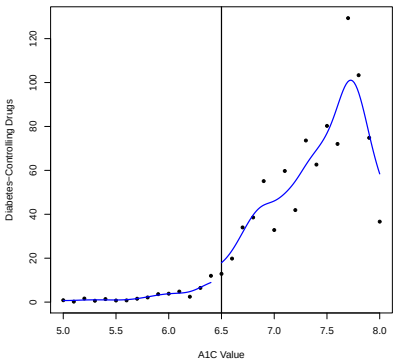
Figure 3: Drug Spending and Other Utilization Measures in 12 Months Following Initial Lab Test



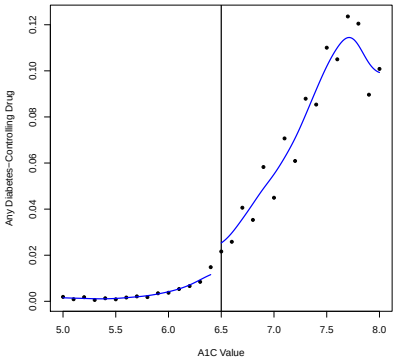
(a) # of Diabetes-Related Office Visits



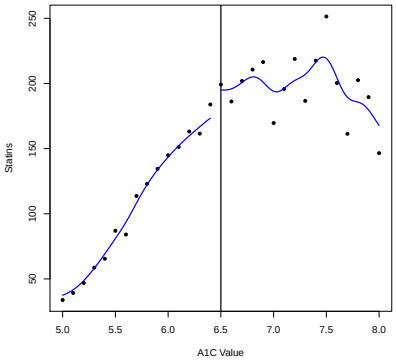
(b) # Glucose/A1c Labs



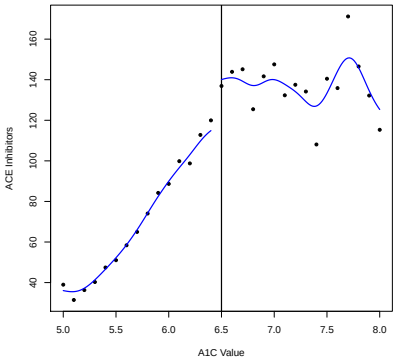
(c) Diabetes-Controlling Drugs



(d) Any Diabetes-Controlling Drug Spending



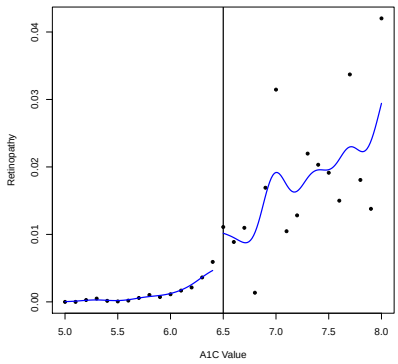
(e) Statins



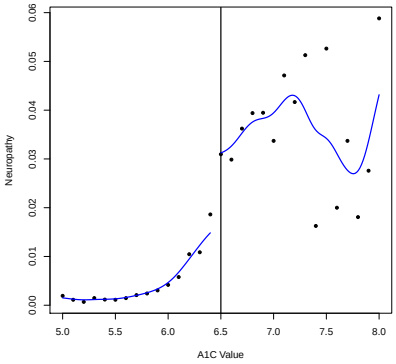
(f) ACE Inhibitors

Source: Authors' calculations from the Optum claims data. Horizontal axis displays the test value. Vertical axis displays within test value averages of outcome variable. Vertical line at 6.5 is the diagnosis cutoff for diabetes.

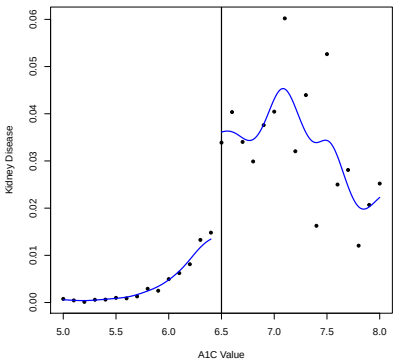
Figure 4: Diagnosis of Complications Related to Diabetes and Optum Health Outcomes in 12 Months Following Initial Lab Test



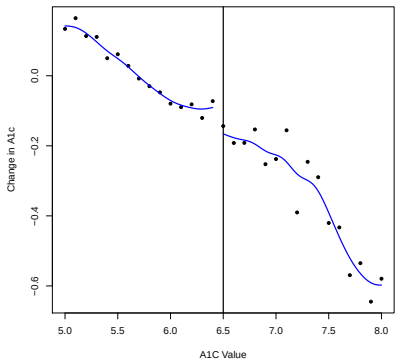
(a) Diabetic Retinopathy



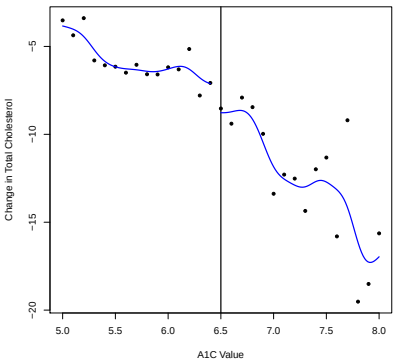
(b) Diabetic Neuropathy



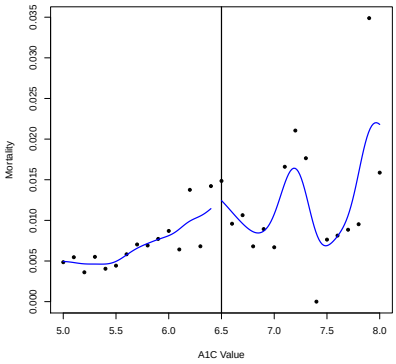
(c) Kidney Disease



(d) Change in A1c



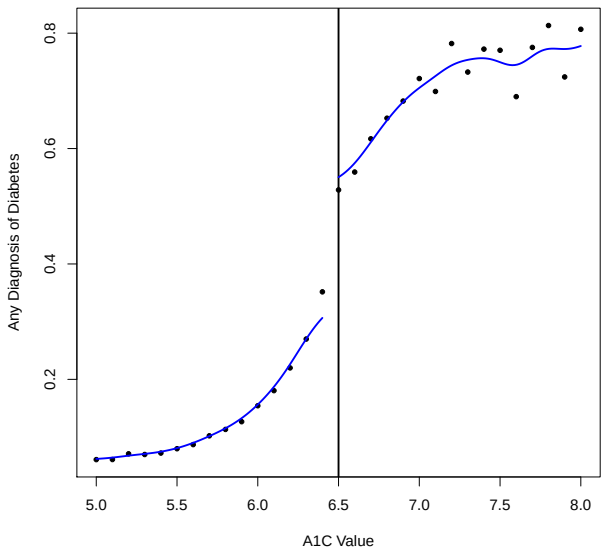
(e) Change in Total Cholesterol



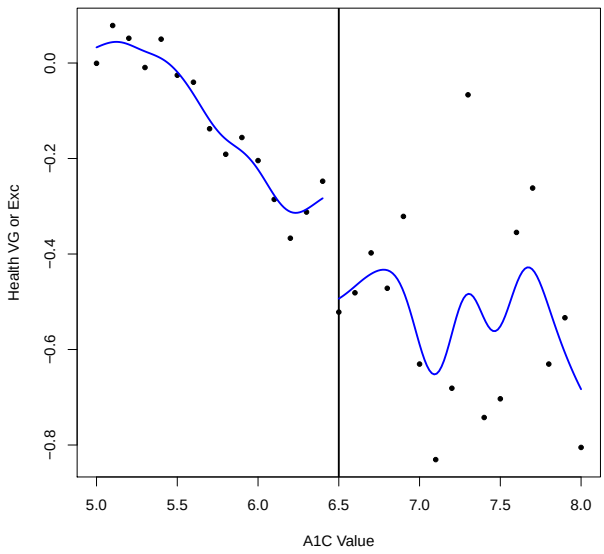
(f) Mortality

Source: Authors' calculations from the Optum claims data. Horizontal axis displays the test value. Vertical axis displays within test value averages of outcome variable. Vertical line at 6.5 is the diagnosis cutoff for diabetes.

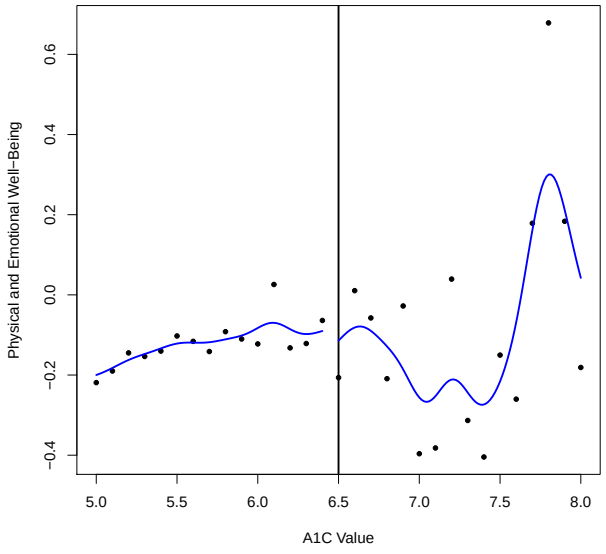
Figure 5: Self-Reported Health and Behavior in 12 Months Following Initial Lab Test



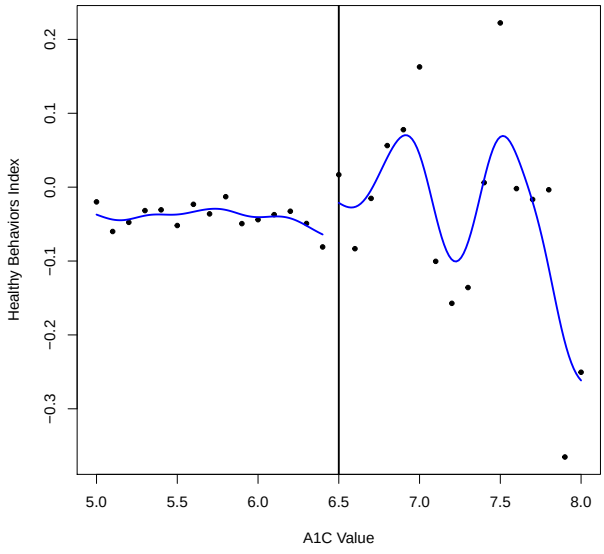
(a) Diabetes Diagnosis



(b) Health VG or Exc



(c) Physical and Emotional Well-Being



(d) Health Behaviors

Source: Authors' calculations from the Optum Health Risk Assessment data. Horizontal axis displays the test value. Vertical axis displays within test value averages of outcome variable. Vertical line at 6.5 is the diagnosis cutoff for diabetes.

Table 1: Demographic Statistics, CDM Claims and HRA Data

	Medical claims	Health Risk Assessment
Female	0.523	0.554
Age	51.13	44.22
Medicare	0.281	0.006
Private	0.719	0.994
Consumer-Driven Health Plan	0.160	0.405
N	142,541	16,916

Note: Authors' calculations from medical claims and Health Risk Assessment data. This table contains descriptive statistics for patients with an A1c test who meet all sample criteria described in the text.

Table 2: RD Estimates of Effects of Diabetes Diagnosis on Spending by Category in First 12 Months Following A1c Test

	Control Mean (Std Dev)	N in Bandwidth	Bandwidth (+/-)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Diabetes Diagnosis	0.183	14,821	0.342	0.108*** (0.019)	NA
Spending on ADA-Recommended Care	362.03 (718.09)	71,050	0.847	207.609*** (22.912)	1,096.894*** (112.295)
Spending on Office-Based Care Related to Diabetes	26.17 (122.28)	42,279	0.637	79.112*** (6.912)	507.661*** (42.690)
Spending on Diabetes-Related Vision Services	2.26 (27.92)	55,892	0.712	5.483*** (1.446)	28.215*** (7.360)
Spending on Diabetes-Related Labs	89.10 (353.69)	42,279	0.637	78.650*** (13.891)	505.367*** (80.798)
Endocrinology Specialist Care Spending	19.95 (313.62)	137,896	1.624	38.741* (23.024)	139.092* (82.668)

Notes: Control means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table 3: RD Estimates of Effects of Diabetes Diagnosis on Drug Spending and Other Utilization Measures in First 12 Months Following A1c Test

	Control Mean (Std Dev)	N in Bandwidth	Bandwidth (+/-)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Number of Diabetes-Related Office Visits	0.177 (0.775)	30,974	0.554	0.450*** (0.045)	3.166*** (0.298)
Number of A1c or Glucose Labs	1.46 (1.01)	42,286	0.663	0.063* (0.032)	0.393** (0.194)
Spending on Diabetes-Controlling Drugs	4.37 (97.04)	100,938	1.030	5.264 (3.497)	24.379 (16.183)
Any Spending on Diabetes Drugs	0.005	100,938	1.043	0.009*** (0.003)	0.043*** (0.014)
Spending on Statins	151.21 (436.84)	113,199	1.149	2.073 (9.799)	8.939 (49.50)
Spending on ACE Inhibitors	96.00 (299.36)	122,886	1.29	13.767** (7.020)	55.420** (28.258)

Notes: Control means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table 4: RD Estimates of Effects of Diabetes Diagnosis on Subsequent Diagnoses of "Preventable" Conditions, Subsequent A1c, and Mortality in First 12 Months Following A1c Test

	Control Mean (Std Dev)	N in Bandwidth	Bandwidth (+/-)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Diabetic Retinopathy	0.0018	86,441	0.918	0.004**	0.021**
				(0.002)	(0.010)
Diabetic Neuropathy	0.007	30,974	0.593	0.010**	0.068***
				(0.004)	(0.029)
Kidney Disease	0.0063	42,279	0.665	0.018***	0.111***
				(0.004)	(0.026)
Change in A1c from Test Year	-0.0564 (0.460)	11469	0.555	-0.059**	-0.342**
				(0.027)	(0.166)
Change in Total Cholesterol from Test Year	-6.429 (32.62)	29,159	1.162	-1.471	-4.993
				(1.212)	(4.114)
Mortality	0.005	12,799	0.727	0.0002	0.001
				(0.001)	(0.007)

Notes: Control means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table 5: RD Estimates of Effects of Diabetes Diagnosis on Self-Reported Health and Behaviors in First 12 Months Following A1c Test (Health Risk Assessment Results)

	N in Bandwidth	Bandwidth (+/-)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Diabetes Diagnosis	7,588	1.053	0.159***	NA
			(0.040)	
Self Reported Health Very Good or Excellent	4,795	1.053	-0.116	-0.728
			(0.083)	(0.558)
Physical and Emotional Well-Being	7,096	1.845	-0.077	-0.292
			(0.058)	(0.221)
Health Behaviors	6,974	1.554	0.041	0.172
			(0.040)	(0.172)

Notes: Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

References

Almond, D., J. J. Doyle, A. E. Kowalski, and H. Williams (2010, May). Estimating Marginal Returns to Medical Care: Evidence from At-risk Newborns. *The Quarterly Journal of Economics 125*(2), 591–634.

Almond, D., A. Lee, and A. E. Schwartz (2016, March). Impacts of classifying New York City students as overweight. *Proceedings of the National Academy of Sciences 113*(13), 3488–3491.

American Diabetes Association (2010, January). Diagnosis and Classification of Diabetes Mellitus. *Diabetes Care 33*(Suppl 1), S62–S69.

American Diabetes Association (2017, January). Standards of Medical Care in Diabetes - 2017. *Diabetes Care 40*(1).

American Diabetes Association (2018a, January). 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2018. *Diabetes Care 41*(Supplement 1), S13–S27.

American Diabetes Association (2018b, May). Economic Costs of Diabetes in the U.S. in 2017. *Diabetes Care 41*(5), 917–928.

Angrist, J. D. and J.-S. Pischke (2009). Instrumental Variables in Action: Sometimes You Get What You Need. In *Most Harmless Econometrics: An Empiricist's Companion*, pp. 113–220. Princeton, NJ: Princeton University Press.

Barreca, A. I., M. Guldi, J. M. Lindo, and G. R. Waddell (2011, November). Saving Babies? Revisiting the effect of very low birth weight classification. *The Quarterly Journal of Economics 126*(4), 2117–2123.

Barreca, A. I., J. M. Lindo, and G. R. Waddell (2016, January). Heaping-induced Bias in Regression-discontinuity designs. *Economic Inquiry 54*(1), 268–293.

Berchick, E. R., J. C. Barnett, and R. D. Upton (2019, November). Health Insurance Coverage in the United States: 2018. Current Population Reports P60-267(RV), U.S. Government Printing Office, Washington, DC.

Bharadwaj, P., K. V. Løken, and C. Neilson (2013, August). Early Life Health Interventions and Academic Achievement. *American Economic Review 103*(5), 1862–1891.

Centers for Disease Control and Prevention (2017). National Diabetes Statistics Report. Technical report, Department of Health & Human Services.

Centers for Disease Control and Prevention (2018, August). All About Your A1C.

Cholesterol Treatment Trialists' (CTT) Collaborators (2012, August). The effects of lowering LDL cholesterol with statin therapy in people at low risk of vascular disease: meta-analysis of individual data from 27 randomised trials. *The Lancet 380*(9841), 581–590. Publisher: Elsevier.

Chyn, E., S. Gold, and J. S. Hastings (2019, April). The Returns to Early-life Interventions for Very Low Birth Weight Children. Working Paper 25753, National Bureau of Economic Research.

Coussens, S. (2018, August). Heuristic Thinking in the Emergency Department. *Columbia School of Public Health Working Paper*.

Currie, J. and D. Slusky (2020, January). Does the Marginal Hospitalization Save Lives? The Case of Respiratory Admissions for the Elderly. Technical Report w26618, National Bureau of Economic Research, Cambridge, MA.

Dahlberg, M., K. Mani, M. Öhman, and A. Wanhainen (2016, January). Health Information and Well-Being: Evidence from an Asymptomatic Disease. Technical Report 2016:2, Uppsala University, Department of Economics.

Doyle, J. J. (2011, July). Returns to Local-Area Healthcare Spending: Evidence from Health Shocks to Patients Far From Home. *American economic journal. Applied economics 3*(3), 221–243.

Doyle, J. J., J. A. Graves, J. Gruber, and S. A. Kleiner (2015, February). Measuring Returns to Hospital Care: Evidence from Ambulance Referral Patterns. *Journal of Political Economy 123*(1), 170–214.

Edwards, R. D. (2018, April). If My Blood Pressure Is High, Do I Take It to Heart? Behavioral Effects of Biomarker Collection in the Health and Retirement Study. *Demography 55*(2), 403–434.

Finkelstein, A., S. Taubman, B. Wright, M. Bernstein, J. Gruber, J. P. Newhouse, H. Allen, and K. Baicker (2012, August). The Oregon Health Insurance Experiment: Evidence from the First Year. *The Quarterly Journal of Economics 127*(3), 1057–1106.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas, and É. L. Pinder (2003a, February). The Implications of Regional Variations in Medicare Spending. Part 1: The Content, Quality, and Accessibility of Care. *Annals of Internal Medicine 138*(4), 273.

Fisher, E. S., D. E. Wennberg, T. A. Stukel, D. J. Gottlieb, F. L. Lucas, and É. L. Pinder (2003b, February). The Implications of Regional Variations in Medicare Spending. Part 2: Health Outcomes and Satisfaction with Care. *Annals of Internal Medicine 138*(4), 288.

Frakes, M. (2013, February). The Impact of Medical Liability Standards on Regional Variations in Physician Behavior: Evidence from the Adoption of National-Standard Rules. *American Economic Review 103*(1), 257–276.

Fuji, D., G. Imbens, and K. Kalyanaraman (2009, September). Notes for Matlab and Stata Regression Discontinuity Software.

Fukuma, S., T. Iizuka, T. Ikenoue, and Y. Tsugawa (2020, December). Association of the National Health Guidance Intervention for Obesity and Cardiovascular Risks With Health Outcomes Among Japanese Men. *JAMA Internal Medicine 180*(12), 1630.

Garber, A. J., T. G. Duncan, A. M. Goodman, D. J. Mills, and J. L. Rohlf (1997). Efficacy of Metformin in Type II Diabetes: Results of a Double-Blind, Placebo-controlled, Dose-Response Trial. *The American Journal of Medicine 103*, 491–497.

Gerteis, J., D. Izrael, D. Deitz, L. LeRoy, R. Ricciardi, T. Miller, and J. Basu (2014, April). Multiple Chronic Conditions Chartbook. Technical Report AHRQ Publications No, Q14-0038, Agency for Healthcare Research and Quality, Rockville, MD.

Greene, J. A. (2007). *Prescribing by Numbers: Drugs and the Definition of Disease*. Baltimore, MD: Johns Hopkins University Press.

Iizuka, T., K. Nishiyama, B. Chen, and K. Eggleston (2021, March). False alarm? Estimating the marginal value of health signals. *Journal of Public Economics 195*, 104368.

Inzucchi, S. and B. Lupsa (2020). Glycemic control and vascular complications in type 2 diabetes mellitus. In D. Nathan and J. Mulder (Eds.), *UptoDate*.

Inzucchi, S. E., B. Lupsa, and J. E. Mulder (2021). Clinical presentation, diagnosis, and initial evaluation of diabetes mellitus in adults. In *UpToDate*.

Jutel, A. G. (2011). Driving Diagnosis: Peddlers and Pushers. In *Putting a Name to It: Diagnosis in Contemporary Society*. Baltimore, MD: Johns Hopkins University Press.

Kim, H. B., S. A. Lee, and W. Lim (2019, May). Knowing is not half the battle: Impacts of information from the National Health Screening Program in Korea. *Journal of Health Economics 65*, 1–14.

Kim, J. M. and D.-J. Kim (2015). The Optimal Cutoff Value of Glycated Hemoglobin for Detection of Diabetic Retinopathy. *Diabetes & Metabolism Journal 39*(1), 16.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental Analysis of Neighborhood Effects. *Econometrica 75*(1), 83–119.

Kolesár, M. and C. Rothe (2018, August). Inference in Regression Discontinuity Designs with a Discrete Running Variable. *American Economic Review 108*(8), 2277–2304.

Levitsky, L., M. Misra, and A. G. Hoppin (2021). Epidemiology, presentation, and diagnosis of type 1 diabetes mellitus in children and adolescents. In *UpToDate*.

Mayo Foundation for Medical Education and Research (2021, January). A1c Test.

Myerson, R. M., D. Lakdawalla, L. D. Colantonio, M. Safford, and D. Meltzer (2018, February). Effects of Expanding Health Screening on Treatment - What Should We Expect? What Can We Learn?

Nakagami, T., K. Takahashi, C. Suto, J. Oya, Y. Tanaka, M. Kurita, C. Isago, Y. Hasegawa, A. Ito, and Y. Uchigata (2017, February). Diabetes diagnostic thresholds of the glycated hemoglobin A1c and fasting plasma glucose levels considering the 5-year incidence of retinopathy. *Diabetes Research and Clinical Practice 124*, 20–29.

Oster, E. (2018, October). Diabetes and Diet: Purchasing Behavior Change in Response to Health Information. *American Economic Journal: Applied Economics 10*(4), 308–348.

Papanicolas, I., L. R. Woskie, and A. K. Jha (2018, March). Health Care Spending in the United States and Other High-Income Countries. *JAMA 319*(10), 1024.

Phillipov, G. and P. J. Phillips (2001, October). Components of Total Measurement Error for Hemoglobin A1c Determination. *Clinical Chemistry 47*(10), 1851–1853.

Ponssen, H. H., J. W. F. Elte, P. Lehert, J. P. Schouten, and D. Bets (2000). Combined Metformin and Insulin Therapy for Patients with vpe 2 Diabetes Mellitus. *Clinical Therapeutics 22*(6), 709–718.

The International Expert Committee (2009, July). International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes. *Diabetes Care 32*(7), 1327–1334.

Tsugawa, Y., O. Takahashi, J. B. Meigs, R. B. Davis, F. Imamura, T. Fukui, W. C. Taylor, and C. C. Wee (2012, December). New Diabetes Diagnostic Threshold of Hemoglobin A1c and the 3-Year Incidence of Retinopathy. *Diabetes 61*(12), 3280–3284.

Vandorsten, J. P., W. C. Dodson, M. A. Espeland, W. A. Grobman, J. M. Guise, B. M. Mercer, H. L. Minkoff, B. Poindexter, L. A. Prosser, G. F. Sawaya, J. R. Scott, R. M. Silver, L. Smith, A. Thomas, and A. T. N. Tita (2013, March). NIH consensus development conference: diagnosing gestational diabetes mellitus. *NIH consensus and state-of-the-science statements 29*(1), 1–31.

Welch, H. G., L. M. Schwartz, and S. Woloshin (2011). *Overdiagnosed: Making People Sick in the Pursuit of Health*. Boston, MA: Beacon Press.

Wexler, D. (2021). Overview of general medical care in nonpregnant adults with diabetes mellitus. In D. Nathan and J. Mulder (Eds.), *UptoDate*.

Zhao, M., Y. Konishi, and P. Glewwe (2013, March). Does information on health status lead to a healthier lifestyle? Evidence from China on the effect of hypertension diagnosis on food consumption. *Journal of Health Economics 32*(2), 367–385.

What Difference Does a Diagnosis Make? Evidence from Marginal Patients

Appendix

A Health Risk Assessment Questions by Category

Alcohol and Tobacco Use

- How many alcoholic drinks do you consume per week?
- Do you currently use cigarettes?
- Do you currently use cigars or pipes?
- Do you currently use smokeless tobacco?

Diet and Exercise

- How many servings of high-fat foods do you consume per day?
- During a typical week, how many days do you do physical activity outside your job?
- Have you made changes recently to manage your weight?
- Have you made an effort recently to change physical activity?

Health Status

- How would you rate your health?

Physical and Emotional Well-Being

- In general, how often is stress a problem for you?
- In the past year, how much effect has stress had on your health
- In the past year, how many days have you missed an entire workday due to physical or mental health problems?
- In the past two weeks, how often did physical/emotional problems make it difficult to work the required hours?
- In the past two weeks, how often were you able to repeat the same motions working without difficulty cause by physical health
- In the past two weeks, how often did physical/emotional problems make it difficult to concentrate on work?
- In the past two weeks, how often did physical/emotional problems make it difficult to handle your workload?

- In the past two weeks, how often did physical/emotional problems make it difficult to help people get work done?
- In the past two weeks, how often did physical/emotional problems make it difficult to feel capable?
- In the past four weeks, how much did your health problems affect your ability to do your regular daily activities?

B Construction of Analytic Samples

1 CDM Data

We identify each patient's initial A1c lab test in the CDM lab file. The lab file contains test descriptions and result values for lab tests provided by CDM affiliated laboratories. Patients who never receive A1c tests are not included in the analytic sample. We identified A1c tests as those with LOINC codes 4548-4, 27353-2, 17856-6, and 17855-8. We used the test description text to drop remaining inappropriate tests. We include only patients who were 18 or older at the time of the initial test. We do not include patients whose first A1c test occurred before 2009.

We drop patients who have any inpatient, outpatient, or drug spending related to diabetes prior to the date of the first observed lab test (see Appendix Table A20. We treat inpatient and outpatient spending as related to diabetes if any of the associated ICD-9 diagnosis codes begin with "249," or "250," or if the associated ICD-10 codes begin with "E08," "E09," "E10," "E11," "E13." Individuals with diabetes related lab tests (e.g. fasting plasma glucose, and random glucose tests) prior to the date of the first A1c test are also excluded. We also drop patients with pregnancy related diagnosis codes within 365 days before or after the initial A1c test so that we can focus on standard diabetes diagnoses separately from gestational diabetes. We treat drug spending as related to diabetes if the prescription's National Drug Code (NDC) matches any of those in Appendix Table A21. We only observe the quarter in which a drug was prescribed, and we treat the first date of the quarter as the time of prescription.

We impose restrictions on how long patients must be present in the CDM data to ensure that we are identifying each patient's first A1c test. Without requiring that patients be present in the data prior to the date of the first test, we would likely mis-classify individuals with previously diagnosed diabetes (diagnosed and treated prior to their first CDM claim). The CDM files provide records of insurance enrollment dates for each patient. We drop patients from our analysis if they were not enrolled in an CDM-covered insurance plan for at least 366 days prior to and following the date of the initial lab test. Many patients changed insurance plans during the sample period. We included these patients as long as they were covered under some CDM plan for the full 732 day period around the initial lab test, no matter how many times they switched plans. Some patients had brief breaks in coverage in the dates before and after the initial lab test. We dropped patients who either had multiple breaks in coverage or had at least one break of more than 60 days.

We identify providers who ordered a patient's first A1c test by linking the first A1c test in the lab file to claims in the outpatient file. We link claims by date, patient, and procedure code. We consider a claim linked to the lab file for the same patient if the first date on the claim is within the seven days prior to the first date on the lab, and the procedure code is either 83036 or 83021 (hemoglobin A1c), or if

the type of service code is "office visit." When multiple claims link to a lab by this approach, we choose one claim to identify the provider using the following ranking: 1) claims where the procedure code is either 83036 or 83021 and the provider category is not in 0005, 0054, 0055, 0057, 1211, 1294, or 4834 (independent laboratory). 2) claims where the type of service code is "office visit," and the provider category is not in 0005, 0054, 0055, 0057, 1211, 1294, or 4834. 3) claims where the provider category is in 0005, 0054, 0055, 0057, 1211, 1294, or 4834. If multiple claims meet one of these conditions, we select the claim with the first date closest to the first date of the A1c test in the lab file. If multiple claims meet one of these conditions and share the same first date, we select a claim at random.

For each provider linked to a patient in our analytic sample, we calculate the fraction of patients who had a claim with that provider for whom the provider ordered an A1c test. To calculate this measure, we use claims from the outpatient file. Claims with a procedure code of 83036 or 83021 are included as A1c tests. We determine the number of patients for whom the provider ordered at least one A1c test in the six months starting from the first date the provider ordered an A1c test for a patient in our analytic file. We divide the number of patients with an A1c test by the number of patients who had any claim from the provider during that same six-month period.

C Heterogeneity Across Subgroups based on Similarity to Never Takers

The group of "never takers"–that is, the population who falls above the cutoff but does not receive a diagnosis of diabetes as the guideline recommends–may be of special interest to policymakers. While this group is clinically recommended to receive interventions related to diabetes, we do not observe evidence of this care in the claims data. If the foregone care would be highly effective for this group, policymakers may wish to design interventions that would improve take up for the never takers.

It is not possible to directly learn about the effects of a diabetes diagnosis on this group since they do not receive one. However, we can ask if populations that look similar to this group on observable characteristics and *do* receive a diagnosis of diabetes experience positive effects. We identify a subpopulation that resembles the never taker group in the following way. First, we estimate a propensity score using a logit model where the dependent variable equals 1 if the patient has an initial A1c test value above 6.5 but for whom no diagnosis of diabetes is observed in the first year. The predictors we include in this model are patient age, age squared, sex, spending on endocrinology care, ACE inhibitors, and recommended care in the year prior to the test, and whether the patient was enrolled in a consumer-directed health plan in the year before the test. Then, we use the estimates of that model to split the sample into those with below- and above-median estimated propensity scores. Those with below-median propensity scores are less similar to never takers based on observable characteristics than others in the sample, and those with above-median propensity scores are more similar to never takers. Finally, we conduct a subgroup analysis that estimates the effects of a diabetes diagnosis in these two groups.

The results are reported in Appendix Table A22. The group that is similar to the never takers experiences a smaller first stage and a smaller increase in spending at the cutoff compared to the group that is dissimilar. However, they experience larger improvements in A1c. This is suggestive that the never takers, who are by definition "under diagnosed," may experience beneficial effects of their "missed" diagnoses could be realized.

Figure A1: Details on Sample Construction

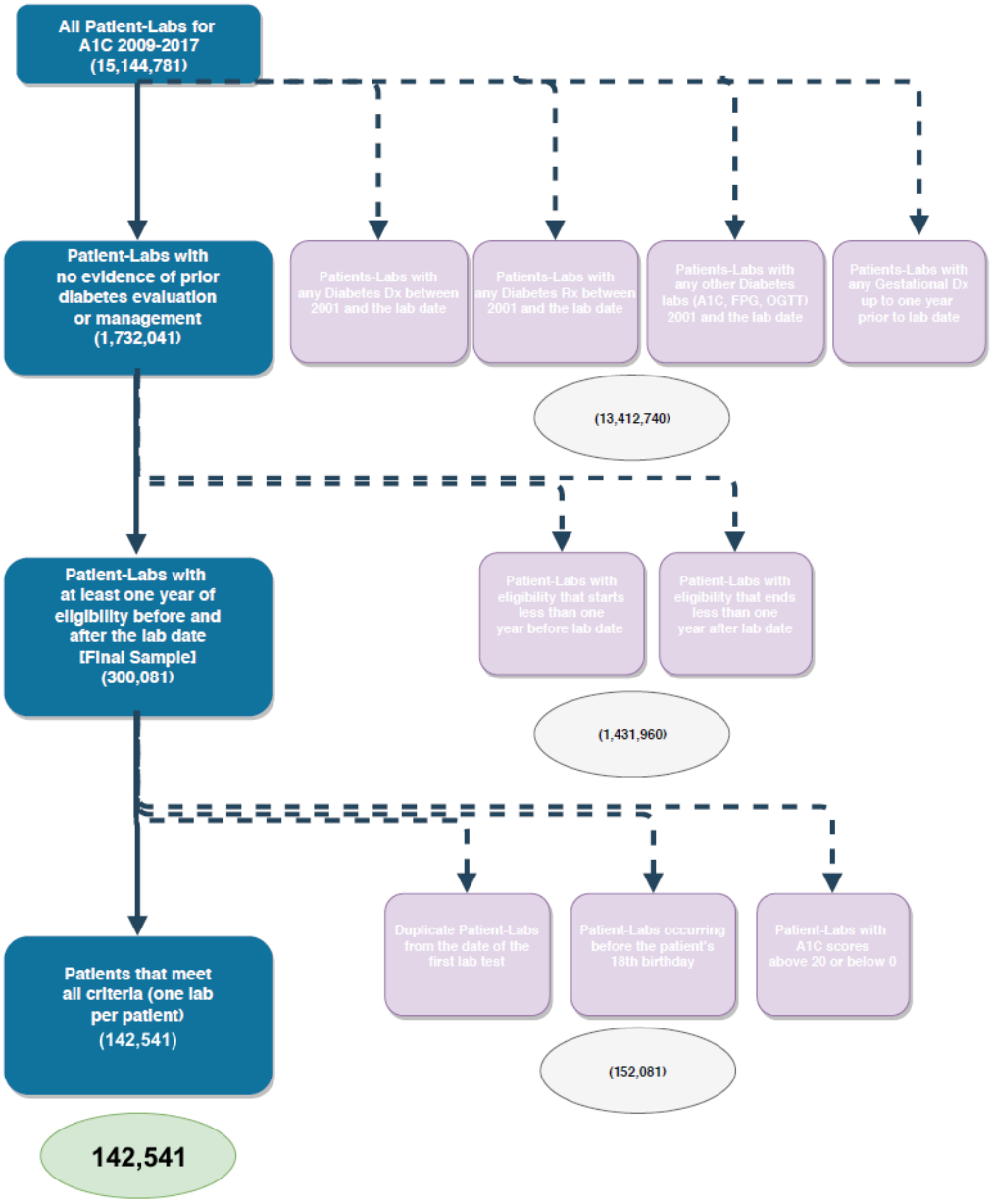
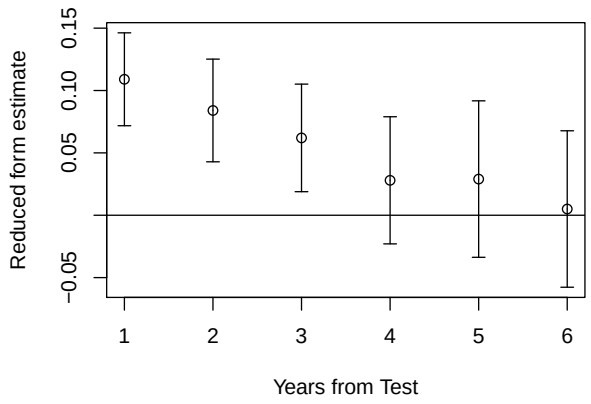
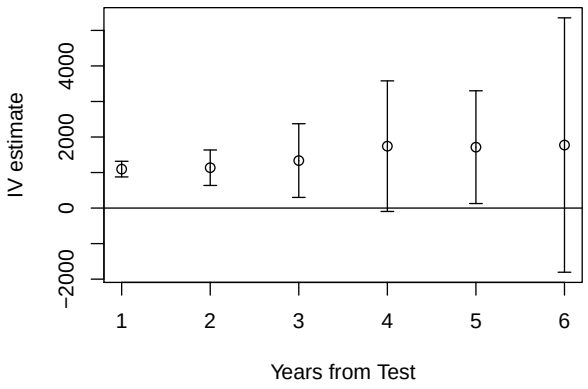


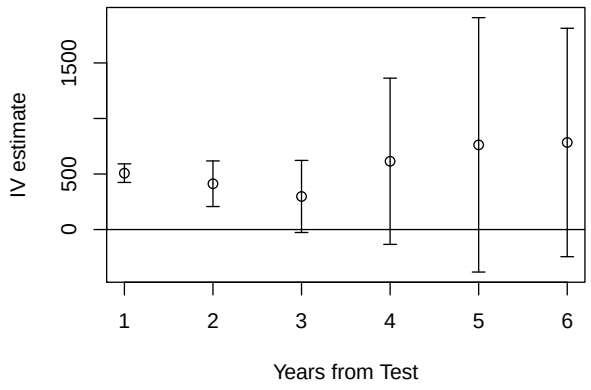
Figure A2: Diagnosis and Spending Outcomes by Year After Initial Lab Test



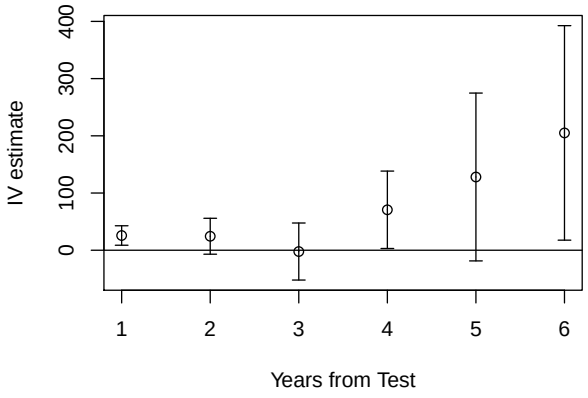
(a) Any Diagnosis of Diabetes



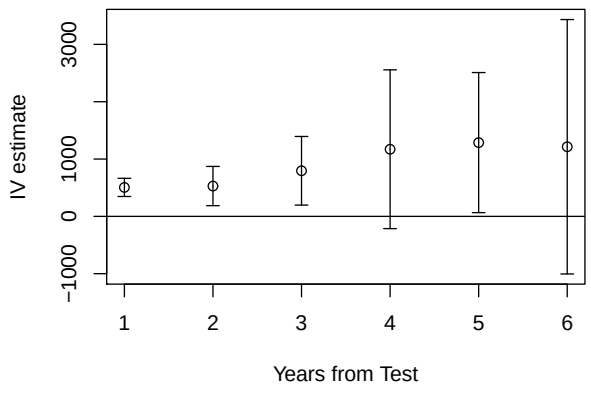
(b) Recommended Care



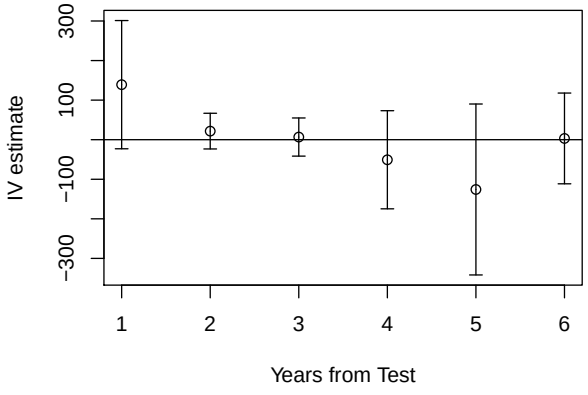
(c) Diabetes-related Office Visits



(d) Diabetes-related Vision Care



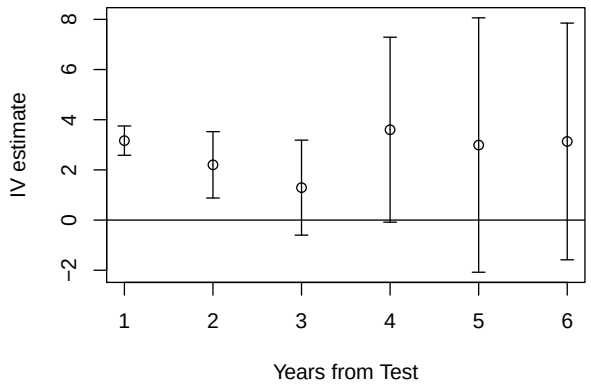
(e) Diabetes-related Labs



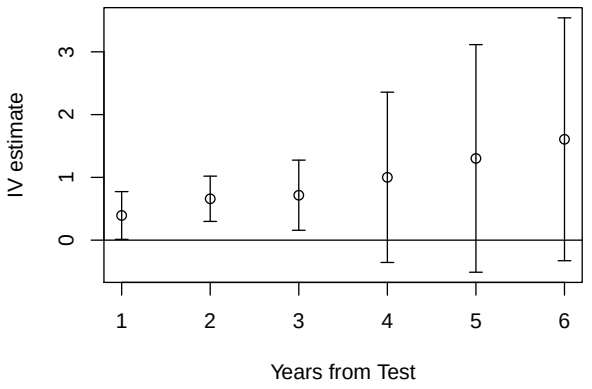
(f) Endocrinology Specialist Care

Source: Authors' calculations from the CDM claims data. Horizontal axis indicates the year relative to the initial lab test. Vertical axis indicates size of the regression discontinuity effect estimated using local linear regression. 95 percent confidence intervals displayed.

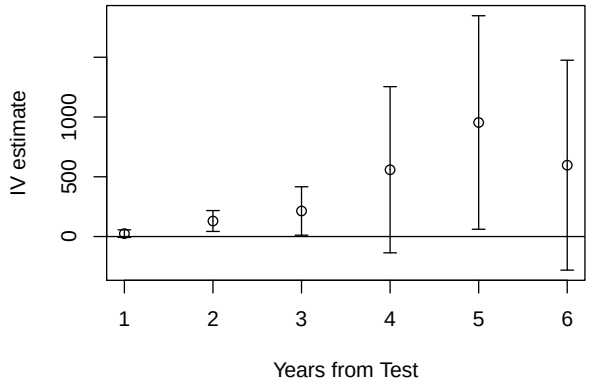
Figure A3: Utilization Outcomes by Year After Initial Lab Test



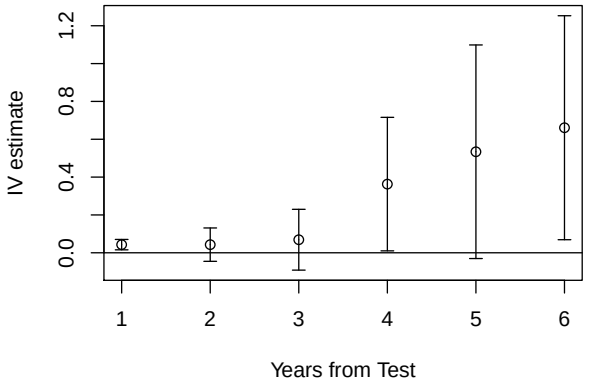
(a) # of Diabetes-Related Office Visits



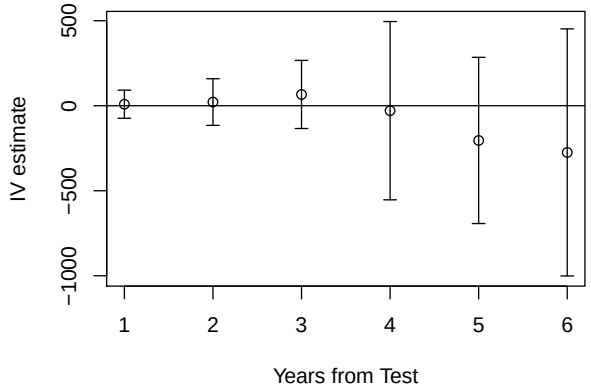
(b) # Glucose or A1c Labs



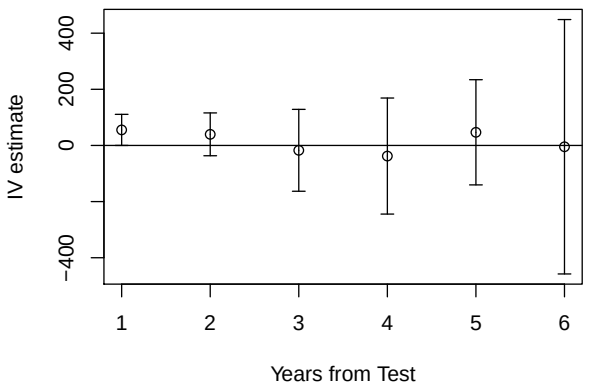
(c) Diabetes-Controlling Drugs



(d) Any Diabetes-Controlling Drug



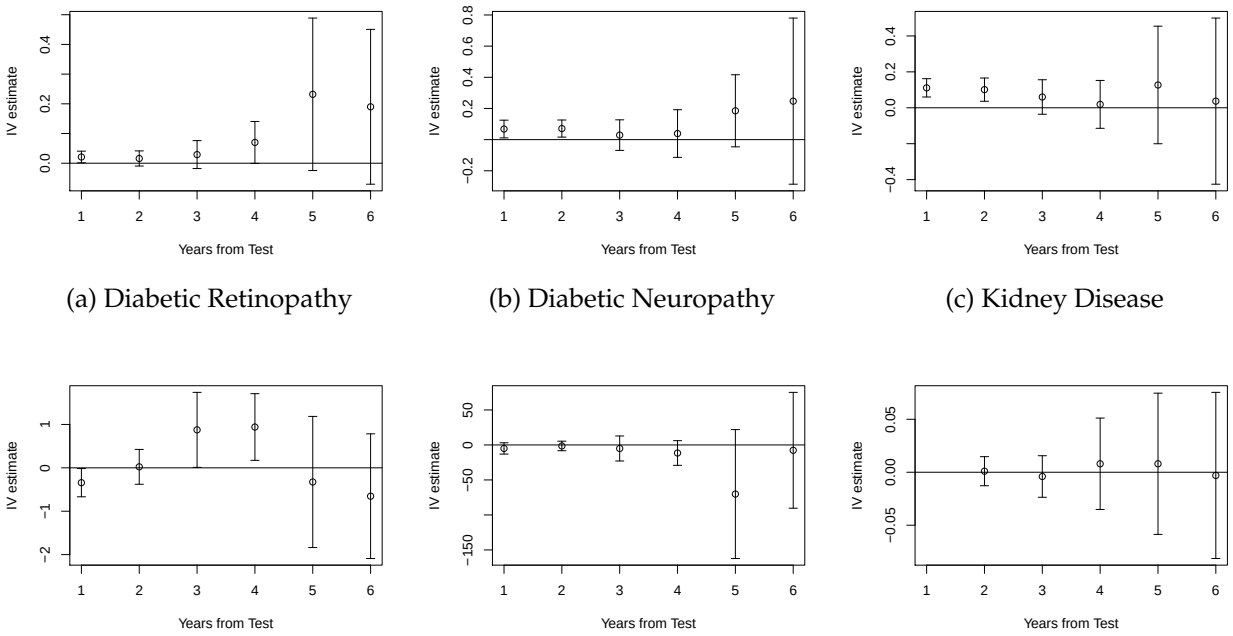
(e) Statins



(f) ACE Inhibitors

Source: Authors' calculations from the CDM claims data. Horizontal axis indicates the year relative to the initial lab test. Vertical axis indicates size of the regression discontinuity effect estimated using local linear regression. 95 percent confidence intervals displayed.

Figure A4: Incidence of Complications by Year After Initial Lab Test

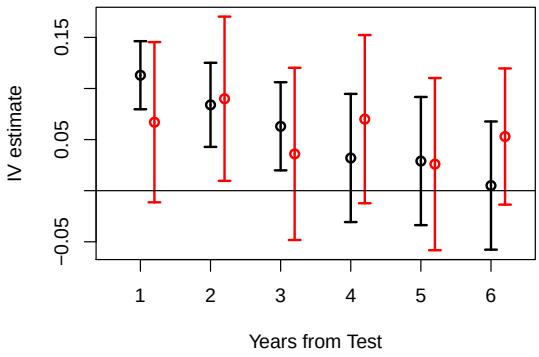


(a) Diabetic Retinopathy (b) Diabetic Neuropathy (c) Kidney Disease

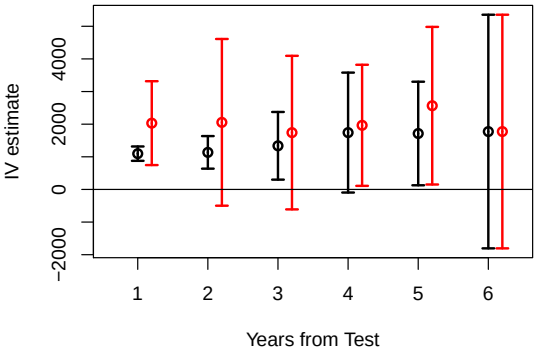
(d) Change in A1c (e) Change in Total Cholesterol (f) Mortality

Source: Authors' calculations from the CDM claims data. Horizontal axis indicates the year relative to the initial lab test. Vertical axis indicates size of the regression discontinuity effect estimated using local linear regression. 95 percent confidence intervals displayed. Note that ketoacidosis and diabetic coma are rare conditions. Because we have smaller samples in later years, we are unable to estimate an effect of diabetes diagnosis on these conditions after five years (for ketoacidosis) or four years (for diabetic coma) following the initial test due to insufficient variation to compute the optimal bandwidth.

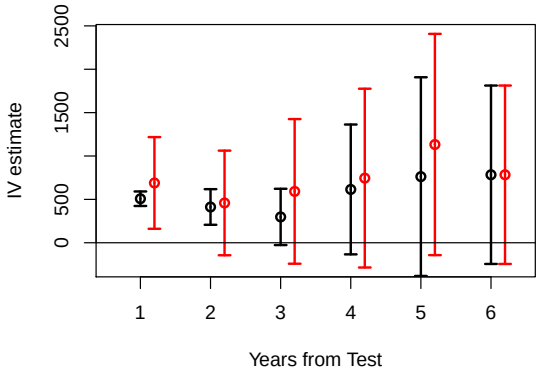
Figure A5: Spending Outcomes Using a Balanced (red) vs. Unbalanced (black) Panel



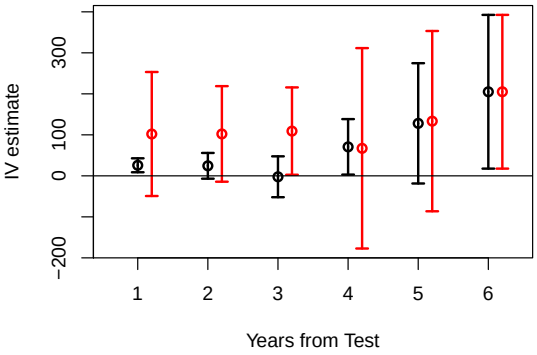
(a) Any Diagnosis of Diabetes



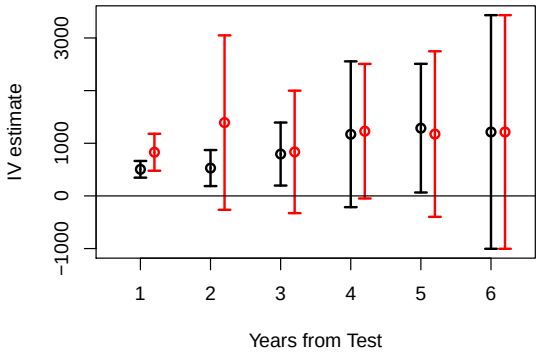
(b) Recommended Care



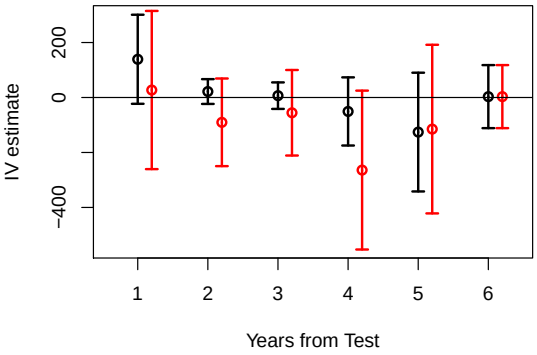
(c) Diabetes-related Office Spending



(d) Diabetes-related Vision Care



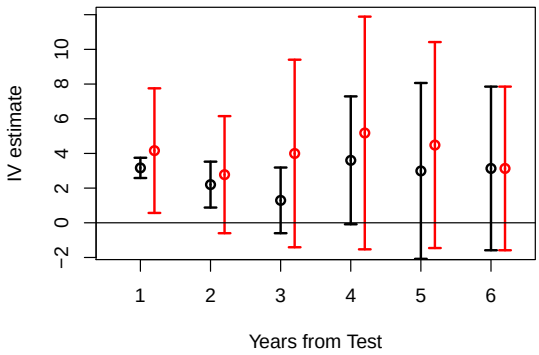
(e) Diabetes-related Labs



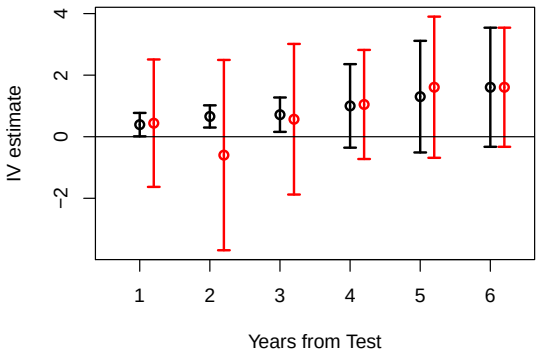
(f) Endocrinology Specialist Care

Source: Authors' calculations from the Health Risk Assessment data. Horizontal axis indicates the year relative to the initial lab test. Vertical axis indicates size of the regression discontinuity effect estimated using local linear regression. 95 percent confidence intervals displayed.

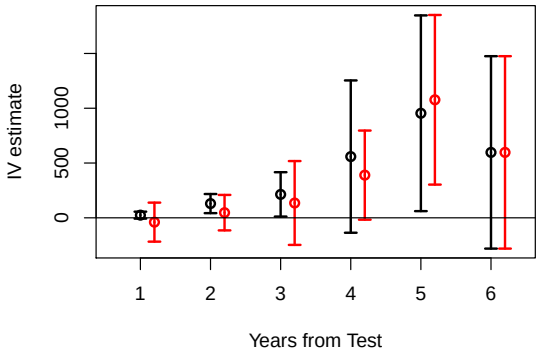
Figure A6: Utilization Outcomes Using a Balanced (red) vs. Unbalanced (black) Panel



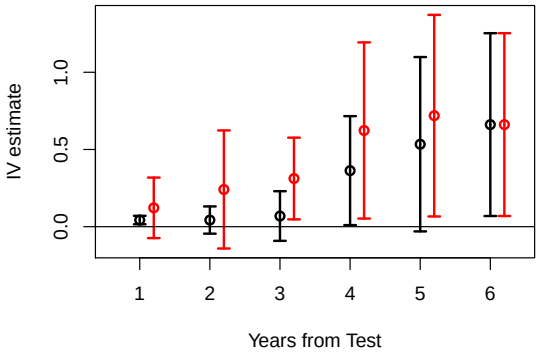
(a) # Diabetes-Related Office Visits



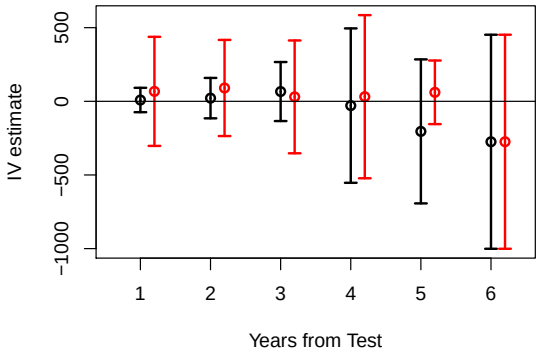
(b) # Glucose/A1c Labs



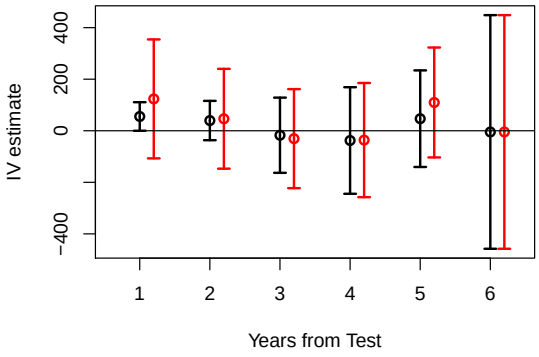
(c) Diabetes-controlling Drug Spending



(d) Any diabetes-controlling drug spending



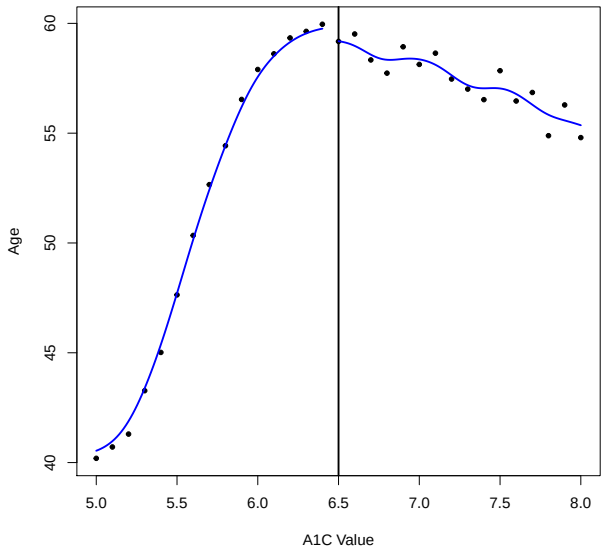
(e) Statins



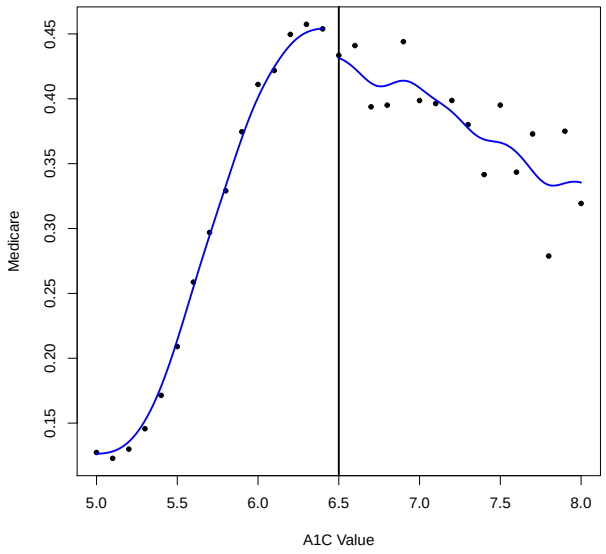
(f) ACE Inhibitors

Source: Authors' calculations from the CDM claims data. Horizontal axis indicates the year relative to the initial lab test. Vertical axis indicates size of the regression discontinuity effect estimated using local linear regression. 95 percent confidence intervals displayed.

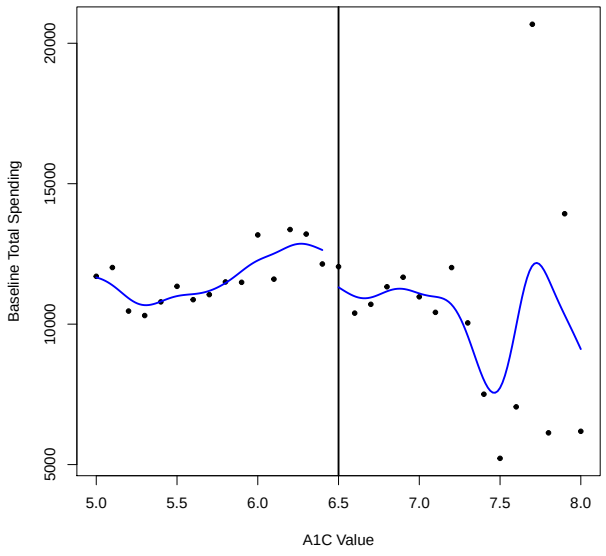
Figure A7: RD Graphs of Observable Baseline Characteristics at the Cutoff



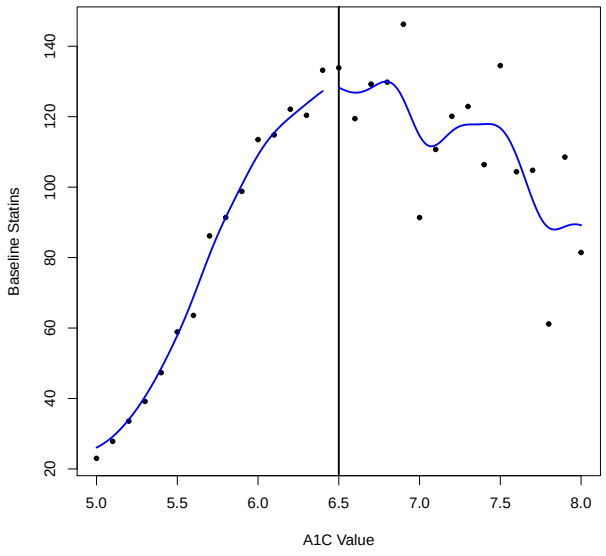
(a) Age



(b) Medicare Advantage



(d) Total Spending



(e) Spending on Statins

Source: Authors' calculations from CDM data. Horizontal axis indicates the year relative to the initial lab test. Vertical axis indicates size of the regression discontinuity effect estimated using local linear regression. 95 percent confidence intervals displayed.

Table A1: RD Estimates of Effects of Diabetes Diagnosis in First 12 Months on Lab Tests During the First Year

	Control Mean (Std Dev)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Number of A1c Tests at CDM-Affiliated Lab	0.436	0.046**	0.205**
		(0.019)	(0.081)
Any A1c Test at CDM-Affiliated Lab	0.326	0.004	0.018
		(0.011)	(0.054)
Number of Cholesterol Tests at CDM-Affiliated Lab	0.509	0.024	0.120
		(0.020)	(0.099)
Any Cholesterol Test at CDM-Affiliated Lab	0.369	0.003	0.019
		(0.011)	(0.070)

Notes: Control means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A2: RD Estimates of Effects of Diabetes Diagnosis in First 12 Months on Probability of Observable A1c Test During Subsequent Years

	Control Mean (Std Dev)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Year 2:			
Any A1c Test at	0.283	0.002	0.010
CDM-Affiliated Lab		(0.011)	(0.052)
Year 3:			
Any A1c Test at	0.247	0.001	0.005
CDM-Affiliated Lab		(0.016)	(0.118)
Year 4:			
Any A1c Test at	0.229	-0.010	-0.096
CDM-Affiliated Lab		(0.020)	(0.199)
Year 5:			
Any A1c Test at	0.240	0.010	0.130
CDM-Affiliated Lab		(0.026)	(0.340)
Year 6:			
Any A1c Test at	0.267	0.040	-0.096
CDM-Affiliated Lab		(0.029)	(0.286)

Notes: Control means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A3: Diagnosis and Procedure Codes Used to Measure Spending Outcomes

Description	ICD 9 Codes	ICD 10 Codes	CPT Codes
Diabetes Related DX	249, 250	E08, E09, E10, E11, E13	N/A
Kidney Disease	2504	E102, E112	N/A
Diabetic Retinopathy	2505	E103, E113	N/A
Diabetic Neuropathy	2506	E104, E114	N/A
Consultations	N/A	N/A	99201-99215, 99241-99245
Nutritional Therapy	N/A	N/A	98960, 98961, 98962, 99078
Monitoring	N/A	N/A	95250, 95251
Self Management Training	N/A	N/A	98960, 98961, 98962, 99078
Recommended Care	N/A	N/A	All Codes in "Monitoring," "Consultations," "Nutritional Therapy," and "Self Management and Training"
A1c or Random Glucose Labs	N/A	N/A	82947, 83036

Table A4: RD Estimates of Effects of Diabetes Diagnosis in First 12 Months Following A1c Test (Health Risk Assessment Results)

	Baseline Mean	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Self-Reported Health			
Health is Very Good or Excellent	0.514	-0.058	-0.363
		(0.041)	(0.276)
Physical and Emotional Well-Being			
Health Interferes with Activities	0.131	0.012	0.135
of Daily Life		(0.039)	(0.446)
Stress Intereferes with Health	0.101	0.005	0.026
Quite a Bit or A Lot		(0.029)	(0.167)
Stress is a problem Sometimes,	0.371	0.022	0.174
Almost Always or Always		(0.050)	(0.396)
In the past 2 weeks, physical/emotional			
problems made it difficult... (some of the time or more)			
...to feel capable	0.147	-0.014	-0.091
		(0.036)	(0.237)
...to handle workload	0.089	-0.008	-0.043
		(0.028)	(0.149)
...to help others with work	0.152	-0.001	-0.003
		(0.035)	(0.157)
...to start work right away	0.069	-0.035	-0.16
		(0.022)	(0.103)
...to concentrate at work	0.18	0.034	0.201
		(0.041)	(0.248)
...to do repetitive tasks	0.151	-0.033	-0.13
		(0.034)	(0.141)
...to work the number of hours	0.076	-0.007	-0.054
required by your job		(0.030)	(0.235)
Number of days missed work due to physical	0.151	0.023	0.118
or mental health problems		(0.036)	(0.192)
Health Behaviors			
Eat 3+ Servings of High Fat Food	0.193	-0.036	-0.138
per Day		(0.031)	(0.120)
Recently made Changes to	0.884	0.019	0.189
Improve Weight		(0.044)	(0.467)
Engage in Physical Activity 3+	0.519	0.044	0.192
times per week		(0.042)	(0.195)
Recently Made Changes to	0.903	0.008	0.051
Increase Physical Activity		(0.037)	(0.250)
Currently Smoke Cigarettes	0.084	0.045*	0.208
		(0.027)	(0.130)
Currently Smoke Cigars	0.018	-0.006	-0.028
		(0.009)	(0.040)
Currently Use Smokeless Tobacco	0.014	-0.003	-0.022
		(0.005)	(0.044)
Number of Alcoholic Drinks Weekly	1.358	0.022	0.091
		(0.181)	(0.758)

Notes: Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A5: Coding All Individuals with A1c of 6.5% or Higher As Receiving a Diagnosis

	IV Estimate of Effect of Diagnosis
Spending on ADA Recommended Care	242.314***
	(30.471)
Spending on Office-Based Care Related to Diabetes	134.602***
	(9.796)
Spending on Diabetes-Related Labs	126.778***
	(21.583)
Spending on Diabetes-Related Vision Services	5.437***
	(1.906)
Spending on Diabetes-Controlling Drugs	7.574*
	(4.352)
Spending on Statins	0.826
	(11.404)
Spending on ACE Inhibitors	19.058*
	(9.950)
Endocrinology Specialist Care Spending	54.409*
	(31.843)
Change in A1c from Previous Year	-0.135**
	(0.067)
Mortality^	0.001
	(0.002)

Notes: Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). ^Note that mortality is measured 2 years after the initial test as we require individuals to be present in the data and alive for the first 12 months after the initial test. Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A6: Selection on Observable Characteristics at the Cutoff

	Female	Age	Medicare	Consumer-Driven Health Plan	
Medical claims data	-0.019	-0.928*	-0.083***	0.007	
	(0.010)	(0.556)	(0.011)	(0.007)	
Health Risk Assessment	0.029	-1.185*	-0.001	-0.106	
	(0.026)	(0.637)	(0.005)	(0.067)	
Spending in Year Before Test:	Total spending	Recommended Care	Statins	ACE Inhibitors	Endocrinology Specialty Care
Medical claims data	-1,808.679**	-13.21	-14.173*	4.791	1.545
	(808.22)	(10.338)	(7.902)	(5.787)	(1.916)
Health Risk Assessment	251.64	-17.542	-42.771*	26.489	-3.299
	(1,578.118)	(30.655)	(23.423)	(18.638)	(3.240)

Notes: Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A7: RD Estimates of Effects of Diabetes Diagnosis in First 12 Months Following A1c Test, Controlling for Medicare Advantage Enrollment and Baseline Spending

	Control Mean (Std Dev)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Diabetes Diagnosis	0.183	0.109***	
		(0.019)	
Spending on ADA Recommended Care	362.03 (718.09)	196.928***	1,025.114***
		(22.797)	(108.940)
Spending on Office-Based Care Related to Diabetes	26.17 (122.28)	79.876***	505.539***
		(6.910)	(42.050)
Spending on Diabetes-Related Labs	89.10 (353.69)	79.354***	502.903***
		(13.914)	(79.820)
Spending on Diabetes-Related Vision Services	2.26 (27.92)	4.595***	27.031***
		(1.469)	(8.559)
Endocrinology Specialist Care Spending	19.95 (313.62)	37.613	133.221
		(22.987)	(81.415)
Number of Diabetes-Related Office Visits	0.177 (0.775)	0.456***	3.160***
		(0.045)	(0.294)
Number of A1c/Glucose Labs	1.46 (1.01)	0.064**	0.397**
		(0.032)	(0.192)
Spending on Diabetes-Controlling Drugs	4.37 (97.04)	4.843	22.096
		(3.503)	(15.966)
Any Spending on Diabetes-Controlling Drugs	0.005	0.009***	0.041***
		(0.003)	(0.014)
Spending on Statins	151.21 (436.84)	-7.354	-31.242
		(9.649)	(41.094)
Spending on ACE Inhibitors	96.00 (299.36)	7.996	31.717
		(6.966)	(27.601)

Notes: Control means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A8: RD Estimates of Effects of Diabetes Diagnosis in First 12 Months Following A1c Test, Controlling for Medicare Advantage Enrollment And Baseline Spending

	Control Mean (Std Dev)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Diabetic Retinopathy	0.0018	0.005**	0.023**
		(0.002)	(0.010)
Diabetic Neuropathy	0.007	0.011**	0.075***
		(0.004)	(0.029)
Kidney Disease	0.0063	0.019***	0.118***
		(0.004)	(0.026)
Change in A1c from Previous Year	-0.056 (0.460)	-0.054**	-0.293**
		(0.025)	(0.141)
Change in Total Cholesterol from Previous Year	-6.429 (32.62)	-0.928	-3.169
		(1.208)	(4.123)
Mortality	0.005	0.001	0.004
		(0.001)	(0.007)

Notes: Control means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). ^Note that mortality is measured 2 years after the initial test as we require individuals to be present in the data and alive for the first 12 months after the initial test. Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A9: RD Estimates for Sample Attrition

	Control Mean	Reduced Form Estimate
Year 2	0.754	-0.017
		(0.012)
Year 3	0.484	0.016*
		(0.0097)
Year 4	0.300	0.003
		(0.010)
Year 5	0.200	-0.005
		(0.009)
Year 6	0.123	-0.001
		(0.008)

Notes: Control means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A10: RD Estimates of Effects of Diabetes Diagnosis on Having Some Spending in All Categories of Recommended Care

	Control Mean (Std Dev)	N in Bandwidth	Bandwidth (+/-)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Received All Categories of Recommended Care	0.003	42,286	0.688	0.009*** (0.003)	0.057*** (0.018)

Notes: Control means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A11: Control, Complier, Always Taker and Never Take Characteristics

	Control Mean	Compliers	Always Takers	Never Takers
Female	0.528	0.489	0.507	0.497
Age $<$ 40	0.130	0.181	0.096	0.123
Consumer-Driven Health Plan	0.126	0.080	0.115	0.130
Any Statin Use Prior to Test	0.255	0.446	0.287	0.260
Any Endocrinology care Prior to Test	0.015	0.014	0.015	0.012

Note: Authors' calculations from CDM claims data. This table contains descriptive statistics for all patients with an A1c test with values placing them just below the diagnosis threshold (column 1) and for compliers (column 2). In column 3, we report means for those whose A1c places them just below the threshold but who nevertheless receive a diagnosis ("always takers"). In column 4, we report means for those whose A1c places them just above the threshold, but who nevertheless do not receive a diagnosis ("never takers").

Table A12: Alternative models and bandwidths: Spending and Utilization Outcomes

	Main model	Global quadratic model	Optimal Bandwidth x 2	Optimal Bandwidth x 0.5
Diabetes Diagnosis	0.108***	0.292***	0.161***	0.178***
	(0.019)	(0.011)	(0.013)	(0.016)
Spending on ADA Recommended Care	207.609***	270.396***	268.938***	137.849***
	(22.912)	(18.260)	(18.671)	(32.218)
Spending on Office-Based Care Related to Diabetes	79.112***	140.478***	115.549***	52.189***
	(6.912)	(5.615)	(5.566)	(10.033)
Spending on Diabetes-Related Labs	78.650***	139.370***	114.484***	34.774*
	(13.891)	(9.964)	(10.831)	(20.660)
Spending on Diabetes-Related Vision Services	4.245***	9.249***	9.060***	3.214
	(1.489)	(1.279)	(1.196)	(2.157)
Endocrinology Specialist Care Spending	38.741*	30.525*	33.576*	44.18
	(23.024)	(16.665)	(19.218)	(27.185)
Number of Diabetes-Related Office Visits	0.442***	0.899***	0.260***	0.673***
	(0.046)	(0.034)	(0.075)	(0.035)
Number of A1c/Glucose Labs	0.063**	0.102***	0.109***	0.06
	(0.032)	(0.024)	(0.024)	(0.048)
Spending on Diabetes-Controlling Drugs	5.264	-1.01	8.232**	-1.157
	(3.497)	(5.168)	(3.422)	(4.627)
Any Spending on Diabetes-Controlling Drugs	0.010***	0.010***	0.013***	0.005
	(0.003)	(0.003)	(0.003)	(0.004)
Spending on Statins	2.073	-13.013	-0.991	9.522
	(9.799)	(7.967)	(8.209)	(13.486)
Spending on ACE Inhibitors	13.767**	5.975	16.678***	13.999
	(7.020)	(5.825)	(5.949)	(9.447)

Notes: Baseline means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A13: Alternative model and bandwidths: Preventable Conditions and Health Outcomes

	Main model	Global quadratic model	Optimal Bandwidth x 2	Optimal Bandwidth x 0.5
Diabetic Retinopathy	0.004**	0.006***	0.006***	0.004
	(0.002)	(0.002)	(0.002)	(0.003)
Diabetic Neuropathy	0.010**	0.018***	0.017***	0.004
	(0.004)	(0.003)	(0.003)	(0.007)
Kidney Disease	0.018***	0.024***	0.023***	0.017***
	(0.004)	(0.003)	(0.003)	(0.006)
Change in A1c	-0.062**	0.061**	-0.014	-0.105***
from Previous Year	(0.028)	(0.024)	(0.019)	(0.039)
Change in Total Cholesterol	-5.26	-2.006*	3.686	-5.212
from Previous Year	(4.03)	(1.172)	(2.757)	(8.796)
Mortality^	0.000	0.000	0.000	0.000
	(0.001)	(0.0004)	(0.001)	(0.001)

Notes: Baseline means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). ^Note that mortality is measured 2 years after the initial test as we require individuals to be present in the data and alive for the first 12 months after the initial test. Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A14: Alternative model and bandwidths: Health Risk Assessment outcomes

	Main model	Global quadratic model	Optimal Bandwidth x 2	Optimal Bandwidth x 0.5
Self Reported Health	-0.116	-0.048	-0.115*	-0.242**
Very Good or Excellent	(0.083)	(0.090)	(0.067)	(0.121)
Physical and Emotional	-0.077	-0.023	-0.097*	-0.044
Well-Being	(0.058)	(0.067)	(0.050)	(0.073)
Health Behaviors	0.172	0.062*	0.084	0.522
	(0.172)	(0.035)	(0.109)	(0.563)

Notes: Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A15: RD Estimates of Effects of Diabetes Diagnosis in First 12 Months Following A1c Test Including Only Those Enrolled for at least 2 Years Prior to Initial Test

	Baseline Mean (Std Dev)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Diabetes Diagnosis	0.172	0.108***	
		(0.022)	
Spending on ADA Recommended Care	351.46 (706.71)	211.666***	1,050.503***
		(26.851)	(124.805)
Spending on Office-Based Care Related to Diabetes	25.13 (120.05)	70.649***	499.217***
		(9.316)	(63.702)
Spending on Diabetes-Related Labs	80.22 (341.30)	80.121***	449.530***
		(14.223)	(71.981)
Spending on Diabetes-Related Vision Services	2.17 (28.01)	3.953**	16.988**
		(1.550)	(6.599)
Endocrinology Specialist Care Spending	19.65 (207.01)	33.653	111.773
		(28.627)	(95.055)
Number of Diabetes-Related Office Visits	0.170 (0.770)	0.427***	3.223***
		(0.062)	(0.444)
Number of A1c/Glucose Labs	1.48 (0.996)	0.054	0.28
		(0.036)	(0.179)
Spending on Diabetes-Controlling Drugs	4.51 (99.59)	2.644	11.177
		(4.371)	(18.462)
Any Spending on Diabetes-Controlling Drugs	0.005 (0.072)	0.006	0.027
		(0.004)	(0.018)
Spending on Statins	151.7 (436.06)	21.72	117.688
		(14.213)	(77.162)
Spending on ACE Inhibitors	94.79 (296.56)	15.313*	54.366*
		(8.003)	(28.455)

Notes: Baseline means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A16: RD Estimates of Effects of Diabetes Diagnosis in First 12 Months Following A1c Test Including Only Those Enrolled for at least 2 Years Prior to Initial Test

	Baseline Mean (Std Dev)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Diabetic Retinopathy	0.002	0.004	0.021
		(0.003)	(0.013)
Diabetic Neuropathy	0.006	0.016***	0.097***
		(0.005)	(0.030)
Kidney Disease	0.007	0.022***	0.121***
		(0.005)	(0.029)
Change in A1c from Previous Year	-0.08 (0.441)	-0.085**	-0.374**
		(0.040)	(0.186)
Change in Total Cholesterol from Previous Year	-6.691 (32.47)	-1.915	-5.663
		(1.478)	(4.378)
Mortality^	0.001	0.001	0.005
		(0.001)	(0.007)

Notes: Baseline means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). ^Note that mortality is measured 2 years after the initial test as we require individuals to be present in the data and alive for the first 12 months after the initial test. Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A17: RD Estimates of Effects of Diabetes Diagnosis in First 12 Months Following A1c Test Including Only Those Enrolled for at least 2 Years Prior to Initial Test

	Control Mean (Std Dev)	Reduced Form Estimate	IV Estimate of Effect of Diagnosis
Diabetes Diagnosis	0.031	0.410***	–
		(0.044)	
BMI	28.49 (5.54)	0.321	0.872
		(0.329)	(0.891)
Normal BMI	0.262	-0.001	-0.003
		(0.029)	(0.070)
Overweight	0.337	0.02	0.049
		(0.035)	(0.088)
Obese	0.363	-0.006	-0.017
		(0.028)	(0.080)
Diastolic blood pressure	76.19 (8.43)	-0.677	-1.969
		(0.762)	(2.189)
Systolic blood pressure	129.67 (14.10)	0.51	1.248
		(1.117)	(2.745)
High blood pressure	0.23	0.009	0.021
		(0.041)	(0.100)

Notes: Baseline means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A18: Spending on Elements of Recommended Care Not Identified with a Diabetes Diagnosis Code

	IV Estimate of Effect of Diagnosis
Spending on All Office Care	22.23
	(84.39)
Number of Evaluation and Management Visits	-0.339
	(0.479)
Spending on Statins, ACE inhibitors, diabetes-controlling drugs, glucose labs and endocrinology specialty care	254.07**
	(102.30)

Notes: Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A19: Heterogeneity in RD Estimates of Effects of Diabetes Diagnosis on Spending

	Primary Care		Specialist	
	Control Mean	Estimate	Control Mean	Estimate
Reduced Form RD Estimate				
Diabetes Diagnosis	0.196	0.121***	0.184	0.193***
		(0.025)		(0.046)
IV Estimate of Effect of Diagnosis				
Spending on ADA-Recommended Care	778.3	1,200.072***	929.7	1,458.434***
		(185.481)		(413.089)
Spending on Office-Based Care Related to Diabetes	76.49	531.150***	44.15	636.954**
		(79.383)		(266.801)
Spending on Diabetes-Related Vision Services	9.668	22.146*	4.770	18.279
		(12.059)		(18.682)
Spending on Diabetes-Related Labs	204	550.138***	216.7	636.954**
		(103.936)		(266.801)
Endocrinology Specialist Care Spending	15.61	89.584	59.63	168.179
		(75.718)		(107.060)
Spending on Diabetes-Controlling Drugs	5.842	8.737*	6.456	-14.025
		(5.229)		(25.799)
Change in A1c	-0.0503	-0.111	-0.06816	-0.651
		(0.223)		(0.504)

Notes: Control means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.

Table A20: Diagnosis and Procedure Codes used in Sample Selection

Description	ICD 9 Codes	ICD 10 Codes	CPT Codes	LOINC Codes
Diabetes Related DX	249, 250	E08, E09, E10, E11, E13	N/A	N/A
Pregnancy Related Diagnosis	630-672, V22-V23	O00-O99	N/A	N/A
Diabetes Lab Tests	N/A	N/A	82947, 83036, 82947, 82952	1558-6, 20435-2, 20437-0, 20438-8, 2345-7, 74774-1, 4548-4, 27353-2, 17856-6, 17855-8

Table A21: Drugs used to Determine Diabetic Drug Spending and for Sample Selection

Diabetes Related Drugs by Generic Name
acarbose
albiglutide
alogliptin
alogliptin-metFORMIN
alogliptin-pioglitazone
canagliflozin
canagliflozin-metFORMIN
chlorproPAMIDE
dapagliflozin
empagliflozin
exenatide
glimepiride
glimepiride-pioglitazone
glimepiride-rosiglitazone
glipiZIDE
glipiZIDE-metFORMIN
glyBURIDE
glyBURIDE micronized
glyBURIDE-metFORMIN
insulin aspart
insulin aspart-insulin
insulin detemir
insulin glargine
insulin glulisine
insulin isophane
insulin lispro
insulin lispro-insulin lispro protamine
insulin regular
linagliptin
linagliptin-metFORMIN
liraglutide
metFORMIN-pioglitazone
metFORMIN-repaglinide
metFORMIN-rosiglitazone
metFORMIN-saxagliptin
metFORMIN-sitaGLIPtin
miglitol
nateglinide
pioglitazone
pramlintide
repaglinide
rosiglitazone
saxagliptin
simvastatin-sitaGLIPtin
sitaGLIPtin
TOLAZamide
TOLBUTamide

Table A22: RD Estimates of Effects of Diabetes Diagnosis on Select Spending and Health Outcomes, by Similarity to Never Takers

		Similar to Never Takers		*Dissimilar to Never Takers*	
	Control Mean (Std Dev)	Reduced Form Estimate	IV Estimate Effect of Diagnosis	Reduced Form Estimate	IV Estimate Effect of Diagnosis
Diabetes Diagnosis	0.183	0.109***	NA	0.129***	NA
		(0.022)		(0.027)	
Spending on ADA-Recommended Care	362.03	151.89***	954.88***	193.759***	1,326.174***
	(718.09)	(37.53)	(200.43)	(34.158)	(227.537)
Change in A1c from Test Year	-0.0564	-0.065**	-0.341**	0.019	0.133
	(0.460)	(0.031)	(0.174)	(0.047)	(0.337)
Change in Total Cholesterol from Test Year	-6.429 (32.62)	-1.342	-5.309	-1.914	
	(1.502)	(1.502)	(5.946)	(2.806)	(10.719)

Notes: Control means calculated using observations with A1c test values below 6.5 and greater than or equal to 6. Coefficient estimates from local linear regression with optimal bandwidth calculation from Fuji, Imbens, and Kalyanaraman (2009). Asterisks denote level of statistical significance: * $>$ 0.10, ** $>$ 0.05, *** $>$ 0.01.