Birkbeck
UNIVERSITY OF LONDON

BIROn - Birkbeck Institutional Research Online

Kalsiki, Daniel (2023) Identifying the impact of health insurance on subgroups with changing rates of diagnosis. Health Economics , ISSN 1099-1050.

Downloaded from: https://eprints.bbk.ac.uk/id/eprint/51429/

Usage Guidelines:
Please refer to usage guidelines at https://eprints.bbk.ac.uk/policies.html or alternatively contact lib-eprints@bbk.ac.uk.

Check for updates

Received: 5 August 2022 | Revised: 18 April 2023 | Accepted: 22 May 2023

DOI: 10.1002/hec.4718

RESEARCH ARTICLE

Health Economics WILEY

Identifying the impact of health insurance on subgroups with changing rates of diagnosis

Daniel Kaliski

Department of Economics, Mathematics and Statistics, Birkbeck, University of London, London, UK

Correspondence
Daniel Kaliski.
Email: d.kaliski@bbk.ac.uk

Abstract

Expanded access to health care often leads to new diagnoses for previously undetected conditions. New diagnoses make it difficult to identify the causal effect of expanding health insurance on individuals with particular diagnoses: the newly diagnosed in the treatment group are likely to differ in unobserved ways from the control group. This paper provides two methods for dealing with this problem depending on the data available to the researcher and diagnosis-specific knowledge. If there is no panel dimension to the data, then the causal effect for the subgroup of interest can be bounded from either above or below depending on the condition in question. If panel data are available, then the newly diagnosed can be identified, and their treated outcomes subtracted from the overall effect of interest. I apply these methods to find that the difference-in-discontinuities estimator underestimates the effect of Medicare prescription drug coverage on the uptake of insulin by first-time users by 20%.

KEYWORDS
diabetes, health insurance, insulin, Medicare, panel data, partial identification

JEL CLASSIFICATION
H51, I12, I13, I18, J14

1 | INTRODUCTION

A central issue in the literature on health insurance is how its impact on recipients varies across individuals. Causal estimates of this impact are obtained by comparing a randomly or quasi-randomly assigned treatment and control group, with one group receiving differential access to health insurance, and then examining heterogeneity in the estimated treatment effect. Identifying how the effect of access varies by subgroup is difficult since membership of that subgroup might itself be affected by better access to medical services. For example, better access to testing improves the rate at which SARS-COV2 infections are detected. If we naively compared the death rate from these infections among insured individuals to that among uninsured individuals, we will be overestimating the effect of access to insurance. This will be because uninsured individuals will have fewer detected cases of SARS-COV2, artificially shrinking the denominator when dividing the number of deaths by the number of cases.

In this paper, I propose methods for bounding or point-identifying these subgroup-specific treatment effects, depending on the data available to the researcher. Bounds can be obtained via imposing monotonicity assumptions in the spirit of Manski and

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2023 The Authors. Health Economics published by John Wiley & Sons Ltd.

Pepper (2000). These amount to assuming that the sign of the unobserved selection bias is known, which is plausible in many contexts. For instance, prior research has found that enrollees in supplemental health insurance for Medicare are healthier than the overall Medicare-eligible population (Fang et al., 2008). Point identification can be obtained by exploiting repeated observations of the same individuals over time. The advantages of this second method are its simplicity and its relatively undemanding data requirements. Many publicly available data sets contain the repeated observations it requires for point identification, including the Medical Expenditures Panel Survey (MEPS) and Health and Retirement Study (HRS). Even when panel data are not available, monotonicity assumptions can produce a lower or upper bound on the treatment effect of interest, which may nonetheless be informative.

I apply the method proposed in this paper to study the effect of Medicare prescription drug coverage under Part D on insulin use. Using a standard difference-in-discontinuities estimator, and ignoring the effect of new diagnoses, I find a 3% point increase in initiation of insulin use among individuals with diabetes when they turn 65 in 2006–2009 relative to those who turn 65 in 1998–2005. Accounting for the increase in diagnoses of diabetes that occurs at age 65 in 2006–2009 (Geruso & Layton, 2020), I find that the true effect among those who already had been diagnosed before age 65 is likely to be at least as large as the point estimate; exploiting panel data to identify the rate of initiation among the newly diagnosed at age 65, I find that the true effect is 0.6% points larger, 20% larger in relative terms. Intuitively, this underestimation of the diagnosis-constant effect occurs because at the beginning of a diagnosis of Type II diabetes, it is rare for a physician to immediately recommend the use of insulin Nathan et al. (2009). Newly diagnosed individuals will therefore mostly add zeroes to the Medicare-eligible group, reducing the sample estimate of insulin use in that group, and leading the estimated treatment effect to be smaller than the diagnosis-constant effect of interest on those who had diabetes prior to qualifying for Medicare. The data are consistent with the theoretical arguments above regarding the lower rate of initiation of insulin use among the newly diagnosed at age 65: the proportion is 0.033, accounting for less than half of total new initiations of insulin use. Overall insulin use in this group at age 65 is also 3.3%, due to the non-use of insulin by undiagnosed individuals; this is around one-eighth the value of the overall average among individuals with diabetes.

This paper contributes to the following literature. First, it extends the literature on estimating Conditional Average Treatment Effects (CATEs) (Abrevaya et al., 2015; Crump et al., 2008; MaCurdy et al., 2011) by considering conditioning on endogenous variables (rather than exogenous covariates and endogenous treatment assignment, as is considered explicitly in e.g., MaCurdy et al. (2011), Nguimkeu et al. (2019) and Liu et al. (2020)) and a treatment effect - the average diagnosis-constant effect - that is not, to my knowledge, explicitly considered in this literature. Second, it contributes to the literature on partial identification (Manski & Pepper, 2000; Molinari, 2020; Tamer, 2010) by using partial identification techniques to bound a treatment effect not typically considered in this literature. Third, it contributes to the literature on using panel data for either bounding or point-identifying causal effects of interest (Athey & Imbens, 2006; Callaway, 2021; Callaway & Li, 2019; Imbens & Wooldridge, 2009), since the second method proposed uses a differences-in-differences-style trend stability assumption: the treated outcomes of those who are observed joining the subgroup of interest between period $t-1$ and t, given that they are treated in t and not in $t-1$, is due to assignment to treatment and not an unobserved time trend that is particular to those joining the subgroup. The most similar papers to this one are Huber and Lafférs (2022), Anderson et al. (2012), Manski and Pepper (2000), and Callaway (2021). This paper is distinguished from these papers in the following ways: The first is its new empirical results on the sensitivity of insulin use to prescription drug coverage. Second, this paper is concerned with subgroup-specific treatment effects when membership in that subgroup is not randomly or quasi-randomly assigned, which is not explicitly discussed in Manski and Pepper (2000) or most of the Conditional Average Treatment Effect literature. An exception is Huber and Lafférs (2022), who develop bounds for the direct and indirect effects of an intervention in the presence of endogenous mediating variables. These are different objects of interest from the one studied in this paper, the diagnosis-specific average treatment effect, which is the average effect of treatment for a particular pair of values for a covariate of interest (the subgroup membership indicator). Moreover, Huber and Lafférs (2022) do not provide analytic bounds as is done in this paper, using instead the implied deviation of the propensity score from its value with exogenous treatment assignment and exogenous covariates to compute numerical bounds on the direct and indirect effects of treatment. Perhaps the closest paper to this one is Anderson et al. (2012); that paper is concerned with selection on their outcome variable, emergency department visits, due to the data being based on visits rather than individuals, and hence the outcome of interest not being observed for individuals who do not visit the emergency department. Similarly to this paper, Anderson et al. (2012) make use of a monotonicity assumption to achieve identification: they assume that all of any observed change in emergency department visits comes from loss of health insurance coverage. This paper differs from Anderson et al. (2012) in four ways: first, by examining selection on a covariate, rather than selection on the outcome variable as in their paper; second, the assumptions used for partial identification in this study differ from the monotonicity assumption in Anderson et al. (2012), which restricts the effect of exposure to treatment on the outcome of interest to have the same sign for every individual. By contrast, this paper's assumptions allow the effect of exposure to treatment to differ in sign across individuals for the outcome but not for its effect on subgroup membership (Monotone Subgroup Selection), while the effect on the average outcome is restricted (Monotone Diagnosis Response); third, by the object

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

of interest considered - the diagnosis-constant treatment effect - which is different from theirs; and fourth, by its use of the panel dimension of data. Lastly, Callaway (2021), similarly to this paper, exploits the availability of panel data, but has a more ambitious goal: identifying the distribution of the unconditional treatment effect. Consequently, their method requires more demanding assumptions than the one used in this paper, which is focused on a particular conditional average treatment effect rather than the distribution of treatment effects. In addition, while Callaway (2021) combines panel data and partial identification to create bounds on the object of interest, this paper uses each separately, and so its first method can be applied even when repeated observations on the same units might not be available.

2 | IDENTIFYING CAUSAL EFFECTS WITH NEW DIAGNOSES

Let $Y_i(1)$, $Y_i(0)$ be the potential outcomes for individual i in the treated and untreated states respectively. Let $D_i = 1$ if individual i is assigned to the treated group and 0 otherwise. Let $X_i = 1$ if i belongs to a subgroup of interest and 0 otherwise. Let $X_i(1)$ denote individual i's subgroup membership indicator in the treated state and $X_i(0)$ the same in the untreated state. The object of interest, assuming the effect of treatment is homogeneous conditional on X, is

$$E[Y_i(1) - Y_i(0)|D_i = 1, X_i(1) = X_i(0) = 1] \tag{1}$$

which is the average effect of treatment on the treated for the subgroup with $X = 1$, for members of the subgroup whose membership does not depend on treatment status. This cannot be directly identified from the data since we observe each individual in at most one of the treated and untreated states.

The assumptions necessary for identifying this subgroup-specific treatment effect may be plausible or implausible depending on the subgroup. Consider the following assumption:

Assumption A1. (Exogenous Subgroup Selection). For every individual i, $X_i(1) = X_i(0)$.

Under this assumption $X_i(1) = 1 \Rightarrow X_i(0) = 1$ and so $E[Y_i(1) - Y_i(0)|D_i = 1, X_i(1) = X_i(0) = 1] = E[Y_i(1) - Y_i(0)|D_i = 1, X_i(1) = 1] = E[Y_i(1)|D_i = 1, X_i(1) = 1] - E[Y_i(0)|D_i = 1, X_i(0) = 1]$, and so conditioning on subgroup membership poses no additional problems for identifying the effect of treatment on the treated if $D_i \perp Y_i$, since then the unobserved quantity $E[Y_i(0)|D_i = 1, X_i(0) = 1]$ can be replaced with its sample analog $E[Y_i(0)|D_i = 0, X_i(0) = 1]$.

For some subgroups, such as age brackets or race, this assumption is plausible. It seems unlikely that assignment to the treatment or control group affects racial self-identification. In other cases, this assumption is less plausible. In the motivating example for this paper, better access to health insurance can increase the frequency with which previously undiagnosed conditions are discovered by health care providers. If the treatment group is assigned to expanded health insurance coverage, and a subgroup with a particular diagnosis is of interest, then Exogenous Group Membership is unlikely to hold. In this paper, I consider instead a weaker assumption, analogous (but not identical) to the Monotone Treatment Selection assumption in Manski and Pepper (2000):

Assumption A2. (Monotone Subgroup Selection). For every individual i, $X_i(1) \geq X_i(0)$.

This assumption is relatively plausible in the motivating example of being assigned to receive health insurance in the treated state. In the case where $X_i = 1$ indicates diagnosis with some underlying latent condition such as diabetes, Monotone Subgroup Selection says that no one who received a diagnosis with worse coverage is undiagnosed when they have better coverage, and undiagnosed cases in the control group are weakly more likely to be diagnosed if assigned to treatment.

Monotone Subgroup Selection is useful for the following reason. If we abandoned Exogenous Subgroup Selection and left the sign of the selection effect unspecified, then we would have that

$$E[Y_i(0)|D_i = 0, X_i(0) = 1] \neq E[Y_i(0)|D_i = 0, X_i(1) = X_i(0) = 1],$$

which would necessitate having to disentangle the untreated outcomes of those whose diagnoses are rescinded when assigned to treatment from those whose diagnoses are made more likely by assignment to treatment. From this it is clear that this assumption is most directly applicable to long-lasting chronic conditions that may go undiagnosed in their early stages, such as Type II diabetes. By contrast, conditions that are either chronic or transitory depending on the individual will violate Monotone Subgroup Selection. For example, iron deficiency may either be more likely to be detected and persist, or be treated and reversed,

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

depending on the underlying cause. A chronic cause such as chronic kidney disease can lead to persistent iron deficiency (Nemeth & Ganz, 2023), whereas dietary insufficiency can be remedied with the appropriate lifestyle changes. Hence greater contact with healthcare professionals can either diagnose chronic iron deficiency with greater regularity or eliminate more cases of treatable iron deficiency. The differing signs of these two effects would violate the assumption of Monotone Subgroup Selection.

To assess whether the subgroup-specific treatment effect is a lower or upper bound for the true subgroup-specific effect, we need to make an additional assumption. This assumption says whether the expected outcome Y is larger or smaller for those individuals who are undiagnosed in the untreated state but diagnosed in the treated state.

Assumption A3. (Monotone Diagnosis Response). $E[Y_i(1)|X_i(1) = X_i(0) = 1] \geq E[Y_i(1)|X_i(1) = 1, X_i(0) = 0]$.

Monotone Diagnosis Response is likely to be more plausible when diagnosis occurs at a similar point in the progression of a chronic illness across individuals. For example, individuals with undiagnosed and uncontrolled diabetes suffer increasingly extreme health consequences the longer they have the disease, making undetected diabetes less likely with the passage of time. (I discuss this case in more detail in the empirical application section, which uses data on individuals with diabetes). Cancer, by contrast, is diagnosed at various stages, depending on the individual and the type of cancer, and so some individuals may be made *less* likely to undergo surgery as a result of their diagnosis (by detecting the cancer early enough to allow for less invasive procedures, such as radiotherapy or chemotherapy), while others will be more likely to undergo surgery. For an example of the latter case, pancreatic cancer, if detected early enough, will typically require surgery (Li et al., 2004). So if the subgroup of interest in the data is "individuals diagnosed with cancer", Monotone Diagnosis Response will not typically hold if the outcome of interest is surgical intervention, as more frequent contact with healthcare practitioners may make surgery either more or less likely depending on the particular cancer in question and the stage at which it is diagnosed.

Monotone Diagnosis Response allows us to bound the subgroup-specific treatment effect in an analogous manner to the way the Monotone Treatment Selection and Monotone Treatment Response assumptions of Manski and Pepper (2000) allow for an upper or lower bound on the average treatment effect more generally. In particular, under Monotone Subgroup Selection and Monotone Diagnosis Response, we have

$$\begin{aligned} &E[Y_i(1)|D_i = 1, X_i(1) = 1] \\ &= E[Y_i(1)|D_i = 1, X_i(1) = X_i(0) = 1]\Pr(X_i(0) = 1|X_i(1) = 1) \\ &+E[Y_i(1)|D_i = 1, X_i(1) = 1, X_i(0) = 0]\Pr(X_i(0) = 0|X_i(1) = 1) \\ &\leq E[Y_i(1)|D_i = 1, X_i(1) = X_i(0) = 1] \end{aligned}$$

where the first equality follows from the Law of Total Probability, and the second equality follows from Monotone Diagnosis Response. Combined with Monotone Subgroup Selection, so that $E[Y_i(0)|D_i = 0, X_i(0) = 1] = E[Y_i(0)|D_i = 0, X_i(1) = X_i(0) = 1]$ (since $X_i(1) \geq X_i(0)$ and $X_i(0) = 1$ together imply $X_i(1) = 1$), the above inequality implies

$$E[Y_i(1) - Y_i(0)|D_i = 1, X_i(1) = X_i(0) = 1] \geq E[Y_i(1)|D_i = 1, X_i(1) = 1] - E[Y_i(0)|D_i = 0, X_i(0) = 1]$$

so that the "diagnosis-constant" subgroup-specific effect of treatment on the treated is at least as large as the sample estimate of the subgroup-specific treatment effect.

If we have a proxy in the data for the average outcome among those who are undiagnosed in the control group, but receive a diagnosis in the treated group, then we can point-identify the diagnosis-constant treatment effect. The following assumption is sufficient:

Assumption A4. Either (i) $E[Y_i|D_i = 1, X_i(1) = X_i(0) = 1] = E[Y_{it}|D_{it} = 1, X_{it}(1) = X_{it-1}(0) = 1]$ or (ii) $E[Y_i|D_i = 1, X_i(1) = 1, X_i(0) = 0] = E[Y_{it}|D_{it} = 1, X_{it}(1) = 1, X_{it-1}(0) = 0]$.

This assumption says that either (i) those who would be in the subgroup of interest regardless of exposure to treatment or (ii) the newly diagnosed, when exposed to the treatment that causes their new diagnosis, are not selected for idiosyncratic time trends. Suppose that the causal relationships in question are stable over time. Then A4 (ii) would be violated if, for example, $E[Y_{it}|D_{it} = 1, X_{it}(1) = 1, X_{it}(0) = 0] \neq E[Y_{it}|D_{it} = 1, X_{it}(1) = 1, X_{it-1}(0) = 0]$. This would be the case if there were other time-varying shocks particular to period t or $t - 1$ that would make the value of the outcome variable in the undiagnosed state ($X = 0$) different in the two periods. Assumption A4 (ii) is likely to hold for new diagnoses of individuals with diabetes, for instance, since undiagnosed individuals are not recommended to use insulin at all, regardless of age, cohort, or time period. A

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

potential threat to this assumption would be if the two time periods $t-1$ and t were sufficiently far apart to make the outcomes of those who change diagnosis over that time period subject to unobserved age trends as well as the fact of new diagnosis. We would not expect this assumption to hold if the gap between $t-1$ and t were 20 years, as the treated outcomes for those who switch status between ages 40 and 60 is likely to be driven partly by the aging process, and so would be a bad proxy for the treated outcomes for hypothetical individuals who are randomly assigned to the treatment group (and as a consequence are diagnosed when they would otherwise not have been at the *same* age).

For example, if we have repeated observations of the same individuals in different time periods, indexed by t, then we can in principle observe what happens to the same individual before and after receiving a diagnosis, and apply A4 (ii). Consider again the decomposition above, via the Law of Total Probability, of the observable quantity $E[Y_i(1)|D_i=1, X_i(1)=1]$ into the unobserved quantity $E[Y_i(1)|D_i=1, X_i(1)=X_i(0)=1]\Pr(X_i(0)=1|X_i(1)=1) + E[Y_i(1)|D_i=1, X_i(1)=1, X_i(0)=0]\Pr(X_i(0)=0|X_i(1)=1)$. If we replace $E[Y_i(1)|D_i=1, X_i(1)=1, X_i(0)=0]$ with $E[Y_{it}(1)|D_{it}=1, X_{it}(1)=1, X_{it-1}(0)=0]$, then we can rearrange the equality to obtain

$$E[Y_i(1)|D_i=1, X_i(1)=X_i(0)=1] = \frac{E[Y_{it}(1)|D_{it}=1, X_{it}(1)=1] - E[Y_{it}(1)|D_{it}=1, X_{it}(1)=1, X_{it-1}(0)=0]\Pr(X_{it-1}(0)=0|X_{it}(1)=1)}{\Pr(X_{it-1}(0)=1|X_{it}(1)=1)},$$

which expresses the single unobservable quantity $E[Y_i(1)|D_i=1, X_i(1)=X_i(0)=1]$ in terms of sample quantities, which together with Monotone Subgroup Selection point-identifies the diagnosis-constant treatment effect, since then

$$\begin{aligned} &E[Y_i(1)-Y_i(0)|D_i=1, X_i(1)=X_i(0)=1] \\ &= E[Y_i(1)|D_i=1, X_i(1)=X_i(0)=1] - E[Y_i(0)|D_i=0, X_i(0)=1] \\ &= \frac{E[Y_{it}(1)|D_{it}=1, X_{it}(1)=1] - E[Y_{it}(1)|D_{it}=1, X_{it}(1)=1, X_{it-1}(0)=0]\Pr(X_{it-1}(0)=0|X_{it}(1)=1)}{\Pr(X_{it-1}(0)=1|X_{it}(1)=1)} \\ &\quad - E[Y_{it}(0)|D_{it}=0, X_{it}(0)=1], \end{aligned}$$

and everything on the right hand side of the second equality is identified from the data.

An alternative to the strategy for point identification above is to both use panel data and make weaker assumptions that produce bounds rather than a point estimate. I do not deal with this case in this paper - instead, the first method produces bounds regardless of whether the researcher has panel data, and the other yields point identification in case the researcher does have access to repeated observations of the same units over time. This increases the range of cases to which this paper's methods are applicable. Callaway (2021) discusses one method for combining the two (repeated observations and partial identification) to obtain bounds on a parameter of interest using panel data.

The following subsection applies the strategy above to identify the causal effect of prescription drug coverage on insulin use.

3 | APPLICATION: THE IMPACT OF MEDICARE PRESCRIPTION DRUG COVERAGE ON INSULIN USE

3.1 | Background, empirical strategy and data

Diabetes is a disorder where the cells of the body do not respond to insulin (insulin resistance), and/or the pancreas produces insufficient insulin to reduce blood sugar levels in the body. Since insulin decreases blood sugar levels, inability to absorb insulin results in both higher levels of blood sugar and higher volatility of blood sugar levels, both of which are corrosive to the blood vessels within the human body.

Injecting insulin is not typically recommended in the early stages of a new diagnosis of Type II diabetes, which ordinarily occurs among older adults (Nathan et al., 2009). In these early stages, oral medication (usually Metformin, available as a generic medication since 1985 in the US) and elimination of refined sugar from the individual's diet are more likely to be recommended. This is because the pancreas may still secrete insulin, but either not enough is secreted and/or the cells of the body do not respond to it. By contrast, those with Type I diabetes (10% of the total population of individuals with diabetes), who develop the disease in childhood, cannot produce insulin at all. Accordingly, poor control of their blood sugar levels will quickly result in life-threatening complications for this group. Unlike Type Is, Type IIs are most likely to be recommended to use insulin only once their disease has progressed to the point where intermediate methods for controlling blood sugar levels such as dieting or oral medication have become relatively, though not completely, ineffective (Karter et al., 2010; Nathan et al., 2009).

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

The setting for this empirical application is the United States, since the data are drawn from the Health and Retirement Study (HRS) survey data, using waves from 1998 to 2009 (the pre-Affordable Care Act era). In the United States during this period, every resident (subject to the mild requirement that they or their spouse have worked for 40 quarters over their working life, which almost all U.S. residents meet) becomes eligible in the month that they turn 65 for health insurance via Medicare. Prior to 2006, Medicare Parts A and B covered the majority of out-of-pocket costs for enrollees for outpatient and inpatient services, but without purchasing supplemental insurance these default components of the program did not cover prescription drug costs. Individuals who lacked health insurance could face significant out-of-pocket costs when purchasing insulin, in part because there was (and still is as of the time of writing) no generic form of insulin. For example, Eli Lilly's insulin, Humalog, cost $34.81 per vial (which would typically contain a month's worth of insulin) in 2001. The yearly cost of insulin at this intensity of usage amounts to $416.72. This is a modest estimate since many individuals with diabetes will require more than one vial's worth of insulin per month. Individuals who use a "basal-bolus" regime, so called because it combines a baseline daily dose of insulin (the "basal" part) with regular injections before mealtimes (the "bolus" part), will require 9 vials every 2 months on average. The yearly cost of insulin at this intensity of usage amounts to $1879.74 in 1998 dollars. In 1998 200% of the federal poverty line (which would exclude the possibility of qualifying for Medicaid) outside of Alaska and Hawaii for a two-person household was $21 700. Therefore an uninsured married couple with one member with diabetes could expect to spend 7% of total household income on insulin alone if they were at 200% of the federal poverty line in 1998.

The fact that individuals just above and just below the cutoff age of 65 are otherwise comparable apart from qualifying for health insurance coverage has motivated regression-discontinuity design-based estimates of Medicare on a range of outcomes. After 2006, Medicare beneficiaries could enroll in a private prescription-drug benefit separately to Medigap or Medicare Advantage under Medicare Part D (though Part D plans were often bundled with Medicare Advantage plans (Geruso & Layton, 2020)). This motivates a difference-in-discontinuities estimator of the impact of prescription drug coverage on insulin use, with 65 as the regression discontinuity cutoff age, comparing the effect of turning age 65 in 2006–2009 to its effect in 1998–2005.[1] The objects of interest are the net impacts of Medicare coverage on individuals with diabetes (with i subscripting individuals, a subscripting age and t subscripting the year in which the survey wave was collected),

$$E[Y_{iat}|Diab_{iat}=1, Medicare_{iat}=1]-E[Y_{iat}|Diab_{iat}=1, Medicare_{iat}=0], \quad \text{for } 2006\le t\le 2009, 1998\le t<2006, \tag{2}$$

Which is unobservable since every individual above age 65 is eligible for Medicare, while individuals eligible for Medicare prior to 65 are selected for worse health (as they typically qualify via having been on SSDI for two years), and so aren't comparable to non-enrollees younger than 65. I exploit the discontinuity in Medicare eligibility at age 65 to identify the reduced-form effect of Medicare on those made eligible for it (an Intention to Treat (ITT) effect)[2]:

$$\begin{aligned} Y_{it} &= \beta_0+\beta_1 Medicare_{iat}+f_0(a_{it}-\bar{a})+f_1(a_{it}-\bar{a})\times Medicare_{iat}+\beta_2 1_{[t\ge 2006]} \\ &+\beta_3 Medicare_{iat}\times 1_{[t\ge 2006]}+\beta_4 1_{[t\ge 2004]}+\beta_5 Medicare_{iat}\times 1_{[t\ge 2004]}+\delta X_{iat}+\upsilon_{iat}, \left|\frac{a_{it}-\bar{a}}{h}\right|<1, \end{aligned} \tag{3}$$

where $Medicare_{iat}=1$ if individual i's age in months a in period t is greater or equal to 780 ($=65\times 12$). f_0, f_1 are polynomial functions of order p, according to the order of the local polynomial regression that is used (in practice, either local linear ($p=1$) or local quadratic ($p=2$) due to recent work demonstrating that higher-order polynomials can lead to estimators with undesirable properties (Gelman & Imbens, 2018)).

The standard identifying assumptions in this design are that

$$\lim_{a_{it}\to\bar{a}^+}E[\upsilon_{iat}|a_{it}]=\lim_{a_{it}\to\bar{a}^-}E[\upsilon_{iat}|a_{it}], \tag{4}$$

$$\begin{aligned} &\lim_{a_{it}\to\bar{a}^+}E[\upsilon_{iat}|a_{it}, t\ge 2006]-\lim_{a_{it}\to\bar{a}^-}E[\upsilon_{iat}|a_{it}, t\ge 2006] \\ &=\lim_{a_{it}\to\bar{a}^+}E[\upsilon_{iat}|a_{it}, t<2006]-\lim_{a_{it}\to\bar{a}^-}E[\upsilon_{iat}|a_{it}, t<2006] \end{aligned} \tag{5}$$

that is, that unobserved factors trend smoothly at the cutoff age $\bar{a}$, and that there is no difference in the discontinuities after 2006 relative to before 2006. In the Appendix A, I test for the presence of discontinuities in observable outcomes other than diagnosis, which would cast doubt on this assumption. As in Card et al. (2008) and Card et al. (2009), I find that the main threat to identification - discontinuous changes in labor force participation - is not borne out in the data (see Appendix A). Intuitively, individuals just above and below the age-eligibility criterion are comparable to each other apart from their eligibility

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

for Medicare, and those close to the cutoff age in 2006–2009 are comparable to their older counterparts in 1998–2005 in their reaction to Medicare except for the expanded options for prescription drug coverage on Medicare in 2006–2009.

I also include an indicator for the post-2004 period and its interaction with the Medicare eligibility indicator for two reasons. The first is to capture announcement effects, since Part D was signed into law in December 2003, some months after the end of the 2002 wave of the HRS and a few months before the beginning of the interviews for the 2004 wave of the HRS. The second is that in the period between Part D's announcement and implementation, Medicare beneficiaries were provided with subsidies via the Prescription Drug Card and Transitional Assistance programs (Cubanski et al., 2004; Huh & Reif, 2017). Any effect of these interim measures that is distinct from the impact of the full implementation of Part D will be absorbed by this post-2004 indicator and its interaction with the Medicare eligibility indicator.

Low-income, low-asset individuals with diabetes who qualified for Medicaid coverage in the period before 2006 were guaranteed full coverage for the costs of using insulin. This motivates a placebo test: estimating Equation (3) restricting attention to Medicaid recipients. I implement and discuss this placebo test, which shows no change in the behavior of those who transition to being dually eligible for both Medicaid and Medicare, in the Appendix A.

I use data from three sources: the Health and Retirement Study (HRS), a cleaned version of the HRS called the RAND-HRS (Chien et al., 2013), and a cleaned version of a survey administered to a subset of Health and Retirement Study participants called the Consumption and Activities Mail Survey (CAMS), the RAND-CAMS.

3.1.1 | The Health and Retirement Study (HRS)

The Health and Retirement Study is a nationally representative longitudinal survey administered by the Institute for Social Research at the University of Michigan. In 1998, the individuals are drawn from four birth cohorts: the Oldest Old (born pre-1924), the Children of the Depression (born 1924–31), the original cohort from 1992 (born 1931–41) and the War Babies (born 1942–47), plus their co-habitants in the households in which they resided at the time of the survey. The HRS followed up respondents every two years, providing data on surviving individuals from the 1998 wave in 2000, 2002, 2004, 2006 and 2008. Individuals are included in the sample if they self-report a diagnosis of diabetes. In 1998, this amounts to 3043 respondents to the HRS, rising to 3771 respondents in the 2008 wave of the HRS.

Since the 2006 wave of the survey was collected between March 2006 and February 2007, which was after the initial difficulties with the rollout of Medicare Part D had been overcome, I code the post-Part D era as inclusive of both the 2006 and 2008 waves. The interview period for the 2008 wave extended into early 2009. Not all of the variables used in the analysis were available in the RAND HRS data. In particular, the insulin usage indicator variable had to be merged from the original HRS data.

3.1.2 | RAND-HRS

The RAND-HRS data is a subset of the HRS data with cleaning and imputations performed by a team of researchers from the RAND Corporation. Further details of its construction can be found in Chien et al. (2013). This data provides the following variables used in the analysis: individual-level characteristics such as education, gender, race, marital status, self-reported health, and age in months at the time of the interview (based on the difference between the month and year of the interview and reported month and year of birth). This last variable is the running variable for the difference-in-discontinuities design employed in this paper.

3.1.3 | Consumption and Activities Mail Survey (RAND-CAMS)

The RAND-CAMS data is a cleaned subset of the Consumption and Activities Mail Survey (CAMS), a survey administered to a subset of HRS respondents in years that fall in between survey years of the HRS. I use the waves that correspond most closely to the sample period since this is not administered at the same time as the main HRS interviews: 2001, 2003, 2005, 2007 and 2009. This provides data on expenditure on a wide array of goods, both durable and nondurable. This allows for an additional test for discontinuities in expenditure around the age 65 threshold (see Appendix A).

3.2 | Baseline results

Figure 1 displays local linear regressions fit either side of age 65, the age at which individuals qualify for Medicare coverage, for a regression with self-reported diagnosis of diabetes as the dependent variable. Diagnoses appear to trend smoothly at age 65

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

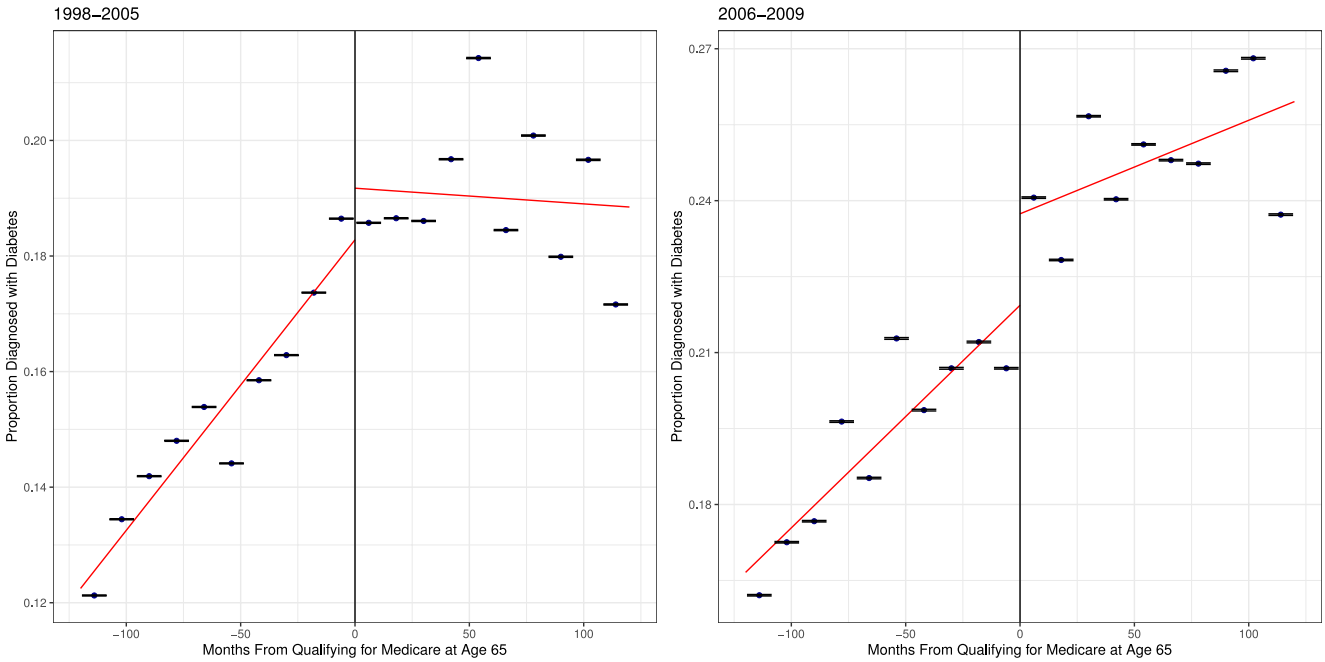

FIGURE 1 Changes in self-reported diabetes diagnoses at age 65, 1998–2009. The dependent variable is self-reported, coded as 1 if a respondent answers "yes" to the question of whether they have been diagnosed with diabetes, and 0 otherwise. The running variable is months of Medicare eligibility, calculated as age in months—780, so that 0 is the month in which individuals turn 65. *Source*: Author's own calculations using a combination of the 1998–2008 waves of the Health and Retirement Study and RAND-HRS data, via the R package "rdrobust".

TABLE 1 Changes in diabetes diagnoses at age 65 in 2006–2009.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Bandwidth (h)	**30**	**60**	**120**	∞	**30**	**60**	**120**	∞
$Medicare \times 1_{[t \geq 2006]}$	0.027*	0.010	0.018*	0.017**	0.027*	0.010	0.018*	0.017**
	(2.092)	(0.838)	(2.267)	(3.001)	(2.080)	(0.836)	(2.212)	(2.998)
N	19,221	37,648	70,342	114,254	19,221	37,648	70,342	114,254

Note: Estimates are from interaction between the treatment indicator for qualifying for Medicare in the month individuals turn 65 and the indicator for the period 2006–2009. Standard errors are clustered at the individual level. Columns (1–4) report local linear regression results; Columns (5–8) report local quadratic regression results. The optimal bandwidth for the local linear regression, selected by the MSE criterion of Calonico et al. (2014), is $h = 60$. t statistics in parentheses.

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

before 2006, when Part D prescription drug coverage is introduced, and thereafter jump slightly at age 65 in 2006–2009. Table 1 reports estimates from local linear and local quadratic regressions of the size of the post-2006 difference-in-discontinuities in diabetes diagnoses. The estimates for both sets of regressions are estimated for four cases: (i) the "optimal" bandwidth for the local linear specification, $h = 60$, (ii) half this bandwidth, $h = 30$, (iii) double this bandwidth, $h = 120$, and (iv) including all observations. Though the difference-in-discontinuities is not statistically significant for the "optimal" local linear bandwidth of $h = 60$, there does appear to be a statistically significant increase in the rate of diagnosis of diabetes at both narrower and wider bandwidths. A researcher presenting the above table as a robustness check for main results examining outcomes among individuals with diabetes is unlikely to persuade her audience that there is no problem with endogenous sample selection on the basis of diabetes diagnoses, which would motivate her to use one of the two methods proposed above to either bound or point-identify the impact of Medicare on outcomes measured in this subgroup. Consistent with Figure 1, but not calculations made using the "optimal" (for the local-linear case, MSE-minimizing) bandwidth of Calonico et al. (2014) of $h = 60$, an increase of between 1.7 and 2.7% points in the frequency of diabetes diagnosis may occur at age 65 after 2006. This suggests that individuals in the sample are more likely to report being diagnosed with diabetes once they turn 65 once prescription drug coverage is made available on Medicare.

That the increase in diagnoses after age 65 happens only after 2006 can be rationalized via two mechanisms. First, therapeutic measures and screenings are complements (Kenkel, 2000). If individuals anticipate that they are more likely to afford treating the consequences of living with diabetes via prescription drug coverage, this may encourage them to be screened for

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

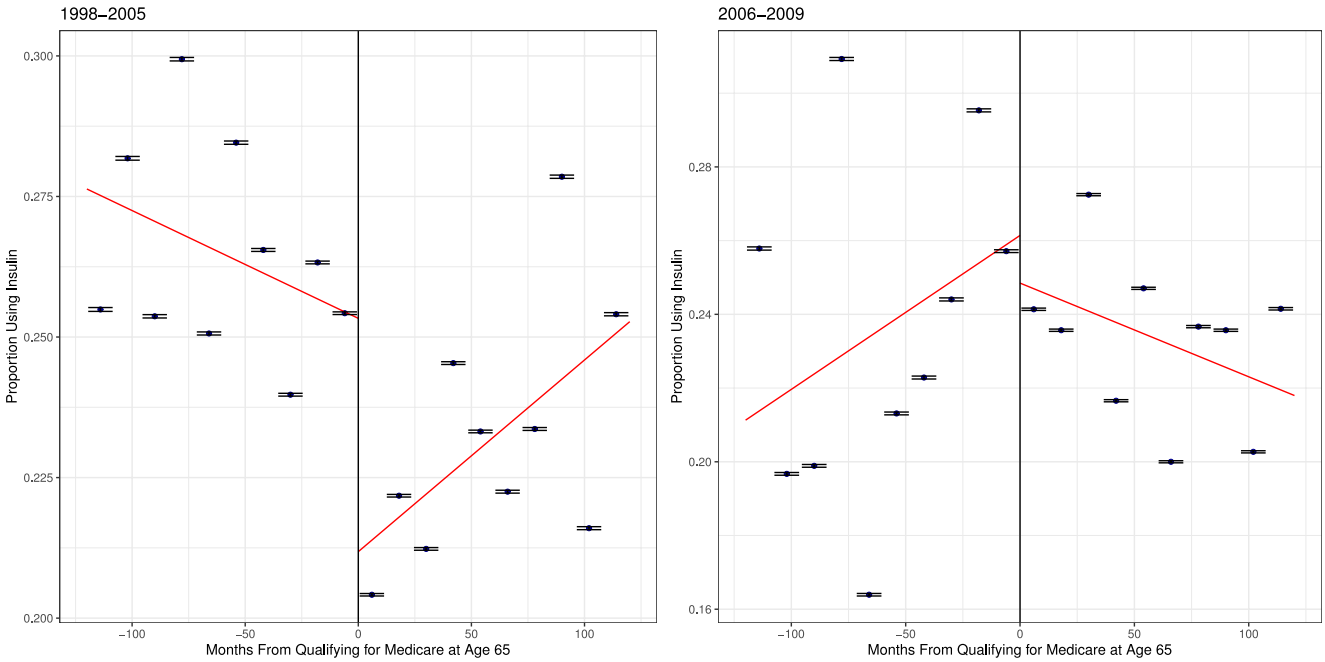

FIGURE 2 Changes in insulin use at age 65 among individuals with diabetes. The dependent variable is self-reported, coded as 1 if a respondent answers "yes" to the question of whether they use insulin to manage their diabetes, and 0 otherwise. The running variable is months of Medicare eligibility, calculated as age in months—780, so that 0 is the month in which individuals turn 65. *Source*: Author's own calculations using a combination of the 1998–2008 waves of the Health and Retirement Study and RAND-HRS data, via the R package "rdrobust".

diabetes relative to the case where they do not have prescription drug coverage and face higher cost-sharing for antidiabetic medications. Second, at least one previous study has found evidence of "upcoding" of diagnoses due to risk adjustment after the implementation of Part D (Geruso & Layton, 2020). This is due to the bundling of many Part D plans with Medicare Advantage plans, the latter of which provided incentives for physicians to make diagnoses of chronic conditions at higher rates in order to earn greater transfer payments for enrolling riskier individuals. Geruso and Layton (2020) find a statistically significant increase in the probability of Medicare enrollees being diagnosed with a chronic condition in 2006 relative to their pre-2006 counterparts. Consistent with this explanation, I find higher and increasing proportion of individuals with diabetes who report being enrolled in a Medicare HMO in 2006–2009 relative to 1998–2005, in which this proportion declined (see Appendix A).

Consider a researcher who has run the above robustness check that tests whether inclusion in the subsample of people with diabetes changes at age 65. The null results for the "optimal" bandwidth (Table 1) are unlikely to convince a skeptical reviewer that her main results are not subject to sample selection on the basis of diagnosis, especially given the pattern visible in Figure 1. What this paper's proposed method allows her to do is to argue for the informativeness of her main results regardless of whether diagnosis rates actually do increase at the cutoff. If they do not, she has point-identified the effect of interest. If diagnosis rates do increase, and Monotone Diagnosis Response and Monotone Subgroup Selection hold, she can argue that her estimates represent lower bounds for the treatment effect of interest. In addition, if she has a panel dimension in the data, and the necessary stability assumption holds, she can use within-person variation to estimate the extent of the bias due to changing diagnosis rates and correct for them.

Figure 2 and Table 2 both exhibit evidence that self-reported insulin use *decreased* at age 65 among individuals with diabetes prior to 2006, with this decrease diminishing or closing completely with the advent of prescription drug coverage under Medicare Part D. There is an overall decline in self-reported insulin use of 4% points prior to 2006 when individuals with diabetes qualify for Medicare (Table 2, column 1). In the Appendix A, I provide some evidence that this is due to individuals dropping their employer-provided health insurance when they qualify for Medicare. It is unlikely that this is due to employer-based health plans disenrolling individuals at age 65 as this contradicts provisions of both COBRA and HIPAA regulations.

Though the increase of 3% points at the same age once Part D prescription drug coverage is introduced is not statistically significant, a Wald test of the sum of the two effects canceling out cannot reject the null hypothesis that $\beta_1 + \beta_3 = 0$ ($p = 0.6446$). The decline pre-2006 and subsequent increase after Part D prescription drug coverage is introduced appears to be driven primarily by changes in initiation of insulin therapy by individuals who were not previously using insulin (column 2), which is consistently statistically significant across the two specifications. The magnitude of these effects on initiations of

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 2 Changes in insulin use at age 65 for individuals with diabetes, pre- and post-2006.

	$p = 1$				$p = 2$			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Local polynomial order, p	**Insulin**	**Began insulin**	**Stopped insulin**	**Didn't fill: Cost**	**Insulin**	**Began insulin**	**Stopped insulin**	**Didn't fill: Cost**
Medicare	−0.06**	−0.09***	−0.00	0.04*	−0.06**	−0.09***	−0.00	0.03
	(−3.07)	(−5.10)	(−0.32)	(2.35)	(−2.74)	(−5.28)	(−0.62)	(1.50)
Medicare × $1_{[t \geq 2006]}$	0.03	0.03*	−0.01	−0.08***	0.02	0.03**	−0.01	−0.07***
	(0.90)	(2.18)	(−1.28)	(−3.76)	(0.99)	(2.79)	(−1.92)	(−4.20)
N	7285	7288	7288	7282	11,402	11,405	11,405	11,389

Note: Estimates are from the treatment indicator for qualifying for Medicare in the month individuals turn 65 and its interaction with the indicator for the period 2006–2009. Standard errors are clustered at the individual level. All specifications use the Uniform kernel and include controls for race, gender, education and smoking status. The bandwidth used is 60 months for the local linear specification, and 100 months for the local quadratic specification, selected by the MSE criterion of Calonico et al. (2014). *t* statistics in parentheses.

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

insulin use is also consistent across the two specifications: a 9-percentage point decline at age 65 in the rate of take-up among non-users prior to 2006, offset by a 3 percentage-point increase in this rate post-2006.

Both the empirical results above and the diagnosis-constant treatment effects (in the next subsection) suggest a greater sensitivity to prescription drug coverage than some of the prior literature for insulin use decisions; for example, the estimated own-price elasticity for insulin in Einav et al. (2018) is −0.02 (not reported in the main text of their paper, but in Table A14 of their online Appendix A). There are two potential reasons for this.

First, since the insulin use question in the HRS is just a yes/no response to whether or not an individual currently uses insulin, the above estimates correspond to the extensive margin of insulin use. The fact that the responses estimated here are at the extensive rather than intensive margin has the advantage of making it easier to reconcile the large extensive-margin response with the smaller intensive margin price elasticity of demand for insulin[3] identified by Einav et al. (2018), since extensive- and intensive-margin responses need not be similar. Their relative size will depend on which individuals are at the margin and which are infra-marginal, as in the case of the Marginal Treatment Effect and Average Effect of Treatment on the Treated (cf. Heckman et al. (2006)).

Second, the identifying variation used in this paper induces a different subset of marginal individuals to change their behavior. Einav et al. (2018) exploit bunching at the kink created by Medicare Part D's "donut hole" in coverage: beyond $2510 in prescription drug costs, individuals face 100% cost-sharing until they reach the "catastrophic" coverage threshold of $5726 in out-of-pocket expenditures. Users of insulin who face costs of these magnitudes - almost certainly due to heavy required usage - are likely much less able to maintain their immediate well-being than individuals who require less intensive regimens of insulin usage to control their blood sugar levels. This is likely why the responsiveness of these individuals to the changes in cost-sharing around the donut hole is much smaller than the responsiveness of individuals who are at the margin of deciding whether or not to begin using insulin. The results suggest that the extensive margin of insulin usage - the start/don't start insulin decision - accounts for a large share of the price elasticity of −0.25 for all antidiabetic medications estimated by Goldman et al. (2004).[4]

3.3 | The diagnosis-constant effect of prescription drug coverage on insulin use

The baseline results indicate a negative impact of qualifying for Medicare on insulin use prior to prescription drug coverage being included in Medicare in 2006. Of the coefficients that measure the positive, offsetting impact of Part D prescription drug coverage, the positive coefficient for overall insulin use is not statistically significant, but the one for initiations of insulin use by previous non-users is. At the same time, there is some evidence that the rate at which diabetes is diagnosed for new Medicare beneficiaries increased after 2006, consistent with prior studies (for example, Geruso and Layton (2020)), and that cost-related nonadherence to medication decreased among individuals with diabetes at age 65 after Part D was introduced.

This combination of results, in line with the method outlined in Section 2, suggests that the estimated effect of Medicare prescription drug coverage in 2006–2009 on initiations of insulin use is an underestimate relative to the diagnosis-constant effect. Both Monotone Subgroup Selection and Monotone Diagnosis Response are plausible due to the increasing severity of

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

TABLE 3 Rates of progression to insulin use in the Health and Retirement Study (HRS) data for individuals with diabetes, by length of time for which a diagnosis is reported.

	(1)	(2)	(3)	(4)	(5)	(6)
	Overall	**0–2 years**	**2–4 years**	**4–6 years**	**6–8 years**	**8–10 years**
Insulin use rate	0.24***	0.04***	0.04***	0.05***	0.07***	0.09***
	(43.07)	(10.04)	(10.94)	(11.54)	(12.21)	(12.46)
N	20,327	2516	3536	3492	2913	1664

Note: Estimates are mean rates of insulin use, conditioning on the length of time in the HRS data for which the respondent reports having diabetes. Increments are in 2 years, reflecting that the HRS survey data is collected every 2 years. *t* statistics in parentheses.

*$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

the illness over time. Undiagnosed individuals with diabetes are likely to only be more likely to be diagnosed when in more frequent contact with medical practitioners, which supports the assumption of Monotone Subgroup Selection in the case of the effect of Medicare eligibility on diabetes diagnosis. Rates of insulin use are likely to be lower for individuals who are only diagnosed in the counterfactual with greater access to healthcare since those on the margins of a diagnosis will tend to be healthier, in the early stages of the disease, or otherwise be capable of adequate control of their blood sugar levels without injecting insulin. This is borne out by Table 3, which documents rates of insulin use by time from first diagnosis in the HRS data (cf. Inzucchi et al. (2015), who find that in a sample of individuals with diabetes who had been diagnosed on average 2 years prior to the start of the study period, one of their subgroups had a cumulative insulin use initiation rate of 9.9%, similar to the overall rate found in this data for individuals who were between 8 and 10 years from diagnosis).

Otherwise, an undiagnosed individual is likely to undergo a severe medical episode that necessitates contact with health care providers, during which a diagnosis is more likely to be made. Thus, Monotone Diagnosis Response is also plausible in this setting. This combination of assumptions allows us to conclude, as in Section 2, that the estimated effect of prescription drug coverage is likely to be a lower bound for the diagnosis-constant effect. Formally:

Monotone Subgroup Selection: For every individual i,

$$Diab_i(1) \geq Diab_i(0) \Leftrightarrow Diab_i(Medicare = 1) \geq Diab_i(Medicare = 0),$$

Monotone Diagnosis Response:

$$E\left[\text{Began Insulin}_i(1) | Diab_i(1) = Diab_i(0) = 1\right] \geq E\left[\text{Began Insulin}_i(1) | Diab_i(1) = 1, Diab_i(0) = 0\right],$$

and so,

$$\begin{aligned} &E\left[\text{Began Insulin}_i(1) - \text{Began Insulin}_i(0) | Medicare_i = 1, Diab_i(1) = Diab_i(0) = 1\right] \\ \geq\ &E\left[\text{Began Insulin}_i(1) | Medicare_i = 1, Diab_i(1) = 1\right] - E\left[\text{Began Insulin}_i(0) | Medicare_i = 0, Diab_i(0) = 1\right], \end{aligned}$$

the left hand side of which inequality is the diagnosis-constant effect of treatment on the treated, and the right hand side of which is estimated for the pre- and post-2006 regimes. Since we have evidence that rates of diagnosis of diabetes are affected by the advent of the post-2006 regime, we can conclude that the positive impact on initiations of insulin use estimated above post-2006 is a lower bound for its impact on initiations among individuals who had already been diagnosed with diabetes prior to qualifying for Medicare in the period 2006–2009.

The HRS data also follows the same respondents from wave to wave, allowing us to examine within-person outcomes among the newly diagnosed. Since the regression-discontinuity framework assumes that individuals close to the cutoff are comparable, under those assumptions it follows that

$$\begin{aligned} &E\left[\text{Began Insulin}_{65+} | Diab_{65+} = 1, Diab_{65-} = 0\right] \\ &= E[\text{Began Insulin} | Medicare = 1, Diab(1) = 1, Diab(0) = 0], \end{aligned}$$

which allows us to calculate

$$= \frac{E\left[\text{Began Insulin}_{65+} | Diab_{65+} = 1\right] - E\left[\text{Began Insulin}_{65+} | Diab_{65+} = 1, Diab_{65-} = 0\right] \dfrac{[\Pr(Diab_{65+} = 1) - \Pr(Diab_{65-} = 1)]}{\Pr(Diab_{65+} = 1)}}{\dfrac{\Pr(Diab_{65-} = 1)}{\Pr(Diab_{65+} = 1)}},$$

$$E[\text{Began Insulin} | Medicare = 1, Diab(1) = 1, Diab(0) = 1]$$

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

via the Law of Total Probability (as in Section 2) and Bayes' Rule.

Since the HRS data includes information on whether individuals are newly diagnosed, we can calculate $E\left[\text{Began Insulin}_{65+} | Diab_{65+} = 1, Diab_{65-} = 0\right]$, which turns out to be 0.033, just under half the overall average of 0.079. Using the most pessimistic of the estimates of the discontinuous increase in individuals with diabetes in 2006–2009 of 0.027, together with the pre-65 proportion of individuals with diabetes of 0.207, produces $\frac{\Pr(Diab_{65-}=1)}{\Pr(Diab_{65+}=1)} = \frac{0.207}{0.207+0.027} = 0.88$ and hence $\frac{[\Pr(Diab_{65+}=1)-\Pr(Diab_{65-}=1)]}{\Pr(Diab_{65+}=1)} = 0.12$. Together with the sample proportion $E\left[\text{Began Insulin}_{65+} | Diab_{65+} = 1\right] = 0.079$, this gives $E\left[\text{Began Insulin}_{65+} | Diab_{65+} = 1 \& Diab_{65-} = 1\right] = \frac{0.079 - 0.033 \times 0.12}{0.88} = 0.085$. Under Monotone Subgroup Selection, the untreated outcome estimated for the under-65s is common to the estimated treatment effect and its diagnosis-constant counterpart, and so the diagnosis-constant treatment effect exceeds the "naive" estimate that ignores new diagnoses by $E\left[\text{Began Insulin}_{65+} | Diab_{65+} = 1 \& Diab_{65-} = 1\right] - E\left[\text{Began Insulin}_{65+} | Diab_{65+} = 1\right] = 0.085 - 0.079 = 0.006$. Since the main estimate of the positive effect of Part D on initiations of insulin use is an increase of 3% points, this implies that the *diagnosis-constant* estimate of this effect is an increase of 3.6% points. Under the maintained assumptions, the main results could be understating the impact of Part D prescription drug coverage on initiation of insulin use by 20% relative to the true increase.

4 | DISCUSSION

Identifying the causal impact of health insurance on particular subgroups is often of interest to applied researchers. When those subgroups themselves change - as the result of new diagnoses, for example, - this can result in biased estimators of the objects of interest. This paper exploits panel data to estimate the extent of this bias by identifying the average outcome among the newly diagnosed and using the Law of Total Probability and Bayes' Rule to subtract it from the overall estimate. This results in a "diagnosis-constant" estimate of the causal effect of interest. The differences between the two can be substantial: the "diagnosis-constant" treatment effect is one-sixth larger in the empirical application considered in this paper. Even in the absence of panel data, relatively mild assumptions based on subject matter knowledge allow us to conclude that the estimated treatment effect is a lower bound for its diagnosis-constant counterpart.

There are two main mechanisms that could produce the increase in diabetes diagnoses observed above. The first is increased contact with health care providers, which increases the probability that a latent condition is diagnosed. The second is "upcoding" (Geruso & Layton, 2020), whereby Medicare Advantage plans' primary care providers are encouraged to loosen the criteria for making a diagnosis due to the benefits to such plans of enrolling individuals with diabetes that accrue via risk adjustment payments. Both mechanisms have similar implications for insulin use among the newly diagnosed. Individuals with Type II diabetes who are able to live without being conscious of an official diagnosis are typically in the early stages of the disease, as mismanagement of the disease in its later stages can induce a hyperglycemic ("high blood sugar") coma. They are therefore unlikely to be recommended to begin insulin therapy to manage their condition by a primary care provider. This is borne out by the lower-than-average estimated rate of insulin use among those newly diagnosed with diabetes in the HRS data. As a result, the estimated causal effect of prescription drug coverage on insulin use via the difference-in-discontinuities design is biased toward zero, and is a lower bound for the true causal effect. Exploiting the panel dimension of the HRS data allows us to say by how much this estimator underestimates the "diagnosis-constant" causal effect: it is as much as 20% smaller in relative terms.

The empirical results of this paper contribute to a growing literature that finds that individuals are equally responsive to the price of high-value care as they are to the price of low-value care, despite economic theory predicting that the latter should exceed the former (Abaluck et al., 2018; Baicker et al., 2015; Chandra et al., 2021; Einav et al., 2018; Goldman et al., 2004, 2007; Gross et al., 2022). Despite physicians agreeing that individuals with Type II diabetes should begin insulin use as soon as it is recommended to them by their primary care provider (Nathan et al., 2009), I find significant evidence of delayed initiation of insulin therapy. By contrast, I find much weaker evidence that worse prescription drug coverage leads to cessation of insulin therapy once it has already begun. While there has been significant recent attention devoted to insulin rationing among users due to its cost (Herkert et al., 2019), this study's results suggest that an additional, underappreciated problem with the high price of insulin is that it deters the marginal individual from beginning to use it in a timely manner.

The main policy implication of this paper's methodology is that even when "causal" estimates of the sensitivity of health care consumption to its price are available, policymakers may still be underestimating the extent to which individuals may forgo important investments in their health due to cost considerations. This will be the case if the treatment effect is examined

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

separately for costly subgroups of interest (such as individuals with diabetes) without accounting for the fact that membership in that subgroup in self-reported data is endogenous. The asymmetry between current users and non-users' responses to price changes is evidence that current use of insulin makes future use more likely (i.e., there is persistence in insulin use, possibly via habit formation). As a result, one option available to policymakers that doesn't appear to be used in any OECD countries as of the time of writing is a "first *X* fills free" policy that lowers the price for first-time users for a pre-specified amount of time. Since current users are less sensitive to the price than those not yet using insulin, this lower price for non-users is likely to capture most of the increase in take-up that would result from a lower price for all individuals with diabetes. Similarly, once the price increases again after the initial subsidy period expires, the decline in insulin use due to higher out-of-pocket costs is, as per the empirical results, likely to be significantly smaller than the initial increase in the proportion using insulin. I leave further study of these short-run subsidies to future research.

While the paper uses health insurance and diagnosis of latent conditions as the setting for these methods throughout, in principle they could be applied to any situation where a covariate is affected by treatment and a subgroup defined by a particular value of that covariate is of interest. For example, some welfare payments such as the United States' Earned Income Tax Credit (EITC) treat single and married recipients differently in the calculation of benefits for which they are eligible (Dickert-Conlin & Houser, 2002). Recent work by Ortigueira and Siassi (2022) finds that the overall tax and transfer system in the United States incentivizes cohabitation for single mothers. If a researcher wants to examine the separate impact of the EITC on single recipients relative to married or cohabiting recipients, then the method outlined in this paper could help them to account for the impact of the EITC on household formation, which would ordinarily prevent the researcher from isolating the subgroup-specific causal effects.

Finally, this paper deals only with binary indicators of inclusion in a subgroup. Many endogenously determined subgroups of interest will vary continuously with some characteristic instead. In the empirical application used in this paper, data on the severity of the diabetes for any particular individual (as measured by their HbA1C levels, for instance) was not available. Had it been available, different subgroups within the subgroup of individuals with diabetes could be considered, such as the subgroup with blood sugar levels high enough to make it likely that their physician would recommend that they begin using insulin. These data would allow the researcher to speak not just to the question of take-up but adherence to medical recommendations (the latter of which are not observed in the data used above). Future work could extend the above analysis to examine "severity-constant" treatment effects that condition on some variable being above a threshold value that triggers different treatment regimes.

ACKNOWLEDGMENTS

I have benefited greatly from the comments and advice of three anonymous referees, James Best, Gabriella Conti, Ian Crawford, Tal Gross, Rury Holman, Johannes Jaspersen, Michael P. Keane, Hamish Low, and Olivia S. Mitchell, as well as seminar participants at the 2017 National Tax Association Meetings (and my discussant there, David Powell), the 2018 European Health Economics Association (EuHEA) meetings, the 2018 European Economics Association meetings, the 2018 European Winter Meetings of the Econometric Society, the 2019 Royal Economic Society Meetings, the 2020 Allied Social Science Association Meetings, Ludwig Maximilian University, Oxford, the University of Pennsylvania, the University of Siena, Singapore Management University and the National University of Singapore. Material in this study previously circulated as part of the working papers "Does Insurance for Treatment Crowd Out Prevention? Evidence from Diabetics' Insulin Usage" and "Missing Prescriptions: Private Prescription Drug Coverage and the Heterogeneous Effects of Medicare on Insulin Use".

CONFLICT OF INTEREST STATEMENT

The author has no relevant financial or non-financial interests to disclose. No funding from grants or external organizations was used to conduct this study.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are publicly available to registered users (requiring only an email address) in the HRS Public Access Files and RAND-HRS Derived Files at https://hrsdata.isr.umich.edu/dataproducts/public-survey-data and https://hrsdata.isr.umich.edu/data-products/rand. Additional data for calculating CPI deflators were derived from the following resources available in the public domain, via the Federal Reserve Bank of St Louis: https://fred.stlouisfed.org/series/CPIAUCSL and https://fred.stlouisfed.org/series/CPIMEDSL.

ORCID

Daniel Kaliski https://orcid.org/0000-0003-1230-600X

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

ENDNOTES

[1] Studies using a similar research design include Card et al. (2008, 2009), Anderson et al. (2012, 2014) and Asfaw (2019). Studies that also exploit the age of Medicare eligibility as a source of exogenous variation, but in a "triple-difference" design, include Dave and Kaestner (2009) and Wettstein (2020).

[2] Note that the object of interest is not the causal effect of access to health insurance per se as in the motivating example for this paper, as Medicare both gives the previously uninsured access to health insurance *and* expands coverage options for those who were already insured prior to age 65 (Card et al., 2009). This paper therefore focuses on the net effect of qualifying for Medicare coverage, rather than attempting to identify the separate contributions of the extensive margin (health insurance vs. no insurance) and intensive margin (more generous insurance) of health insurance to this effect.

[3] The measure of quantity demanded in Einav et al. (2018) is whether an individual fills a prescription of a particular drug in the month of December. This measures the intensive margin of the demand for prescription drugs. The question in the Health and Retirement Study regarding insulin usage asks whether or not individuals currently use insulin to manage their condition, that is, whether they use it at all.

[4] Goldman et al. (2004) estimate that a doubling of co-payments for antidiabetic medications results in a 25% (relative, not percentage point) reduction in their usage. This is not reported as a price elasticity in their paper; I take the implied point elasticity to be $\frac{-25\%}{100\%} = -0.25$.

REFERENCES

Abaluck, J., Gruber, J., & Swanson, A. (2018). Prescription drug use under Medicare Part D: A linear model of nonlinear budget sets. *Journal of Public Economics*, *164*, 106–138. https://doi.org/10.1016/j.jpubeco.2018.05.005

Abrevaya, J., Hsu, Y.-C., & Lieli, R. P. (2015). Estimating conditional average treatment effects. *Journal of Business and Economic Statistics*, *33*(4), 485–505. https://doi.org/10.1080/07350015.2014.975555

Anderson, M., Dobkin, C., & Gross, T. (2012). The effect of health insurance coverage on the use of medical services. *American Economic Journal: Economic Policy*, *4*(1), 1–27. https://doi.org/10.1257/pol.4.1.1

Anderson, M. L., Dobkin, C., & Gross, T. (2014). The effect of health insurance on emergency department visits: Evidence from an age-based eligibility threshold. *The Review of Economics and Statistics*, *96*(1), 189–195. https://doi.org/10.1162/rest_a_00378

Asfaw, A. A. (2019). The effect of prescription drug insurance on health behavior: Evidence from Medicare Part D. *Health Economics*, *28*(3), 403–418. https://doi.org/10.1002/hec.3853

Athey, S., & Imbens, G. W. (2006). Identification and inference in nonlinear difference-in-differences models. *Econometrica*, *74*(2), 431–497. https://doi.org/10.1111/j.1468-0262.2006.00668.x

Baicker, K., Mullainathan, S., & Schwartzstein, J. (2015). Behavioral hazard in health insurance. *Quarterly Journal of Economics*, *130*(4), 1623–1667. https://doi.org/10.1093/qje/qjv029

Callaway, B. (2021). Bounds on distributional treatment effect parameters using panel data with an application on job displacement. *Journal of Econometrics*, *222*(2), 861–881. https://doi.org/10.1016/j.jeconom.2020.02.005

Callaway, B., & Li, T. (2019). Quantile treatment effects in difference in differences models with panel data. *Quantitative Economics*, *10*(4), 1579–1618. https://doi.org/10.3982/qe935

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, *82*(6), 2295–2326. https://doi.org/10.3982/ecta11757

Card, D., Dobkin, C., & Maestas, N. (2008). The impact of nearly universal insurance coverage on health care utilization and health: Evidence from Medicare. *The American Economic Review*, *98*(5), 2242–2258. https://doi.org/10.1257/aer.98.5.2242

Card, D., Dobkin, C., & Maestas, N. (2009). Does Medicare save lives? *Quarterly Journal of Economics*, *124*(2), 597–636. https://doi.org/10.1162/qjec.2009.124.2.597

Chandra, A., Flack, E., and Obermeyer, Z. (2021). The health costs of cost-sharing. Technical report.

Chien, S., Campbell, N., Hayden, O., Hurd, M., Main, R., Mallett, J., Martin, C., Meijer, E., Miu, A., Moldoff, M., Rohwedder, S., & St. Clair, P. (2013). *RAND HRS data documentation*. Version M. RAND Center for the Study of Aging.

Crump, R. K., Hotz, V. J., Imbens, G. W., & Mitnik, O. A. (2008). Nonparametric tests for treatment effect heterogeneity. *The Review of Economics and Statistics*, *90*(3), 389–405. https://doi.org/10.1162/rest.90.3.389

Cubanski, J., Frank, R. G., & Epstein, A. M. (2004). Savings from drug discount cards: Relief for Medicare beneficiaries? *Health Affairs*, *23*(Suppl1), W4–W198. https://doi.org/10.1377/hlthaff.w4.198

Dave, D., & Kaestner, R. (2009). Health insurance and ex ante moral hazard: Evidence from Medicare. *International Journal of Health Care Finance and Economics*, *9*(4), 367–390. https://doi.org/10.1007/s10754-009-9056-4

Dickert-Conlin, S., & Houser, S. (2002). EITC and marriage. *National Tax Journal*, *55*(1), 25–40. https://doi.org/10.17310/ntj.2002.1.02

Einav, L., Finkelstein, A., & Polyakova, M. (2018). Private provision of social insurance: Drug-specific price elasticities and cost sharing in Medicare Part D. *American Economic Journal: Economic Policy*, *10*(3), 122–153. https://doi.org/10.1257/pol.20160355

Fang, H., Keane, M. P., & Silverman, D. (2008). Sources of advantageous selection: Evidence from the Medigap insurance market. *Journal of Political Economy*, *116*(2), 303–350. https://doi.org/10.1086/587623

Gelman, A., & Imbens, G. (2018). Why high-order polynomials should not be used in regression discontinuity designs. *Journal of Business and Economic Statistics*, *37*(3), 1–10. https://doi.org/10.1080/07350015.2017.1366909

Geruso, M., & Layton, T. (2020). Upcoding: Evidence from Medicare on squishy risk adjustment. *Journal of Political Economy*, *128*(3), 984–1026. https://doi.org/10.1086/704756

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License

Goldman, D. P., Joyce, G. F., Escarce, J. J., Pace, J. E., Solomon, M. D., Laouri, M., Landsman, P. B., & Teutsch, S. M. (2004). Pharmacy benefits and the use of drugs by the chronically ill. *JAMA*, *291*(19), 2344–2350. https://doi.org/10.1001/jama.291.19.2344

Goldman, D. P., Joyce, G. F., & Zheng, Y. (2007). Prescription drug cost sharing: Associations with medication and medical utilization and spending and health. *JAMA*, *298*(1), 61–69. https://doi.org/10.1001/jama.298.1.61

Gross, T., Layton, T. J., & Prinz, D. (2022). The liquidity sensitivity of healthcare consumption: Evidence from social security payments. *The American Economic Review: Insights*, *4*(2), 175–190. https://doi.org/10.1257/aeri.20200830

Heckman, J. J., Urzua, S., & Vytlacil, E. (2006). Understanding instrumental variables in models with essential heterogeneity. *The Review of Economics and Statistics*, *88*(3), 389–432. https://doi.org/10.1162/rest.88.3.389

Herkert, D., Vijayakumar, P., Luo, J., Schwartz, J. I., Rabin, T. L., DeFilippo, E., & Lipska, K. J. (2019). Cost-related insulin underuse among patients with diabetes. *JAMA Internal Medicine*, *179*(1), 112–114. https://doi.org/10.1001/jamainternmed.2018.5008

Huber, M., & Lafférs, L. (2022). Bounds on direct and indirect effects under treatment/mediator endogeneity and outcome attrition. *Econometric Reviews*, *41*(10), 1141–1163. https://doi.org/10.1080/07474938.2022.2127077

Huh, J., & Reif, J. (2017). Did Medicare Part D reduce mortality? *Journal of Health Economics*, *53*, 17–37. https://doi.org/10.1016/j.jhealeco.2017.01.005

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*(1), 5–86. https://doi.org/10.1257/jel.47.1.5

Inzucchi, S., Tunceli, K., Qiu, Y., Rajpathak, S., Brodovicz, K., Engel, S., Mavros, P., Radican, L., Brudi, P., Li, Z., Fan, C. P. S., Hanna, B., Tang, J., & Blonde, L. (2015). Progression to insulin therapy among patients with type 2 diabetes treated with sitagliptin or sulphonylurea plus metformin dual therapy. *Diabetes, Obesity and Metabolism*, *17*(10), 956–964. https://doi.org/10.1111/dom.12489

Karter, A. J., Subramanian, U., Saha, C., Crosson, J. C., Parker, M. M., Swain, B. E., Moffet, H. H., & Marrero, D. G. (2010). Barriers to insulin initiation: The translating research into action for diabetes insulin starts project. *Diabetes Care*, *33*(4), 733–735. https://doi.org/10.2337/dc09-1184

Kenkel, D. (2000). Prevention. In A. J. Culyer & J. P. Newhouse (Eds.), *Handbook of health economics* (Vol. 1, pp. 1675–1720). Elsevier. chapter 31.

Li, D., Xie, K., Wolff, R., & Abbruzzese, J. L. (2004). Pancreatic cancer. *The Lancet*, *363*(9414), 1049–1057. https://doi.org/10.1016/s0140-6736(04)15841-8

Liu, S., Mourifié, I., & Wan, Y. (2020). Two-way exclusion restrictions in models with heterogeneous treatment effects. *The Econometrics Journal*, *23*(3), 345–362. https://doi.org/10.1093/ectj/utaa013

MaCurdy, T., Chen, X., & Hong, H. (2011). Flexible estimation of treatment effect parameters. *The American Economic Review*, *101*(3), 544–551. https://doi.org/10.1257/aer.101.3.544

Manski, C. F., & Pepper, J. V. (2000). Monotone instrumental variables: With an application to the returns to schooling. *Econometrica*, *68*(4), 997–1010. https://doi.org/10.1111/1468-0262.00144

Molinari, F. (2020). Microeconometrics with partial identification. *Handbook of Econometrics*, *7*, 355–486.

Nathan, D. M., Buse, J. B., Davidson, M. B., Ferrannini, E., Holman, R. R., Sherwin, R., & Zinman, B. (2009). Medical management of hyperglycemia in type 2 diabetes: A consensus algorithm for the initiation and adjustment of therapy: A consensus statement of the American diabetes association and the European association for the study of diabetes. *Diabetes Care*, *32*(1), 193–203. https://doi.org/10.2337/dc08-9025

Nemeth, E., & Ganz, T. (2023). Hepcidin and iron in health and disease. *Annual Review of Medicine*, *74*(1), 261–277. https://doi.org/10.1146/annurev-med-043021-032816

Nguimkeu, P., Denteh, A., & Tchernis, R. (2019). On the estimation of treatment effects with endogenous misreporting. *Journal of Econometrics*, *208*(2), 487–506. https://doi.org/10.1016/j.jeconom.2018.10.005

Ortigueira, S., & Siassi, N. (2022). The US tax-transfer system and low-income households: Savings, labor supply, and household formation. *Review of Economic Dynamics*, *44*, 184–210. https://doi.org/10.1016/j.red.2021.02.010

Tamer, E. (2010). Partial identification in econometrics. *Annu. Rev. Econ.*, *2*(1), 167–195. https://doi.org/10.1146/annurev.economics.050708.143401

Wettstein, G. (2020). Retirement lock and prescription drug insurance: Evidence from Medicare Part D. *American Economic Journal: Economic Policy*, *12*(1), 389–417. https://doi.org/10.1257/pol.20160560

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Kaliski, D. (2023). Identifying the impact of health insurance on subgroups with changing rates of diagnosis. *Health Economics*, 1–15. https://doi.org/10.1002/hec.4718

10991050, 0, Downloaded from https://onlinelibrary.wiley.com/doi/10.1002/hec.4718, Wiley Online Library on [13/06/2023]. See the Terms and Conditions (https://onlinelibrary.wiley.com/terms-and-conditions) on Wiley Online Library for rules of use; OA articles are governed by the applicable Creative Commons License