UCD CENTRE FOR ECONOMIC RESEARCH

WORKING PAPER SERIES

2023

Effectiveness of Population-Based Hypertension Screening: A Multidimensional Regression Discontinuity Design

Fabrice Kämpfen
University College Dublin

Xavier Gómez-Olivé
University of the Witwatersrand

Owen O'Donnell
Erasmus University Rotterdam

Carlos Riumallo Herl
Erasmus University Rotterdam

WP23/20

August 2023

UCD SCHOOL OF ECONOMICS
UNIVERSITY COLLEGE DUBLIN
BELFIELD
DUBLIN 4

Effectiveness of Population-Based Hypertension Screening: A Multidimensional Regression Discontinuity Design*

Fabrice Kämpfen†
Owen O'Donnell‡
Carlos Riumalló Herl§
Xavier Gómez-Olivé¶

October 10, 2023

Abstract

Population-based screening can prevent disease but also induce false positives to use low-value healthcare. Using data on individuals aged 40+ in rural South Africa and a multidimensional regression discontinuity design, we estimate effects of clinical referral based on blood pressure (BP) above diagnostic thresholds for hypertension. Referral increases hypertension treatment but has no effect on BP after four years, on average. However, for screens that are less likely to be false positives—based on the time of day and air temperature at which BP was measured—we estimate that referral reduces mean systolic BP by 5 mm Hg (3.6%) and raises the probability of achieving BP control by 22 percentage points (44%). These results demonstrate the potential for false positives to lower the average effect of screening.

Keywords: Clinical referral | Blood pressure | False positive | Sub-Saharan Africa

JEL: C21, I12, I18

*We are grateful for the generous support provided by the the Swiss Programme for Research on Global Issues for Development (SNF r4d Grant 400640_160374). We would like to thank Anirban Basu, Jere Behrman, Alberto Ciancio, Pascal Geldsetzer, Christopher Jepsen, Iliana Kohler, Hans-Peter Kohler, and participants at various seminars and conferences for comments.

†University College Dublin. E-mail:fabrice.kampfen@ucd.ie, Corresponding author.

‡Erasmus University Rotterdam. E-mail:odonnell@ese.eur.nl

§Erasmus University Rotterdam. E-mail:riumalloherl@ese.eur.nl

¶MRC/Wits Rural Public Health and Health Transitions Research Unit (Agincourt), University of the Witwatersrand, Faculty of Health Sciences, Johannesburg, South Africa. E-mail:F.Gomez-OliveCasas@wits.ac.za

1 Introduction

Almost 1.3 billion people around the globe suffer from hypertension (Zhou *et al.*, 2021b). Around 8.5 million deaths per year are attributed to this condition (Zhou *et al.*, 2021a), which is the principal risk factor for cardiovascular disease (CVD) (Stanaway *et al.*, 2018)—the largest contributor to the global burden of disease (Vos *et al.*, 2020). The CVD burden is largest, and continues to grow, in low- and middle-income countries (LMICs), which have particularly low rates of diagnosis, treatment and control of hypertension (Gómez-Olivé *et al.*, 2017; Zhou *et al.*, 2021b). Among Sub-Saharan Africans living with hypertension, which is often asymptomatic, only 48% of women and 34% of men are diagnosed, only 29% of women and 22% of men are treated, and only 13% of women and 9% of men have controlled blood pressure (Zhou *et al.*, 2021b). Improving the diagnosis and management of hypertension is of first order importance to global health, with the potential to generate substantial economic benefits (Olsen *et al.*, 2016). We evaluate the effectiveness of population-based screening for hypertension in South Africa, where prevalence is among the highest in the world (Zhou *et al.*, 2021b).

Screening for hypertension requires no more than blood pressure (BP) measurement and is therefore feasible in the community (population-based) at relatively low cost. If people respond to the information provided through screening, then healthcare costs may rise immediately, due to follow-up clinical assessment and prescribed medication, but lifetime costs may fall, due to early detection of disease risk and avoidance of complications from hypertension. There are a number of threats to this positive scenario. First, those screened to be at risk for hypertension may not visit a clinic for formal diagnosis. Second, false positives raise healthcare utilisation and short-run costs without reducing long-run costs through disease avoidance. Third, clinicians may fail to diagnose, counsel, and medicate asymptomatic patients who present with no more than a one-time high BP

reading at screening. Fourth, antihypertensive medicines may be unavailable, unaffordable or imperfectly adhered to. Cost-effectiveness of population-based hypertension screening is not yet established (Schmidt *et al.*, 2020).

Measurement of BP in surveys and referral of respondents with high BP for clinical assessment mimics population-based screening. Like others (Chen *et al.*, 2019; Ciancio *et al.*, 2021; Pedron *et al.*, 2022; Sudharsanan *et al.*, 2020; Zhao *et al.*, 2013), we take advantage of this similarity to estimate effects of referral on BP, as well as diagnosis and treatment, four years after BP measurement in a survey of people aged 40 years and older in a rural province of South Africa (Gómez-Olivé *et al.*, 2018). Applying this general empirical strategy in Malawi, Ciancio *et al.* (2021) estimate that referral raises the probability of hypertension diagnosis by 20 percentage points (pp) and of BP control by 22 pp. This gives grounds for optimism about the effectiveness of population-based hypertension screening in Sub-Saharan Africa (SSA). However, the Malawi study referred people at BP thresholds (systolic/diastolic BP $\geq$ 160/110 mm Hg) that are considerably higher than those (140/90) used in general screening programmes to refer at-risk cases and in clinical settings to diagnose hypertension. It remains to be seen whether, in SSA, it is effective to undertake population-based screening using the conventional 140/90 BP thresholds that can be expected to generate many more referrals and false positives. At these thresholds, 82% of cases initially identified as potentially hypertensive were false positives in a US study (Handler *et al.*, 2015).

Over time, relaxation of criteria used to diagnose chronic conditions, such as hypertension, diabetes and hyperlipidemia, has shifted the marginal diagnosed patient down the respective biomarker distribution (Alalouf *et al.*, 2023; Whelton *et al.*, 2018b). Since there is within individual variation in any biomarker, screening for a condition at the threshold used for diagnosis inevitably produces false positives, as well as false negatives. In Japan, screening for diabetes at the blood sugar threshold used for diagnosis (after multiple measures) is not cost-effective (Iizuka *et al.*, 2021). In more constrained health systems, there is an even higher

opportunity cost of screening at the threshold used for diagnosis and so potentially referring a high proportion of the population to primary care clinics that are less than adequately equipped to manage chronic diseases (Kämpfen *et al.*, 2018).

We use systematic variation in BP with time of day and air temperature (Brook *et al.*, 2011; O'Brien *et al.*, 2018; Whelton *et al.*, 2018a) to distinguish between survey respondents with higher and lower likelihoods of being false positives when screened for hypertension at the 140/90 thresholds. We find that referral at these thresholds is effective in reducing BP only for those who are less likely to be false positives because their BP was measured in conditions associated with lower BP—around the middle of the day (12-3pm) or at higher temperatures ($\geq 80^o F$).

We estimate effects by comparing outcomes on either side of a BP threshold at which a referral letter is issued to a survey respondent. With one exception (Pedron *et al.*, 2022), other evaluations of hypertension screening that use this empirical strategy deploy a unidimensional regression discontinuity design (RDD) (Chen *et al.*, 2019; Ciancio *et al.*, 2021; Dai *et al.*, 2022; Rodriguez-Lesmes, 2021; Sudharsanan *et al.*, 2020; Zhao *et al.*, 2013). One limitation of this approach is that it estimates a local effect around one or each of the systolic BP (SBP) and diastolic BP (DBP) thresholds. Neither effect corresponds to the effect of screening as implemented, which involves using both BP measures simultaneously. Another limitation is that to maintain plausibility of the identification assumption of continuity of the outcome through each threshold under the no-treatment counterfactual, observations that are treated by crossing the other threshold must be dropped. This further compromises external validity and leads to a loss of power.

We use both BP measures and thresholds simultaneously. The only other study to have done this combines the two measures into one running variable (Pedron *et al.*, 2022), which has three limitations. First, the outcome is constrained to change with the running variable in the same way irrespective of whether that variable changes because of SBP or DBP. Second,

the effect at the SBP threshold is not allowed to depend on the level of DBP, and vice versa. Third, any difference in the effects of being referred because each of SBP and DBP crosses the respective threshold is not revealed. We overcome these limitations with semi-parametric multidimensional RDD (MRDD) (Papay *et al.*, 2011; Reardon and Robinson, 2012; Wong *et al.*, 2013). This estimates an effect that is specific to each threshold but also combines the two effects into a weighted average along the frontier formed by the two thresholds. Moreover, this approach allows us to estimate heterogeneity of the treatment effects *along* a specific treatment frontier. As far as we know, this is the first application of MRDD in health economics and one of only a few in any field.

For the full population (aged 40+), we find no evidence that referral reduces BP four years later. Estimated effects on reported hypertension diagnosis and treatment are consistently positive, but they are imprecise and not close to conventional levels of significance. These findings hold when we stratify by sex, age, and proximity to health facilities.

Stratifying by the likelihood of obtaining a BP measurement that is a false positive for hypertension, we find effects only for those with a lower likelihood. For this group, our MRDD estimates indicate that a referral lowers mean SBP by around 5 mm Hg (3.6%, p-value = 0.084) and raises the probability of achieving controlled BP by 22 pp (44%, p-value = 0.013). While our MRDD estimate of the effect on mean DBP is smaller in magnitude (-2.9 mm Hg) and not significant (p-value = 0.129), we estimate a 4.8 mm Hg (5.7%, p-value = 0.089) reduction in this outcome using an alternative estimator that combines the two BP measures into a single running variable. For the same group, we obtain point estimates of referral increasing probabilities of hypertension diagnosis by 12.2 pp (28.3%, p-value = 0.243) and current hypertension treatment by 15.5 pp (55.8%, p-value = 0.119), although neither estimate is statistically significant at conventional levels. Overall, the estimates obtained for this group suggest that the null effects estimated for the full population reflect ineffective referrals of people with a high likelihood of giving a false positive measurement.

A Cochrane Review of evidence on the effectiveness of hypertension screening failed to find any studies that met the inclusion criteria (Schmidt *et al.*, 2020). The evidence that exists is mixed. In Germany, a referral issued to survey respondents above the 140/90 BP thresholds had no effect on each of hypertension diagnosis, behaviour (smoking, alcohol, physical exercise, and body mass index (BMI)) and BP over an average follow-up of 8 years during the 1990s (Pedron *et al.*, 2022). There was no effect on CVD mortality or morbidity over 17 years (ibid). In the US, a warning of high BP issued to Health and Retirement Study (HRS) respondents not previously diagnosed with hypertension is estimated to have raised the probability of hypertension diagnosis and medication by 17 pp (Edwards, 2018). However, such a diagnosis is estimated to have little or no effect on diet (Hut and Oster, 2022; Slade and Kim, 2014). In the UK, advice to consult a family doctor following a survey measurement of BP above 140/85 caused an increase in hypertension diagnosis and medication after two years but had no significant effects on these outcomes or on BP after four years (Rodriguez-Lesmes, 2021).

Evidence from LMICs is somewhat more positive. Besides the estimates from Malawi of large effects on diagnosis and BP of referral at high BP thresholds (Ciancio *et al.*, 2021), there is evidence from China and South Africa of more modest effects on health behaviours and BP (Chen *et al.*, 2019; Dai *et al.*, 2022; Sudharsanan *et al.*, 2020; Zhao *et al.*, 2013). Using a unidimensional RDD, Zhao *et al.* (2013) estimate that crossing the SBP threshold of 140 mm Hg has no effect on the probability of being diagnosed with hypertension but improves diet in China.[1] Another Chinese study uses unidimensional RDD to estimate that crossing the SBP = 140 threshold, and so receiving advice to seek hypertension care and change health

[1]In this study, survey respondents were told their BP measurements but were not systematically issued with a warning or referral at any particular threshold (Zhao *et al.*, 2013). The authors presume that the SBP = 140 threshold is sufficiently well known, such that information on BP level would imply a warning about high BP. Similarly, Iizuka *et al.* (2021) use unidimensional RDD to estimate effects of warnings about pre-diabetes and diabetes, despite the fasting blood sugar thresholds at which warnings were given not necessarily being the same across all observations.

behaviours, caused a reduction in SBP two years later of 6.3-8.3 mm Hg (Chen *et al.*, 2019). There is no significant effect on DBP. The study lacks power to precisely estimate effects on behaviours, although the point estimates suggest that there may have been effects on weight loss, smoking, alcohol consumption and physical exercise but not diet. A third Chinese study finds that a hypertension diagnosis due to SBP above the 140 threshold caused a reduction in fat intake, and an increased likelihood of quitting smoking (Dai *et al.*, 2022). Hypertension diagnosis due to DBP $\geq$ 90 had no such effects.

Previous evidence from South Africa suggests that crossing the SBP = 140 threshold, which triggered a warning of high BP and its health consequences along with advice to seek care, reduced SBP by 4.7 mm Hg for women but had no effect for men, and there was no effect on DBP (Sudharsanan *et al.*, 2020). Both this study and Chen *et al.* (2019) use the SBP threshold to estimate the effect on SBP and the DBP threshold to estimate the effect on DBP. There is no obvious reason for this since the treatment—a warning and advice to seek care—is the same irrespective of which threshold is crossed.

We use MRDD to contribute evidence on the effectiveness of population-based hypertension screening that does not suffer from the limitations of estimates previously obtained either with unidimensional RDD or by combining the two BP measures into one running variable. Using MRDD allows us to estimate more effects—at each threshold and along the frontier of the two—with greater flexibility and power. Our second contribution is to deliver evidence of false positives lowering the average effect of population-based screening at the BP threshold used for hypertension diagnosis.

Beyond these additions to evidence on the effectiveness of population-based hypertension screening, we contribute to a broader evidence base on behavioural response to health information. Increasingly popular prevention and wellness programmes presume that health behaviour, and so healthcare costs, responds to information on disease risk. And yet, evidence from high-income countries points to muted and short-lived responses of behaviour

and biomarkers to diagnoses of diabetes (Alalouf *et al.*, 2023; Gaggero *et al.*, 2022; Iizuka *et al.*, 2021; Kim *et al.*, 2019; Oster, 2018; Slade, 2012), hyperlipidemia (Carrera *et al.*, 2020; Kim *et al.*, 2019) and overweight (Almond *et al.*, 2016; Cook, 2019; Kim *et al.*, 2019). In LMICs, information has been shown to impact health behaviour provided it is comprehensive, credible and delivered to people who are sufficiently educated to process it (Dupas, 2011; Prina and Royer, 2014; Thornton, 2008). We show that in a low-income rural setting with high prevalence of hypertension, information on hypertension risk induces change in behaviour and medication sufficient to lower BP only among those less likely to be false positives.

2 Data

2.1 Setting

Like other countries in SSA, South Africa has high prevalence and low awareness of hypertension (Mills *et al.*, 2016; Murray *et al.*, 2020). It is estimated that almost half (45-48%) of South African adults were hypertensive in 2016 (Kandala *et al.*, 2021; Peer *et al.*, 2021). Of those with hypertension, only 18% of men and 29% of women were aware of their condition. The percentage of those diagnosed with hypertension who were treated has increased substantially, to reach 85% of men and 82% of women in 2016 (Peer *et al.*, 2021). However, only 26% of men and 30% of women who were treated had controlled BP (Peer *et al.*, 2021).

We use data from the first two waves of the *Health and Ageing in Africa: A Longitudinal Study of an Indepth Community (HAALSI)*. This is a sister of the Health and Retirement Study (HRS) and follows a cohort of individuals aged 40 years and older living in the Agincourt Health and Demographic Surveillance Site (Kahn *et al.*, 2012). This is a rural

location in northeast South Africa covering 420 squared kilometers (km), with 31 villages and a predominantly poor population of slightly more than 100,000 people, with 12,875 permanent residents aged 40+ at the 2013 census. Life expectancy was 64.4 years for females and 55.7 years for males in 2009 (Kahn *et al.*, 2012). Hypertension prevalence is similar to the South African average (Jardim *et al.*, 2017) and substantially higher than in other SSA countries (Gómez-Olivé *et al.*, 2017).

One health centre and six clinics are located within the Agincourt surveillance site. Three district hospitals are located 25-60km from the site. Following the Integrated Chronic Diseases Management model (Department of Health, 2010), a nurse is required to assess the medical needs and measure the BP of anyone attending one of these facilities (Ameh *et al.*, 2017; Chang *et al.*, 2019; Mahomed and Asmall, 2015; Mahomed *et al.*, 2014). Patients with a high BP reading on one occasion should be entered into the system as hypertensive and referred to a physician for prescription of medication and lifestyle advice, with the aim of bringing BP below 140/90. While all facilities are expected to have sufficient manpower and stocks to provide hypertension therapy, the care provided can be suboptimal due to non-operational equipment, staff shortages and erratic supply of drugs (Ameh, 2020; Connor *et al.*, 2006; Limbani *et al.*, 2019). However, in the last decade, there have been substantial improvements, particularly with respect to hypertension awareness and the proportion of those with the condition under treatment and receiving appropriate management (Abrahams-Gessel *et al.*, 2023; Ferro *et al.*, 2022; Houle *et al.*, 2021).

2.2 Sample

HAALSI randomly selected a representative sample of 6,281 individuals aged 40+. Of those selected, 391 were ineligible due to death or out-migration, leaving 5,890 potential respondents. Of these, 5,059 (85.9%) participated in the first wave between November 2014

and November 2015. In each selected household, all members aged 40+ were included in the sample. Spouses younger than 40 were not included. In the second wave, conducted between October 2018 and November 2019, 82.6% of the initial sample was contacted and interviewed (Gómez-Olivé *et al.*, 2018). Attrition is not associated with receipt of a referral letter for elevated BP at baseline (Appendix A Table A1).

In each wave, prior to measurement of BP, each respondent was asked if they (a) had ever had their BP measured by a doctor, nurse or other healthcare worker, (b) had ever been told by a doctor, nurse or other healthcare worker that they have high BP or hypertension, (c) have been newly-diagnosed with high BP in the last 12 months, (d) had ever received any type of prescribed treatment for high BP, (e) are currently receiving any type of prescribed treatment for high BP, and, finally, (f) are currently taking any herbal or traditional remedy for high BP.

BP was measured according to the HRS protocol. After a five-minute rest, a trained enumerator used an Omron BP Monitor applied to the left arm to measure BP three times in intervals of two minutes. If the first measurement was above 140/90, the wait was extended to 5 minutes. For each measurement, the enumerator recorded SBP, DBP, pulse, time of the day and reason for any failure to obtain a reading.

The analysis sample consists of 3,304 respondents for whom we have valid BP measurements for both waves. Exclusion from this sample due to refusal to have BP measured at follow-up is also not associated with receipt of a referral letter at baseline (Appendix A Table A1). A majority (56%) of the sample is female and the average age at baseline is about 62 years (Table 1).

2.3 Intervention

In each wave, enumerators were instructed and trained to give a respondent a clinical referral letter (Appendix A Figure A1) for high BP if SBP was at least 140 mm Hg or DBP was at least 90 mm Hg on either the second or the third measurement. A referral could also be issued if rapid tests indicated a potential problem with blood glucose or cholesterol.[2] The enumerator gave the referral letter at the end of the interview if any criterion was fulfilled. If referral was for high BP, the enumerator explained to the respondent that their BP was high and that they should go to a (non-specified) clinic with the letter for further assessment. The letter, which included the contact information of the study manager, indicated the risk factor that was observed but did not give the exact measurement.[3]

The data do identify respondents who were given referral letters. We assume that all enumerators followed the study protocol and so identify respondents with SBP $\geqslant$ 140 or DBP $\geqslant$ 90 on either the second or third measurement in the first wave (baseline) as treated. If there are enumerators who did not follow the protocol, then we are estimating the intention-to-treat effect for a BP screening programme design that is consistent with that protocol.

On presentation at a clinic with a referral letter, a nurse should measure a respondent's BP, register them as hypertensive if it was high and refer them to a doctor for treatment or advice.

[2]In our analysis sample, 218 respondents (6.6%) received a referral letter because of abnormal blood glucose: plasma glucose concentration > 7.0mmol/L if fasting and > 11.0mmol/L if not fasting. Moreover, 121 respondents (3.66%) received a referral because of abnormal cholesterol: total cholesterol > 5.0mmol/L (200 mg/dl) when fasting. In total, 317 respondents (9.59%) received a referral because of either of these conditions. We show our estimates are robust to restricting the sample to respondents who received a referral only because of their BP. The referral letter also mentions Haemoglobin, although this biomarker was not used to trigger issue of a letter.

[3]From the second wave, all respondents received an information sheet with all their biomarker measurements. Separate referral letters were issued to those who fulfilled any criterion, as in the first wave.

2.4 Outcomes

Primary outcomes are mean SBP and mean DBP calculated from the last two of the three BP measurements taken in the second wave (follow-up) and an indicator of hypertension derived from those measurements, $\mathbb{1}(\text{mean SBP} \geqslant 140 \vee \text{mean DBP} \geqslant 90)$. Secondary outcomes are intermediate between the issue of a referral letter at baseline and any effect on the primary outcomes at follow-up. They are indicators constructed from follow-up self-reports of (a) ever having been diagnosed with hypertension, (b) ever been treated for hypertension, and (c) currently taking treatment for hypertension.[4]

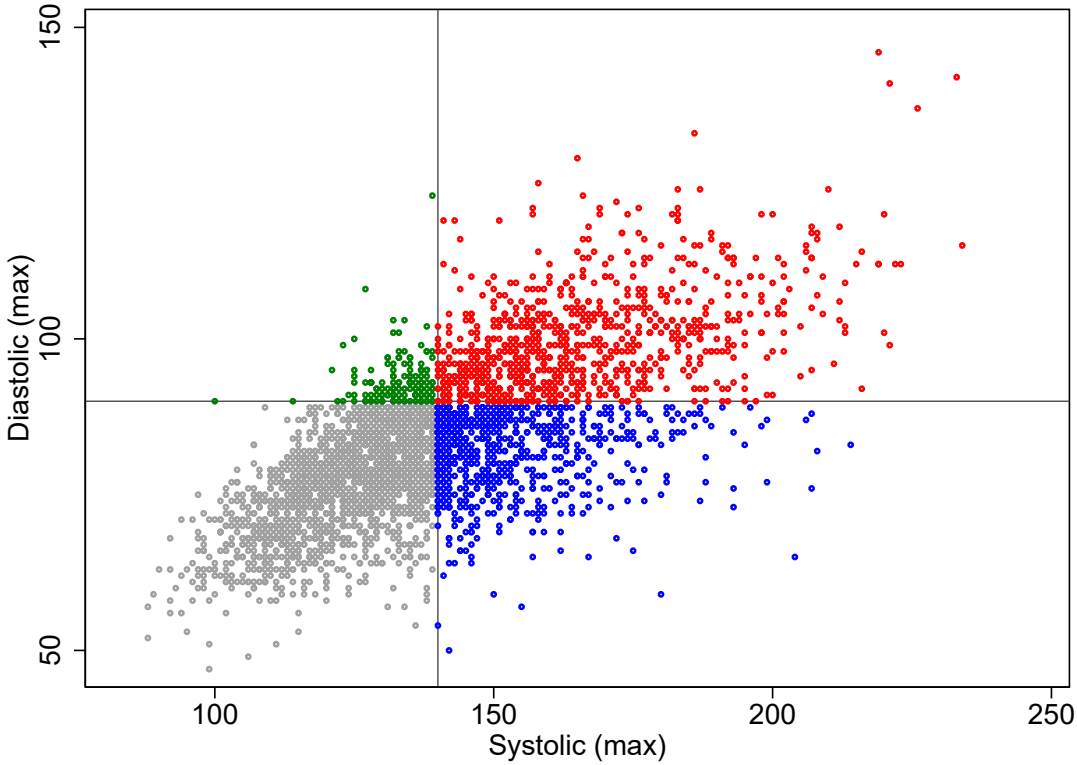
2.5 Descriptives

Figure 1 shows a scatter plot of the maximum of the last two DBP measurements at baseline against the maximum of the last two SBP measurements at baseline. Respondents with a maximum DBP at least 90 or a maximum SBP at least 140 received a referral letter for high BP (according to the protocol). Colours distinguish those who were referred because of (a) DBP only (green), (b) SBP only (blue), and (c) both DBP and SBP (red) from those who were not referred (grey).

Table 1 shows that around one half (50.6%) of the analysis sample received a referral letter at baseline because of high BP. Less than one third (30.3%) had a maximum DBP at or above the 90 mm Hg threshold that was sufficient to qualify for referral (green and red dots in Figure 1). A larger percentage (46.5%) had a maximum SBP at or above the respective 140 mm Hg threshold (blue and red dots in Figure 1). On the basis of mean SBP and DBP

[4]Reporting of hypertension diagnosis and treatment is conditional on previously reporting ever having had BP measured. Any errors in the latter reports will add noise to the reports of diagnosis and treatment. Moreover, there appear to be inconsistencies in these BP-related self-reports across waves. About 14% (N = 324) of the respondents who reported ever having had their BP measured by a medical professional in Wave 1 reported otherwise in Wave 2. Among those who reported having ever been diagnosed with hypertension in Wave 1, 29.5% (N = 413) reported otherwise in Wave 2. In view of such measurement errors, we consider these outcomes as secondary and focus on those that are objectively measured.

Figure 1: Scatter of maximum systolic BP and diastolic BP at baseline



Notes: $N = 3,304$. The x-axis (y-axis) shows maximum SBP (DBP) from last two measurements at baseline. Blue, green, and red dots identify respondents who received a referral letter at baseline because their SBP only, DBP only, and both SBP and DBP were above the respective thresholds (SBP = 140 mm Hg, DBP = 90 mm Hg). Grey dots identify respondents who did not receive a referral letter at baseline.

compared with the respective threshold, 44.3% would be categorized as hypertensive. This probably overestimates hypertension prevalence since a clinical diagnosis may be made on the basis of repeated measurements on multiple occasions. Nonetheless, it is close to the (single-occasion) survey-based prevalence estimate (45-48%) for all of South Africa (Kandala *et al.*, 2021; Peer *et al.*, 2021). Hypertension is extremely common in the population over 40 years in the study setting.

The middle panel of Table 1 shows means of the outcomes at follow-up. Mean SBP and DBP was 128.4 mm Hg and 79.6 mm Hg, respectively. On the basis of these measurements, 29.6% of the sample would be categorised as hypertensive at follow-up. This is a substantial

Table 1: Sample descriptives

	Mean	Std. dev.	N
Baseline			
Mean SBP	137.577	22.699	3304
Mean DBP	82.068	12.280	3304
$\mathbb{1}$(Max DBP $\geqslant 90$)	0.303	0.460	3304
$\mathbb{1}$(Max SBP $\geqslant 140$)	0.465	0.499	3304
Referral = $\mathbb{1}$(Max DBP $\geqslant 90 \vee$ Max SBP $\geqslant 140$)	0.506	0.500	3304
Hypertensive = $\mathbb{1}$(Mean DBP $\geqslant 90 \vee$ Mean SBP $\geqslant 140$)	0.443	0.497	3304
Outcomes (at follow-up)			
Mean SBP	128.410	20.179	3304
Mean DBP	79.596	11.497	3304
Hypertensive = $\mathbb{1}$(Mean DBP $\geqslant 90 \vee$ Mean SBP $\geqslant 140$)	0.296	0.456	3304
Ever told have high BP	0.413	0.492	3296
Ever treated for high BP	0.354	0.478	3296
Currently treated for high BP	0.317	0.466	3295
Controls (at baseline)			
Female	0.559	0.497	3304
Age (years)	61.638	12.534	3281

Notes: SBP and DBP are systolic and diastolic blood pressure, respectively. Max SBP (DBP) represents the maximum and Mean SBP (DBP) is the mean from the last two of three measurements. $\mathbb{1}(.)$ is the indicator function. Main analysis sample consists of 3,304 individuals with non-missing SBP and DBP at baseline and follow-up. Sample sizes are smaller for some variables because of missing values.

decrease in prevalence from the baseline estimate. One explanation is a general increase in diagnosis and treatment (Abrahams-Gessel *et al.*, 2023; Ferro *et al.*, 2022; Houle *et al.*, 2021). Among those who reported at baseline that they had never been told they had high BP, around one fifth (381/1900) reported at follow-up being told they have high BP. Moreover, among those who reported never having received treatment for high BP at baseline, around 16% (359/2196) reported that they had been treated at follow-up.

2.6 False positives

Separation of true positives from false positives would require that all respondents given a referral letter at baseline undergo diagnostic testing for hypertension at a clinic and observation of the results. In the absence of such data, we use systematic variation in BP with time of day (O'Brien *et al.*, 2018) and outdoor air temperature (Barnett *et al.*, 2007; Brook *et al.*, 2011) to categorize respondents by the likelihood that BP measurement at baseline gives a false positive indication of hypertension. BP tends to increase in the late afternoon (Bhalotra *et al.*, 2020; Kawano, 2011; Mancia *et al.*, 1983) and at lower temperatures (Alpérovitch *et al.*, 2009; Barnett *et al.*, 2007; Bhalotra *et al.*, 2020; Modesti, 2013). In the sample, average BP is close to the hypertension thresholds (mean SBP=137.6 and mean DBP=82.1 at baseline—see Table 1). Hence, even small perturbations in BP caused by time of day and temperature could result in misclassification and a high rate of false positives.[5]

The time of each interview is recorded. We obtain air temperature data from the Terrestrial Hydrology Research Group at Princeton University.[6] We use gridded 3-hourly data points (0.5 by 0.5 degrees, which corresponds roughly to 27km by 27km) that we match to each respondent's household location and time of interview.

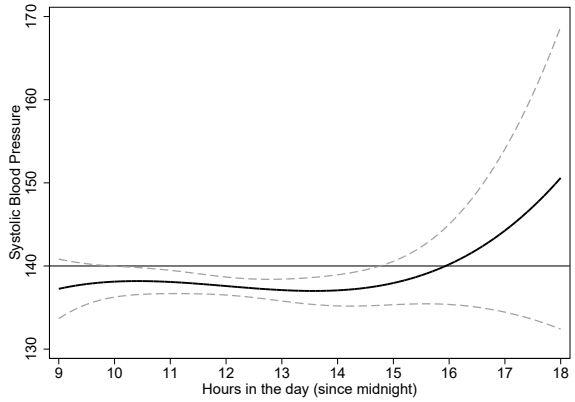
There is substantial variation in both the time of interview and temperature (Appendix A Figure A2). Figure 2 shows BP variation with each of these factors that is similar to the evidence cited above. Each sub-figure plots predicted SBP or DBP against time of interview or air temperature. Predictions are obtained from a linear regression of the mean BP of the last two measurements on a third order polynomial of each of time of interview and

[5]SBP is more variable than DBP (Musini and Wright, 2009). There is likely to be a higher proportion of false positives above the SBP threshold than there is above the DBP threshold. There is discussion about whether higher BP variability, conditional on BP level, is an independent risk factor for CVD (Ebinger *et al.*, 2022; Parati *et al.*, 2018; Schutte *et al.*, 2022).

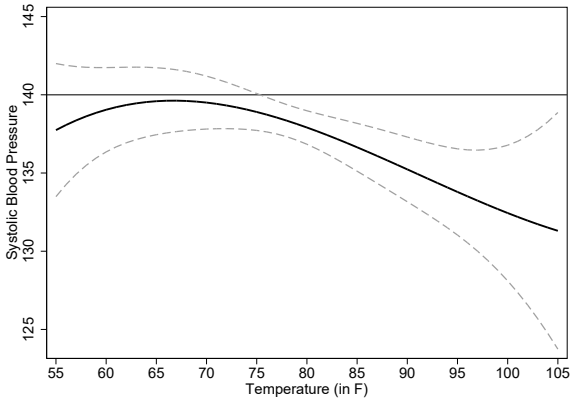
[6]http://hydrology.princeton.edu/data/pgf/v3/. For additional details on how the data are constructed, including the bias correction and downscaling methodology, see Sheffield *et al.* (2006).

Figure 2: BP variation by time of interview and air temperature at baseline

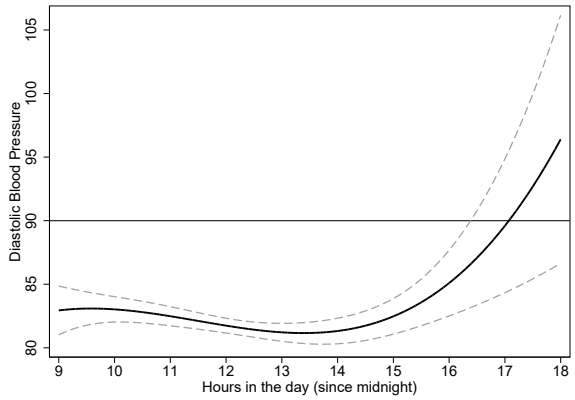
(a) Mean SBP by time of interview



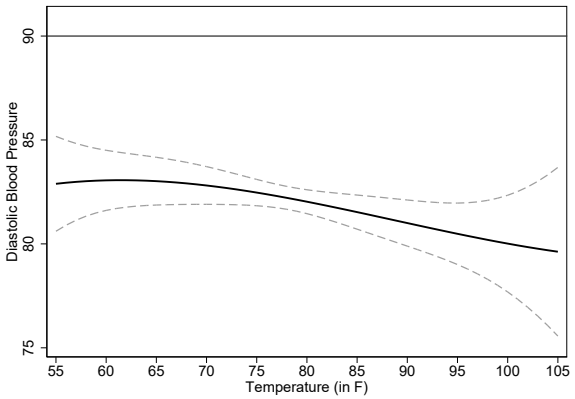
(b) Mean SBP by air temperature



(c) Mean DBP by time of interview



(d) Mean DBP by air temperature



Notes: Predicted mean BP (SBP or DBP) against time of interview and air temperature (Fahrenheit). Predictions obtained from a linear regression of each mean BP measure on a third order polynomial of each of time of interview and air temperature, plus year and month fixed effects. Dashed lines show 95% confidence intervals around estimate of expected value. Horizontal lines indicate thresholds at which a referral letter is issued.

temperature, plus year and month fixed effects.

Both mean SBP and DBP fall slightly from late morning until early afternoon, and then rise. Both types of BP fall as temperature rises. For the average individual, whether the point or interval estimate of expected SBP or DBP lies above the respective hypertension threshold at which a referral letter was issued depends on the time of the interview. For SBP, the interval estimate of the expected value for the average individual lies above the threshold

at low temperatures but not at high temperatures. There is also a strong association between DBP and temperature. However, because the average DBP is some distance below the 90 mm Hg threshold, the covariation with temperature is never sufficient for even the interval estimate of the expected value of DBP for the average individual to cross the threshold. For individuals with above average BP risk factors, the strong DBP-temperature correlation may well be sufficient for expected DBP to be above 90 mm Hg at low temperatures but not at high temperatures.

The association between the predicted probability of receiving a referral letter (by crossing either the SBP = 140 mm Hg or the DBP = 90 mm Hg threshold) and each of time of day and air temperature (Appendix A Figure A3) is very similar to the respective association shown in Figure 2.

We categorise respondents as more likely to give a false positive indication of hypertension if their BP was measured at baseline (a) before noon or after 3pm, OR (b) when the outside air temperature was below 80 degrees Fahrenheit. Respondents are categorised as less likely to give a false positive indication when BP was measured (a) between noon and 3pm, AND (b) when the outside air temperature was above 80 degrees Fahrenheit. Since BP tends to be lower in the early afternoon and at temperatures above 80^o Fahrenheit, if in these environmental conditions BP is above one of the hypertension thresholds, it is less likely to be a false positive. We expect a referral letter to be more effective—raising the probability of being diagnosed with hypertension and being prescribed antihypertensives, and reducing BP—when issued to a respondent who is less likely to be a false positive. Importantly, per the study protocol, the timing of interviews—days and hours—was randomly determined.[7]

[7] Appendix A Table A2 confirms that receipt of a referral letter at baseline does not have any causal effect on the time of interview or the probability of being a false positive at follow-up. This suggests no confounding through sorting along these two dimensions. In the study site, there is unlikely to be systematic variation in interview time arising from certain types of respondents being unavailable during work hours. The unemployment rate is extremely high. A large majority of individuals aged 40+ remain at home throughout the day because they are unemployed or they conduct their business from home.

This is confirmed by observation of only marginal differences in characteristics between the two groups. The normalized (standardized) differences in sex, age and distance to the nearest clinic between those who are less and more likely to be false positive are 0.0585 (0.0827), 0.0289 (0.0410) and -0.0326 (-0.0461), respectively. Each is well below the threshold 0.1 (0.25) often taken as indicative of imbalance between groups (Austin, 2009; Imbens and Rubin, 2015; Imbens and Wooldridge, 2009).

3 Empirical Strategy

We use a regression discontinuity design (RDD) to estimate effects on follow-up outcomes of receiving a referral letter at baseline because BP measurements reach thresholds. The general identification assumption is that respondents in a narrow bandwidth around a threshold are sufficiently similar in observed and unobserved characteristics such that an outcome would evolve smoothly and continuously through the threshold if no referral letters were issued. Any discontinuity in the outcome at the threshold can then be attributed to the causal effect of receiving a referral, which is triggered by BP crossing the threshold. For each outcome, we estimate various treatment effects that differ in the BP threshold used—systolic, diastolic or both—as well as the precise identification assumption and sample selection.

We observe respondents ($i = 1, 2, .., n$) whose BP was measured three times ($t = 1, 2, 3$) at baseline. For each respondent, we have three measurements each of SBP (s_{ti}) and DBP (d_{ti}). A referral letter—the treatment (T)—was given to a respondent if the maximum of the last two SBP measurements was at least 140 mm Hg or the maximum of the last two DBP measurements was at least 90 mm Hg: $T_i = \mathbb{1}(s_i^{max} \geqslant 140 \vee d_i^{max} \geqslant 90)$, where $x_i^{max} = max\,(x_{2i}, x_{3i})$, $x \in \{s, d\}$.

3.1 Unidimensional RDD

As far as we know, all previous RDD evaluations of hypertension screening, except one (Pedron *et al.*, 2022), estimate treatment effects either at only one threshold or separately at the systolic and diastolic thresholds using unidimensional RDD (Chen *et al.*, 2019; Ciancio *et al.*, 2021; Dai *et al.*, 2022; Rodriguez-Lesmes, 2021; Sudharsanan *et al.*, 2020; Zhao *et al.*, 2013). This is one of the strategies we adopt. It is feasible because BP crossing at least one of the thresholds is sufficient for deterministic assignment to treatment—receipt of a referral letter. To ensure that treatment status (but nothing else) differs on each side of the respective threshold, respondents who cross the other threshold must be excluded from the sample used for unidimensional RDD. Exclusion extends to respondents who cross both thresholds since otherwise the continuity assumption that is critical to identification would be implausible. Hence, the respondents used for unidimensional RDD are a subset of those around the respective threshold.

The effect of receiving a referral solely because SBP crosses the 140 mm Hg threshold, τ_s, is estimated along the frontier defined by that threshold and below the DBP threshold, $F_s = (s^{max} = 140, d^{max} < 90)$ (Appendix Figure A4). The respective effect of a referral that is triggered only by DBP crossing its threshold value, τ_d, is estimated along the respective diastolic frontier, $F_d = (d^{max} = 90, s^{max} < 140)$ (Appendix Figure A4). These effects are defined as follows:

$$\begin{aligned} \tau_x &= lim_{x_i^{max} \downarrow c_x} \mathbb{E}[Y_{1i} \mid x_i^{max}] - lim_{x_i^{max} \uparrow c_x} \mathbb{E}[Y_{0i} \mid x_i^{max}] \\ &= \mathbb{E}[Y_{1i} - Y_{0i} \mid x_i^{max} = c_x] \end{aligned} \tag{1}$$

where Y_{1i} and Y_{0i} are potential outcomes with and without treatment, respectively, and $c_x \in \{c_s, c_d\} = \{140, 90\}$ is the relevant threshold.

We estimate these effects non-parametrically. For the main estimates, we use local linear

regression and triangular weights determined by kernel functions centered at the threshold to put more weights on observations closer to it. We assess robustness to using local quadratic regression and uniform weights. We use the Mean Square Error (MSE) optimal bandwidth selector to set the bandwidths that can differ on each side of the threshold (Calonico *et al.*, 2014a,b, 2015). We estimate standard errors using the heteroskedasticity-robust plug-in residuals variance estimator (Calonico *et al.*, 2017; Kolesár and Rothe, 2018).

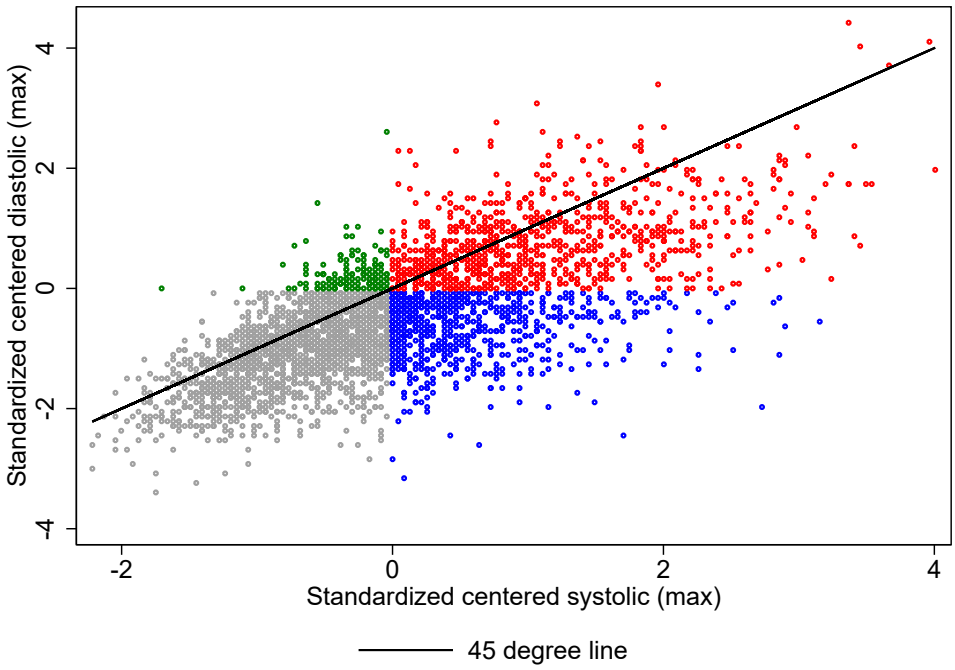
With this approach, there is a substantial loss of information and power resulting from the exclusion of a considerable number of observations from each unidimensional RDD. Moreover, the effects identified are frontier-specific and do not correspond to the overall average effect on those referred on the basis of their SBP or DBP. We deal with both limitations with two alternative strategies that make use of the change in treatment status at both thresholds simultaneously and so use observations in the four quadrants of Figure 1.

3.2 Binding-score RDD

Binding-score RDD (Reardon and Robinson, 2012) is a centering approach (Wong *et al.*, 2013) that reduces the dimensions over which treatment is determined from two to one by creating a single running variable from the two assignment variables, s_i^{max} and d_i^{max}. This has previously been used to evaluate hypertension screening in Germany (Pedron *et al.*, 2022). It is particularly appealing for this purpose because s_i^{max} and d_i^{max} are in the same measurement units (mm Hg), which facilitates interpretation of the meaning and magnitude of the estimated effects.

First, we center each assignment variable on its respective threshold: $\tilde{x}_i^{max} = x_i^{max} - c_x$. Then, we standardize each centered variable on its standard deviation (sd_x) to adjust for differences in the scales of the two measures: $\tilde{x}_i^{std} = \tilde{x}_i^{max} \big/ sd_x$. Figure 3 shows a scatterplot of the two centered and standardized variables—$\tilde{d}_i^{std}$ against $\tilde{s}_i^{std}$. A 45-degree line comes

Figure 3: Scatter of centered and standardized maximum diastolic BP and systolic BP at baseline



Notes: $N = 3,304$. The x-axis and y-axis show centered and standardized maximum systolic BP $\left(\tilde{s}_i^{std}\right)$ and diastolic BP $\left(\tilde{d}_i^{std}\right)$, respectively. Maximums from last two BP measurements at baseline are used. See text for further definitions. Blue, green, and red dots identify respondents who received a referral letter due to systolic BP, diastolic BP, and both above thresholds, respectively. Grey dots identify respondent who did not received a referral letter.

close to cutting the cloud of observations in two. Standardization appears to ensure that a 1-unit deviation from one frontier has similar "intensity" as a 1-unit deviation from the other frontier.

Next, we calculate the maximum centered distance from the two thresholds, $r_i = max(\tilde{s}_i^{std}, \tilde{d}_i^{std})$. This binding score determines treatment status: $T_i = \mathbb{1}(r_i \geqslant 0)$. We then use r_i as the single running variable in a unidimensional RDD to estimate the effect of a referral that is local to the frontier that runs along and connects the SBP and DBP thresholds. That

is,

$$\begin{aligned}\tau_r &= lim_{r_i \downarrow 0}\mathbb{E}[Y_{1i} \mid r_i] - lim_{r_i \uparrow 0}\mathbb{E}[Y_{0i} \mid r_i] \\ &= \mathbb{E}[Y_{1i} - Y_{0i} \mid r_i = 0] \end{aligned} \tag{2}$$

While still a local average effect, this parameter is not as local as the effects obtained from unidimensional RDD applied around each of the systolic and diastolic thresholds separately. It is an overall average treatment effect at the frontier running along the two thresholds. Using observations from all four quadrants defined by the thresholds (Figure 3), including those who are treated because both their SBP and DBP are above the respective thresholds, improves both power and external validity compared with the unidimensional approach.

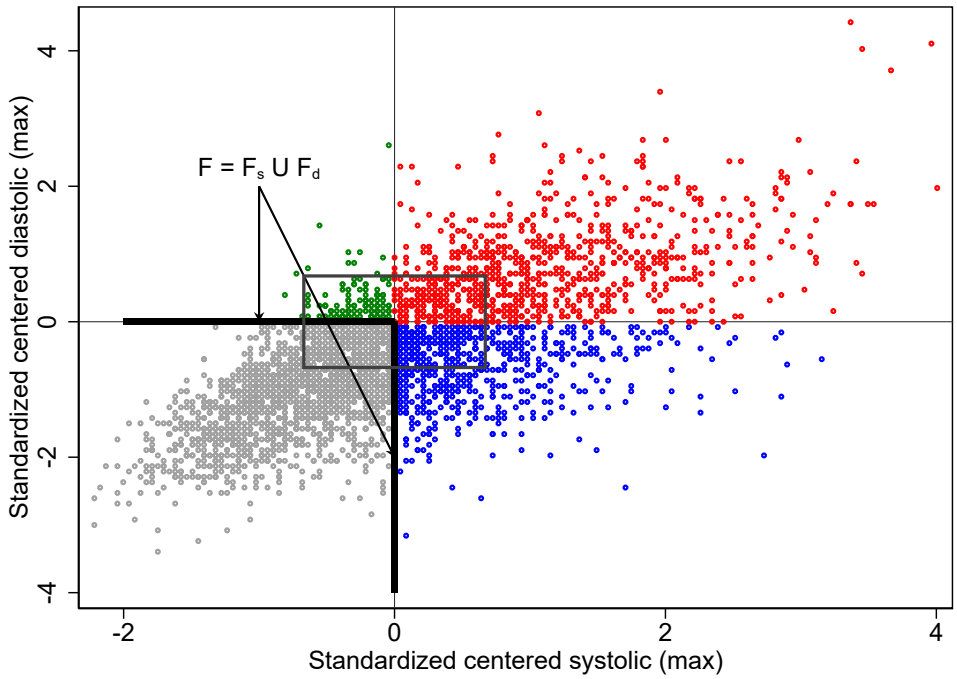
To implement the binding-score approach, we choose the same estimator (local linear regression), kernel function, optimal bandwidths, and standard error estimator as for unidimensional RDD.

3.3 Multidimensional RDD

Derivation of a single running variable from SBP and DBP has some disadvantages. First, it requires that standardization deals with differences in scale, such that a unit change in the running variable has the same consequence for the outcome irrespective of whether it arises from a change in SBP or DBP. Second, it does not capture any difference that may exist between effects of crossing the SBP threshold and the DBP threshold. Third, it assumes that the effect at the SBP threshold is constant irrespective of DBP, and vice versa.

We relax these restrictions with a semi-parametric multidimensional RDD (MRDD) approach that involves parametric estimation of discontinuities in the outcome response surface at frontiers determined by the two thresholds and non-parametric estimation of weights that are used to combine the frontier-specific effects into an overall average effect

Figure 4: MRDD estimation frontier and sample selection window



Notes: $N = 3,304$. The x-axis and y-axis show centered and standardized maximum systolic BP $\left(\tilde{s}_i^{std}\right)$ and diastolic BP $\left(\tilde{d}_i^{std}\right)$, respectively. Maximums from last two BP measurements at baseline. Blue, green, and red dots identify respondents who received a referral letter due to systolic BP, diastolic BP, and both above thresholds, respectively. Grey dots identify respondents who were not referred. MRDD analysis is restricted to observations that are 0.7 standard deviation from the s_i^{max} threshold and the d_i^{max} threshold. These observations are within the small black rectangle. F shows the treatment frontier defined by both the systolic BP (F_s) and the diastolic BP (F_d) frontiers.

(Papay *et al.*, 2011; Reardon and Robinson, 2012; Wong *et al.*, 2013). We are not aware of any other application of a MRDD in health economics.[8]

We use MRDD to estimate the average treatment effect of the referral letter at the frontier $F = F_s \cup F_d = \{(\tilde{s}_i^{std} = 0, \tilde{d}_i^{std} < 0) \cup (\tilde{s}_i^{std} < 0, \tilde{d}_i^{std} = 0)\}$, as well as effects at F_d and F_s separately (Figure 4).

[8] Dell (2010) uses MRDD to estimate effects of forced mining labour on child stunting, as well as on household consumption. She does not deploy the MRDD strategy we use. In part, this is because treatment is determined by a necessary and sufficient condition on two assignment variables (longitude *and* latitude). In our case, treatment is determined by a sufficient condition on SBP *or* another sufficient condition on DBP. There are a few MRDD applications in education (Heller and Slungaard Mumma, 2019; Jepsen *et al.*, 2017).

The average treatment effect at the frontier F, $\tau_m = E[Y_{1i} - Y_{0i} \mid (\tilde{s}_i^{std}, \tilde{d}_i^{std}) \in F]$, is a weighted average of the treatment effects at the frontiers F_s and F_d:

$$\begin{aligned}\tau_m &= E[g_i \mid (\tilde{s}_i^{std}, \tilde{d}_i^{std}) \in F] = \omega_s E[g_i \mid \tilde{s}_i^{std} \in F_s] + \omega_d E[g_i \mid \tilde{d}_i^{std} \in F_d] \\ &= \omega_s \tau_s + \omega_d \tau_d \,, \end{aligned} \tag{3}$$

where $g_i = Y_{1i} - Y_{0i}$, and ω_s and ω_d are probabilities of observing a treated unit on the F_s and F_d frontiers, respectively (Wong *et al.*, 2013). These weights are determined by the joint density of the two assignment variables, $f(\tilde{s}_i^{std}, \tilde{d}_i^{std})$ (Appendix B).

We derive the treatment effects τ_s and τ_d, which are average effects at the respective frontiers and would be the targets of estimation with separate unidimensional RDD, as

$$\begin{aligned}\tau_x &= \frac{\int_{\tilde{z}_i^{std}<0} g_i(\tilde{x}_i^{std}=0, \tilde{z}_i^{std}) f(\tilde{x}_i^{std}=0, \tilde{z}_i^{std}) \mathrm{d}\tilde{z}_i^{std}}{\int_{\tilde{z}_i^{std}<0} f(\tilde{x}_i^{std}=0, \tilde{z}_i^{std}) \mathrm{d}\tilde{z}_i^{std}} \\ &= \frac{\int_{\tilde{z}_i^{std}<0} g_i(\tilde{x}_i^{std}=0, \tilde{z}_i^{std}) f(\tilde{z}_i^{std} \mid \tilde{x}_i^{std}=0) \mathrm{d}\tilde{z}_i^{std}}{\int_{\tilde{z}_i^{std}<0} f(\tilde{z}_i^{std} \mid \tilde{x}_i^{std}=0) \mathrm{d}\tilde{z}_i^{std}} \,, \end{aligned} \tag{4}$$

where, as above, $x \in \{s, d\}$ and $z = d$ if $x = s$ and vice versa, and $g_i(\tilde{x}_i^{std}=0, \tilde{z}_i^{std}) = Y_{1i}(\tilde{x}_i^{std}=0, \tilde{z}_i^{std}) - Y_{0i}(\tilde{x}_i^{std}=0, \tilde{z}_i^{std})$ is the difference in the potential outcomes that are now explicitly allowed to depend on the values of the assignment variables. The second line of eq.(4) follows from writing the joint density, $f(\tilde{x}_i^{std}=0, \tilde{z}_i^{std})$, as the product of the conditional density, $f(\tilde{z}_i^{std} \mid \tilde{x}_i^{std}=0)$, and the marginal density, and cancelling the latter from the numerator and denominator.

We use the non-parametric bivariate kernel density estimator and numerical integration to estimate τ_x and so the weights, ω_s and ω_d (Wong *et al.*, 2013) (Appendix B). We parametrically estimate the outcome response surfaces, $g_i(\tilde{x}_i^{std}, \tilde{z}_i^{std})$, and the discontinuities

in them at $g_i(\tilde{x}_i^{std} = 0, \tilde{z}_i^{std})$, with the following regression of the observed outcome (y_i):

$$y_i = \alpha + \gamma_1 \tilde{s}_i^{std} + \gamma_2 \tilde{d}_i^{std} + \beta_1 \mathbb{1}_{SD} + \beta_2 \mathbb{1}_S \times \tilde{d}_i^{std} + \beta_3 \mathbb{1}_S \times \tilde{d}_i^{std} \times \tilde{s}_i^{std} \tag{5}$$
$$+ \beta_4 \mathbb{1}_D \times \tilde{s}_i^{std} + \beta_5 \mathbb{1}_D \times \tilde{d}_i^{std} \times \tilde{s}_i^{std} + X_i'\delta + \epsilon_i \,,$$

where $\mathbb{1}_{SD} = \mathbb{1}(\tilde{s}_i^{std} \geqslant 0 \vee \tilde{d}_i^{std} \geqslant 0)$, $\mathbb{1}_S = \mathbb{1}(\tilde{s}_i^{std} \geqslant 0 \wedge \tilde{d}_i^{std} < 0)$, $\mathbb{1}_D = \mathbb{1}(\tilde{s}_i^{std} < 0 \wedge \tilde{d}_i^{std} \geqslant 0)$, and X_i is a vector of exogenous and predetermined covariates that includes age (years) and sex. Although these controls are not necessary for identification, they potentially increase precision.

The discontinuities in the expected outcome at the two frontiers are as follows (Appendix B)[9]:

$$g(\tilde{s}_i^{std} = 0, \tilde{d}_i^{std}) = E(Y_i \mid \mathbb{1}_S = 1, \mathbb{1}_D = 0, \tilde{s}_i^{std} = 0, \tilde{d}_i^{std}) - E(Y_i \mid \mathbb{1}_S = 0, \mathbb{1}_D = 0, \tilde{s}_i^{std} = 0, \tilde{d}_i^{std})$$
$$= \beta_1 + \beta_2 \tilde{d}_i^{std} \,, \tag{6}$$
$$g(\tilde{s}_i^{std}, \tilde{d}_i^{std} = 0) = E(Y_i \mid \mathbb{1}_S = 0, \mathbb{1}_D = 1, \tilde{s}_i^{std}, \tilde{d}_i^{std} = 0) - E(Y_i \mid \mathbb{1}_S = 0, \mathbb{1}_D = 0, \tilde{s}_i^{std}, \tilde{d}_i^{std} = 0)$$
$$= \beta_1 + \beta_4 \tilde{s}_i^{std} \,. \tag{7}$$

From equations (6) and (7), it is clear that the specification allows (a) the discontinuities at the two frontiers to differ, and (b) each discontinuity to be heterogeneous along the respective frontier—the outcome shift at the SBP frontier can depend on DBP, and vice-versa. For a small $\mu > 0$, the specification restricts the discontinuities from the point $(\tilde{s}_i^{std} = 0 - \mu, \tilde{d}_i^{std} = 0 - \mu)$ to the points $(\tilde{s}_i^{std} = 0, \tilde{d}_i^{std} = 0)$, $(\tilde{s}_i^{std} = 0 - \mu, \tilde{d}_i^{std} = 0)$, and $(\tilde{s}_i^{std} = 0, \tilde{d}_i^{std} = 0 - \mu)$ to be identical and equal to β_1. Imposition of this restriction makes sense because the treatment is the same—receipt of a referral letter—along the two frontiers.

To estimate the response surface, we use observations within 0.7 of a standard deviation

[9] In the first line of eq.(6), there may appear to be a contradiction in the term that is subtracted since the expectation is taken at $\tilde{s}_i^{std} = 0$ while $\mathbb{1}_S = 0$. This term should be thought of as the expected outcome without treatment if the (SBP) assignment variable were to be extremely close to crossing the treatment threshold. In any case, the effect of the level of the assignment variable cancels out from the two terms of the subtraction whatever the level is. The same reasoning applies to the interpretation of eq.(7).

of each assignment variable from the respective threshold—those within the small black rectangle at the intersection of the two thresholds in Figure 4. We use this selection because 0.7 of a standard deviation roughly corresponds to the average size of the bandwidths optimally derived for the binding-score RDD. This facilitates comparison of the estimates obtained with the two approaches. Moreover, the selection ensures an approximately equal number of observations in all four quadrants defined by the intersection of the two thresholds. Limiting estimation to a narrow boundary around the intersection makes the linear specification more plausible. We show robustness to imposing other sample restrictions.

We feed the regression estimates of the outcome response surface discontinuities—eq.(6) and eq.(7)—into eq.(4), along with the kernel estimates of the conditional densities, and use numerical integration to obtain estimates of the two frontier-specific treatment effects, τ_s and τ_d. We then combine these with the non-parametric estimates of the weights, ω_s and ω_d, in eq.(3) to get an estimate of the overall average effect of a referral letter along the treatment frontier, F. We use a bootstrap (1000 replications) over the whole procedure to estimate standard errors.

3.4 Tests of identification assumption

All three empirical strategies rely on the assumption of no sorting around the thresholds. That is, respondents cannot manipulate receipt of a referral letter through their BP measurements. This is highly plausible given the nature of that measurement and the lack of incentive for either the respondent or the survey enumerator to adjust the measure recorded. Nonetheless, we assess the validity of the assumption by examining histograms of the SBP and DBP assignment variables and conducting density tests (McCrary, 2008). This reveals no graphical evidence of heaping of the densities at the thresholds and no statistical evidence

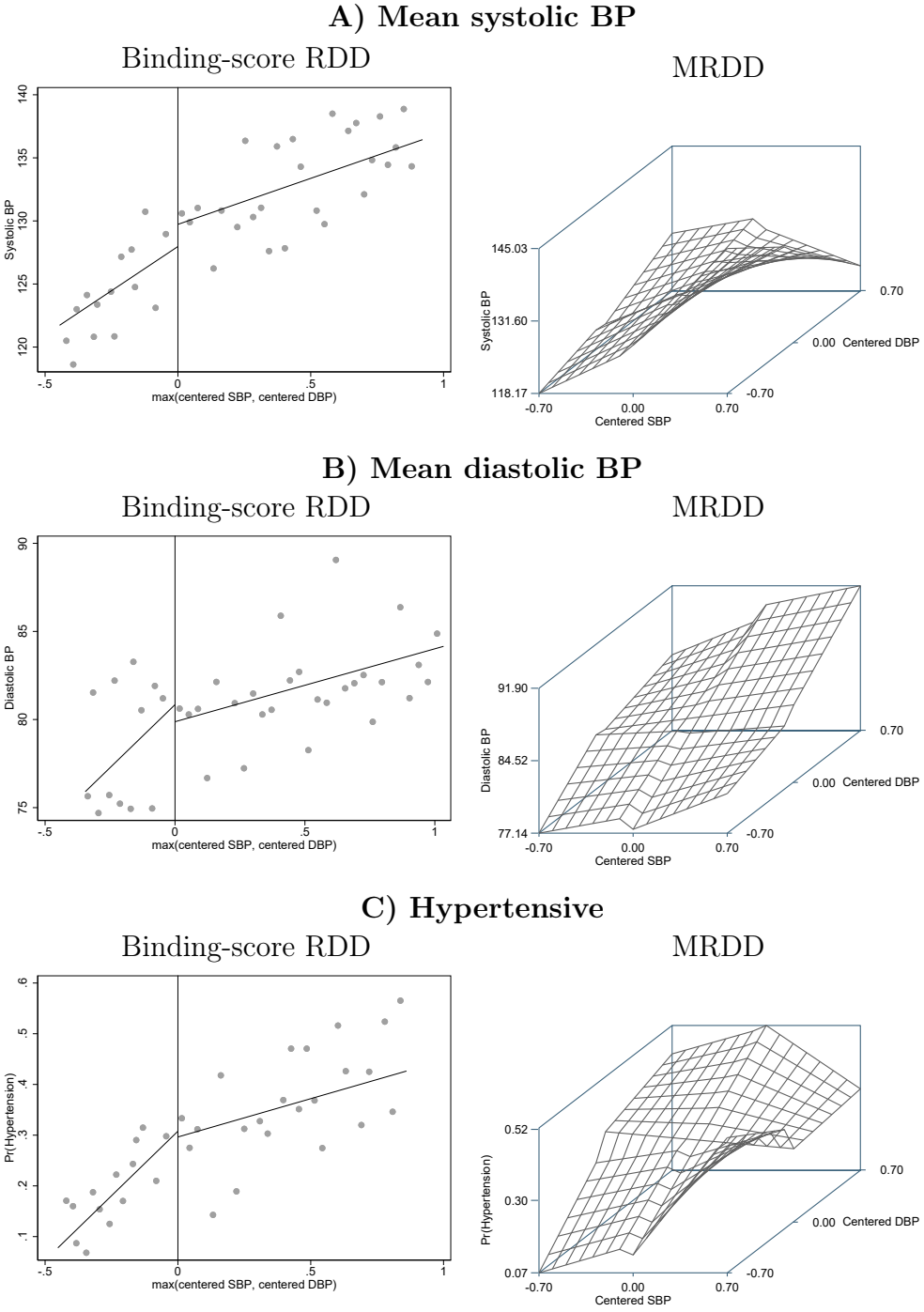
of sorting around them (Appendix A Figure A5). A heat plot of the joint density of SBP and DBP also gives no evidence that suggests any manipulation of recorded BP (Appendix A Figure A6). Furthermore, MRDD, binding-score RDD, and unidimensional RDD all show no discontinuity in age or sex at the thresholds (Appendix A Table A3).

4 Results

4.1 Main estimates

Figure 5 shows that at the BP thresholds for receipt of a referral letter at baseline, there are no clear discontinuities in the primary outcomes—mean SBP, mean DBP and the probability of being hypertensive, all at follow-up. In each binding-score RDD plot (left), a referral is received when the single running variable (r_i)—derived from baseline SBP and DBP—is non-negative. In each MRDD plot (right), a referral is received when either the SBP running variable on the x-axis ($\tilde{s}_i^{std}$) or the DBP running variable on the y-axis ($\tilde{d}_i^{std}$) is non-negative. The (mean) outcome is on the y-axis in each binding-score RDD plot and the z-axis of each MRDD plot.

Figure 5: Discontinuity plots for primary outcomes



Notes: Plots of average outcomes conditional on the running variable(s). Outcomes are defined in Table 1. Non-negative values of the running variable(s) indicate observations that receive a referral letter. For binding-score RDD, on the x-axis is the running variable $r_i = max\left(\tilde{s}_i^{std}, \tilde{d}_i^{std}\right)$. For these plots, we use local linear regression, triangular kernels, and the MSE optimal bandwidth selector. Each dot is the mean of the respective outcome in a given bin. We use optimal bins obtained with variance evenly-spaced method using spacing estimators (Calonico *et al.*, 2014a,b, 2015, 2017). For MRDD, running variables on x-axis and y-axis are $\tilde{s}_i^{std}$ and $\tilde{d}_i^{std}$, respectively. For MRDD plots, each point on the grid represents the predicted value from estimates of eq.(5) at the respective values of the running variables. Predicted values are computed for each 0.1 increment in the two running variables. In all plots, we control for sex and age.

Table 2: Effects of a referral letter on outcomes

	Binding-score RDD		*MRDD*			
	Combined frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)
Primary outcomes						
Mean systolic BP	1.541	1566	-0.530	-0.265	-1.153	945
	(2.164)		(1.819)	(2.021)	(2.603)	
Mean diastolic BP	-1.380	1568	-0.316	-0.391	-0.144	945
	(1.535)		(1.050)	(1.132)	(1.674)	
Hypertensive	-0.025	1537	-0.007	-0.017	0.015	945
	(0.060)		(0.053)	(0.059)	(0.076)	
Secondary outcomes						
Ever told have high BP	0.043	2073	0.081	0.097	0.045	943
	(0.052)		(0.054)	(0.062)	(0.077)	
Ever treated for high BP	0.045	1995	0.089*	0.100*	0.062	943
	(0.053)		(0.052)	(0.059)	(0.075)	
Currently treated for high BP	0.046	1813	0.100**	0.099*	0.102	943
	(0.055)		(0.050)	(0.058)	(0.071)	
Estimated weights (ω_s and ω_d)				0.698	0.302	

Note: See Table 1 for definitions of outcomes. Binding-score RDD estimates of eq.(2) obtained from local linear regression, with triangular kernels and optimal bandwidths on each side of the threshold using the MSE optimal bandwidth selector. MRDD estimates obtained from linear regression estimates of the outcome response surface, eq.(5) and bivariate kernel estimates of the conditional densities and weights that are combined with the outcome response estimates using eq.(4) and, for column (3), (3). Standard errors in parentheses are bootstrapped (1000 repetitions) for MRDD and heteroscedasticity-robust for binding-score RDD. N is the effective number of observations used in estimation. All estimates are from specifications that control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Estimates in the top panel of Table 2 confirm that there is no significant effect of a referral letter on any of the primary outcomes. The binding-score RDD point estimates in column 1 are positive for the effect on SBP and negative for DBP and the probability of being hypertensive. None of these estimates approaches a conventional level of significance.[10] MRDD point estimates of average effects at the combined systolic-diastolic frontier ($F = F_s \cup F_d$) are all negative for the primary outcomes (column 3). Each is smaller in magnitude than the respective binding-score RDD estimate. Again, none is significant. MRDD point estimates differ at the two frontiers (columns 4 and 5). But disaggregation does not reveal

[10]Results are similar when using local quadratic regressions, uniform weights, or when restricting the analysis to individuals who were not diagnosed for hypertension at baseline (Appendix C Figure C1).

any significant estimate of an effect on a primary outcome. Estimates of the weights (ω_s and ω_d), shown in the bottom row of the table, reveal that an estimated effect at the systolic frontier accounts for around 70% of the MRDD estimate of the average effect at the combined frontier. The remainder is the contribution of the estimated effect at the diastolic frontier.

The second panel of Table 2 shows estimated effects on secondary outcomes.[11] Notably, all estimates from both estimators are positive, which suggests that receipt of a referral letter may have been effective in raising the probability of hypertension diagnosis and treatment. The MRDD estimates of average effects at the combined frontier are larger and, despite reliance on a smaller estimation sample, as precise as the respective binding-score RDD estimates. The MRDD estimates of effects on hypertension treatment at the combined and systolic frontiers are marginally significant (p-value < 0.1). These estimates indicate that referral increases the probabilities of ever having been treated for hypertension and of receiving treatment at the time of the follow-up survey by 9 pp and 10 pp, respectively. The estimated effect on ever being treated is larger at the systolic frontier than it is at the diastolic frontier, although the estimates at the two frontiers do not differ significantly.

When we use an alternative parametric specification of the outcome response surfaces, MRDD estimates continue to indicate null effects on primary outcomes and, even more than before, positive effects on secondary outcomes (Appendix C Table C2).[12] To produce the main MRDD estimates, we include observations within 0.7 of a standard deviation (SD) of each running variable from the respective threshold. When we widen this interval by including those up to 1 SD below the diastolic threshold, we continue to get estimates of null effects on the primary outcomes, while the positive estimated effects on the secondary

[11] Corresponding discontinuity plots are in Appendix C Figure C1.

[12] The alternative specification is

$$y_i = \alpha + \gamma_1 \tilde{s}_i^{std} + \gamma_2 \tilde{d}_i^{std} + \beta_1 \mathbb{1}_{SD} + \beta_2 \mathbb{1}_S \times \tilde{d}_i^{std} + \beta_3 \mathbb{1}_D \times \tilde{s}_i^{std} + \beta_4 \mathbb{1}_{SD} \times \tilde{d}_i^{std} \times \tilde{s}_i^{std} + X_i'\delta + \epsilon_i \,.$$

outcomes get smaller in magnitude and weaken in significance (Appendix C Table C2).[13]

Null effects also persist when we restrict the analysis sample to individuals who received a referral letter only because of their BP and not because of their blood glucose or cholesterol (Appendix C Table C7). This suggests that the null estimated effects are not due to dilution of true effects through any BP treatment that may be given to respondents referred on the basis of another biomarker.

Full sample estimates of null effects on primary outcomes do not hide significant effects in sub-populations defined by sex, age (below/above 60 years) and distance from the nearest health facility (below/above median) (Appendix C Tables C3, C4, C5 and C6).[14]

MRDD estimates of effects on secondary outcomes are generally larger and more precise for females and older people (Appendix C Tables C4 and C5). These estimates indicate large and significant effects on secondary outcomes for those living closer to health facilities, which is as expected (Appendix C Table C6). However, positive and significant effects in this sub-population are not obtained with binding-score RDD.

Unidimensional RDD at each of the systolic and diastolic thresholds separately also gives null effects on the primary outcomes (Appendix C Table C8). At the diastolic threshold, there are positive and significant effects on secondary outcomes.

4.2 Heterogeneity by false positive likelihood

For respondents who are more likely to be false positives, there is no clear evidence of discontinuities in primary outcomes at the thresholds for receipt of a referral letter (Figure

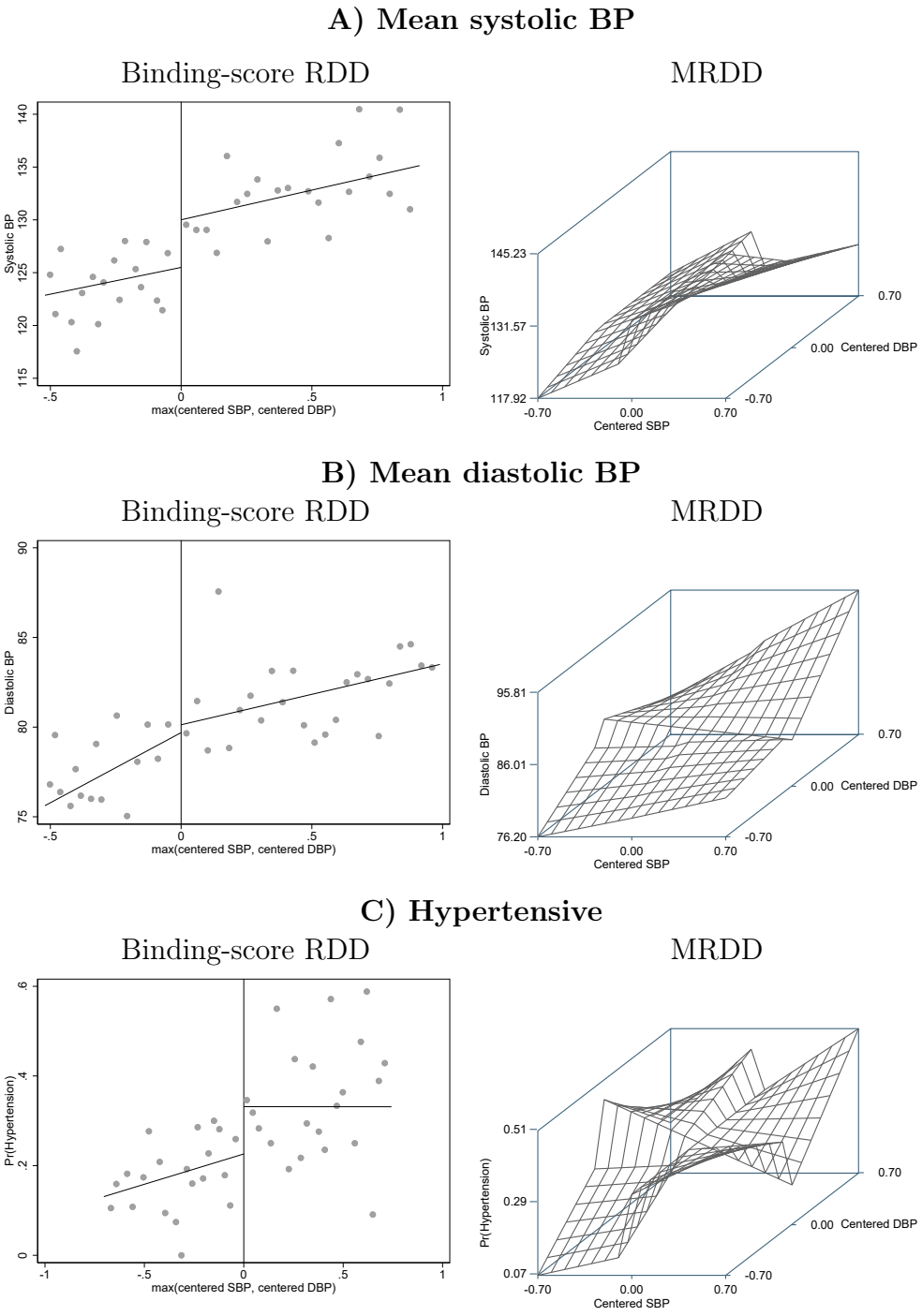
[13]This sample selection is shown by the black rectangle in Appendix A Figure A7. Since relatively few individuals are treated only because their DBP is above the respective threshold (green dots in Figures 4 and A7), if we were to widen the window to also include observations 1 SD below the SBP frontier and/or 1 SD above the DBP frontier, it would be difficult to consistently and convincingly estimate the response surface of all quadrants of Figure A7.

[14]The binding-score RDD estimator gives a counter intuitive positive and significant effect on SBP of those located above the median distance from a health facility. However, this is not confirmed by the respective MRDD estimate and the binding-score RDD point estimate on DBP is negative for this sub-population.

6).[15] The estimates in the top panel of Table 3 confirm that there is no significant negative effect on any of these outcomes using this sample. In fact, binding-score RDD gives counterintuitive positive and significant effects on mean SBP and the probability of being above the hypertension thresholds at follow-up. One possible explanation is that if false positives do indeed comprise a large proportion of the sub-sample used to obtain these estimates, then many in the sample may have relaxed their health behavior, contributing to a rise in BP, on learning that they are not hypertensive after all. However, this may read too much into the estimates since the respective MRDD estimates are not significantly positive. The estimated effects on all the secondary outcomes are positive, but none are significant.

[15] Appendix C Figure C2 in the Appendix shows corresponding graphs for the secondary outcomes.

Figure 6: Discontinuity plots for primary outcomes – high false positive likelihood



Notes: Plots as for Figure 5 except here we use only respondents for whom, at baseline, (a) time of interview was $<$ 12 noon or $>$ 3pm or (b) outdoor air temperature $<$ 80^o Fahrenheit.

Table 3: Effects of a referral letter on outcomes – high false positive likelihood

	Binding-score RDD		*MRDD*			
	Combined frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)
Primary outcomes						
Mean systolic BP	5.000**	1089	2.159	3.106	-0.164	622
	(2.535)		(2.315)	(2.613)	(3.312)	
Mean diastolic BP	0.565	1140	1.096	0.809	1.798	622
	(1.579)		(1.292)	(1.388)	(2.066)	
Hypertensive	0.114*	1133	0.102	0.104	0.096	622
	(0.065)		(0.064)	(0.073)	(0.092)	
Secondary outcomes						
Ever told have high BP	0.036	1307	0.051	0.069	0.006	622
	(0.065)		(0.064)	(0.076)	(0.084)	
Ever treated for high BP	0.053	1319	0.089	0.094	0.077	622
	(0.063)		(0.062)	(0.074)	(0.082)	
Currently treated for high BP	0.047	1176	0.086	0.085	0.088	622
	(0.067)		(0.060)	(0.073)	(0.081)	

Note: As Table 2 (see notes to that table), except here we use only respondents for whom, at baseline, (a) time of interview $<$ 12 noon or $>$ 3pm or (b) outdoor air temperature $<$ 80^o Fahrenheit.

Figure 7 shows discontinuity plots obtained from respondents who we categorise as less likely to be false positives. Both the binding-score RDD and MRDD approaches reveal drops in the three primary outcomes at the BP frontiers where a referral letter was issued. From the MRDD plots, it appears that the effects are relatively constant along the systolic frontier, whereas the discontinuity along the diastolic frontier decreases as systolic BP increases (Panel B).[16]

The estimates in Table 4 confirm that receipt of a referral letter at baseline appears to have lowered the BP of individuals who are less likely to have been false positives. The MRDD estimates indicates a referral reduced mean systolic BP by about 5 mm Hg (p-value=0.084), or 3.6% of the baseline mean for this sample. This effect appears to mainly come from the effect of referral as a result of crossing the systolic frontier (-6.5 mm Hg, p-value=0.012). The binding-score RDD estimate of the effect on systolic BP is smaller

[16]See Appendix C Figure C3 for corresponding graphs for the secondary outcomes.

in magnitude and not statistically significant. Binding-score RDD indicates a reduction of about 4.8 mm Hg in DBP, while the MRDD gives a smaller and not significant estimate.

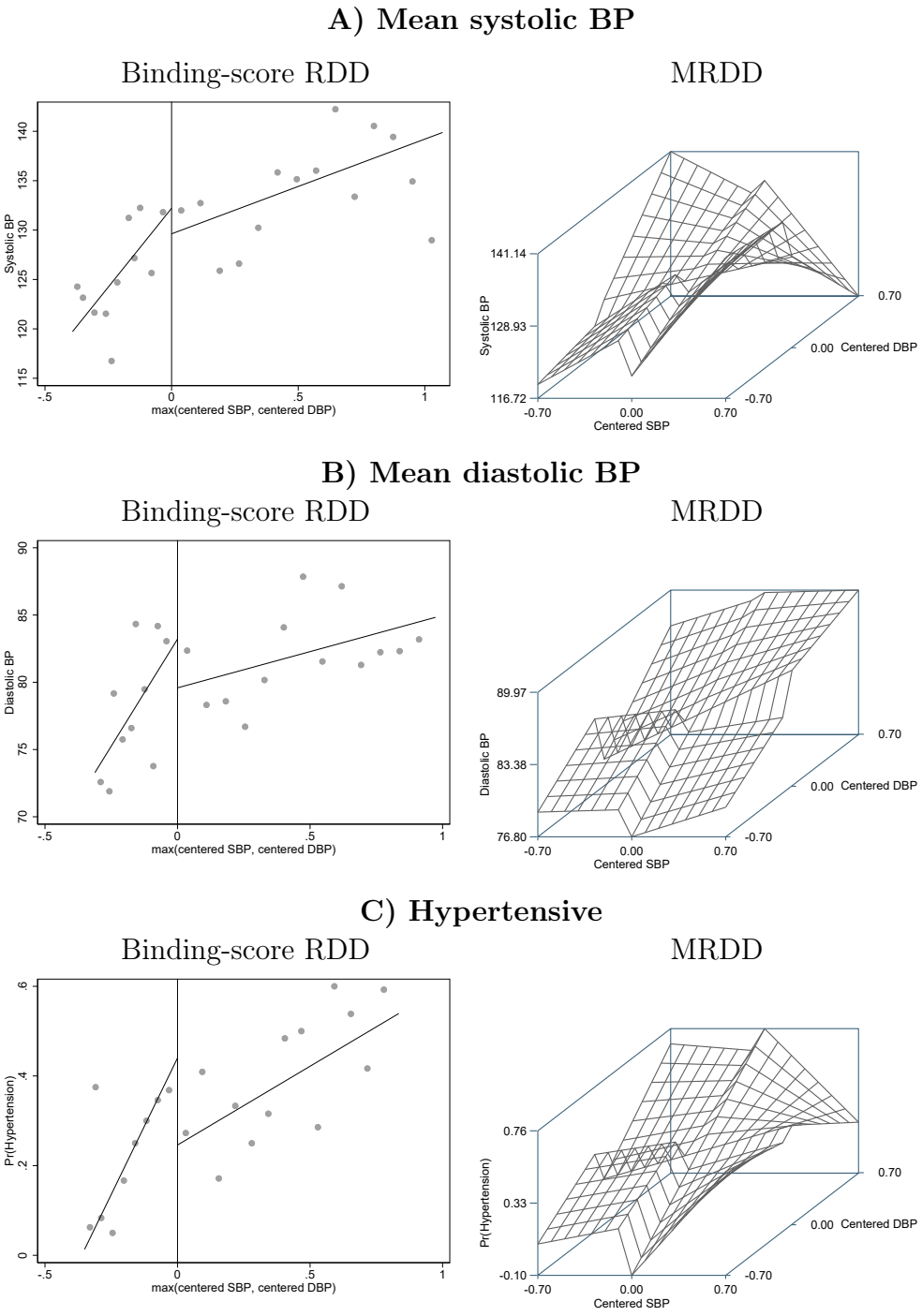
The two estimators both estimate a significant reduction of about 22-24 pp in the probability of being hypertensive at follow-up (p-values = 0.013 (MRDD) and 0.043 (Binding-score)). This is around 44-48% of the mean prevalence of hypertension in this sample at baseline. The MRDD estimate again reveals that most of this effect comes from referral as a result of crossing the systolic frontier (β = −0.239, p-value= 0.013). These estimated effects on BP control are very similar to the estimate from Malawi of a 22 pp drop in the probability of being hypertensive caused by referral due to crossing a higher systolic BP threshold of 160 mm Hg (Ciancio *et al.*, 2021).

The estimated effects on the secondary outcomes obtained using respondents who are less likely to be false positives are consistently positive, larger at the systolic frontier (using MRDD) and larger using MRDD than binding-score RDD (Table 4). However, none of these estimates are statistically significant at conventional levels.

The magnitudes of the MRDD estimates obtained from both samples—those more and less likely to be false positives—are robust to using an alternative specification of the outcome response surfaces (Appendix C Tables C9 and C10)[17] and to widening the windows of observations around the frontier (Appendix C Tables C11 and C12). In both robustness checks, significance of the estimated effect on mean systolic BP falls below conventional levels, but the reduction in the probability of being hypertensive remains statistically significant. Finally, estimates of effects in both sub-samples defined by false positive likelihood are robust to excluding respondents who received a referral letter (also) because their blood glucose or cholesterol exceeded some threshold (Appendix C Tables C13 and C14).

[17] See specification reported in footnote 12

Figure 7: Discontinuity plots for primary outcomes – low false positive likelihood



Notes: Plots as for Figure 5 (see notes to that figure) except here we use only respondents for whom, at baseline, (a) time of interview was ≥ 12 noon and ≤ 3pm and (b) outdoor air temperature ≥ 80^o Fahrenheit.

Table 4: Effects of a referral letter on outcomes – low false positive likelihood

	Binding-score RDD		*MRDD*			
	Combined frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)
Primary outcomes						
Mean systolic BP	-3.802	536	-4.988*	-6.488**	-1.755	299
	(3.750)		(2.890)	(3.274)	(4.446)	
Mean diastolic BP	-4.795*	475	-2.878	-2.582	-3.518	299
	(2.818)		(1.897)	(2.150)	(3.162)	
Hypertensive	-0.242**	466	-0.220**	-0.239**	-0.181	299
	(0.121)		(0.089)	(0.096)	(0.149)	
Secondary outcomes						
Ever told have high BP	0.053	633	0.122	0.128	0.111	297
	(0.098)		(0.105)	(0.125)	(0.158)	
Ever treated for high BP	0.058	632	0.097	0.125	0.037	297
	(0.094)		(0.101)	(0.121)	(0.155)	
Currently treated for high BP	0.073	639	0.155	0.145	0.176	297
	(0.091)		(0.099)	(0.119)	(0.151)	

Note: As Table 2 (see notes to that table), except here use only respondents for whom, at baseline, (a) time of interview was ≥ 12 noon and ≤ 3pm and (b) outdoor air temperature ≥ 80^o Fahrenheit.

5 Discussion

We find that referral for clinical assessment after a measurement of blood pressure (BP) that is above the diagnostic thresholds used for hypertension, as would happen in a population-based screening programme, does not lower BP averaged over all those screened. Referral does appear to raise the likelihood of being treated for hypertension, and possibly of being diagnosed, although the point estimate of the latter effect is not significant at conventional levels. The lack of effect on BP despite the increase in hypertension treatment suggests that average treatment effectiveness may be muted by treatment of false positives, although imperfect treatment adherence could also explain this finding.

When we limit attention to baseline BP measurements that are not taken at times of the day and temperatures associated with higher BP, and so are less likely to generate false

positives, we find clear evidence that referral reduces mean BP and substantially increases the likelihood of having controlled BP. The magnitude of the estimated effect on BP control is very close to the estimate obtained in another low-income, rural setting in Sub-Saharan Africa with high prevalence of hypertension, and where referral was made at BP thresholds above those used for diagnosis (Ciancio *et al.*, 2021).

While raising the BP thresholds at which screened cases are referred for clinical assessment is likely to increase the average effect of a programme, it is not necessarily an optimal policy as it would reduce false positives but raise false negatives. The optimal threshold depends on the relative cost of each type of error (Phelps and Mushlin, 1988). This study does not deliver evidence on these costs, and so we cannot draw conclusions about the optimal threshold. We can infer that a null average effect of screening may hide a positive effect among those less likely to be false positives. It would be a mistake to conclude that a screening programme is not (cost-)effective because it does not reduce a risk factor on average. If it were to reduce the risk factor for some and health benefits are disproportionate to the risk factor reduction, then the programme may be cost-effective if the benefits to those now facing lower risks are large relative to programme costs.

While raising a BP screening threshold would reduce false positives at the cost of increasing false negatives, organising a programme to avoid screening for hypertension at times of day and temperatures associated with higher BP could reduce the first type of error without increasing the second. Costs arising from interruptions to testing would need to be set against the benefits of reducing false positives. But this cost-benefit analysis would be relatively straightforward.

Our findings of a null average effect in the full population and positive effects only for those who are less likely to be false positives are consistent with results from other evaluations of screening for chronic disease. Iizuka *et al.* (2021) find that mandatory health checkups in Japan improve health outcomes, and are cost-effective, only for high-risk individuals. Kim

et al. (2019) find that information provided through the Korean National Health Screening Programme on risk of diabetes, obesity, and hyperlipidemia has limited or no impact on healthcare utilisation and health behaviours at medium risk thresholds for all three conditions and at high risk thresholds for obesity and hyperlipidemia. There are impacts on those identified to be at high risk of diabetes, which is the only threshold at which risk information is supplemented by active prompting to go for diagnostic testing and treatment. Exposure to a general health screening programme in Austria is found to increase healthcare costs in the short run, reduce them in the medium run, and have no impact in the long run (Hackl *et al.*, 2015). Costs increase overall without any effect on health, suggesting that screening induces false positives to use healthcare.

The use of MRDD allows us to identify effects from variation in exposure to referral around both the systolic and diastolic BP thresholds simultaneously. This adds power and external validity compared with previous evaluations of hypertension screening that have adopted unidimensional RDD (with one exception) and so estimate an effect that is local to only one of the thresholds. With MRDD, we also avoid restrictions that binding-score RDD—used by the one evaluation that does not use unidimensional RDD—imposes as a result of combining both BP measures into one running variable. The method also allows us to estimate an effect at each threshold, as well as their weighted average effect. In addition to heterogeneity across frontiers, MRDD makes it possible to uncover heterogeneity along a specific treatment frontier.

Application of MRDD pays off. Estimates obtained from this design reveal that, among those who are less likely to be false positives, the effects of crossing the diastolic BP frontier on diastolic BP decrease as systolic BP increases, whereas the effects of the screening on the probability of being hypertensive at follow-up are relatively constant along the two frontiers (Figures 7 Panels B and C). Such heterogeneity in screening effects cannot be detected using the point-estimation binding-score approach. MRDD gives estimates of effects on secondary

outcomes that are approximately twice as large and more precise than the corresponding binding-score RDD estimates. MRDD reveals that the negative effects on BP for those less likely to be false positives are mainly driven by effects at the systolic frontier.

One limitation of this study is that we do not know if each referred respondent attended a clinic, as advised. Consequently, we cannot assess the extent to which the null effect in the full population is due to insufficient response of individuals to information or insufficient clinical advice, diagnosis and medication given to those who do respond. Information may fail to generate health gains if it is not accompanied by active encouragement for high risks to use healthcare and supply-side efforts to deliver high-value care (Iizuka *et al.*, 2021). A randomised experiment in the Philippines found that going for a check-up at a clinic responsible for CVD risk screening increased the probabilities of having BP measured and receiving medical advice, but it did not increase the likelihood of diagnosis or medication of hypertension (Capuno *et al.*, 2021). Kim *et al.* (2019) argue that health effects of the information provided by screening programmes are likely to be modest without follow-up supply-side interventions. There is evidence from the UK that a diagnosis of hypertension reduces CVD morbidity, has a large negative effect on smoking and a smaller effect on improved diet (Bhalotra *et al.*, 2020). Importantly, these effects are obtained in the context of a primary healthcare system that incentivizes general practitioners to actively manage chronic conditions.

A second limitation is that we cannot identify false positives, nor can we estimate the proportion of false positives among those who are referred. We use systematic variation in BP with time of day and temperature to categorise screens by the likelihood of generating a false positive. This categorisation proves useful in revealing heterogeneity in the effects of referral. But it does not allow calculation of a false positive rate that could be used, with other parameters, to design the optimal screening programme (Phelps and Mushlin, 1988).

A third limitation is that our secondary outcomes (diagnosis and treatment) are self

reported and so are likely to contain measurement error that may explain why we consistently obtain positive estimates that are only significant at conventional levels when using MRDD at the combined systolic-diastolic frontier with the full sample. The positive but not statistically significant effects obtained from other analyses, particularly that using the sample that is likely to include a smaller proportion of false positives, may be due to imprecision arising from a smaller sample combined with large measurement error (see footnote 4).

6 Conclusion

Notwithstanding its limitations, this study contributes needed evidence on the effectiveness of population-based hypertension screening in LMICs. Using MRDD to estimate the effect of screening as it is implemented through reference to systolic and diastolic BP simultaneously, we show that a null average effect over the full target population plausibly results from ineffective screens that are more likely to produce false positives offsetting a statistically and clinically significant effect of screens that are less likely to give false positives. This evidence can potentially be used to improve the design of hypertension screening programmes and so slow the growing burden of cardiovascular disease in LMICs.

References

Abrahams-Gessel, S., Gómez-Olivé, F.X., Tollman, S. et al (2023). Improvements in hypertension control in the rural longitudinal haalsi cohort of South African adults aged 40 and older, from 2014 to 2019. *American Journal of Hypertension*, **36**(6), 324–332.

Alalouf, M., Miller, S. and Wherry, L.R. (2023). What difference does a diagnosis make? Evidence from marginal patients. *American Journal of Health Economics*.

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

Alpérovitch, A., Lacombe, J.M., Hanon, O. et al (2009). Relationship between blood pressure and outdoor temperature in a large sample of elderly individuals: the three-city study. *Archives of Internal Medicine*, **169**(1), 75–80.

Ameh, S. (2020). Evaluation of an integrated hiv and hypertension management model in rural South Africa: a mixed methods approach. *Global Health Action*, **13**(1), 1750216.

Ameh, S., Klipstein-Grobusch, K., D'ambruoso, L. et al (2017). Quality of integrated chronic disease care in rural South Africa: user and provider perspectives. *Health Policy and Planning*, **32**(2), 257–266.

Austin, P.C. (2009). Using the standardized difference to compare the prevalence of a binary variable between two groups in observational research. *Communications in Statistics-Simulation and Computation*, **38**(6), 1228–1234.

Barnett, A.G., Sans, S., Salomaa, V. et al (2007). The effect of temperature on systolic blood pressure. *Blood Pressure Monitoring*, **12**(3), 195–203.

Bhalotra, S.R., Delavande, A., Fisher, P. et al (2020). The impact of a personalised blood pressure warning on health outcomes and behaviours. Technical report, ISER Working Paper Series.

Brook, R.D., Weder, A.B. and Rajagopalan, S. (2011). "Environmental hypertensionology" the effects of environmental factors on blood pressure in clinical practice and research. *The Journal of Clinical Hypertension*, **13**(11), 836–842.

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

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

Calonico, S., Cattaneo, M.D. and Titiunik, R. (2015). rdrobust: An R package for robust nonparametric inference in regression-discontinuity designs. *R Journal*, **7**(1), 38–51.

Calonico, S., Cattaneo, M.D., Farrell, M.H. et al (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal*, **17**(2), 372–404.

Calonico, S., Cattaneo, M.D. and Farrell, M.H. (2018). Coverage error optimal confidence intervals for local polynomial regression. *arXiv preprint arXiv:1808.01398*.

Capuno, J., Kraft, A. and O'Donnell, O. (2021). Effectiveness of clinic-based cardiovascular disease prevention: A randomized encouragement design experiment in the Philippines. *Social Science & Medicine*, **283**, 114194.

Carrera, M., Hasan, S.A. and Prina, S. (2020). Do health risk assessments change eating habits at the workplace? *Journal of Economic Behavior & Organization*, **172**, 236–246.

Cattaneo, M.D., Jansson, M. and Ma, X. (2018). Manipulation testing based on density discontinuity. *The Stata Journal*, **18**(1), 234–261.

Cattaneo, M.D., Jansson, M. and Ma, X. (2020). Simple local polynomial density estimators. *Journal of the American Statistical Association*, **115**(531), 1449–1455.

Chang, A.Y., Gómez-Olivé, F.X., Manne-Goehler, J. et al (2019). Multimorbidity and care for hypertension, diabetes and HIV among older adults in rural South Africa. *Bulletin of the World Health Organization*, **97**(1), 10.

Chen, S., Sudharsanan, N., Huang, F. et al (2019). Impact of community based screening for hypertension on blood pressure after two years: regression discontinuity analysis in a national cohort of older adults in China. *BMJ*, **366**, l4064.

Ciancio, A., Kämpfen, F., Kohler, H.P. et al (2021). Health screening for emerging non-communicable disease burdens among the global poor: Evidence from sub-Saharan Africa. *Journal of Health Economics*, **75**, 102388.

Connor, M., Hopkins, T., Tollman, S. et al (2006). Blood pressure-measuring devices in rural South Africa: an audit conducted by the SASPI team in the Agincourt field site: cardiovascular topics. *Cardiovascular Journal of South Africa*, **17**(4), 192–196.

Cook, W. (2019). The effect of personalised weight feedback on weight loss and health behaviours: Evidence from a regression discontinuity design. *Health Economics*, **28**(1), 161–172.

Dai, T., Jiang, S., Liu, X. et al (2022). The effects of a hypertension diagnosis on health behaviors: A two-dimensional regression discontinuity analysis. *Health Economics*, **31**(4), 574–596.

Dell, M. (2010). The persitent effects of Peru's mining Mita. *Econometrica*, **78**(6), 1863–1903.

Department of Health (2010). Integrated Chronic Disease Management Model Manual. Technical report, Republic of South Africa.

Dupas, P. (2011). Health behavior in developing countries. *Annu. Rev. Econ.*, **3**(1), 425–449.

Ebinger, J.E., Driver, M., Ouyang, D. et al (2022). Variability independent of mean blood pressure as a real-world measure of cardiovascular risk. *Eclinicalmedicine*, **48**, 101442.

Edwards, R.D. (2018). 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**, 403–434.

Ferro, E.G., Abrahams-Gessel, S., Kapaon, D. et al (2022). Significant improvement in blood pressure levels among older adults with hypertension in rural South Africa: findings from a prospective 5,000-patient cohort, 2014-2019. *European Heart Journal*, **43**(Supplement 2), ehac544.2542.

Gaggero, A., Gil, J., Jiménez-Rubio, D. et al (2022). Does health information affect lifestyle behaviours? The impact of a diabetes diagnosis. *Social Science & Medicine*, **314**, 115420.

Gómez-Olivé, F.X., Ali, S.A., Made, F. et al (2017). Regional and sex differences in the prevalence and awareness of hypertension: an H3Africa AWI-Gen study across 6 sites in sub-Saharan Africa. *Global heart*, **12**(2), 81–90.

Gómez-Olivé, F.X., Montana, L., Wagner, R.G. et al (2018). Cohort profile: health and ageing in Africa: a longitudinal study of an indepth community in South Africa (HAALSI). *International Journal of Epidemiology*, **47**(3), 689–690j.

Hackl, F., Halla, M., Hummer, M. et al (2015). The effectiveness of health screening. *Health Economics*, **24**(8), 913–935.

Handler, J., Mohan, Y., Kanter, M.H. et al (2015). Screening for high blood pressure in adults during ambulatory nonprimary care visits: opportunities to improve hypertension recognition. *The Journal of Clinical Hypertension*, **17**(6), 431–439.

Heller, B. and Slungaard Mumma, K. (2019). Who benefits from the GED? New regression discontinuity evidence from Massachusetts. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3118546. SSRN Working paper.

Houle, B., Gaziano, T.A., Angotti, N. et al (2021). Hypertension incidence among middle-aged and older adults: findings from a 5-year prospective study in rural south africa, 2010–2015. *BMJ open*, **11**(12), e049621.

Hut, S. and Oster, E. (2022). Changes in household diet: Determinants and predictability. *Journal of Public Economics*, **208**, 104620.

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

Imbens, G.W. and Rubin, D.B. (2015). *Causal inference in statistics, social, and biomedical sciences*. Cambridge University Press.

Imbens, G.W. and Wooldridge, J.M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, **47**(1), 5–86.

Jardim, T.V., Reiger, S., Abrahams-Gessel, S. et al (2017). Hypertension management in a population of older adults in rural South Africa. *Journal of Hypertension*, **35**(6), 1283.

Jepsen, C., Mueser, P. and Troske, K. (2017). Second chance for high school dropouts? a regression discontinuity analysis of postsecondary educational returns to the GED. *Journal of Labor Economics*, **35**(S1), S273–S304.

Kahn, K., Collinson, M.A., Gómez-Olivé, F.X. et al (2012). Profile: Agincourt health and socio-demographic surveillance system. *International Journal of Epidemiology*, **41**(4), 988–1001.

Kämpfen, F., Wijemunige, N. and Evangelista, B. (2018). Aging, non-communicable diseases, and old-age disability in low-and middle-income countries: a challenge for global health. *International Journal of Public Health*, **63**.

Kandala, N.B., Nnanatu, C.C., Dukhi, N. et al (2021). Mapping the burden of hypertension in South Africa: A comparative analysis of the National 2012 SANHANES and the 2016 Demographic and Health Survey. *International Journal of Environmental Research and Public Health*, **18**(10), 5445.

Kawano, Y. (2011). Diurnal blood pressure variation and related behavioral factors. *Hypertension Research*, **34**(3), 281–285.

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

Kolesár, M. and Rothe, C. (2018). Inference in regression discontinuity designs with a discrete running variable. *American Economic Review*, **108**(8), 2277–2304.

Limbani, F., Thorogood, M., Gómez-Olivé, F.X. et al (2019). Task shifting to improve the provision of integrated chronic care: realist evaluation of a lay health worker intervention in rural South Africa. *BMJ Global Health*, **4**(1), e001084.

Mahomed, O.H. and Asmall, S. (2015). Development and implementation of an integrated chronic disease model in South Africa: lessons in the management of change through improving the quality of clinical practice. *International Journal of Integrated Care*, **15**.

Mahomed, O.H., Asmall, S. and Freeman, M. (2014). An integrated chronic disease management model: a diagonal approach to health system strengthening in South Africa. *Journal of Health Care for the Poor and Underserved*, **25**(4), 1723–1729.

Mancia, G., Ferrari, A., Gregorini, L. et al (1983). Blood pressure and heart rate variabilities in normotensive and hypertensive human beings. *Circulation Research*, **53**(1), 96–104.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics*, **142**(2), 698–714.

Mills, K.T., Bundy, J.D., Kelly, T.N. et al (2016). Global disparities of hypertension prevalence and control. *Circulation*, **134**(6), 441–450.

Modesti, P.A. (2013). Season, temperature and blood pressure: a complex interaction. *European Journal of Internal Medicine*, **24**(7), 604–607.

Murray, C.J., Aravkin, A.Y., Zheng, P. et al (2020). Global burden of 87 risk factors in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The lancet*, **396**(10258), 1223–1249.

Musini, V.M. and Wright, J.M. (2009). Factors affecting blood pressure variability: lessons learned from two systematic reviews of randomized controlled trials. *PLoS One*, **4**(5), e5673.

O'Brien, E., Kario, K., Staessen, J.A. et al (2018). Patterns of ambulatory blood pressure: clinical relevance and application. *The Journal of Clinical Hypertension*, **20**(7), 1112–1115.

Olsen, M.H., Angell, S.Y., Asma, S. et al (2016). A call to action and a lifecourse strategy to address the global burden of raised blood pressure on current and future generations: the Lancet commission on hypertension. *The Lancet*, **388**(10060), 2665–2712.

Oster, E. (2018). Diabetes and diet: Purchasing behavior change in response to health information. *American Economic Journal: Applied Economics*, **10**(4), 308–48.

Papay, J.P., Willett, J.B. and Murnane, R.J. (2011). Extending the regression-discontinuity approach to multiple assignment variables. *Journal of Econometrics*, **161**(2), 203–207.

Parati, G., Stergiou, G.S., Dolan, E. et al (2018). Blood pressure variability: clinical relevance and application. *The Journal of Clinical Hypertension*, **20**(7), 1133–1137.

Pedron, S., Hanselmann, M., Burns, J. et al (2022). The effect of population-based blood pressure screening on long-term cardiometabolic morbidity and mortality in Germany: A regression discontinuity analysis. *PLOS Medicine*, **19**(12), 1–18.

Peer, N., Uthman, O.A. and Kengne, A.P. (2021). Rising prevalence, and improved but suboptimal management, of hypertension in South Africa: A comparison of two national surveys. *Global Epidemiology*, **3**, 100063.

Phelps, C.E. and Mushlin, A.I. (1988). Focusing technology assessment using medical decision theory. *Medical Decision Making*, **8**(4), 279–289. PMID: 3185180.

Prina, S. and Royer, H. (2014). The importance of parental knowledge: Evidence from weight report cards in Mexico. *Journal of Health Economics*, **37**, 232–247.

Reardon, S.F. and Robinson, J.P. (2012). Regression discontinuity designs with multiple rating-score variables. *Journal of Research on Educational Effectiveness*, **5**(1), 83–104.

Rodriguez-Lesmes, P. (2021). Estimating the gains of early detection of hypertension over the marginal patient. *PLOS ONE*, **16**(7), 1–18.

Schmidt, B.M., Durao, S., Toews, I. et al (2020). Screening strategies for hypertension. *Cochrane Database of Systematic Reviews*, **5**.

Schutte, A.E., Kollias, A. and Stergiou, G.S. (2022). Blood pressure and its variability: classic and novel measurement techniques. *Nature Reviews Cardiology*, **19**(10), 643–654.

Sheffield, J., Goteti, G. and Wood, E.F. (2006). Development of a 50-year high-resolution global dataset of meteorological forcings for land surface modeling. *Journal of Climate*, **19**(13), 3088–3111.

Slade, A.N. (2012). Health investment decisions in response to diabetes information in older Americans. *Journal of Health Economics*, **31**(3), 502–520.

Slade, A.N. and Kim, H. (2014). Dietary responses to a hypertension diagnosis: Evidence from the National Health and Nutrition Examination Survey (NHANES) 2007-2010. *Behavioral Medicine*, **40**(1), 1–13. PMID: 24512360.

Stanaway, J.D., Afshin, A., Gakidou, E. et al (2018). Global, regional, and national comparative risk assessment of 84 behavioural, environmental and occupational, and metabolic risks or clusters of risks for 195 countries and territories, 1990–2017: a systematic analysis for the global burden of disease study 2017. *The Lancet*, **392**(10159), 1923–1994.

Sudharsanan, N., Chen, S., Garber, M. et al (2020). The effect of home-based hypertension screening on blood pressure change over time in South Africa. *Health Affairs*, **39**(1), 124–132.

Thornton, R.L. (2008). The demand for, and impact of, learning HIV status. *American Economic Review*, **98**(5), 1829–63.

Vos, T., Lim, S.S., Abbafati, C. et al (2020). Global burden of 369 diseases and injuries in 204 countries and territories, 1990–2019: a systematic analysis for the global burden of disease study 2019. *The Lancet*, **396**(10258), 1204–1222.

Whelton, P.K., Carey, R.M., Aronow, W.S. et al (2018a). 2017 ACC/AHA/AAPA/ABC/ACPM/AGS/APhA/ASH/ASPC/NMA/PCNA guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: a report of the American College of Cardiology/American Heart Association Task Force on Clinical Practice Guidelines. *Journal of the American College of Cardiology*, **71**(19), e127–e248.

Whelton, P.K., Carey, R.M., Aronow, W.S. et al (2018b). 2017 acc/aha/aapa/abc/acpm/ags/apha/ash/aspc/nma/pcna guideline for the prevention, detection, evaluation, and management of high blood pressure in adults: Executive summary: A report of the american college of cardiology/american heart association task force on clinical practice guidelines. *Hypertension*, **71**(6), 1269–1324.

Wong, V.C., Steiner, P.M. and Cook, T.D. (2013). Analyzing regression-discontinuity designs with multiple assignment variables: A comparative study of four estimation methods. *Journal of Educational and Behavioral Statistics*, **38**(2), 107–141.

Zhao, M., Konishi, Y. and Glewwe, P. (2013). 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.

Zhou, B., Perel, P., Mensah, G.A. et al (2021a). Global epidemiology, health burden and effective interventions for elevated blood pressure and hypertension. *Nature Reviews Cardiology*, **18**, 785–802.

Zhou, B., Carrillo-Larco, R.M., Danaei, G. et al (2021b). Worldwide trends in hypertension prevalence and progress in treatment and control from 1990 to 2019: a pooled analysis of 1201 population-representative studies with 104 million participants. *The Lancet*, **398**(10304), 957–980.

APPENDICES

Appendix A Additional descriptives

Figure A1: Referral letter

P.O. Box 2, Acornhoek 1360, South Africa
Telephone: +27 13 7955076 (Acornhoek) or +27 13 708 0003 (Agincourt)
Fax: +27 13 7955076 (Acornhoek) or +27 13 708 1540 (Agincourt)

REFERAL TO THE CLINIC

As part of the survey "Health and Aging in Africa: Longitudinal studies in INDEPTH communities – HI KURILE" that the MRC/Wits Agincourt Research Unit is currently running in the area, we have identified that _____ _____ has a potential problem with his/her:

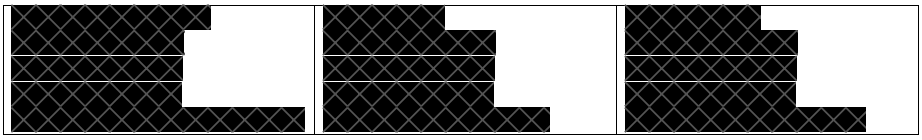
a) **Blood pressure**

b) **Blood glucose**

c) **Cholesterol**

d) **Haemoglobin**

In case you need to contact us please do it using the following contacts:



Field worker signature: _____

Date |___|___|/|___|___|___|/|___|___|___|___|

[1]

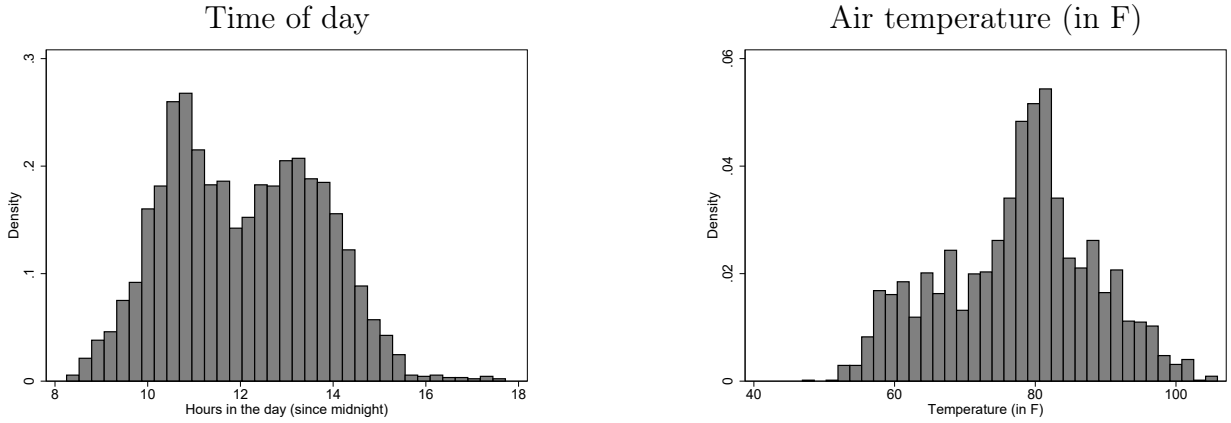
[1] HAALSI_referralclinicletter_V2_13032015.docx

Table A1: Effects of a referral letter on selection (missing at follow-up)

	Binding-score RDD		*MRDD*			
	Combined frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)
Overall sample						
Missing at follow-up	0.041	2467	-0.025	-0.054	0.035	1491
	(0.048)		(0.041)	(0.047)	(0.060)	
Consent for BP not given at follow-up	0.042	2275	-0.020	-0.034	0.011	1300
	(0.045)		(0.040)	(0.047)	(0.055)	
More likely to be FP						
Missing at follow-up	-0.020	1752	0.013	0.006	0.029	990
	(0.062)		(0.051)	(0.059)	(0.070)	
Consent for BP not given at follow-up	0.010	1454	-0.031	-0.016	-0.061	861
	(0.054)		(0.047)	(0.055)	(0.059)	
Less likely to be FP						
Missing at follow-up	0.164*	657	-0.090	-0.140	0.027	460
	(0.096)		(0.077)	(0.087)	(0.116)	
Consent for BP not given at follow-up	0.122	690	0.007	-0.063	0.174	407
	(0.079)		(0.076)	(0.087)	(0.113)	

Note: "FP" stands for false positive. Binding-score RDD estimates of eq.(2) obtained from local linear regression, with triangular kernels and optimal bandwidths on each side of the threshold using the MSE optimal bandwidth selector. MRDD estimates obtained from linear regression estimates of the outcome response surface, eq.(5) and bivariate kernel estimates of the conditional densities and weights that are combined with the outcome response estimates using eq.(4) and, for column (3), eq.(3). "Missing at follow-up" is a dichotomous variable that takes the value one if we have complete blood pressure information about a respondent in Wave 1 but not in wave 2, and zero otherwise. "Consent for BP not given at follow-up" is a dichotomous variable that takes the value one if we have complete blood pressure information about a respondent in Wave 1 but not in wave 2 because the respondent did not consent to have their blood pressure measurement taken, and zero otherwise. Standard errors in parentheses are bootstrapped (1000 repetitions) for MRDD and heteroscedasticity-robust for binding-score RDD. N is the effective number of observations used in estimation. All estimates are from specifications that control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

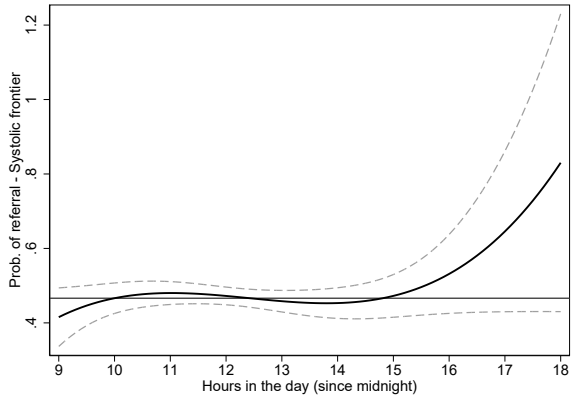
Figure A2: Distributions of time of day and air temperature at baseline interview



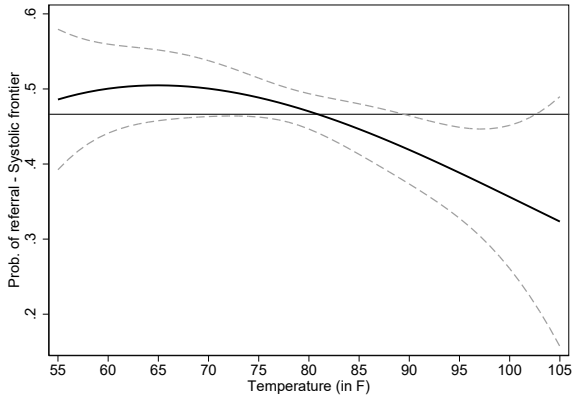
Notes: On left, x-axis shows hour since midnight at baseline interview. On right, x-axis shows outside air temperature (in Fahrenheit) at baseline interview. Temperature data from http://hydrology.princeton.edu/data/pgf/v3/. For additional details, including the bias correction and downscaling methodology, see Sheffield *et al.* (2006). We use gridded 3-hourly data points (0.5 by 0.5 degrees, which corresponds roughly to 27km by 27km) that we match to each individual's household location and time of interview. Sample sizes are $N = 3,300$ and $N = 3,241$ for the "Hours of the day" and "Temperature" plots, respectively.

Figure A3: Variation in the probability of receiving a referral letter by time of interview and air temperature at baseline

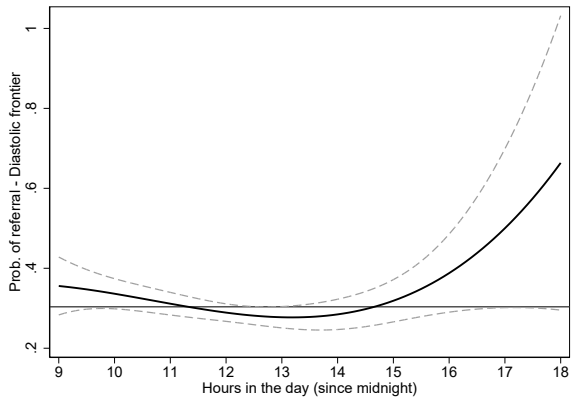
(a) Referral letter because of SBP by time of interview



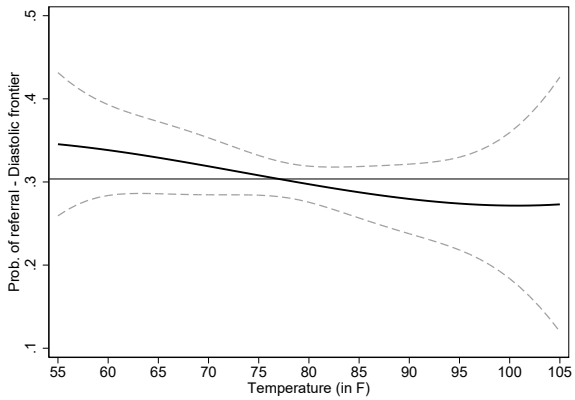
(b) Referral letter because of SBP by air temperature



(c) Referral letter because of DBP by time of interview



(d) Referral letter because of DBP by air temperature



Notes: Predicted probability of receiving a referral letter (because of SBP and DBP) against time of interview and air temperature (in Fahrenheit). Predictions obtained from a linear regression of each probability on a third order polynomial of each of time of interview and temperature, plus year and month fixed effects. Dashed lines show 95% confidence intervals around estimated expected value. Horizontal lines indicate mean probability in sample.

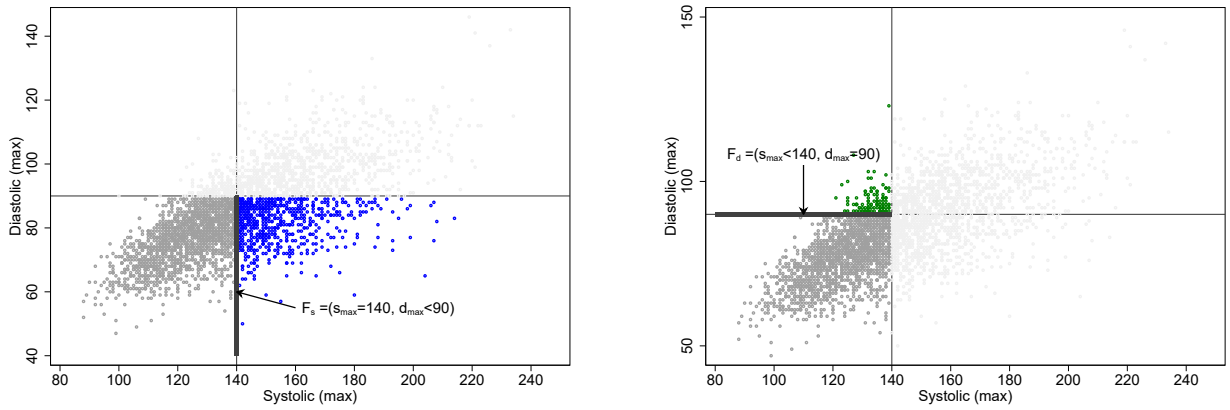
Table A2: Effects of the referral letter received in 2014/2015 on time of interview and probability of being more likely to be false positive at follow-up

	Binding-score RDD		*MRDD*			
	Combined frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)
More likely to be FP[a]	0.067	1399	-0.038	-0.056	0.002	937
	(0.080)		(0.058)	(0.066)	(0.080)	
Time of interview	0.114	1776	0.094	0.151	-0.037	937
	(0.212)		(0.181)	(0.204)	(0.253)	

Note: [a]: based on time of interview only. "FP" stands for false positive. Outcome variables are measured at follow-up. Binding-score RDD estimates of eq.(2) obtained from local linear regression, with triangular kernels and optimal bandwidths on each side of the threshold using the MSE optimal bandwidth selector. MRDD estimates obtained from linear regression estimates of the outcome response surface, eq.(5) and bivariate kernel estimates of the conditional densities and weights that are combined with the outcome response estimates using eq.(4) and, for column (3), (3). Standard errors in parentheses are bootstrapped (1000 repetitions) for MRDD and heteroscedasticity-robust for binding-score RDD. N is the effective number of observations used in estimation. All estimates are from specifications that control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

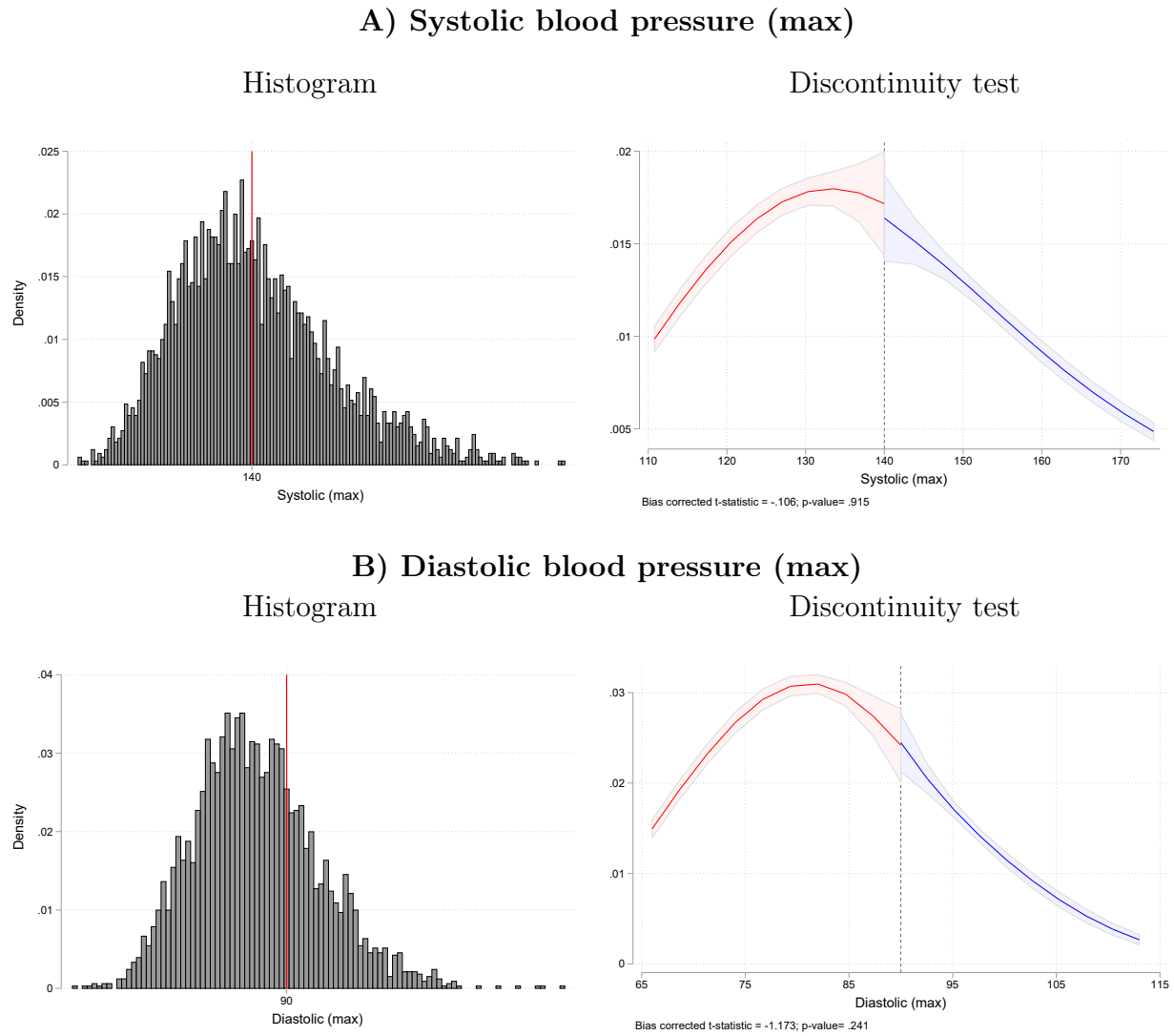
Figure A4: Sample restrictions for unidimensional RDD

(a) Systolic frontier **(b) Diastolic frontier**



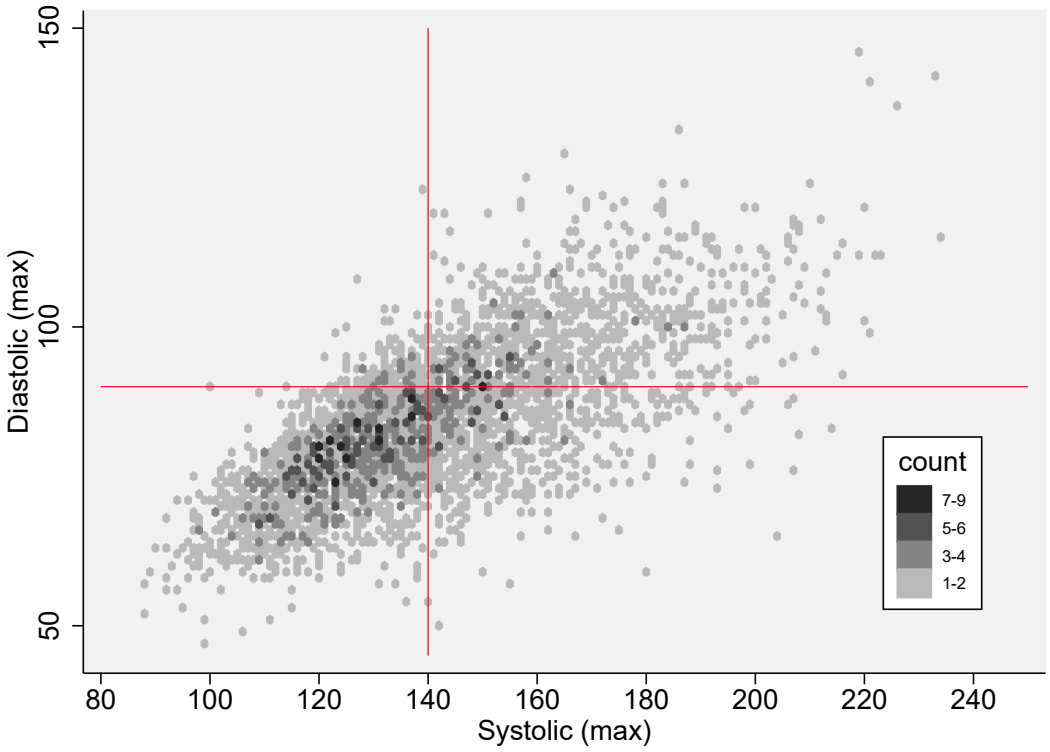
Notes: $N = 3,304$. x-axis (y-axis) shows the maximum systolic (diastolic) blood pressure from the last two measurements at baseline. F_s and F_d are the systolic BP and diastolic BP treatment frontiers, respectively. Blue (green) dots identify respondents who received a referral card solely because their systolic (diastolic) BP was at or above the threshold and their diastolic (systolic) BP was below the respective threshold. Grey dots identify control group respondents who did not receive a referral letter. Lighter grey dots identify those who got a referral letter because the other BP measure was also above the respective threshold.

Figure A5: Histograms and tests of threshold discontinuities in running variables



Notes: Left panels show histograms of maximum systolic and diastolic blood pressure. Maximum over last two measurements at baseline. Right panels show there is no statistically significant discontinuity in either running variables at the respective threshold (140 mm Hg for SBP and 90 mm Hg for DBP). These RDD plots are generated using third order polynomial, triangular weights, and a different optimal bandwidth on each side of the threshold (based on the MSE of each density estimator separately) (Calonico *et al.*, 2015, 2017, 2018; Cattaneo *et al.*, 2018, 2020). Below each RDD plot is the bias-corrected t-statistic and corresponding p-value for test of no discontinuity at the threshold (McCrary, 2008).

Figure A6: Joint density of the running variables – heat plot



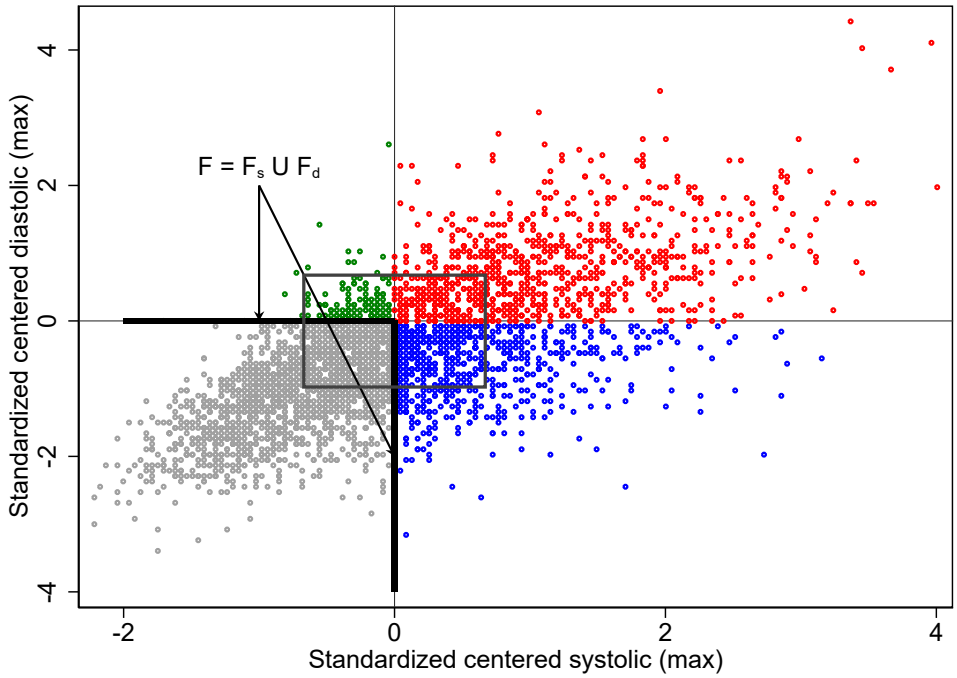
Notes: $N = 3,304$. x-axis (y-axis) shows maximum systolic (diastolic) blood pressure. Maximum is over last two measurements at baseline. The red lines identify treatment assignment thresholds (140 mm Hg for systolic and 90 mm Hg for diastolic). Darker dot indicates a larger number of observations (greater density) at that cell.

Table A3: Tests for threshold discontinuities in predetermined variables – unidimensional RDD, MRDD, and binding-score RDD

	Unidimensional RDD				*MRDD*				*Binding-score RDD*	
	Systolic frontier	N	Diastolic frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N	Combined frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Female	0.111	1109	-0.079	764	-0.031	-0.034	-0.026	945	0.030	1730
	(0.073)		(0.100)		(0.057)	(0.064)	(0.080)		(0.064)	
Age	-2.375	1116	1.296	764	-0.547	-1.273	1.063	945	-1.927	1871
	(1.673)		(2.209)		(1.170)	(1.364)	(1.613)		(1.429)	

Note: Unidimensional RDD and binding-score RDD estimates of eqs. (1) and (2), respectively, obtained from local linear regression, with triangular kernels and optimal bandwidths on each side of the threshold using the MSE optimal bandwidth selector. MRDD estimates obtained from linear regression estimates of the outcome response surface, eq.(5) and bivariate kernel estimates of the conditional densities and weights that are combined with the outcome response estimates using eq.(4) and, for column (5), (3). Standard errors in parentheses are bootstrapped (1000 repetitions) for MRDD and heteroscedasticity-robust for unidimensional RDD and binding-score RDD. N is the effective number of observations used in estimation. * $p<0.1$, ** $p<0.05$, *** $p<0.01$.

Figure A7: MRDD estimation frontier with wider sample selection window



Notes: The main MRDD estimates use observations within 0.7 of a standard deviation (SD) of the centralized and standardized running variable from the respective frontier. This figure shows a wider sample selection window that is used in the robustness analysis reported in Appendix C Table C2. The window is extended to include observations up to 1 SD below the diastolic frontier. Otherwise, it is the same as the window used for the main estimates.

Appendix B Details of estimators

Weights

The weights in eq.(3) are defined as follows (Wong *et al.*, 2013):

$$\omega_s = \frac{\int_{\tilde{d}_i^{std}<0} f(\tilde{s}_i^{std}=0,\tilde{d}_i^{std})\mathrm{d}\tilde{d}_i^{std}}{\int_{\tilde{d}_i^{std}<0} f(\tilde{s}_i^{std}=0,\tilde{d}_i^{std})\mathrm{d}\tilde{d}_i^{std}+\int_{\tilde{s}_i^{std}<0} f(\tilde{s}_i^{std},\tilde{d}_i^{std}=0)\mathrm{d}\tilde{s}_i^{std}},$$

$$\omega_d = \frac{\int_{\tilde{s}_i^{std}<0} f(\tilde{s}_i^{std},\tilde{d}_i^{std}=0)\mathrm{d}\tilde{s}_i^{std}}{\int_{\tilde{d}_i^{std}<0} f(\tilde{s}_i^{std}=0,\tilde{d}_i^{std})\mathrm{d}\tilde{d}_i^{std}+\int_{\tilde{s}_i^{std}<0} f(\tilde{s}_i^{std},\tilde{d}_i^{std}=0)\mathrm{d}\tilde{s}_i^{std}},$$

where $f(\tilde{s}_i^{std},\tilde{d}_i^{std})$ is the joint density of the assignment variables $\tilde{s}_i^{std}$ and $\tilde{d}_i^{std}$.

Write each joint density as the product of the conditional and marginal densities,

$$f(\tilde{s}_i^{std}=0,\tilde{d}_i^{std}) = f(\tilde{d}_i^{std} \mid \tilde{s}_i^{std}=0)\times f_S(\tilde{s}_i^{std}=0),$$
$$f(\tilde{s}_i^{std},\tilde{d}_i^{std}=0) = f(\tilde{s}_i^{std} \mid \tilde{d}_i^{std}=0)\times f_D(\tilde{d}_i^{std}=0).$$

Define,

$$A = \int_{\tilde{d}_i^{std}<0} f(\tilde{d}_i^{std} \mid \tilde{s}_i^{std}=0)\mathrm{d}\tilde{d}_i^{std}\times f_S(\tilde{s}_i^{std}=0),$$
$$B = \int_{\tilde{s}_i^{std}<0} f(\tilde{s}_i^{std} \mid \tilde{d}_i^{std}=0)\mathrm{d}\tilde{s}_i^{std}\times f_D(\tilde{d}_i^{std}=0).$$

Then, $\omega_s = \frac{A}{A+B}$ and $\omega_d = \frac{B}{A+B}$. We estimate these weights with the bivariate kernel density estimator and numerical integration (Wong *et al.*, 2013).

Discontinuities in outcome response surface

Given the specification of the outcome regression eq.(5), the expected outcomes when treated by crossing the systolic and diastolic frontiers are given by equations (B1) and (B2), respectively.

$$E(Y_i \mid \mathbb{1}_S = 1, \mathbb{1}_D = 0, \tilde{s}_i^{std}, \tilde{d}_i^{std}) = \alpha + \gamma_1 \tilde{s}_i^{std} + \gamma_2 \tilde{d}_i^{std} + \beta_1 + \beta_2 \tilde{d}_i^{std} + \beta_3 \tilde{d}_i^{std} \times \tilde{s}_i^{std} + X_i' \delta \tag{B1}$$

$$E(Y_i \mid \mathbb{1}_S = 0, \mathbb{1}_D = 1, \tilde{s}_i^{std}, \tilde{d}_i^{std}) = \alpha + \gamma_1 \tilde{s}_i^{std} + \gamma_2 \tilde{d}_i^{std} + \beta_1 + \beta_4 \tilde{s}_i^{std} + \beta_5 \tilde{d}_i^{std} \times \tilde{s}_i^{std} + X_i' \delta \tag{B2}$$

The expected outcome response surface for the control group that is below both frontiers is given by:

$$E(Y_i \mid \mathbb{1}_S = 0, \mathbb{1}_D = 0, \tilde{s}_i^{std}, \tilde{d}_i^{std}) = \alpha + \gamma_1 \tilde{s}_i^{std} + \gamma_2 \tilde{d}_i^{std} + X_i' \delta \tag{B3}$$

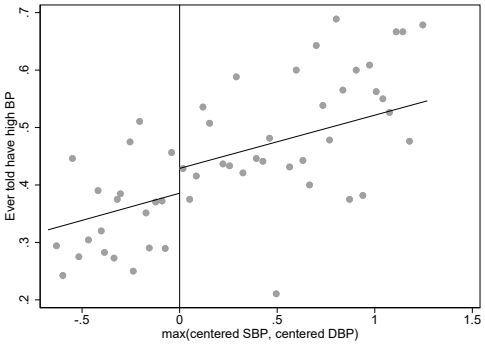
The difference between eq.(B1) and eq.(B3) evaluated at the systolic frontier $F_s = (s_i^{max} = 140, d_i^{max} < 90) = (\tilde{s}_i^{std} = 0, \tilde{d}_i^{std} < 0)$ gives the outcome surface discontinuity at that frontier, eq.(6). Similarly, the difference between eq.(B2) and eq.(B3) evaluated at the diastolic frontier $F_d = (s_i^{max} < 140, d_i^{max} = 90) = (\tilde{s}_i^{std} < 0, \tilde{d}_i^{std} = 0)$ gives eq.(7).

Appendix C Additional results

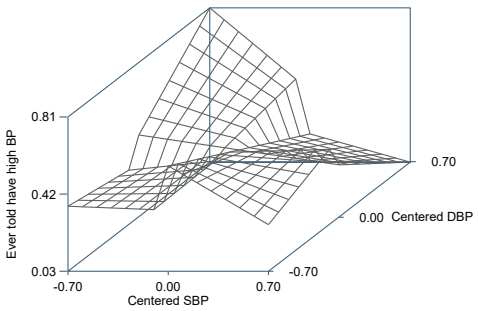
Figure C1: Discontinuity plots for secondary outcomes

A) Ever told have high BP

Binding-score RDD

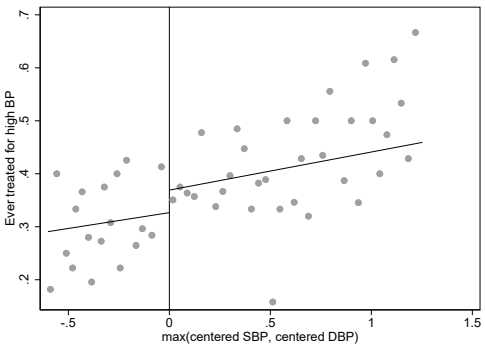


MRDD

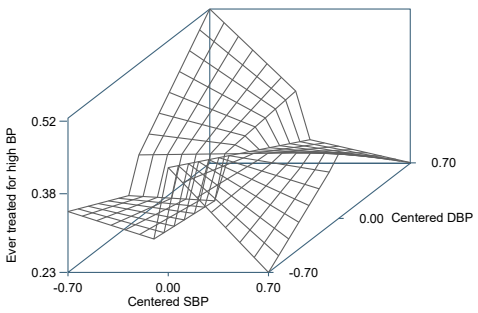


B) Ever treated for high BP

Binding-score RDD

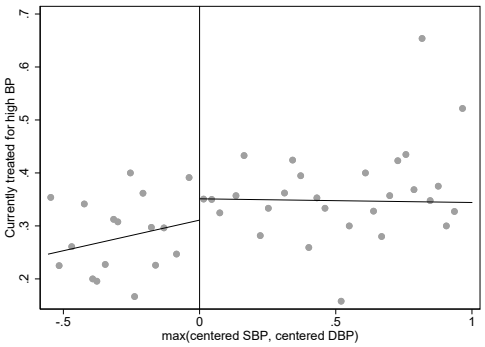


MRDD

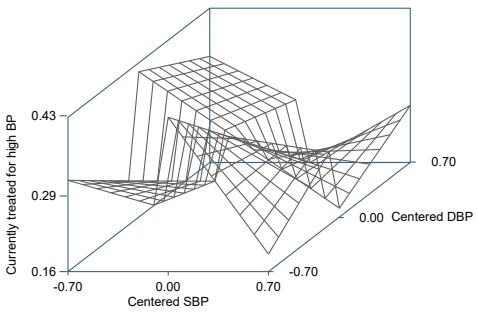


C) Currently treated for high BP

Binding-score RDD

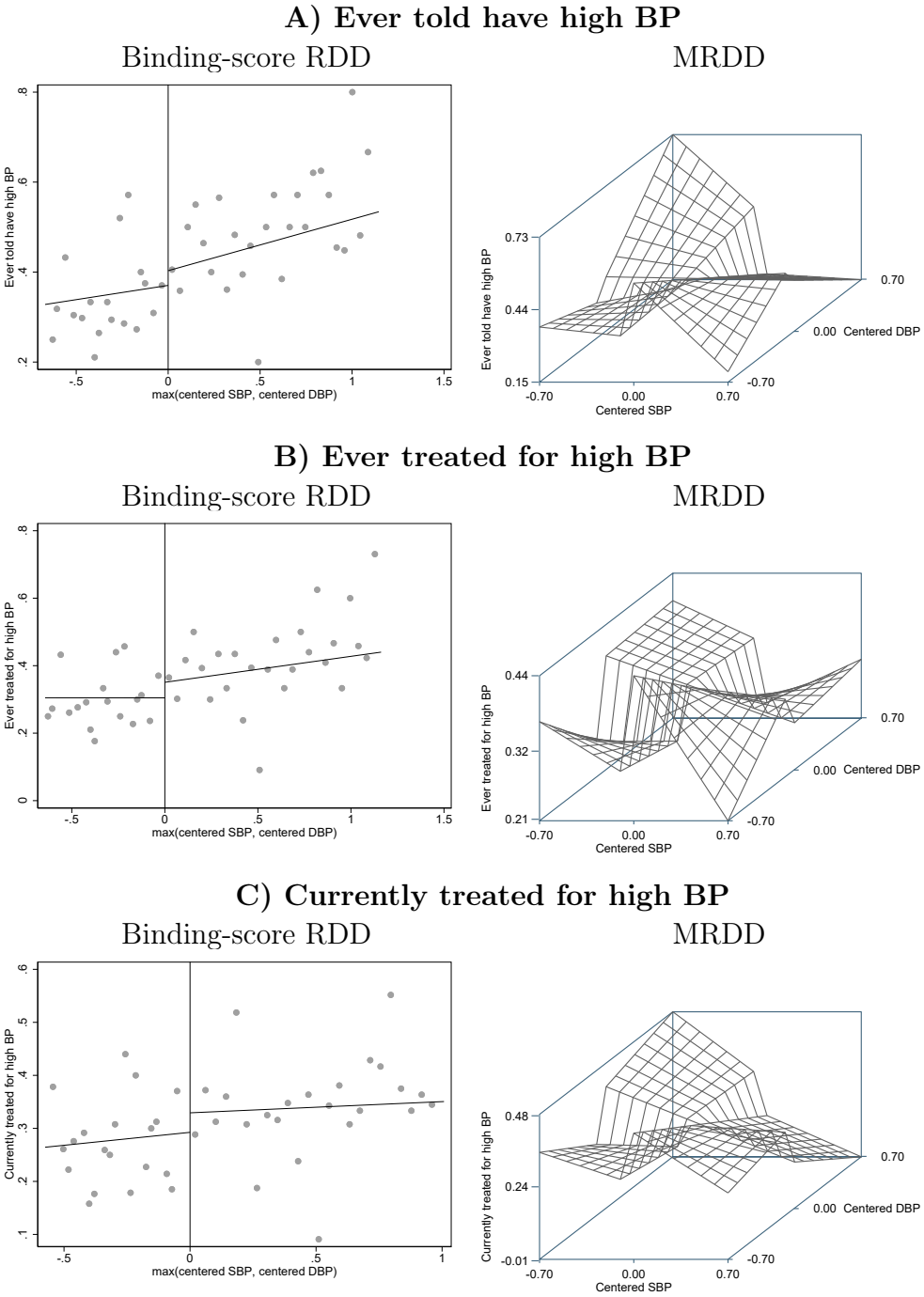


MRDD



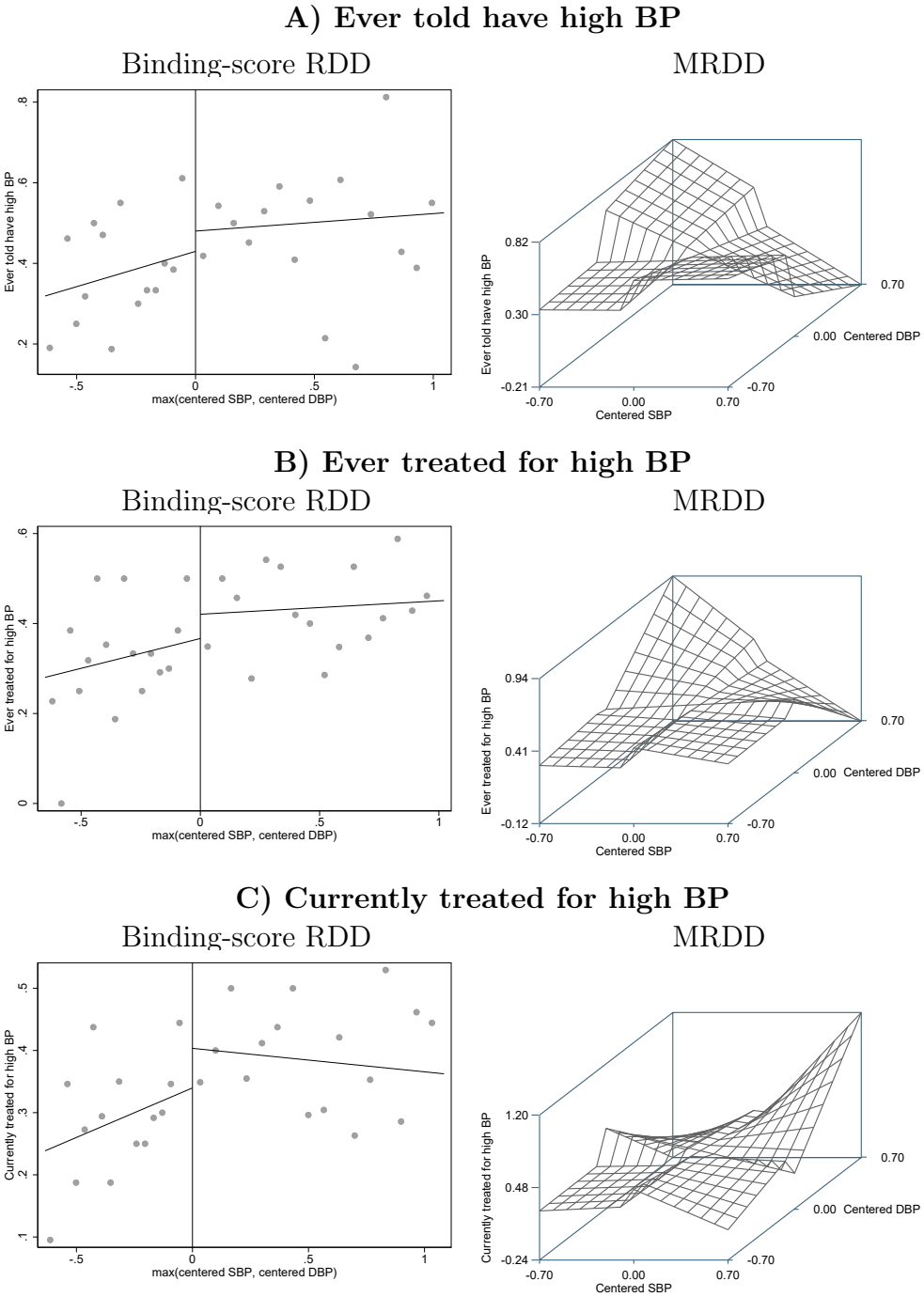
Notes: Plots of average outcomes conditional on the running variable(s). Outcomes are defined in Table 1. Non-negative values of the running variable(s) indicate observations that receive a referral letter. For binding-score RDD, on the x-axis is the running variable, $r_i = max\left(\tilde{s}_i^{std}, \tilde{d}_i^{std}\right)$. For these plots, we use local linear regression, triangular kernels, and the MSE optimal bandwidth selector. Each dot is the mean of the respective outcome in a given bin. We use optimal bins obtained with variance evenly-spaced method using spacing estimators (Calonico *et al.*, 2014a,b, 2015, 2017). For MRDD, running variables on x-axis and y-axis are $\tilde{s}_i^{std}$ and $\tilde{d}_i^{std}$, respectively. For MRDD plots, each point on the grid represents the predicted value from estimates of eq.(5) at the respective values of the running variables. Predicted values are computed for each 0.1 increment in the two running variables. In all plots, we control for sex and age.

Figure C2: Discontinuity plots for secondary outcomes using respondents more likely to be false positives



Notes: Sample restricted to respondents with high likelihood of being false positives for hypertension at baseline on the basis of time of day and air temperature when BP was measured. Plots of average outcomes conditional on the running variable(s). Outcomes are defined in Table 1. Non-negative values of the running variable(s) indicate observations that receive a referral letter. For binding-score RDD, on the x-axis is the running variable, $r_i = max\left(\tilde{s}_i^{std}, \tilde{d}_i^{std}\right)$. For these plots, we use local linear regression, triangular kernels, and the MSE optimal bandwidth selector. Each dot is the mean of the respective outcome in a given bin. We use optimal bins obtained with variance evenly-spaced method using spacing estimators (Calonico *et al.*, 2014a,b, 2015, 2017). For MRDD, running variables on x-axis and y-axis are $\tilde{s}_i^{std}$ and $\tilde{d}_i^{std}$, respectively. For MRDD plots, each point on the grid represents the predicted value from estimates of eq.(5) at the respective values of the running variables. Predicted values are computed for each 0.1 increment in the two running variables. In all plots, we control for sex and age.

Figure C3: Discontinuity plots for secondary outcomes using respondents less likely to be false positives



Notes: Sample restricted to respondents with low likelihood of being false positives for hypertension at baseline on the basis of time of day and air temperature when BP was measured. Plots of average outcomes conditional on the running variable(s). Outcomes are defined in Table 1. Non-negative values of the running variable(s) indicate observations that receive a referral letter. For binding-score RDD, on the x-axis is the running variable, $r_i = max\left(\tilde{s}_i^{std}, \tilde{d}_i^{std}\right)$. For these plots, we use local linear regression, triangular kernels, and the MSE optimal bandwidth selector. Each dot is the mean of the respective outcome in a given bin. We use optimal bins obtained with variance evenly-spaced method using spacing estimators (Calonico *et al.*, 2014a,b, 2015, 2017). For MRDD, running variables on x-axis and y-axis are $s_{i,c}^{std}$ and $d_{i,c}^{std}$, respectively. For MRDD plots, each point on the grid represents the predicted value from estimates of eq.(5) at the respective values of the running variables. Predicted values are computed for each 0.1 increment in the two running variables. In all plots, we control for sex and age.

Table C1: Binding-score RDD estimates of effects of a referral letter - robustness

	Alternative specifications				***Sample restrictions***					
					Treatment group if					
	Local quadratic		*Uniform weights*		*Above both thresholds*		*No mean below threshold*		*Not already diagnosed*	
	Combined frontier	N	Combined frontier	N	Combined frontier	N	Combined frontier	N	Combined frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Primary outcomes										
Mean systolic BP	1.667	2131	1.781	1078	4.145	826	2.214	1281	2.847	1058
	(2.460)		(2.409)		(4.802)		(2.697)		(2.354)	
Mean diastolic BP	-1.496	2145	-1.615	1109	1.332	738	-1.781	1247	-1.377	1007
	(1.562)		(1.562)		(2.600)		(1.712)		(1.579)	
Hypertensive	-0.040	2101	-0.004	1289	0.003	861	-0.003	1240	0.013	1115
	(0.076)		(0.056)		(0.120)		(0.076)		(0.064)	
Secondary outcomes										
Ever told have high BP	0.046	2342	0.059	1378	-0.042	1165	0.028	1775	0.032	1082
	(0.072)		(0.060)		(0.108)		(0.064)		(0.060)	
Ever treated for high BP	0.058	2310	0.054	1337	0.019	1056	0.042	1696	0.040	1089
	(0.067)		(0.058)		(0.103)		(0.063)		(0.051)	
Currently treated for high BP	0.060	2290	0.048	1275	-0.029	1023	0.013	1699	0.007	1060
	(0.065)		(0.057)		(0.098)		(0.062)		(0.051)	

Note: See Table 1 for definitions of outcomes. Our main binding-score RDD estimates reported in Table 2 are obtained from local linear regression with triangular kernel weights. Column 1 shows estimates from local quadratic regression. Columns 3 shows estimates with uniform weights. The remaining columns show estimates from samples that (a) restrict treated individuals to those who were at or above both the SBP and the DBP thresholds for receipt of a referral letter at baseline (column 5), (b) exclude treated individuals who had a mean SBP or mean DBP below the respective threshold at baseline (column 7), and (c) exclude individuals who reported already ever having been diagnosed with high BP at baseline (column 9). All estimates obtained using optimal bandwidths on each side of the threshold using the MSE optimal bandwidth selector. Heteroscedasticity-robust standard errors in parentheses. N is the effective number of observations used in estimation. All estimates are from specifications that control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01^{*}$

Table C2: MRDD estimates of effects of a referral letter – robustness

	Alternative specification of response surface				*Wider interval around frontier*			
	Combined frontier	Systolic frontier	Diastolic frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Primary outcomes								
Mean systolic BP	-0.484	-0.037	-1.517	945	0.787	1.183	-0.532	1238
	(1.855)	(2.041)	(2.493)		(1.654)	(1.827)	(2.539)	
Mean diastolic BP	-0.603	-0.562	-0.699	945	-0.172	-0.231	0.025	1238
	(1.066)	(1.152)	(1.609)		(0.955)	(1.014)	(1.636)	
Hypertensive	-0.018	-0.017	-0.020	945	0.027	0.025	0.033	1238
	(0.055)	(0.061)	(0.074)		(0.046)	(0.051)	(0.070)	
Secondary outcomes								
Ever told have high BP	0.107*	0.105*	0.110	943	0.020	0.011	0.053	1235
	(0.056)	(0.063)	(0.074)		(0.048)	(0.053)	(0.073)	
Ever treated for high BP	0.104*	0.106*	0.099	943	0.026	0.016	0.058	1235
	(0.053)	(0.060)	(0.071)		(0.046)	(0.051)	(0.069)	
Currently treated for high BP	0.116**	0.108*	0.134**	943	0.031	0.009	0.102	1234
	(0.051)	(0.059)	(0.068)		(0.045)	(0.051)	(0.067)	

Note: Main MRDD estimates reported in Table 2 use specification of the outcome response surface given by eq.(5) and observations within 0.7 of a standard deviation (SD) of each centered and standardized running variable from the respective threshold. Columns (1)-(3) show estimates using the specification of the response surface given in footnote 12. Columns (5)-(7) show estimates using the main estimate observations plus those up to 1 SD below the diastolic threshold. Bootstrapped (1000 repetitions) standard errors in parentheses. N is the effective number of observations used for estimation. All specifications control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C3: Binding-score RDD estimates of effects of a referral letter - heterogeneity

	Sex				***Age***				***Distance to health facility***			
	Female		*Male*		<60 *years*		≥ 60 *years*		*< Median*		*≥ Median*	
	Combined frontier	N	Combined frontier	N	Combined frontier	N	Combined frontier	N	Combined frontier	N	Combined frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
Primary outcomes												
Mean systolic BP	0.917	1020	2.964	680	-1.169	672	4.897	833	-2.501	854	6.079**	850
	(2.692)		(3.257)		(2.772)		(3.158)		(3.119)		(2.738)	
Mean diastolic BP	-0.250	938	-1.790	617	-2.311	624	-0.997	706	-0.388	859	-2.225	751
	(1.804)		(2.343)		(1.824)		(2.278)		(1.910)		(2.103)	
Hypertensive	-0.013	1057	-0.007	659	-0.018	800	0.023	678	-0.004	918	0.003	765
	(0.068)		(0.102)		(0.074)		(0.092)		(0.076)		(0.084)	
Secondary outcomes												
Ever told have high BP	0.054	1166	0.032	889	0.092	857	-0.028	1028	0.050	887	0.033	880
	(0.072)		(0.075)		(0.073)		(0.086)		(0.084)		(0.080)	
Ever treated for high BP	0.057	1174	0.006	790	0.120*	811	-0.046	949	0.073	834	0.013	880
	(0.070)		(0.076)		(0.069)		(0.091)		(0.084)		(0.080)	
Currently treated for high BP	0.065	1165	-0.033	747	0.095	889	-0.039	948	0.117	778	0.004	955
	(0.070)		(0.080)		(0.066)		(0.093)		(0.082)		(0.072)	

Note: See Table 1 for definitions of outcomes. Columns 9 and 11 show estimates for individuals located below and above the median distance to the nearest health facility, respectively. As for main binding-score RDD estimates reported in Table 2, we use local linear regression with triangular kernel weights and optimal bandwidths on each side of the threshold using the MSE optimal bandwidth selector. Heteroscedasticity-robust standard errors in parentheses. N is the effective number of observations used in estimation. All estimates are from specifications that control for age and sex, except in columns 1) and 2) where sex is omitted. * $p<0.1$, ** $p<0.05$, *** $p<0.01^{*}$

Table C4: MRDD estimates of effects of a referral letter by sex

	Sex							
	Female				*Male*			
	Combined frontier	Systolic frontier	Diastolic frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Primary outcomes								
Mean systolic BP	-0.911	-1.258	-0.247	538	1.260	2.475	-2.598	407
	(2.422)	(2.704)	(3.491)		(2.771)	(3.025)	(3.868)	
Mean diastolic BP	0.734	0.311	1.544	538	-0.878	-0.326	-2.635	407
	(1.378)	(1.436)	(2.299)		(1.733)	(1.849)	(2.708)	
Hypertensive	-0.039	-0.081	0.040	538	0.074	0.104	-0.024	407
	(0.065)	(0.070)	(0.103)		(0.088)	(0.103)	(0.111)	
Secondary outcomes								
Ever told hypertension	0.114	0.108	0.125	536	0.038	0.073	-0.074	407
	(0.072)	(0.083)	(0.106)		(0.083)	(0.096)	(0.104)	
Ever treated for hypertension	0.087	0.079	0.102	536	0.084	0.109	0.006	407
	(0.070)	(0.082)	(0.101)		(0.073)	(0.085)	(0.090)	
Currently treated for hypertension	0.141**	0.123	0.177*	536	0.043	0.059	-0.008	407
	(0.070)	(0.082)	(0.102)		(0.070)	(0.082)	(0.090)	

Note: See Table 1 for definitions of outcomes. MRDD estimates obtained from linear regression estimates of the outcome response surface, eq.(5) and bivariate kernel estimates of the conditional densities and weights that are combined with the outcome response estimates using eq.(4) and, for columns (1) and (5), eq.(3). Columns 1-3 and 5-7 show estimates for females and males, respectively. Standard errors in parentheses are bootstrapped (1000 repetitions) for MRDD. N is the effective number of observations used in estimation. All estimates are from specifications that control for age. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C5: MRDD estimates of effects of a referral letter by age group (below or above 60)

	Age							
	Age ≥ 60				*Age* < 60			
	Combined frontier	Systolic frontier	Diastolic frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Primary outcomes								
Mean systolic BP	-1.870	-2.948	1.084	382	0.463	2.104	-1.946	563
	(3.250)	(3.361)	(6.179)		(2.283)	(2.542)	(2.914)	
Mean diastolic BP	-0.649	-0.779	-0.291	382	0.033	0.000	0.082	563
	(1.845)	(1.796)	(3.964)		(1.369)	(1.467)	(1.946)	
Hypertensive	-0.056	-0.105	0.080	382	0.033	0.049	0.010	563
	(0.089)	(0.093)	(0.168)		(0.065)	(0.076)	(0.088)	
Secondary outcomes								
Ever told hypertension	0.097	0.111	0.058	380	0.063	0.073	0.050	563
	(0.088)	(0.094)	(0.151)		(0.068)	(0.082)	(0.091)	
Ever treated for hypertension	0.120	0.104	0.165	380	0.066	0.086	0.038	563
	(0.088)	(0.097)	(0.145)		(0.063)	(0.076	(0.084)	
Currently treated for hypertension	0.117	0.081	0.217	380	0.089	0.099	0.073	563
	(0.087)	(0.096)	(0.146)		(0.061)	(0.073)	(0.080)	

Note: See Table 1 for definitions of outcomes. MRDD estimates obtained from linear regression estimates of the outcome response surface, eq.(5) and bivariate kernel estimates of the conditional densities and weights that are combined with the outcome response estimates using eq.(4) and, for columns (1) and (5), eq.(3). Columns 1-3 and 5-7 shows estimates for individuals older than 60 and younger than 60, respectively. Standard errors in parentheses are bootstrapped (1000 repetitions) for MRDD. N is the effective number of observations used in estimation. All estimates are from specifications that control for sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C6: MRDD estimates of effects of a referral letter by distance to nearest health facility

	Distance to nearest health facility							
	< Median				*≥ Median*			
	Combined frontier	Systolic frontier	Diastolic frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Primary outcomes								
Mean systolic BP	-3.588	-2.885	-4.960	441	3.723	2.524	5.948	487
	(2.685)	(3.001)	(3.381)		(2.447)	(2.711)	(3.982)	
Mean diastolic BP	-0.878	-0.394	-1.822	441	0.927	-0.405	3.400	487
	(1.563)	(1.627)	(2.261)		(1.464)	(1.595)	(2.702)	
Hypertensive	-0.050	-0.062	-0.026	441	0.072	0.053	0.109	487
	(0.075)	(0.081)	(0.102)		(0.074)	(0.086)	(0.122)	
Secondary outcomes								
Ever told hypertension	0.144*	0.180*	0.073	440	0.039	0.025	0.065	486
	(0.081)	(0.096)	(0.102)		(0.074)	(0.089)	(0.109)	
Ever treated for hypertension	0.192**	0.226**	0.125	440	0.017	-0.015	0.075	486
	(0.077)	(0.091)	(0.097)		(0.073)	(0.086)	(0.112)	
Currently treated for hypertension	0.203***	0.222**	0.165*	440	0.036	-0.011	0.124	486
	(0.076)	(0.090)	(0.093)		(0.071)	(0.083)	(0.112)	

Note: See Table 1 for definitions of outcomes. MRDD estimates obtained from linear regression estimates of the outcome response surface, eq.(5) and bivariate kernel estimates of the conditional densities and weights that are combined with the outcome response estimates using eq.(4) and, for columns (1) and (5), eq.(3). Columns 1-3 and 5-7 show estimates for individuals located below and above the median distance to the nearest health facility, respectively. Standard errors in parentheses are bootstrapped (1000 repetitions) for MRDD. N is the effective number of observations used in estimation. All estimates are from specifications that control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C7: Effects of a referral letter on outcomes, excluding respondents who received a referral letter because of their blood glucose and/or cholesterol levels

	Binding-score RDD		*MRDD*			
	Combined frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)
Primary outcomes						
Mean systolic BP	1.923	1528	-0.905	-0.483	-1.586	864
	(2.159)		(1.944)	(2.138)	(2.645)	
Mean diastolic BP	-1.068	1430	0.023	0.103	-0.104	864
	(1.521)		(1.130)	(1.186)	(1.738)	
Hypertensive	-0.002	1575	-0.012	-0.016	-0.007	864
	(0.057)		(0.054)	(0.060)	(0.076)	
Secondary outcomes						
Ever told have high BP	0.046	1660	0.082	0.101	0.052	862
	(0.060)		(0.056)	(0.064)	(0.078)	
Ever treated for high BP	0.059	1654	0.083	0.095	0.064	862
	(0.056)		(0.053)	(0.062)	(0.074)	
Currently treated for high BP	0.047	1695	0.096*	0.095	0.098	862
	(0.056)		(0.052)	(0.061)	(0.073)	
Estimated weights (ω_s and ω_d)				0.618	0.382	

Note:See Table 1 for definitions of outcomes. Binding-score RDD estimates of eq.(2) obtained from local linear regression, with triangular kernels and optimal bandwidths on each side of the threshold using the MSE optimal bandwidth selector. MRDD estimates obtained from linear regression estimates of the outcome response surface, eq.(5) and bivariate kernel estimates of the conditional densities and weights that are combined with the outcome response estimates using eq.(4) and, for column (3), (3). Standard errors in parentheses are bootstrapped (1000 repetitions) for MRDD and heteroscedasticity-robust for binding-score RDD. N is the effective number of observations used in estimation. All estimates are from specifications that control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$. These estimates are derived from samples that exclude respondents who received a referral letters because of the blood glucose and/or cholesterol levels.

Table C8: Unidimensional RDD estimates of effects of a referral letter on outcomes

	Systolic frontier		*Diastolic frontier*	
	Effect	N	Effect	N
	(1)	(2)	(3)	(4)
Primary outcomes				
Mean systolic BP	0.532	1034	2.564	764
	(2.374)		(3.599)	
Mean diastolic BP	-2.388	801	-1.088	685
	(1.620)		(2.375)	
Hypertensive	-0.028	1073	0.051	600
	(0.066)		(0.096)	
Secondary outcomes				
Ever told hypertension	-0.004	1239	0.170*	523
	(0.065)		(0.101)	
Ever treated for hypertension	0.017	1188	0.204**	405
	(0.065)		(0.102)	
Currently treated for hypertension	0.022	1162	0.168*	528
	(0.065)		(0.096)	

Note: See Table 1 for definitions of outcomes. Unidimensional RDD estimates of eq.(1) obtained from local linear regression, with triangular kernels and optimal bandwidths on each side of the threshold using the MSE optimal bandwidth selector. Standard errors in parentheses are heteroscedasticity-robust. N is the effective number of observations used in estimation. All estimates are from specifications that control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C9: MRDD estimates of effects of a referral letter using a different specification for estimating the response surfaces – see footnote 12 – with respondents more likely to be false positives

	MRDD			
	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)
Primary outcomes				
Mean systolic BP	1.982	3.132	-0.839	622
	(2.413)	(2.632)	(3.247)	
Mean diastolic BP	0.702	0.642	0.850	622
	(1.322)	(1.390)	(1.985)	
Hypertensive	0.091	0.104	0.060	622
	(0.067)	(0.075)	(0.089)	
Secondary outcomes				
Ever told hypertension	0.085	0.085	0.085	622
	(0.067)	(0.077)	(0.085)	
Ever treated for hypertension	0.115*	0.112	0.121	622
	(0.064)	(0.074)	(0.081)	
Currently treated for hypertension	0.112*	0.103	0.134*	622
	(0.063)	(0.074)	(0.079)	

Note: See Table 1 for definitions of outcomes. Main MRDD estimates using respondents more likely to be false positives and reported in Table 3 use specification of the outcome response surface given by eq.(5). Columns (1)-(3) show estimates using the specification of the response surface given in footnote 12. Bootstrapped (1000 repetitions) standard errors in parentheses. N is the effective number of observations used for estimation. All specifications control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C10: MRDD estimates of effects of a referral letter using a different specification for estimating the response surfaces – see footnote 12 – with respondents less likely to be false positives

	MRDD			
	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)
Primary outcomes				
Mean systolic BP	-4.016	-5.486	-0.847	299
	(2.857)	(3.358)	(4.180)	
Mean diastolic BP	-2.636	-2.223	-3.528	299
	(2.008)	(2.303)	(3.184)	
Hypertensive	-0.215**	-0.217**	-0.213	299
	(0.093)	(0.103)	(0.143)	
Secondary outcomes				
Ever told hypertension	0.132	0.134	0.129	297
	(0.107)	(0.126)	(0.145)	
Ever treated for hypertension	0.091	0.109	0.053	297
	(0.101)	(0.119)	(0.141)	
Currently treated for hypertension	0.132	0.112	0.177	297
	(0.100)	(0.117)	(0.138)	

Note: See Table 1 for definitions of outcomes. Main MRDD estimates using respondents less likely to be false positives and reported in Table 4 use specification of the outcome response surface given by eq.(5). Columns (1)-(3) show estimates using the specification of the response surface given in footnote 12. Bootstrapped (1000 repetitions) standard errors in parentheses. N is the effective number of observations used for estimation. All specifications control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C11: MRDD estimates of effects of a referral letter with respondents more likely to be false positives using a wider interval around the diastolic frontier

	MRDD			
	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)
Primary outcomes				
Mean systolic BP	3.737*	4.435*	0.768	816
	(2.105)	(2.310)	(3.250)	
Mean diastolic BP	0.725	0.373	2.225	816
	(1.183)	(1.275)	(2.048)	
Hypertensive	0.116**	0.113*	0.130*	816
	(0.058)	(0.064)	(0.090)	
Secondary outcomes				
Ever told hypertension	0.008	0.005	0.024	816
	(0.060)	(0.067)	(0.084)	
Ever treated for hypertension	0.017	0.000	0.092	816
	(0.057)	(0.064)	(0.081)	
Currently treated for hypertension	0.011	-0.012	0.107	815
	(0.056)	(0.063)	(0.081)	

Note: See Table 1 for definitions of outcomes. Main MRDD estimates using respondents more likely to be false positives and reported in Table 3 use observations within 0.7 of a standard deviation (SD) of each centered and standardized running variable from the respective threshold. Columns (1)-(3) show estimates using the main estimate observations plus those up to 1 SD below the diastolic threshold. Bootstrapped (1000 repetitions) standard errors in parentheses. N is the effective number of observations used for estimation. All specifications control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C12: MRDD estimates of effects of a referral letter with respondents less likely to be false positives using a wider interval around the diastolic frontier

	MRDD			
	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)
Primary outcomes				
Mean systolic BP	-3.262	-3.659	-2.292	393
	(2.744)	(3.128)	(3.899)	
Mean diastolic BP	-1.818	-0.952	-3.937	393
	(1.617)	(1.696)	(2.843)	
Hypertensive	-0.145*	-0.123	-0.199	393
	(0.082)	(0.089)	(0.142)	
Secondary outcomes				
Ever told hypertension	0.000	-0.037	0.091	390
	(0.093)	(0.102)	(0.156)	
Ever treated for hypertension	0.000	0.006	-0.014	390
	(0.094)	(0.101)	(0.160)	
Currently treated for hypertension	0.050	0.019	0.126	390
	(0.090)	(0.098)	(0.151)	

Note: See Table 1 for definitions of outcomes. Main MRDD estimates using respondents less likely to be false positives and reported in Table 4 use observations within 0.7 of a standard deviation (SD) of each centered and standardized running variable from the respective threshold. Columns (1)-(3) show estimates using the main estimate observations plus those up to 1 SD below the diastolic threshold. Bootstrapped (1000 repetitions) standard errors in parentheses. N is the effective number of observations used for estimation. All specifications control for age and sex. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table C13: Effects of a referral letter on outcomes, excluding respondents who received a referral letter because of their blood glucose and/or cholesterol levels – high false positive likelihood

	Binding-score RDD		*MRDD*			
	Combined frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)
Primary outcomes						
Mean systolic BP	5.616**	979	1.949	2.929	-0.175	567
	(2.581)		(2.492)	(2.844)	(3.292)	
Mean diastolic BP	0.913	1061	1.406	1.197	1.859	567
	(1.632)		(1.442)	(1.568)	(2.106)	
Hypertensive	0.102	1064	0.100	0.111	0.077	567
	(0.067)		(0.067)	(0.075)	(0.092)	
Secondary outcomes						
Ever told have high BP	0.036	1189	0.056	0.072	0.022	567
	(0.068)		(0.068)	(0.077)	(0.089)	
Ever treated for high BP	0.062	1117	0.084	0.085	0.082	567
	(0.066)		(0.063)	(0.072)	(0.085)	
Currently treated for high BP	0.032	1138	0.075	0.069	0.087	567
	(0.069)		(0.063)	(0.072)	(0.085)	

Note: As Table 2 (see notes to that table), except here we use only respondents for whom, at baseline, (a) time of interview $<$ 12 noon or $>$ 3pm or (b) outdoor air temperature $<$ 80^o Fahrenheit. These estimates are derived from samples that exclude respondents who received a referral letters because of the blood glucose and/or cholesterol levels.

Table C14: Effects of a referral letter on outcomes, excluding respondents who received a referral letter because of their blood glucose and/or cholesterol levels – low false positive likelihood

	Binding-score RDD		*MRDD*			
	Combined frontier	N	Combined frontier	Systolic frontier	Diastolic frontier	N
	(1)	(2)	(3)	(4)	(5)	(6)
Primary outcomes						
Mean systolic BP	-3.017	465	-5.285*	-6.224*	-3.371	273
	(3.707)		(2.986)	(3.399)	(4.360)	
Mean diastolic BP	-4.188	463	-2.393	-1.769	-3.665	273
	(2.734)		(1.923)	(2.192)	(3.036)	
Hypertensive	-0.223*	431	-0.227**	-0.231**	-0.219	273
	(0.117)		(0.090)	(0.099)	(0.145)	
Secondary outcomes						
Ever told have high BP	0.041	547	0.129	0.146	0.093	271
	(0.103)		(0.104)	(0.125)	(0.162)	
Ever treated for high BP	0.052	538	0.096	0.119	0.028	271
	(0.101)		(0.101)	(0.122)	(0.160)	
Currently treated for high BP	0.070	546	0.168*	0.169	0.166	271
	(0.101)		(0.097)	(0.118)	(0.148)	

Note: As Table 2 (see notes to that table), except here use only respondents for whom, at baseline, (a) time of interview was $\geq$ 12 noon and $\leq$ 3pm and (b) outdoor air temperature $\geq$ 80^o Fahrenheit. These estimates are derived from samples that exclude respondents who received a referral letters because of the blood glucose and/or cholesterol levels.

UCD CENTRE FOR ECONOMIC RESEARCH – RECENT WORKING PAPERS
WP22/23 Karl Whelan:'Risk Aversion and Favorite-Longshot Bias in a Competitive Fixed-Odds Betting Market' August 2022
WP22/24 Diane Pelly, Orla Doyle: 'Nudging in the workplace: increasing participation in employee EDI wellness events' October 2022
WP22/25 Margaret Samahita, Kevin Devereux: 'Are economics conferences gender-neutral? Evidence from Ireland' October 2022
WP22/26 Michele Gubello, Nora M. Strecker: 'Weakly Progressive: Disproportionate Higher Education Attendance and the Structure of Income Taxes' November 2022
WP22/27 David Madden: 'Mental Health in Ireland During the Covid Pandemic: Evidence from Two Longitudinal Surveys' December 2022
WP22/28 Ronald B. Davies, Zuzanna Studnicka: 'Tariff Evasion, the Trade Gap, and Structural Trade' December 2022
WP23/01 Xidong Guo: 'An Analysis of a Rural Hospital's Investment Decision under Different Payment Systems' January 2023
WP23/02 Emanuele Albarosa, Benjamin Elsner: 'Forced Migration and Social Cohesion: Evidence from the 2015/16 Mass Inflow in Germany' January 2023
WP23/03 Xidong Guo: 'Hospital Efficiency and Consultants' Private Practices: Analysing the Impact of a Voluntary Reform' February 2023
WP23/04 Tadgh Hegarty, Karl Whelan: 'Calculating The Bookmaker's Margin: Why Bets Lose More On Average Than You Are Warned' February 2023
WP23/05 Tadgh Hegarty, Karl Whelan: 'Forecasting Soccer MatchesWith Betting Odds: A Tale of Two Markets' February 2023
WP23/06 Karl Whelan: 'How Do Prediction Market Fees Affect Prices and Participants?' March 2023
WP23/07 Kevin Devereux, Zuzanna Studnicka: 'Local Labour Market Concentration and Wages in Ireland' March 2023
WP23/08 Kevin Devereux, Zuzanna Studnicka: 'When and Where do Minimum Wage Hikes Increase Hours? Evidence from Ireland' April 2023
WP23/09 Karl Whelan 'Fortune's Formula or the Road to Ruin?The Generalized Kelly CriterionWith Multiple Outcomes' April 2023
WP23/10 Manuel E. Lago: '#MeToo...or not? Do salient shocks affect gender social norms?' April 2023
WP23/11 Cormac Ó Gráda, Tyler Anbinder, Simone A. Wegge: 'Gaming theSystem: The Not-So-Poor and Savings Banks in Antebellum New York' April 2023
WP23/12 Tadgh Hegarty, Karl Whelan: 'Disagreement and Market Structure in Betting Markets: Theory and Evidence from European Soccer' May 2023
WP23/13 Tadgh Hegarty, Karl Whelan: 'Do Gamblers Understand Complex Bets? Evidence From Asian Handicap Betting on Soccer' May 2023
WP23/14 Matthew Amalitinga Abagna: 'Special Economic Zones and Local Economic Activities in Ethiopia.' May 2023
WP23/15 Gregory Corcos, Stefanie Haller: 'Importer Dynamics: Do Peers Matter?' July 2023
WP23/16 Karl Whelan: 'How Does Inside Information Affect Sports Betting Odds?' July 2023
WP23/17 Michele Gubello: 'Social Trust and the Support for Universal Basic Income' July 2023
WP23/18 Constantin Bürgi, Zuzanna Studnicka: 'Online Interactions, Trade And The Gravity Model' August 2023
WP23/19 Juan S. Morales, Margaret Samahita: 'Can Social Pressure Stifle Free Speech?' August 2023

UCD Centre for Economic Research
Email economics@ucd.ie