# Can Privatized Health Care Add Value? The Mexico Diabetes Experiment *

Ari Bronsoler† Jonathan Gruber‡ Enrique Seira§

*MIT* *MIT and NBER* *ITAM and JPAL*

August 28, 2021

**Abstract**

We implement a novel deniers randomization evaluation of a private supplement to the free public health system for one of the world's deadliest health problems, diabetes. We estimate enormous impacts of the private supplement, increasing the share of those treated who are under control by 69%. This effect arises through both improved treatment compliance and health behaviors. Diabetes complications fall in the short run. The net costs of this intervention are at most one-third of the gross costs, and the returns to private care do not appear to reflect more productive delivery but rather more attachment to medical care. (JEL C93, I11, I15, I18, H51).

*We are grateful to Eduardo Rivera for outstanding research assistance and to Dante Salazar for impressive field coordination. We also want to thank Mohit Karnani, Diego del Pilar and Andres Santos for their contributions as well as Jishnu Das, participants in the Harris Public Policy and Economics Workshop and the development and public finance lunches at MIT for their insightful comments and suggestions. Lastly, we are very grateful to Clinicas del Azucar for being an amazing partner and Eli Lilly and Company for supporting this project.

†aribro@mit.edu

‡gruberj@mit.edu

§enrique.seira@itam.mx

# 1 Introduction

Most countries of the world feature universal or near universal public provision of either health insurance, or direct health care, to residents. And in all of these countries, this public provision is to some extent supplemented by private options, such as private insurance products or private health care providers. A common feature of all nations with such mixed public-private systems is dissatisfaction with the current mix. Regardless of the relative size of the private alternative, advocates and opponents of privatization argue for a larger or smaller public sector role. A classic example is the debate in the U.S. over allowing Veterans to access private providers outside of the Veterans' Affairs system, as well as debates in the U.K. over tendering of private contracts for their National Health System (see the discussion in Frakes et al. (2021)).

This debate extends to the developing world as well. Consider the case of Mexico, which has a series of largely free public health care systems, the largest of which covers 70.3 million formal sector workers and their families in the nation (IMSS, 2020). There is widespread dissatisfaction with these public health care systems; 22% of those who used the public sector said they would not return for future care if given a choice, and under a quarter rate the service as very good (INSP, 2018). As the result of large waiting times, a nascent private sector has emerged to provide complementary primary care.[1] The share of private medical offices rose from 5% in 1990 to over 30% in 2020 (SSP, 2021). Yet a plan by the economist running Mexico's largest social insurance institution to test public reimbursement for private medical services was strongly opposed by the unionized public sector doctors and ultimately shut down.[2]

These debates highlight the importance of understanding the costs and benefits of private delivery of health care when there is a free public system. Yet evaluating the impact of private alternatives faces a number of challenges. Use of private providers is endogenous, making it difficult to use cross-sectional variation in source of care to learn about the effectiveness of private care. Randomized trials are also very difficult in this context due to the challenges of identifying and recruiting those interested in using private care, and the enormous samples needed to obtain sufficient first stage power. As a result, the critical debate over private alternatives continues largely in an evidence vacuum.

This paper tries to fill that evidence void through a randomized trial evaluating private treatment for one of the most important public health issues facing developing countries: diabetes. After decades of being primarily concerned with undernutrition around the world, policy makers are shifting their focus to this new problem that arises from both improper eating and overconsumption. There were 4.2 million deaths due to diabetes complications in 2019. Worldwide prevalence rates

[1] IMSS has an average of 1 hour of 15 minutes in waiting times while private sector has only a 20 minutes wait (INSP, 2018).

[2] See almomento.mx (2015)

have risen from 4.7% in 1980 to 9.3% in 2019 and the increase has been most rapid for developing nations; the rate of diabetes is now *higher* in low-income nations than in high-income nations (Saeedi et al., 2019).

We focus on Mexico, one of the nations hardest hit by the growth in diabetes. Diabetes prevalence has risen from 6.7% in 1994 to 11% in 2018, and, signaling poor control, the country has almost twice the mean diabetic hospital admission rate among the OECD countries. Indeed, diabetes is the second most common cause of death in Mexico and is among the top five causes of disability. Alongside these poor outcomes, Mexico spends enormous amounts combating this illness: The estimated costs of addressing diabetes and its complications are 2.25% of GDP, and diabetes spending represents 10% of the entire budget of the Ministry of Health. (OECD, 2019; Barraza-Lloréns et al., 2015; INEGI, 2019b; INSP, 2018). Due to dissatisfaction with the diabetes care provided by the free public system, nearly 20% of diabetics report getting treatment at a private institution for their diabetes. This raises the key questions of whether this private care is improving health outcomes, whether it is cost effective in doing so, and whether the public sector could mimic any successes or would be best served by outsourcing to private providers.

In this paper, we implement a novel *deniers randomization* approach to cost-effectively provide a causal estimate of enrollment in private diabetes care. We run this experiment in partnership with a private provider of comprehensive diabetes care in Mexico, Clinicas de Azucar (CdA). CdA runs a chain of clinics that provide a wide range of services to diabetics, ranging from blood sugar measurement to medical interventions to nutritional counseling. This service is fairly expensive, at a cost of 7000 Pesos ($350 USD) per year, or 5% of median Mexican family yearly income, at the time of our experiment.[3]

Our initial sample is individuals who undergo a free initial comprehensive diabetes evaluation at CdA, which we randomize into treatment and control groups and survey at baseline. After their evaluation, these individuals are offered the opportunity to enroll in CdA at full-price, and 31% do so. Among those who decline and are willing to hear additional offers, an additional opportunity to enroll at 40% of the baseline price is presented to treatments but not controls. Among those offered this discount, 49% end up using the CdA service, compared to 21% enrollment for the control group that does not get this offer (but may receive subsequent marketing from CdA). This deniers randomization provides a powerful first stage to investigate the impact of CdA treatment in a cost-effective way – we estimate costs that are one-tenth of what they would have been if we had randomized among all those who went for the evaluation – while imposing a minimal set of restrictions are met in our context. Due partly to the COVID outbreak, we have significant attrition in following up our experimental sample, but the sample is well balanced along all measurable

[3]Estimation based on monthly earnings from the 5th and 6th decile in the distribution in INEGI (2018). The cost of CdA would represent 5.4% of income for the 5th decile and 4.5% for the sixth decile. The cost of CdA has risen recently to 8000 Pesos.

dimensions and we see no evidence of differential attrition.

Using this approach, we find a striking positive effect of the CdA intervention: the implied effect of participation is to lower blood sugar levels (measured by glycated hemoglobin or HbA1c) by a full point (relative to a control mean of 8.5%), and to increase the share of those treated who are under control by 69%. This is a huge impact, which according to the widely cited UKPDS study, could reduce microvascular diabetes complications by 35%. Moreover, this impact is at the upper end of estimated effect sizes from other diabetes interventions reviewed in a recent meta-analysis (Pimouguet et al., 2011a).

We show that this effect arises through a number of changes to behavior, including greater compliance with recommended medications and substitution for less invasive treatments; some change in behavior such as exercise and diet, and importantly more frequent visits to medical providers. We find that diabetic complications fall, even in the short run. We also find significant heterogeneity, with those who have worst baseline blood sugar control seeing the largest benefit.

To consider further the welfare implications of our results, we then extend them in two ways. First, we explore the cost effectiveness of the CdA intervention. The net costs of this intervention are much lower than the gross costs because of savings to the public health system. Some of these savings arise from reduced use of public primary care. But the larger source of savings is the direct positive fiscal externality from improved private care in terms of reduced (publicly paid) hospital spending. Adding these components, we estimate that the net benefit of this supplement in terms of money is between 65 and 105% of its gross costs. At the same time, our estimated health benefits are an order of magnitude greater than the gross costs of the program.

Second, we assess whether these returns to private care reflect a more productive delivery of care per visit or simply more quantity of care that is delivered equally effectively. We use data from the major public sector insurer (IMSS), along with variation coming from the distance to public clinics, to quasi-experimentally estimate the marginal returns per visit to IMSS diabetes care. We find that, after controlling for differential observable selection into private sector care, the returns to public and private care per visit are in fact comparable. This suggests that the returns from CdA reflect the ability of the private vendor to encourage more care. We confirm this conclusion by showing that our treatment effects are in fact highest where public clinics are most crowded.

Taken together, our findings suggest that private delivery of diabetes care had major benefits in Mexico. It led to improved health and significant offsets to public sector hospital expenditures, and at standard values of improved health was highly cost effective. But the source of the improvements was not necessarily better technology for diabetes care, but rather encouraging more care. This suggests multiple paths forward for governments seeking to improve their diabetes care, ranging from outsourcing care, to improving the attractiveness of the public option. Our paper also introduces a new cost-effective approach to randomization in evaluating private alternatives for public services

delivery.

Our findings contribute to the long-standing discussion on public vs private healthcare provision. In a systematic review of the literature, Basu et al. (2012) document that the private sector is usually not more efficient, accountable, or medically effective than the public sector but offers better waiting times and hospitality towards patients. Das et al. (2012) and Das et al. (2016) document through standardized patient comparisons in India that both the public and private sector offer similar (low) quality services, with public sector physicians being more prepared but private sector practitioners compensating with more effort per appointment. In the U.S. context, two recent studies evaluate the role of private options relative to health care for the nation's military and veterans, with mixed outcomes; Frakes et al. (2021) find that children of military personnel born in the private sector have higher costs but better outcomes than those born on military bases, while Chan et al. (2020) find that those receiving emergency care at Veterans Administration hospitals see lower costs and better outcomes. We contribute to this literature by showing that there is a large health benefit of adding a (subsidized) private option, even for those with free public healthcare, and even if the private option is not necessarily of higher quality per unit of delivery. This is especially true in a context where public healthcare is overstretched. Second, we run one of the largest double-blind diabetes field experiments to date, and show that standard diabetes care with simple and cheap technology can have an enormous impact on reaching normal sugar levels.[4] Finally, we show that large fiscal externalities pay many-fold for the subsidy we implemented, contributing to the literature on fiscal externalities from health interventions (see Chandra et al. (2010) for a review of this literature).

Our paper proceeds as follows. Section 2 provides background on diabetes in general and the Mexican context, on the Mexican public health care system, and on the role of private alternatives such as CdA. Section 3 describes CdA and the design of the Mexico Diabetes Experiment. Section 4 presents our basic results on outcomes, mechanisms, and heterogeneity. Section 5 estimates the spillovers onto the public sector and the total social value of the improved care. Section 6 then explores the relative efficacy of public and private sector care. Section 7 concludes.

# 2 Diabetes and the Mexican Health Care Context

## 2.1 Diabetes Consequences and Measurement

Over the past 25 years, one of the fastest growing public health problems around the world has been diabetes. Diabetes is a progressive and often-fatal disease with no known cure. It can attack

[4] Meta reviews of different kinds of interventions can be found in Ismail et al. (2004); Umpierre et al. (2011); Ajala et al. (2013); Pimouguet et al. (2011b). Notable exceptions to small sample size studies are those focused on trying to understand the correct level of HbA1c for diabetics, like the UKPDS 35 study (King et al., 2001)

every organ in the body, resulting in higher risk of heart failure, stroke and poor circulation, which can lead to amputation of extremities, kidney failure, retinopathy and death. Those with Type I diabetes don't produce insulin, which turns glucose (sugar) into energy; those with Type II diabetes don't respond to insulin appropriately and don't make sufficient amounts of insulin. Worldwide, more than 450 million people are estimated to have diabetes.

While diabetes cannot currently be cured, it can be brought under control by following diet and exercise recommendations, closely monitoring blood sugar levels, and adjusting prescriptions accordingly.[5] Unlike several other chronic illnesses such as AIDS or Hepatitis C, diabetes can be easily and cheaply managed with relatively inexpensive medicine; Metformin, which is the most common pill used to control early-stage diabetes, costs under 2 dollars/month.[6]

The gold standard for measuring diabetes status is glycated hemoglobin. A hemoglobin A1c (HbA1c) test measures the amount of blood sugar (glucose) attached to hemoglobin, the part of red blood cells that carries oxygen from lungs to the rest of the body. An HbA1c test shows what the average amount of glucose attached to hemoglobin has been over the past three months; a three month average is used because that's typically how long a red blood cell lives. Individuals are diagnosed as diabetic with an HbA1c level of over 6.5% and diabetic patients are recommended to keep their levels below 7% (NIDDK, 2018). A more accessible and easier to use instrument to monitor blood sugar is the glucometer, which captures the blood sugar levels at any one point in time. While this measure has significantly more variance than HbA1c, it does not require lab processing and patients can use it in the privacy of their own home. Normal levels for this measurement are under 100 mm/hg and a patient will be diagnosed if she gets two fasting measurements over 125 (WHO, 2021).

## 2.2 The Mexican Health Care System

Health care in Mexico is provided primarily by several public sector institutions. The largest is Instituto Mexicano del Seguro Social (IMSS), the single payer insurance plan for formal sector workers in the country. This program covers formal workers and their families as well as students but also offers a voluntary enrollment option which makes up under 1% of beneficiaries. Every private employer that hires a new employee is required to enroll him/her to IMSS. This service is paid for in 3 parts: On average, the government contributes 5.3% of employees base wages, employers contribute 16.5% and employees another 2.5%.[7] IMSS runs its own 1522 primary care clinics, 248 acute care hospitals, and 61 specialty hospitals (IMSS, 2018). Smaller but similar public options for particular sectors such as government workers (ISSSTE), the navy (SEMAR),

[5] Saeedi et al. (2019) and IDF (2019).
[6] Clinic (2020).
[7] Social Security Law (1995)

the army (SEDENA), and for workers of the state-owned oil company (PEMEX).

In 2003, Mexico introduced a new program, Seguro Popular, to extend health care coverage to informal workers. This program recently changed its name to the wellness institute (INSABI) and currently covers 60 million people, although everyone is eligible to enroll (SSP, 2018). While this service expansion benefits workers in the informal sector, survey evidence shows that Mexicans prefer other options.[8] Today, 83% of the population reports being affiliated to one of these systems, with 45% at formal sector systems, 38% at INSABI, and 0.6% holding private insurance. Total health care spending in Mexico is 5.6% of GDP (INEGI, 2017, 2019a).

Despite the availability of free public health care, diabetes remains a major problem in Mexico. The public health care system in Mexico has responded with a number of policy efforts, including a program to encourage annual checkups, large scale programs to encourage active lifestyles, and a new tax in 2014 on sugary drinks and high-caloric foods.[9] Despite these efforts, obesity rates have not receded in Mexico, and diabetes diagnosis rates have remained at 11% of the population since 2012 (INSP, 2018, 2012; Ángel Rivera Dommarco et al., 2018). Diabetes patients who do not have the disease under control face much higher risks of hospitalization and disability. Mexico has twice the rate of hospitalizations per diabetic rate than the OECD average, and diabetes was the second highest cause of death in the country in 2019 (OECD, 2019; Ojeda, 2019).

While public health care is free, widespread dissatisfaction with the quality and waiting times of the public sector has caused the rapid growth of private health care systems in Mexico. This private care is primarily focused at the primary level; the share of private medical offices went from 5% in 1990 to over 30% in 2020. (SSP, 2021; INSP, 2018). In 2018, 18.2% of diabetics reported getting treatment on private institutions. Patients affiliated to either the formal sector or INSABI programs can get their medications for free in the pharmacies of their clinics, but often choose to pay a small amount to reduce wait times by going to private pharmacies instead. [10]

# 3 The Mexico Diabetes Experiment

## 3.1 Clinicas de Azucar

One of the private providers of disease care management is Clinicas de Azucar ("sugar clinics"). This chain of clinics was founded in the state of Monterrey by U.S. educated health care entrepreneur Javier Lozano. The first clinic was established in 2011, and the chain has grown to

[8] Based on data from the health and nutrition survey in 2018 that asks about satisfaction with a service. Data shows that the following percentage of patients believe the service is very good: 40% from private, 26% for IMSS and ISSTE and 20% for SP

[9] Aguilar et al. (2021), and Colchero et al. (2017).

[10] Health Federal Law (1983)

24 clinics in 5 states.

The Sugar Clinics are a chain of specialized, diabetes clinics that provide affordable and comprehensive care. Each patient pays fixed-cost membership fees allowing him/her to have unlimited access to diagnostics, labs and consultations for one year. One of the main selling points of the clinics is that a patient can receive a full diabetes check-up with nutritionist assessments, and recommendations for diet, exercise and medication in under 90 minutes, avoiding several visits and long wait-times in the public sector. Appendix A gives more detail on the Sugar Clinics and shows pictures.

The sugar clinics do not offer a revolutionary type of care nor add many benefits to what is already available for free in the public sector. The main advantage is that a patient can go to any of the branches, whenever it suits them to get the care they need without having to wait. Table A.A.-1 compares the services provided by the largest public healthcare provider IMSS and CdA and examines how much it would cost IMSS to provide the same services that CdA offers, according to their reported per unit cost. We can see that both suppliers offer similar services and that IMSS would spend 20% more to provide the same services.

The approach used by CdA parallels a disease management program generally applied internationally for chronic-obstructive pulmonary disease, certain types of cancer, and diabetes; a similar approach is used by Joslin Diabetes Center in Boston and Apollo Sugar Clinics in India.[11] We therefore view our project as evaluating more generally the provision of privatized diabetes care. Although we cannot say with certainty whether the results from CdA extrapolate to other private providers, our findings on mechanisms in Section V suggest that the effects may be quite general.

Non-causal estimates of the impact of CdA are very promising. Estimates from CdA indicate that enrollees see their HbA1c levels fall by 2 points relative to baseline, and such an effect appears to be lasting. Based on these promising findings, we partnered with CdA to design an experimental intervention to assess the causal impact of their program, with funding from Eli Lilly and Company. We preregistered in the AEA registry.[12]

## 3.2 Deniers randomization

The past literature suggests three natural ways to set up such an experimental intervention. The first, which we call "overall randomization", involves finding a representative sample of Mexico's diabetics and incentivizing a random sub-sample of them to get CdA service. This would have been extremely hard and expensive as it would have required us to test a large sample of individuals to assess whether they are diabetic, and then give incentives them to go to CdA clinics. Given potentially small take up, the sample size would have to be huge to achieve standard statistical

[11] IQEHC (2007)

[12] https://www.socialscienceregistry.org/trials/3589

power. In many contexts such as ours, such population screening and randomization is infeasible.

A second approach to inducing differential use of CdA services would be what what we call "visitors randomization", which would in our case involve randomizing all those who arrived at CdA for a free screening, with treatments receiving a subsidy to ultimately enroll in the program. Properly designed, this would allow us to estimate the LATE for those who are interested in CdA services, and who are convinced to use it by our incentive. This parameter is of inherent policy interest, as only some individuals are potentially interested in private alternatives, so that we estimate the LATE for this group. This visitors randomization approach is standard in many settings as it is simple to implement and provides a reliable causal estimate of the intervention as long as the encouragement works.

The disadvantage of this approach, however, is that it is very underpowered when the private option has a high baseline enrolment rate. At CdA, for example, among those who make an initial visit, 31% voluntarily pay for and use the program immediately. Moreover, another 36% show no interest in enrolling in the program at any discount. This 67% combined rate of identified always and never takers severely limits the power of the first stage; moreover, it also implies that we would be giving an incentive to the 31% of patients that we know would use the program anyway. This is not just a problem of our context, it is a general problem faced by any experiment that is testing the implications of a private supplement which has high levels of interest.

We therefore introduce a third new approach we call *deniers randomization*. It consists of screening out always takers and/or never takers from the sample. Under some (weak) assumptions, we can estimate the same LATE that would be estimated with visitors randomization, but with a fraction of observations needed to achieve a given statistical power. To fix ideas we will explain how this screening worked in our context and then revisit the additional assumptions, arguing why they are likely to hold in our case.

### 3.2.1 Our Experiment

Our experiment proceeds in several steps, as illustrated in Figure 1. When patients enter CdA clinics, they fill out the baseline survey while waiting for the free check up process to begin. After they undergo check-up, individuals met with a physician who discussed their diagnosis and suggested a potential care package at CdA. We removed from our sample people who were not diagnosed with diabetes. Then, individuals met with a sales force associate who offered to enroll in CdA at the standard full price. Some of them did enroll at full price and some refused to enroll. We call those that enroll always-takers, and they do not form part of our experimental sample.

Our experiment began when potential clients indicated that they were not willing to buy. At that point, the sales associate asked the client if she was willing to wait 10 seconds to search in

the system for discounts. Some clients said they were not interested, we call these never-takers.[13] Those that said they could wait were consulted for treatment status in the computer. We observe in the data whether a person was consulted or not and define our experimental sample based on that variable, since always takers had already bought and did not need to be consulted and never takers were not interested in any offers. If the individual showed up as being in the treatment group —which happened 50% of the time— the sales associate would then offer her a 60% discount.[14]

We were particularly concerned that the sales force might not wait until after the enrollee declined the full-price membership to offer them the discount lottery, thereby including most always takers and reducing the power of our deniers randomization. We pursued three strategies to address this concern. First, we offered higher bonuses to the sales force for full price than for discounted sales. Second, we carefully instructed the agents on the importance of first ensuring a lack of interest in the full-price membership before offering the discount. Third, we had bi-weekly meetings with the entire sales force where we reinforced this and presented hypothetical cases where they actively participated.

### 3.2.2 Power gains and cost savings from deniers randomization

The advantage of our deniers randomization approach over entry randomization is an enormous gain in efficiency in the experiment, both because our first stage is more powerful, and because we need not survey all always takers and never takers, which don't contribute to our estimates. Failing to remove the 67% of patients we identify as always or never takers would have made our first stage substantially weaker and we would have required a much larger sample, which would in turn imply more spending on discounts and many more follow-up surveys. Since our follow-up was performed at patients' homes to prevent differential attrition, running extra surveys would be prohibitively expensive.

We illustrate the sample size savings with a simulation displayed in Figure 2. It plots the sample size needed to achieve 80 percent power and 95 percent confidence for different effect sizes. To fit our context, we assume that 38% of the individuals are always takers, 53% are never takers and only 9% are compliers. Moreover, we assume that a filter could exclude 82% of always takers and 68% of never takers, which is what ours does in the field. So for visitors randomization, out of every 100 individuals, there would be 38 always takers, 53 never takers, and only 9 compliers. If instead we use our deniers randomization approach, our sample consists of only 7 (38*0.18) always

[13] CdA offers a free checkup as a marketing device. It is well know to CdA management that many of the people who go do not intend to buy in the first place, and only go for the check up, that is why they are unwilling to listen to any kind of discount. Removing both always takers and never takers saves on sample size and increases our statistical power.

[14] We chose the 60% discount based on a pilot run with CdA which showed that such a discount yielded a 48% take up rate.

takers, 17 (53*0.32) never takers, and 9 compliers. Thus, our first stage coefficient would be 0.27, as opposed to the 0.09 first stage we would obtain from full visitors randomization.

Figure 2 shows that we need around 2500 total observations to detect an effect of 0.4 standard deviations, while visitors randomization would require around 25,000. The differences are even more striking for smaller effects sizes more commonly found in the literature. This allowed us to operate at a budget that was one-tenth of what would have been required by visitors randomization – alternatively at the same budget, deniers randomization delivers three times the power of visitors randomization.

### 3.2.3 Deniers and visitor randomization LATE equivalence

The benefits of deniers randomization are clear, but what assumptions are needed for the LATE estimated by deniers randomization to be the same as that estimated by visitors randomization? We begin by recalling that using the potential outcomes framework, Angrist et al. (1996) show that if the stable unit treatment value assumption, the exclusion restriction, relevance, and monotonicity are satisfied, then an instrumental variable identifies the local average treatment effect:

$$\beta^{\text{Visitors}} = \frac{\mathbb{E}[y_i|Z_i=1] - \mathbb{E}[y_i|Z_i=0]}{\mathbb{E}[D_i|Z_i=1] - \mathbb{E}[D_i|Z_i=0]} = \mathbb{E}[y_i(1) - y_i(0)|D_i(1) - D_i(0) = 1]$$

Note that the LATE uses only compliers for identification. Excluding always takers and never takers from the sample would not change $\beta^{\text{Visitors}}$. Using the set up of Angrist et al. (1996), in Appendix B, we show that the LATE estimated from the deniers randomization approach is equivalent to the LATE estimated from visitors randomization under three additional assumptions:

1. *No exclusion of would-be compliers.* That is, we screen out only sets of always takers and/or never takers from the sample, and in doing this we do not screen out compliers.

2. *No direct causal effect of screening on enrollment.* The screening procedure does not itself directly affect enrollment decisions, say by changing the preferences that the visitor has for CdA.

3. *No direct causal effect of screening procedure on the outcome of interest.* For instance, the screening procedure itself does not have a direct effect on long term blood sugar levels.

In general, whenever the screening procedure is quick and painless, as in our context, these assumptions are likely to hold.

We ensured that all clients who rejected the offer at full prices were consulted for treatment status within 10 seconds (unless the client was not interested), and immediately offered the discount

if they were in the treatment group. We verified that the sales force was not excluding would-be compliers by using mystery shoppers to test our process. This makes assumption one plausible.

Given that neither treatment nor controls experience denial of service, and that the screening is very standard and natural —i.e. a quote of the price— it is unlikely that there is a direct causal effect of screening on enrollment in CdA (Assumption 2). The visitor only needs to say she is not interested to buy in order to trigger the offer for the treatment group immediately. One way our screening procedure could affect the decision to enroll is if, by rejecting the offer at full price, the visitor herself affects her own preferences for CdA. This would certainly not be true in the neo-classical framework where demand has a negative slope and elicitation of demand does not change preferences; it is plausible under particular behavioral models but once again unlikely because we are offering a new price which is lower.

The third assumption is that the screening procedure does not directly affect outcomes of interest, in particular HbA1c. This is satisfied since the deniers randomization certainly does not affect the quality of health services received at CdA; the sales force clicks on a screen to consult status, but they have no further interaction with the patients.

Of course, both visitor randomization and deniers randomization occur once individuals express interest in CdA, and as such they do not necessarily capture the effect for the average diabetic in Mexico. We discuss issues of external validity below.

## 3.3 Sample

We recruited individuals over the period from June 2019 to February 2020, Figure 3 summarizes our recruitment. We approached 7,882 individuals who showed up at CdA offices for a free check up. All of these individuals were asked to fill out a baseline survey while waiting to receive the checkup. The survey contained basic demographic questions, questions on the use of health systems, on personal health, habits, trust, among others. 94% of both the Treatment and Control arms completed at least part of the baseline survey.[15] Moreover, we can see that among patients who filled out the survey, the sales force screened out 67% of the individuals: the 31% who were always takers of the initial offer, and the 36% who were never takers, uninterested in our discount offer. Therefore, 33% were consulted for treatment/control status in the second step.

The last row of Figure 3 shows that the balance of our sample along a number of dimensions is excellent. We ended up with 1226 individuals in the treatment arm, and 1184 individuals in the control arm. They are very well balanced in terms of baseline HbA1c, weight and age. Importantly, there is a significant difference in the odds of using the services provided by CdA. Among the controls, 21% eventually enrolled in CdA, probably because of additional efforts made by CdA's

[15] Overall, we have 87% of surveys with complete demographic information which we use on our main specifications, and our results do not change when restricting the sample to full surveys

marketing.[16] Use was 28 points higher, among the treatment group.

After an eight month enrollment period that ended on February 2020, we conducted a follow-up starting in June 2020 with professional surveyors and nurses visiting each patient's home to implement the questionnaire and take a blood-sample that allowed us to measure HbA1c, the key diabetes control metric. Our survey team was composed of part of the personnel who do Mexico's main health survey (ENSANUT), they are highly trained and complied with stringent health protocols. Our follow up was delayed by a month due to COVID, after which we were able to undertake the in home survey due with proper COVID precautions. We did face sizeable attrition, with only 44% of experimental participants participating in the follow up, due to two problems: incorrect addresses, and rejection of visits due to COVID. These two sources of attrition appear to be random with respect to our intervention, however. Table 1 regresses a dummy for follow up data collection on our treatment indicator – we find no evidence of differential attrition between treatment and control groups.

Table 2 shows also that there are no differences on a large battery of baseline characteristics between treatment and control, which is also comforting. The difference in HbA1c is an insignificant 0.15 points, which is only 1.6% of the mean and roughly one-eighth of our estimated treatment effect; the difference in the share out of control is an insignificant 0.02, which is below 10% of the control mean and roughly one-tenth of our estimated treatment effect. The sample is also very well balanced on demographics and type of insurance coverage.

Table 2 also shows that our treatment and control samples are much fairly unhealthy. Among controls, the mean level of HbA1c is 9.35, well above the control level of 7, and 76% of our sample is out of control. Mean BMI is 31, which is outside the recommended range of [18.5-25] and means that the average person who shows up to CdA is obese. The mean age of the control group is 52.4 years, and 32% are male; the lower male share may reflect the willingness of men to buy at full price, while wives may want to consult with spouses as well as the fact that women are more keen to answer our follow-up. Roughly 75% have access to public health systems. The self-reported level of trust in CdA is higher than that alternative providers.[17]

To assess external validity, we can compare our sample to the the full sample of diabetics surveyed in the Mexican health and nutrition survey (ENSANUT) on table A.C.-1. We find that our sample has a higher HbA1c level and is more likely to be getting treated for their diabetes, but is less likely to have hypertension and has a similar overall level of complications.

[16] Following typical CdA practice, people who did not purchase would be called several times during the next two weeks to encourage enrollment. This follow up is orthogonal to treatment status, as treatment status was blind to the marketing department. These subsequent contacts would also offer additional small (10-20%) discounts over the full price.

[17] The sample that did answer our follow-up is composed of more women, a bit poorer and less educated individuals in general, but both samples appear to have the same health on average and the same access to public institutions. These comparisons are shown in Table A.C.-3 in Appendix C.

# 4 The Causal Impacts of CdA

## 4.1 Effect on Blood Sugar

We implement our evaluation of the experiment through a straightforward regression framework. We initially estimate OLS models of the effect of HbA1c of the form:

$$Y_{i,j,t} = \beta_0 + \beta_1 U_i + \Gamma_j + \psi_t + \chi_i + \varepsilon_{i,j,t} \quad (1)$$

where $Y_{i,j,t}$ is HbA1c values for individual $i$ who enrolled in clinic $j$ on month $t$, $U_i$ is an indicator for using CdA services, $\Gamma_j$ and $\psi_t$ capture clinic and month of enrollment fixed effects and $\chi_i$ includes controls for baseline HbA1c, BMI, gender, age, schooling and income. These controls are missing for about 100 patients, lowering our regression sample size to 939; results are almost identical if we exclude the 5 variables and use the somewhat larger sample. Standard errors are robust. We then use the standard two stage least squares approach of instrumenting CdA usage by being in the treatment group to estimate the local average treatment effect of using CdA on our sample based on the same specification.

Our main results are shown in Table 3. Column 1 shows the OLS estimate of using the Cda services on blood sugar, which indicates that CdA lowers blood sugar by -0.98 points, off a base of 8.54 points. This is much lower than the 2-point estimate that CdA finds following its enrollees across time. An important part of the difference is improvement among those who do not use CdA but may go elsewhere for care – in fact, we see that our control population improves 0.9 points relative to baseline. Another part of the difference could be through selection of those who stay in the program, and therefore continue to be measured by CdA's internal metrics, and those who leave, who are also included in our evaluation.

Additionally, our results do not incorporate the value of the initial evaluation done by CdA. Our estimate compares treatments and controls conditional on the initial evaluation. But if the initial evaluation itself has value in terms of helping potential enrollees to understand how to manage their diabetes, it could lead to some decline relative to their level of blood sugar at entry to the clinic.[18] As a result, our experimental effect is a lower bound on the total impact of interacting with CdA.

The rest of the table shows the coefficients on the control variables. By far the most important control variable is baseline HBA1c, with each point in baseline HbA1c associated with a 0.5 percentage point level of HbA1c ex-post. The other significant relationship is with income, where being in the second lowest income group is associated with HbA1c that is 0.57 points higher. Interestingly, we find little impact of BMI conditional on baseline blood sugar. The other coefficients

[18] Note that this is not a particular limitation of our deniers randomization approach – any approach which conditioned on individuals arriving at CdA and being evaluated would suffer from this problem.

are also insignificant.

Column 2 shows the corresponding IV estimate, where we instrument by the treatment indicator. We find an effect that is comparable, but about 10% bigger than the naive OLS, at -1.1 points. Another way to illustrate the effects of the intervention is to look at the share of individuals who have their blood sugar under control, defined as a level of HbA1c below 7. Columns 3 and 4 present these results. For this outcome, our IV estimate shows that 22% of individuals are brought under control by the intervention, which is more than two thirds of the baseline rate of control in our sample.

These are very large effects. For example, a set of recent meta-analysis shows that, on average psychological interventions reduce HbA1c by 0.32, physical activity interventions by 0.67-0.89, self-monitoring of blood-sugar by 0.39, dietary approaches between 0.12 and 0.47 and disease management programs similar to the one we are working with show reductions in HbA1c between 0.38 and 0.45 points, less than half of what we find. This effect is also due to something more than increased adherence as the papers estimating the effect of adherence on outcomes report an effect that is about 40% of the one we find.[19] A widely cited study notes that a reduction of the magnitude we find in Table 3 is sufficient to reduce complications by 35% and reduce mortality by 4% (King et al., 2001; Arnold and Wang, 2014). Moreover, increasing the fraction of patients with HbA1c under control by 69% makes an enormous difference for life expectancy; each year a patient spends with diabetes out of control is estimated to reduce life expectancy by 100 days (Heald et al., 2020).

## 4.2 Mechanisms

Our survey results allow us to explore a variety of mechanisms through which CdA may have had its effects. The results of this analysis are shown in Table 4, using the IV version of equation 1. All dependent variables come from the responses to the follow up survey; the number of observations varies.

We find that CdA enrollment leads to a very large increase in the odds of receiving medical care. Total physician visits increase by 2.6, which is 40% of the control mean. Column 2 shows the number of visits that involve specific check-ups on potential complications from diabetes, and we find that these almost double relative to baseline. The number of visits to diabetes specialists rises by almost 50%. Clearly, an important mechanism for our HbA1c results is that patients are getting a higher *quantity* of care.

We also find evidence for the key mechanism of drug compliance. In the follow up survey we ask individuals if they are likely to stop medication if they "feel good". The right answer to this question is clearly no: diabetes cannot be cured, and ongoing medication remains very inexpensive

[19] Ismail et al. (2004), Umpierre et al. (2011) Ajala et al. (2013), Pimouguet et al. (2011b) Krapek et al. (2004) Krass et al. (2015)

compared to the underlying health risks. Yet 25% of enrollees in the control group answer yes, indicating inappropriate use of medication. We find that the estimated effect of CdA is a reduction of this response by 22%. Thus, by this measure of proper disease management, CdA causes an 88% improvement with respect to the mean.

Corresponding to this finding, we find a dramatic shift in how individuals intervene to manage their blood sugar. We see a significant rise in the use of cost effective (typically generic) blood sugar-reducing medication, with use rising by 25% from a baseline of 73%. Correspondingly, we see a 15% decline in the use of insulin, or 100% of baseline. This is an important finding for the efficiency of diabetes management, since insulin is more costly than sugar-reducing medication and is used at later stages of disease progression.

As noted earlier, CdA provides a full suite of interventions, including nutritional and exercise advice. Behaviors such as eating and exercise are notoriously difficult to influence (Jager, 2003). Yet we find suggestive evidence that, in the short term at least, CdA influenced these behaviors positively. Our follow up survey includes a dichotomous variable for whether individuals report dieting or engaging in exercise. We find that the odds of both exercise and diet rise by 14%, and that there is a marginally significant increase in the odds of either diet or exercise by 26%, from a mean of 73%. We also find insignificant declines in the number of sodas and cigarettes consumed per week.

## 4.3 Heterogeneity

We explored a wide variety of dimensions along which we might find heterogeneous impacts. To model heterogeneity, we estimate the following model:

$$Y_{i,j,t} = \beta_0 + \beta_1 T_i + \beta_2 H_i + + \beta_3 T_i * H_i + \Gamma_j + \chi_i + \varepsilon_{i,j,t} \tag{2}$$

Where $Y_{i,j,t}$ is HbA1c values for individual $i$ who enrolled in clinic $j$ on month $t$, $T_i$ is an indicator for being part of the treatment group, $H_i$ is a dummy for whether the patient has a value higher than the median of the heterogeneity variable and $\Gamma_j$ captures clinic fixed effects. The coefficient of interest for this analysis is $\beta_3$. The results are shown in Table 5.

While our estimates are somewhat imprecise, we find no significant heterogeneity across age, sex, BMI or schooling. We do find that there is significant heterogeneity by baseline level of HbA1c. Sicker individuals at baseline improve more than their healthier counterparts. To further understand how our effects are distributed we apply Athey and Wager (2019)'s method for conditional average treatment effects. Briefly, the method trains a causal forest which enables the estimation of the average treatment effect for each patient in the experiment . To start, a forest is trained using random subsamples of the data; next the algorithm gathers a weighted list of the

sample's neighbors based on what leaf nodes of the tree it falls in; finally, the treatment effect is calculated using the outcomes and treatment status of the neighbor examples.

To use these estimates to assess the impact of our deniers randomization approach versus visitors randomization, we show the CATE for both our enrollees and the set of always takers we screened out in the first step of the experiment's protocol. Figure 4 compares the CATE of both distributions. Here we can see that the effect of CdA is slightly bigger for most always takers than that of our complier sample as the right tail density for always takers is smaller. This is consistent with higher enrollment rates among those with higher returns to CdA treatment.

# 5 Fiscal Externalities on the Public Sector

In many public systems such as Mexico's, privatized care is provided at lower levels of care, but the costs of higher levels of care are primarily borne by the public sector. In Mexico, there is relatively little private hospital care, even among those who use private sources for their primary care; only 24% of beds are in private hospitals. From our sample, 74% of individuals who utilized a hospital report a public option as their main health provider. As a result, improved primary care through private diabetes clinics may impart a significant fiscal externality on the public sector.

Indeed, the hospital costs of diabetes to the Mexican public insurance system are enormous. Total costs amount to 2.25% of GDP, with 1.1% being direct medical costs and 87% of this total is due to complications. These costs primary arise from a series of complications associated with diabetes such as diabetic retinopathy, diabetic foot and diabetes kidney disease.[20] Bringing blood sugar levels under control significantly reduces the risk of such complications – leading to reduced hospital costs (Barraza-Lloréns et al., 2015). The fact that Mexico has twice the OECD average diabetic hospitalization rate highlights that there is much to gain from improving early-stage care.

We investigate this in two ways. First, we directly estimate the effect of CdA on self-reported diabetic complications. This is challenging since we only have a short follow up time for our CdA enrollees, so that we will meaningfully understate the long-term impacts on complications. Second, we use a simulation model based on our estimated reduction in blood sugar, combined with the best estimates of the marginal impacts of reduced blood sugar on future complications.

[20] *Diabetic retinopathy*: is a diabetes complication that affects eyes. It's caused by damage to the blood vessels of the light-sensitive tissue at the back of the eye (retina). At first, diabetic retinopathy may cause no symptoms or only mild vision problems. Eventually, it can cause blindness. *Diabetic foot*: Diabetes can damage your nerves or blood vessels. Nerve damage from diabetes can cause you to lose feeling in your feet. You may not feel a cut, a blister or a sore. Foot injuries such as these can cause ulcers and infections. Serious cases may even lead to amputation. *Kidney disease*: Diabetes can damage the blood vessels in your kidneys. When the blood vessels are damaged, they don't work as well. When your kidneys are damaged, they can't filter blood like they should, which can cause wastes to build up in your body.

Column 1 of Table 6 shows our estimated treatment effects for complications, which are defined as the sum of the early symptoms a person might experience: worsening sight, hands and feet tingling. Despite the short follow up, we find a significant reduction in complications, with enrollment associated with a decline of complications of -0.25 off a base of 1.4.

To put this result in context, we compute the estimated decline in complications that we would expect over time based on our blood sugar reductions. According to the widely cited UKPDS study, a one point reduction in HbA1c could lower the likelihood of complications by 35%- significantly higher than the 18% we find experimentally, which is unsurprising given our short term follow up.

We next turn to estimating the fiscal externality in Figure 5. Because the effects and therefore the savings differ by baseline level of HbA1c, we first classify our sample into 10 baseline HbA1c value bins $j = 5$ to 14, and allocate a person to the bin according to her starting HbA1c. We then calculate a CLATE in terms of HbA1c for each of these bins using Athey and Wager (2019)'s methodology, and call it $\widehat{CLATE}_j$. Second, we calculate the average HbA1c observed at follow up for the control group in each bin $j$, $HbA1C_{C_j}^{followup}$. We sum $\widehat{CLATE}_j$ to $HbA1C_{C_j}^{followup}$ to estimate level of HbA1c for the treatment group in each bin: $HbA1C_{T_j}^{followup} = HbA1C_{C_j}^{followup} + \widehat{CLATE}_j$.

To go from HbA1c effects to health complications and hospitalizations, we use the complications incidence tables by level of HbA1c from the UKPDS 35 study (King et al., 2001).[21]. We define averted hospitalizations as the difference of hospitalizations between treatment and control group for each bin. Finally, to estimate the peso savings to the system from reduced hospitalizations we use the complication-specific hospital cost data from Barraza-Lloréns et al. (2015) for 2013 in Mexico updated for inflation.

Figure 5 shows the results. The bins are represented in the horizontal axis, and the thin blue curve displays the distribution of initial HbA1c. As can be seen, a large fraction of clients are not under control at baseline (Initial HbA1c>7), this fraction is displayed on the vertical axis of the right. The vertical bars represent the savings in pesos (left Y-axis) estimated per patient year for each bin. For instance, those in the 13-HbA1c bin save more than 8,000 pesos per year just in the cost of hospitalizations from complications, while those in the 12-HbA1c bin save almost 7,000. Savings typically rise with HbA1c levels both because (a) the rate of complications is higher for those with higher blood sugar and (b) our CLATE estimates show larger effects of our intervention for those with higher blood sugar at baseline. For comparison, the horizontal red line plots the annual subscription price of CdA of 7,000 pesos. Our results show that at high levels of blood sugar, the program essentially pays for itself in reduced hospital spending as the savings of those between 11 and 14 initial HbA1c averages close to 7,000 pesos; overall, we estimate that reduced

[21] For each group, we interpolate linearly between the 2 HbA1c integer numbers that the incidence tables in the paper report. Since dynamics of the effect of higher blood sugar for each complication vary, we estimate incidence for each main complication separately. We focus on neuropathy, ulcers, amputations, ophthalmic complications, diabetic coma, nephropathy, stroke, and heart attack.

hospital expenditures amount to 55% of the cost of CdA. The bottom of the Figure displays for each bin the average savings as a fraction of the CdA annual subscription. These are 98% for the 12-bin, 116% for the 13-bin, and 97% for the 14-bin.

Of course, these calculations are not precise. On the one hand, we assume that the marginal reductions in blood sugar from our intervention have the same impact as the average reduction in blood sugar used by UKPDS 35. On the other hand, we only consider the direct fiscal externality from hospitalizations averted and do not consider the savings from care substitution (with the CdA visits displacing public visits). To address the latter, we directly measure care substitution in the last two columns of Table 6. The second column measures visits to a public insurance provider. This falls by 0.11 visits, or about one fifth of the control group mean at baseline, but the estimate is insignificant. One problem with this measure, however, is that COVID-19 may have had the impact of reducing all medical care use, for both treatments and controls, mitigating any estimated displacement effect. The next column uses a measure which may be more reliable, whether the respondent considers the public health system to be their main health care provider. While this variable has essentially the same mean, our estimated effect is nearly twice as large, and amounts to 40% of baseline mean.

At baseline, our control group has 5 visits on average to their public provider in the year before joining CdA. The direct cost to IMSS of each visit is 800 pesos. In addition, IMSS data shows that the total cost of maintaining care for an under control diabetic is $9000 pesos per year. So a 20% reduction in visits would save $800-$1800 pesos per year; if we use the larger estimate from the third column of Table 6, we would have savings of $1600-$3600 pesos per year just from lowering visits.

Recall that for our sample the cost of CdA treatment for one year was $7000 pesos. Using our estimate of offsetting hospital spending on complications, as well as our lowest estimate of offsetting primary care expenditures, roughly 65% of the costs of CdA are offset by reduced public sector costs; at the upper bound, the offset is 105%. In either case, the net cost of this incremental care through CdA is much less than the gross costs.

So far this calculus does not consider the value of improved care. In Table 3, we showed that patients enrolled in CdA are 22% more likely to be in control than those who do not enroll. The medical literature estimates that every year that a diabetic patient spends out of control reduces life expectancy by 100 days.[22] Typical estimates of the value of a life-year in international contexts is in the $50,000-$100,000 USD range, which would suggest that this intervention is worth $3000-$6000 USD per person, many multiples of gross or net program costs, which are around $350 USD a year.[23]

[22] Heald et al. (2020).

[23] Lee et al. (2009) argues that the value per year of quality life is $129,00 but $50,000-$100,00 has been de facto international standard. For the calculation we multiply the effect on likelihood of control, times 100/365 days, times

While estimating the value of life in this particular population is beyond our scope, it is worth noting that even this estimate excludes the valuation of the large reduction in morbidity – which does not only save the public sector costs but improves quality of life.

# 6 Why Does the Private Sector Improve Outcomes?

The striking finding that enrolling in CdA dramatically improves health outcomes for diabetics that already have access to free public care raises the key question of *why* the private sector is doing a better job than the public sector addressing the medical needs of diabetics. One potential explanation is that our finding is driven by heterogeneity in the type of public insurance. While there is technically universal public coverage in Mexico, the care delivered by IMSS (the formal sector social insurance program) is typically perceived to be much higher quality than that delivered by the residual public welfare program, Seguro Popular. As a result, if our findings are driven by those individuals in our sample who are not formally employed and have to rely on Seguro Popular, this may reflect the lower quality of care in that public program. However, we find no evidence that the impact of CdA is driven by informal workers – if anything, the opposite appears to be true.[24]

If even the higher quality public programs were not performing as well as private care, the difference must be explained by either quality or quantity differences between the two platforms. On the one hand, it could be the case that CdA is providing a higher quality of service per interaction. On the other hand, perhaps CdA is doing more to attract diabetics to interact with care, improving outcomes through increased quantity of medical interactions.

To separate these hypotheses, we extend our analysis by incorporating data from the IMSS program, the largest formal sector health care system in Mexico, between 2010 and 2015. We worked with IMSS to collect administrative data for every primary care visit and the place of residence for all enrollees; this is a novel data set which has not been previously exploited for economics research. The data combines several large administrative datasets. The first is annual testing data; all IMSS enrollees are supposed to have a check-up that includes blood sugar once per year as part of the PrevenIMSS program and we have administrative data for all checkups recorded in that period.[25] Second, we have administrative data from every primary care visit, which includes ICD-10 diagnosis codes. Third, we have exact location data for families' homes. Fourth, we have an infrastructure dataset that includes geocoded locations for the primary care clinics providing care through IMSS. In order to estimate the marginal return to care, we create a sample of all patients

$50,000-$100,000.

[24] A regression of HbA1c on treatment interacted with having informal insurance shows a positive but insignificant interaction – suggesting that our results are not driven by larger impacts among those informally insured (see last column Table 5).

[25] IMSS (2013)

who have been treated for diabetes at IMSS and who had at least one PrevenIMSS checkup, leaving us with a sizeable sample of 440,000 diabetics. Moreover, 160 thousand have a second PrevenIMSS appointment one year later, which allows us to track their blood sugar dynamics.

We find a variety of evidence that there is a significant role for a higher quantity of care with the private sector option. Table 4 showed directly that the treatment group has more visits to doctors. Table 2 finds a much higher level of trust in CdA than their alternative care provider, suggesting a higher willingness to engage with the provider. Moreover, we find that randomized enrollment in CdA strengthens trust in the program.[26]

Use of private sector clinics may be more frequent due to greater convenience of this source of care. Our follow up survey asked treatments and controls about their waiting time for care. Column 3 of Table A.C.-2 shows our IV estimates for waiting time, and the impacts are striking: a reduction of 30 minutes in waiting time, or more than half of the baseline mean. Note also that this is an estimate per visit; the total time difference will be even larger in the public sector since while CdA offers a single visit one-stop-shop model of care, the public sector usually requires the patient to visit several times: to get tested, to speak with a nutritionist, to get medicines, etc.

To further explore whether making care more accessible is a key mechanism driving our results, we focus on the subsample of our treatment and control groups that is enrolled in IMSS, and we consider the heterogeneity of our treatment effect by the level of IMSS clinic "saturation", a direct measure of how hard it is to access care at IMSS. If indeed more access to care is driving our findings, we would expect our effect to be bigger among the IMSS users who have to utilize a more saturated clinic.[27]

Using data on visits from 2015 for 31 clinics in the region, we match each patient to the closest IMSS clinic to their home and exploit administrative data to define clinic saturation. We begin by dividing the number of patients served at the clinic by the number of medical offices in that clinic. We take as a benchmark the 15 minutes per visit recommended by IMSS guidelines. We then label a clinic as "saturated" if it receives on average more than 85% of the maximum 4 visits per hour they can handle at maximum capacity per year. We then rerun our outcomes regressions, interacting our treatment dummy with a dummy for the closest clinic being saturated; we redo the exercise for

[26] In Table A.C.-2 we explore whether the CdA trust advantage grows as a result of treatment. The first column regresses the difference in the self-reported trust between CdA versus the alternative at follow up against our instrumented "Using CdA" indicator, controlling for the baseline difference. The second column carries out the same exercise for a different variable, which measures whether enrollees trust *the diagnoses* that comes from CdA as opposed to their current health provider. Both variables have a mean difference above 2 points on a ten point scale, and IV regressions of this gap in beliefs on being a CdA user shows that the gap increases significantly with use of CdA. These results highlight that patients are more engaged with the service once they use it – which may lead them to get care more often.

[27] This comparison assumes that the more saturated clinics don't themselves deliver very different quality of care that less saturated clinics. If the care delivered at more saturated clinics is lower than at less saturated clinics, then part of the response we see here may be through differences in quality and not quantity. Of course, if saturation arises because there is more use of the highest quality clinics, the bias would go in the opposite direction.

non-IMSS users as a placebo, matching non-IMSS users to the nearest IMSS clinic.

In the first column of Table 7, we show that our treatment effect is larger for IMSS users when the clinics are more saturated, consistent with the notion that it is those who face the largest barriers to IMSS care who benefit most from CdA. Moreover, in the second column we show that such a relationship does not exist for the patients who are not enrolleed at IMSS; in fact the point estimate goes in the opposite direction. We see this evidence as supporting the hypothesis that improvements due to CdA arise through more care

Finally, we attempt to quantitatively disentangle these two channels with a quasi-experimental estimate of the marginal return to public care. We use our combined CdA and IMSS data to estimate the marginal returns to additional IMSS care versus care from CdA. To assess the marginal returns to IMSS, we use variation in the distance of individuals from their IMSS clinic. Individuals in IMSS are assigned to a local clinic based on fixed geographic designations, and as a result, the distance from homes to an IMSS clinics varies substantially. Figure A.C.-1 shows the distribution of distances from individual homes to IMSS clinics.

We restrict our analysis sample further to patients with two measures of blood sugar from PrevenIMSS, one year apart. We can use these measurements to assess whether more care during the intervening year induces improved outcomes – instrumenting the amount of care received with distance from an IMSS clinic.

In particular, we will estimate models of the following form:

$$Y_{i,j,t} = \beta_0 + \beta_1 N_{t,i} + \beta_2 Y_{t-1,i} + \Gamma_j + \psi_t + \chi_i + \varepsilon_{t,i,j} \tag{3}$$

Where the dependent variable $Y_{i,j,t}$ is the level of blood sugar for individual i who got his checkups at clinic j at time t, $N_{t,i}$ is the number of visits to IMSS clinics in the 12 months after the first blood sugar measurements, $\Gamma_j$ captures clinic fixed effects, $\psi_t$ captures month fixed effects and $\chi_i$ are demographic control variables (gender, age and age-squared). We instrument the number of visits with the distance from residence to the assigned IMSS clinic.

This quasi-experimental approach faces two key identification concerns. The first is that distance is correlated with underlying health. We address this by controlling for baseline blood sugar at time t-1, so that we are assessing the impact of visits on the improvement in blood sugar. Of course, this does not solve the underlying identification problem if those who live near IMSS clinics are on differential underlying health trajectories than are those who live far away. But the inclusion of clinic fixed effects control for any neighborhood factors that might drive such trends.

The second concern is that the measurement itself may be correlated with distance – e.g. those who live farther away may be differentially likely to get their blood sugar measured. This is a particular concern given that only 160,000 out of 440,000 patients have a second yearly check-up. We can address this directly by assessing whether the odds of blood sugar measurement is itself

correlated with distance.

The results of our analysis are shown in Table 8. The first column shows the first stage estimate of the impact of distance on the number of IMSS visits. The coefficient is highly significant, indicating that each 30 kilometers of distance results in .1 fewer visits. The second column tests for selection in having a blood measurement. In fact, we see no evidence of a correlation between likelihood of second check-up and distance. While some people may get less diabetes care because it is too far, whenever they show up at the clinic, for any reason, they are asked to get a PREVENIMSS screening if they have not done so in the last year. Since PREVENIMSS is an independent module of primary care, there is no differential attrition.

We then turn to causal estimates of the impact of visits on blood sugar. Since PREVENIMSS captures capillary blood sugar measurements rather than HbA1c, we utilize that metric instead for our analysis. To compare to our earlier findings, our experimental results from table 3 are equivalent to a reduction from 226 to 197 in capillary blood sugar.[28]

We begin by estimating equation 3 for our CdA intervention. That is, we regress capillary blood sugar levels —which we also measured in our baseline and follow up surveys— on the number of CdA visits, controlling for baseline capillary blood sugar. We instrument number of visits with our treatment indicator, so that we are essentially measuring the total treatment effect as a function of number of visits. In this specification we are assuming linear impacts of each additional CdA visit. The third column of Table 7 shows that each CdA visit reduces capillary blood sugar by 8 points.

The fourth column estimates equation 3, instrumenting by distance to an IMSS clinic, to estimate the return on each marginal visit at IMSS. We find that each additional visit provides a benefit of 5 points. While significant, this is less than two-thirds as large as the estimate for CdA (although the differences between CdA and IMSS are not statistically significant). This suggests that part of the reason for a larger effect for CdA is more effectiveness per visit (although the difference is small); moreover, as noted earlier, we potentially understate the impact of CdA because this treatment-control comparison excludes any impacts of the initial evaluation.

But this result does not account for potential selection on treatment effectiveness. In fact, those who sign up for CdA have considerably higher blood sugar than the typical person in IMSS, while the average baseline blood sugar in Cda is 225, at IMSS it is 135. And we showed earlier that the effect of CdA is larger for those with higher blood sugar – the same may be true for IMSS. To assess this, we re-estimate the regression for IMSS from column 4, but reweighting the sample by baseline blood sugar to make them more comparable. The final column of Table 8 shows that doing so dramatically increases the estimated IMSS treatment effect, which more than doubles. Indeed, this estimate is higher than the comparable CdA estimate, although not significantly so. Thus, these

[28] We utilize a conversion from Nathan et al. (2008) that estimates that each point reduced in HbA1c is equivalent to a 28.7 reduction in capillary blood sugar. This conversion fares well when applied to our experimental sample since we find a reduction of 29 units in capillary blood sugar and a 1.1 point reduction in HbA1c.

results do not indicate that CdA's impacts arise through a better "technology," at least in terms of the returns per visit.

Thus, a suite of evidence supports the notion that the higher quantity of care received at CdA, and not higher quality per unit, is driving the results. We see more use of care at CdA, driven at least partly by shorter wait times. We see that the effects of CdA are largest where access to IMSS is most restricted. And we find that the estimated return to a visit to IMSS is comparable to the estimated return to a CdA visit

# 7 Conclusion

Ongoing debates over private versus public delivery of health care are central throughout the world. Yet these arguments often take place on grounds of political philosophy and not empirical evidence. This is unsurprising, as it is challenging to design studies that compare private versus public options for enrollees in an empirically compelling framework.

This paper introduced such a framework, relying on a novel deniers randomization framework to run a trial of the private provision of diabetes care to a publicly insured population in Mexico. Our findings are striking: supplementary private care causes a highly significant and large reduction in the blood sugar levels of diabetics, increasing blood sugar control by more than two-thirds. We estimate that this occurred through improved use of medication, more frequent medical treatment, and more diet and exercise. This sizeable reduction in blood sugar was associated even within the first year with reduced diabetes complications. These large health effects suggest that this supplemental private service was highly cost effective. We estimate that two-thirds or more of the cost of the private program are offset by reduced public primary and (especially) hospital care, and that the estimated health benefits are many multiples of either gross or net costs.

Interestingly, our results also suggest that the strong performance of this private sector alternative was not because of dramatic improvement in care modality, but rather through stronger attachment of patients to the private alternative due to shorter wait times and other advantages over public clinics. This suggests that much of the gains from privatization in this context could actually be captured by the public sector itself by improving access to care for its enrollees. An open and important question for future work is whether the public sector could best improve care by contracting out to the private sector or by expanding its own resources.

Our approach does have some limitations. We argue that our deniers randomization approach is an innovative means of cost-effectively evaluating an alternative like CdA, and show that based on our estimated heterogeneity in treatment effects, the estimates are likely to apply to always takers as well; but our heterogeneity estimates are noisy and more work is needed to confirm this. We also examine a particular program, CdA, and outcomes may not extend to other alternatives that

are designed differently. Moreover, our conclusion that the gains from privatization are simply in increased care, and not in substantial changes to the technology of care, may reflect the particular case of diabetes, where there is a standard and cost-effective course of treatment. For other diseases with less standardized and/or more expensive treatment modalities, private delivery may or may not offer gains in the quality as well as the quantity of care.

Despite these caveats, however, our study provides a framework for estimating the effects of privatization on medical care. We focus here on diabetes, one of the world's deadliest chronic conditions, in Mexico, one of the countries where this problem is largest. The lessons from diabetes in Mexico could be usefully tested and applied to other diseases and other nations around the world.

# Tables

**Table 1:** Attrition

| | Answered follow up | | |
|---|---|---|---|
| | (1) | (2) | (3) |
| Treatment | 0.03 | 0.03 | 0.01 |
| | (0.02) | (0.02) | (0.02) |
| Branch FE | No | Yes | Yes |
| Enrollment Month Fe | No | Yes | Yes |
| Basic Controls | No | No | Yes |
| Observations | 2,410 | 2,410 | 2,042 |
| R-squared | 0.001 | 0.01 | 0.11 |
| Mean dep. var. | 0.43 | 0.43 | 0.43 |

**Notes:** This table presents the results from running a regression on a dummy of answering follow-up on the treatment group dummy. The first column reports the regression without any controls, the second column controls for branch fixed effects and enrollment month fixed effects and the third column also includes Basic Controls: age, gender, HbA1c, BMI, Schooling and Income. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table 2:** Balance Table for Those Measured at Follow Up

| Variable | Control N | Control Mean/SE | Treatment N | Treatment Mean/SE | Difference C-T |
|---|---|---|---|---|---|
| ***Panel A: Demographics*** | | | | | |
| Age | 509 | 52.39 | 558 | 53.06 | -0.67 |
| | | (0.50) | | (0.52) | (0.73) |
| Male | 509 | 0.32 | 558 | 0.35 | -0.03 |
| | | (0.02) | | (0.02) | (0.03) |
| % High income | 509 | 0.56 | 558 | 0.56 | -0.01 |
| | | (0.02) | | (0.02) | (0.03) |
| % High School or more | 457 | 0.37 | 485 | 0.39 | -0.02 |
| | | (0.02) | | (0.02) | (0.03) |
| ***Panel B: Health and Health services*** | | | | | |
| HbA1c | 508 | 9.35 | 556 | 9.51 | -0.15 |
| | | (0.11) | | (0.11) | (0.16) |
| BMI | 509 | 31.07 | 557 | 30.83 | 0.24 |
| | | (0.27) | | (0.25) | (0.37) |
| Has IMSS, ISSSTE or Seguro Popular | 509 | 0.73 | 558 | 0.74 | -0.01 |
| | | (0.02) | | (0.02) | (0.03) |
| Percentage that use IMSS, ISSSTE or Seguro Popular | 509 | 0.74 | 558 | 0.74 | 0.00 |
| | | (0.02) | | (0.02) | (0.03) |
| HbA1c out of control (HbA1c>7) | 508 | 0.76 | 556 | 0.78 | -0.02 |
| | | (0.02) | | (0.02) | (0.03) |
| ***Panel C: Beliefs*** | | | | | |
| Trust to improve following CdA recomendations (0-10) | 432 | 9.43 | 467 | 9.54 | -0.12 |
| | | (0.07) | | (0.06) | (0.09) |
| Trust to improve following current health provider recomendations (0-10) | 414 | 7.89 | 460 | 7.93 | -0.03 |
| | | (0.13) | | (0.13) | (0.18) |
| Trust in CdA diagnosis (0-10) | 444 | 9.33 | 469 | 9.37 | -0.04 |
| | | (0.07) | | (0.07) | (0.09) |
| Trust in current health provider diagnosis (0-10) | 411 | 7.93 | 450 | 7.88 | 0.05 |
| | | (0.13) | | (0.13) | (0.18) |

**Notes:** This table presents the balance among the patients who did answer our follow-up. *(1) % High income*: Is an indicator variable equal to 1 if the person lived in zip codes that the AMAI classifies as above middle income. This variable is an administrative variable and is based on the place of residency of each potential client at CdA. *(2) Has IMSS, ISSSTE or Seguro Popular*: Is an indicator variable equal to 1 if the person declared to be affiliated to a public service to the medical staff at CdA. *(3) Percentage that use IMSS, ISSSTE or Seguro Popular*: Is an indicator variable equal to 1 if the person declared to go for medical attention to the public services in our baseline survey. *(4) Trust to improve following CdA recommendations (0-10)*: Is variable that measures the trust in improving with CdA. *(5) Trust to improve following current health provider recommendations (0-10)*: Is a variable that measures the trust in improving with their current health provider. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table 3:** Effect on HbA1c

| | HbA1c OLS | HbA1C | Under Control- OLS I(HbA1C<7) | Under Control I(HbA1C<7) |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| Use Cda | -0.98*** | -1.12*** | 0.15*** | 0.22** |
| | (0.13) | (0.41) | (0.03) | (0.09) |
| HbA1c Bl | 0.50*** | 0.50*** | | |
| | (0.03) | (0.03) | | |
| I(HbA1c<7) Bl | | | 0.51*** | 0.52*** |
| | | | (0.04) | (0.04) |
| BMI | -0.02 | -0.02 | 0.00 | 0.00 |
| | (0.01) | (0.01) | (0.00) | (0.00) |
| Age | -0.00 | -0.00 | 0.00 | -0.00 |
| | (0.01) | (0.01) | (0.00) | (0.00) |
| Gender | -0.17 | -0.17 | 0.04 | 0.03 |
| | (0.14) | (0.14) | (0.03) | (0.03) |
| Elementary School | 0.28 | 0.30 | 0.06 | 0.05 |
| | (0.42) | (0.43) | (0.09) | (0.09) |
| Secondary School | 0.43 | 0.44 | 0.04 | 0.03 |
| | (0.42) | (0.43) | (0.09) | (0.09) |
| High School | 0.05 | 0.06 | 0.05 | 0.04 |
| | (0.44) | (0.45) | (0.09) | (0.09) |
| Tecnica o Normal | 0.22 | 0.23 | 0.10 | 0.09 |
| | (0.44) | (0.44) | (0.09) | (0.09) |
| University | -0.08 | -0.06 | 0.14 | 0.12 |
| | (0.47) | (0.48) | (0.10) | (0.10) |
| Income C+ | 0.26 | 0.25 | -0.04 | -0.04 |
| | (0.31) | (0.31) | (0.12) | (0.12) |
| Income C | 0.24 | 0.23 | -0.06 | -0.05 |
| | (0.29) | (0.29) | (0.11) | (0.12) |
| Income D+ | 0.57* | 0.56* | -0.15 | -0.14 |
| | (0.31) | (0.31) | (0.12) | (0.12) |
| Income D or Lower | 0.41 | 0.40 | -0.09 | -0.08 |
| | (0.32) | (0.32) | (0.12) | (0.12) |
| Observations | 939 | 939 | 939 | 939 |
| R-squared | 0.36 | 0.35 | 0.26 | 0.22 |
| F | | 93.26 | | 93.33 |
| First coeff | | 0.301 | | 0.300 |
| Mean dep. var | 8.538 | 8.538 | 0.322 | 0.322 |

**Notes:** This table shows the results of estimating equation 1. The first column captures the OLS regression on the effect of using CdA on HbA1c, the second column captures the local average treatment effect (LATE) on HbA1c from our randomization, the third column captures the OLS regression on a dummy that captures if an individual has controlled his diabetes (HbA1c<7) and the fourth column captures the LATE from our experiment in terms of controlled diabetes.. In all regressions we are controlling for branch and month fixed effects as well as basic demographics control (age, sex, BMI, baseline HbA1c or controlled diabetes at baseline, schooling and income). Our schooling controls are self-reported from our baseline survey and the income controls are from administrative data from CdA; this administrative data use the locations of houses and the Asociación Mexicana de Agencias de Investigación y Opinión Pública A.C. classification: A/B upper class, C+ upper middle class, C middle class, D+ medium-low class and D lower class. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table 4:** Mechanism

| | Visits to Doctor | # Special Check ups | # Specialists | Stop med if feels good | Takes Pills | Takes Insulin | I(Exercise) | I(Diet) | Diet+Exercise | # Cigaretts | # Sodas |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) | (9) | (10) | (11) |
| Use Cda (instrumented) | 2.59** | 1.32*** | 0.64** | -0.22** | 0.25*** | -0.15** | 0.14 | 0.14 | 0.26* | -0.72 | -0.06 |
| | (1.06) | (0.38) | (0.27) | (0.09) | (0.09) | (0.06) | (0.10) | (0.10) | (0.15) | (0.80) | (0.14) |
| Branch FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Months since enrollment FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls Basic | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Observations | 878 | 817 | 940 | 815 | 913 | 913 | 940 | 940 | 940 | 774 | 857 |
| R-squared | 0.04 | 0.03 | 0.10 | 0.05 | 0.09 | 0.14 | 0.05 | 0.06 | 0.07 | 0.45 | 0.15 |
| F | 98.82 | 83.46 | 96.13 | 91.35 | 99.83 | 99.84 | 96.10 | 97.31 | 97.12 | 78.83 | 104.7 |
| First coeff | 0.321 | 0.307 | 0.306 | 0.323 | 0.315 | 0.316 | 0.306 | 0.308 | 0.308 | 0.307 | 0.332 |
| Mean dep. var | 6.502 | 1.408 | 1.383 | 0.249 | 0.727 | 0.134 | 0.289 | 0.436 | 0.725 | 0.911 | 0.841 |

**Notes:** This table presents the results for regressions of the form of equation 1 on different self-reported behaviors. For all of these specifications we are reporting the local average treatment effect (LATE) that we causally estimated from our experiment using instrumental variables. In all regressions we are controlling for branch and month fixed effects as well as basic demographics control (age, sex, BMI, baseline HbA1c, schooling and income). *(1) Visits to Doctor*: Is a variable that captures how many visits to the doctor related with diabetes the person had the last year. *(2) # Special Check Ups*: Is a variable that sum the check ups that the person received during the last year related with diabetes, i.e eyes, kidney, foot and blood check ups. *(3) # Specialists*: Is a variable that captures the number of specialist, related with diabetes, who attended the person in a regular visit to their health provider. *(4) Waiting Time*: Is a variable that captures the waiting time whenever the persons attend to his health provider to have an appointment related with diabetes, we winzorized at 5% the upper tail to omit implausible waiting times. *(5) Stop Medication if feels good*: Is an indicator variable equal to 1 if the person declared that he suspended his medication whenever he start to felling good. *(6) Takes Pills and Takes Insulin*: Both are indicator variables equal to 1 if the person declared that he took pills and/or insulin respectively as a part of his treatment, also we control for their baseline measure. *(7) I(Diet) and I(Exercise)*: Both are indicator variables equal to 1 if the person declared that he did exercise and was involved in a diet as an effort to take care of his health. (8) Diet+Exercise: is a variable that sums the variables I(Diet) and I(Exercise). *(9) # Cigarrets and #Sodas*: are variables that captures the number of cigarettes and sodas consumed in a regular day by the person, also we control for their baseline measure. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table 5:** Heterogeneity Effect

***Dep Var HbA1c follow up***

| | Age | Male | Income | Schooling | HbA1c | BMI | Informal |
|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) |
| Treatment | -0.51** | -0.40** | -0.29* | -0.36* | -0.16 | -0.21 | -0.40** |
| | (0.21) | (0.17) | (0.15) | (0.19) | (0.12) | (0.20) | (0.16) |
| Regressor | -0.66*** | -0.00 | -0.50* | -0.57** | 2.26*** | -0.34 | 0.08 |
| | (0.20) | (0.22) | (0.29) | (0.22) | (0.19) | (0.21) | (0.23) |
| Treatment X regressor | 0.45 | 0.27 | -0.10 | 0.20 | -0.42* | -0.20 | 0.31 |
| | (0.27) | (0.30) | (0.36) | (0.30) | (0.25) | (0.28) | (0.32) |
| Branch FE | Yes | Yes | Yes | Yes | Yes | Yes | Yes |
| Controls | No | No | No | No | No | No | No |
| Observations | 1,067 | 1,067 | 1,067 | 942 | 1,064 | 1,066 | 1,067 |
| R-squared | 0.03 | 0.02 | 0.02 | 0.03 | 0.21 | 0.03 | 0.02 |

**Notes:** This table presents the heterogeneity results from the effect of getting CdA treatment on health, estimated from including in equation 2 an interaction with the regressor listed at the top of each column. For each of these, we report the treatment coefficient, the baseline effect on the explored dimension and the interaction between the 2. In all regressions we are controlling for branch fixed effects. For the variables age, HbA1c and BMI we split by the median. For the variables income and schooling we split for those considered middle or upper class and those with high school or higher education respectively. Informal: is an indicator variable equal to 1 if the person declared to be affiliated to the Seguro Popular or to not have social insurance to the medical staff at CdA. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table 6:** Averted complications and Health Provider Substitution

| | Total Diabetes Complications | Use of Public Service | Public Service as Principal Provider |
|---|---|---|---|
| | (1) | (2) | (3) |
| Use Cda | -0.25* | -0.11 | -0.25** |
| | (0.13) | (0.10) | (0.10) |
| Branch FE | Yes | Yes | Yes |
| Enrollment Month Fe | Yes | Yes | Yes |
| Basic Controls | Yes | Yes | Yes |
| Observations | 939 | 940 | 909 |
| R-squared | 0.05 | 0.09 | 0.14 |
| F | 93.31 | 96.06 | 95.34 |
| First coeff | 0.301 | 0.306 | 0.310 |
| Mean dep. var | 1.383 | 0.611 | 0.640 |

**Notes:** This table presents the results from the effect of getting CdA treatment on complications, the effect of using CdA on utilizing the public sector and the effect of CdA on saying that the public sector is your main provider. We are controlling for branch and month fixed effects as well as basic demographics control (age, sex, BMI, baseline HbA1c, schooling and income). *(1) Total diabetes complications*: Is a variable that sum the short run complications related to diabetes experienced by the person, i.e eyes, feet and hand tingling, we also control by the total complications at baseline. *(2) Use of Public Service*: Is an indicator variable equal to one if the person declared to attended to IMSS, Seguro Popular or ISSSTE in the previous year for any medical reason, we also control for the baseline value of this variable. *(3) Public Service as Main Provider*: Is an indicator variable equal to one if the person declared a public service (IMSS, Seguro Popular, ISSSTE) as their principal health provider, we control for their health affiliation at baseline. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table 7:** Effect on HbA1c by saturation for IMSS enrollees

| | HbA1c IMSS Members | HbA1c non-IMSS-members |
|---|---|---|
| | (1) | (2) |
| Treatment | 0.22 | -0.17 |
| | (0.37) | (0.54) |
| High saturated clinic | 0.88*** | -0.15 |
| | (0.31) | (0.47) |
| Treat x High sat. Clinic | -0.85** | 0.05 |
| | (0.41) | (0.60) |
| Branch FE | Yes | Yes |
| IMSS only | Yes | No |
| Observations | 515 | 292 |
| R-squared | 0.28 | 0.30 |

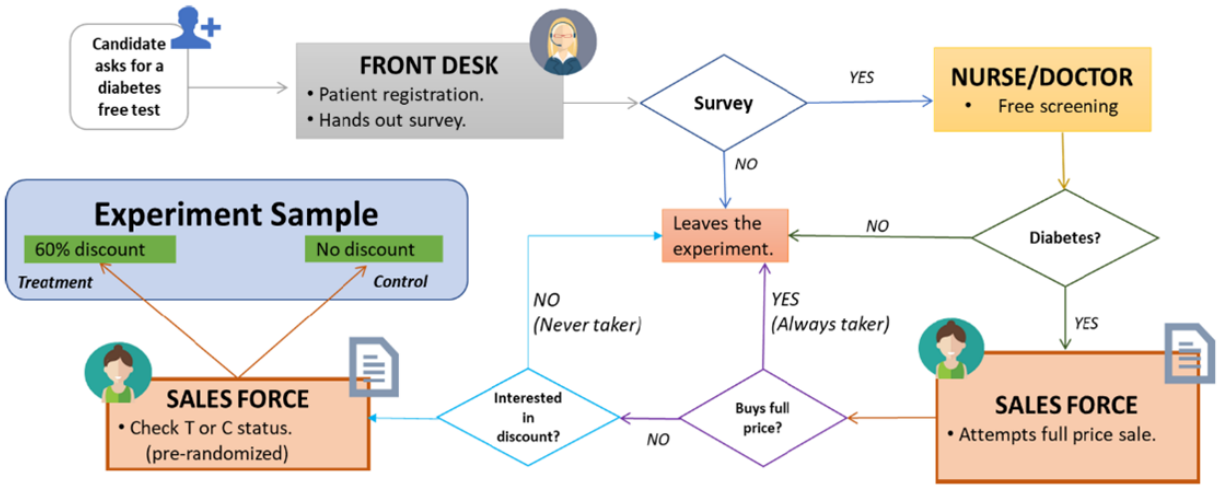
**Notes:** This table presents the heterogeneity results on the effect of CdA on HbA1c by saturation of IMSS clinics in Nuevo León. We define a clinic as saturated if they have at least an 85% flow of maximum capacity on average. That is, if they have at least 3.4 patients per hour per office open on average over a full year. The first column reports the heterogeneity estimates for IMSS population while the second column reports the same estimates for patients that do not report getting access to IMSS, which serves as a placebo. We control by HbA1c and branch fixed effects. We focus only in Nuevo León because Coahuila has very few IMSS clinics in Torreón and Saltillo, so we could not exploit the clinics heterogeneity there. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table 8:** Comparison IMSS vs Cda Effect

| | IMSS | | CdA | IMSS | |
|---|---|---|---|---|---|
| | Number of medical visits | I(12 months follow up) | Capillary Glucose | Capillary Glucose | Capillary Glucose |
| | (1) | (2) | (3) | (4) | (5) |
| Distance (km) | -0.0030*** | 0.0001 | | | |
| | (0.0005) | (0.0000) | | | |
| Number of medical visits | | | -8.39** | -5.07* | -12.69** |
| | | | (4.07) | (2.69) | (6.43) |
| Observations | 160,035 | 439,287 | 1,067 | 160,035 | 137,308 |
| R-squared | 0.11 | 0.05 | 0.17 | -0.06 | -0.28 |
| F | | | 60.31 | 35.29 | 12.96 |
| First coeff | | | 1.284 | -0.003 | -0.002 |
| Instrument | | | Discount | Distance | Distance (W) |

**Notes:** All columns except (3) estimated in IMSS data. First column shows regression of number of visits at IMSS on distance from the clinic. Second column shows regression of dummy for having a follow up blood sugar measurement on distance. Third column shows IV regression in CdA data where we regress capillary glucose on number of visits, instrumented by treatment indicator. Fourth column shows an IV regression of capillary glucose on number of IMSS visits, instrumented by distance. Fifth column repeats this exercise but reweighting the sample so that the baseline distribution for capillary blood sugar matches that of CdA. All specifications have branch and month fixed effects and basic controls (age, age squared, initial capilarity glucose and gender). All specifications have branch and month fixed effects and basic controls (age, age squared, initial capilarity glucose and gender). Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

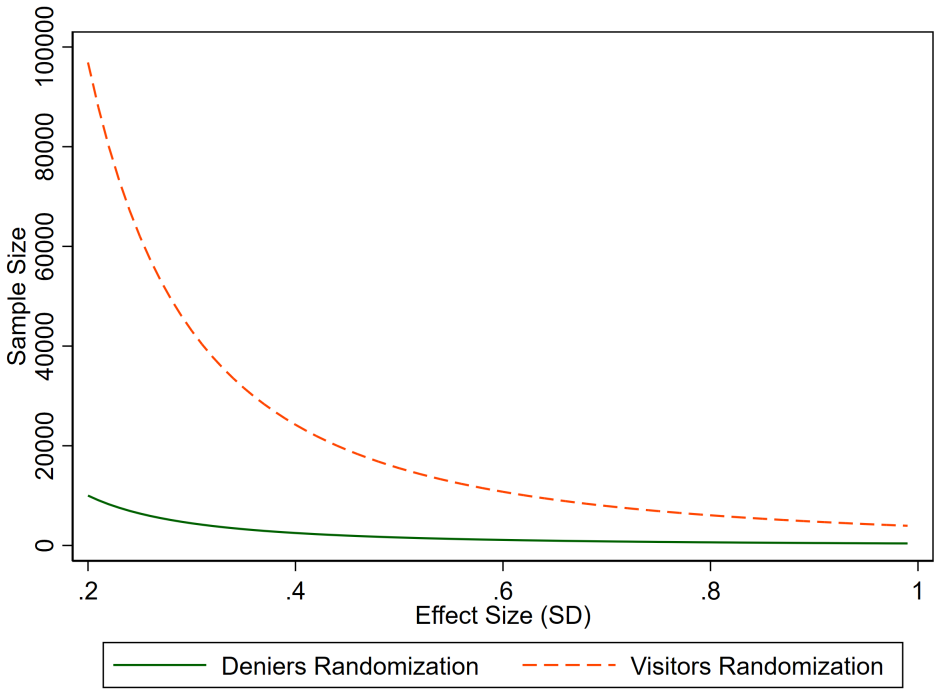
**Figure 1:** Recruitment Process




**Notes:** This figure represents the process through which a patient was included in our experiment. The patient would first go through the regular free-screening that CdA ususally offers and continue as a potential candidate unless she refused to fill out a survey. Then, if the patient was diagnosed as diabetic, the salesforce would try to sell a membership at full-price to that person. If the patient bought at full price, the person would leave the experiment, since we would know that person is an always taker. If the person was not interested in buying a full-price membership, then the salesforce would offer the chance to win a 60% discount from our study. This is the first point in which we would modify the regular flow of patients within CdA. If the person said they were not interested at that price either, then we would know that such a patient was a never taker. However, if the person said she was interested, then a button on the computer would reveal the treatment status to the salesforce and they would be able to offer the 60% discount if the patient was in the treatment group.

# Figures

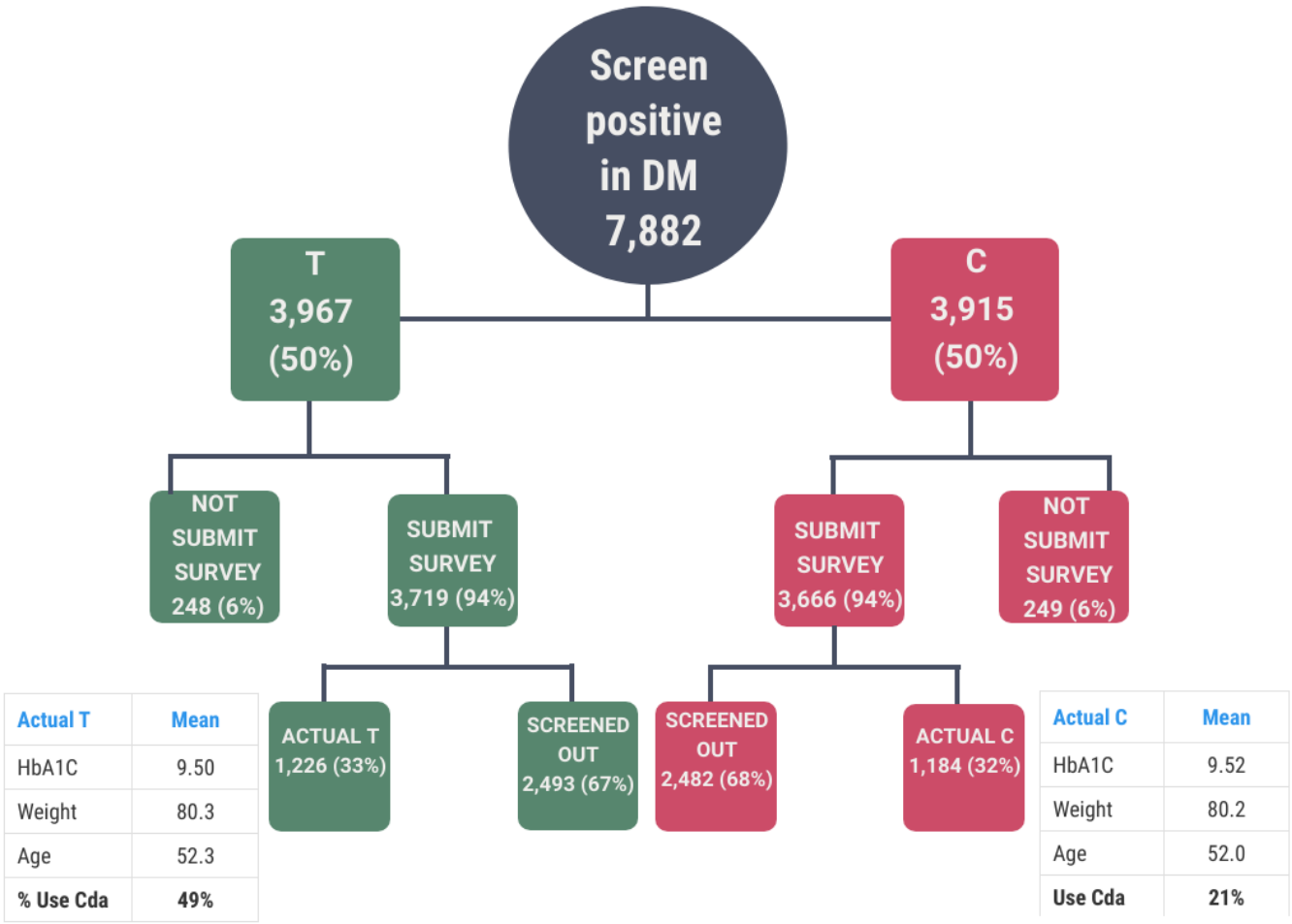
**Figure 2:** Power Calculations




**Notes:** This figure highlights the sample needed to capture an effect with 80% power and an $\alpha = 0.05$ under visitors randomization and through a deniers randomization. To fit our context, we assume that 38% of the individuals are always takers, 53% are never takers and only 9% are compliers. Moreover, we assume that a filter could exclude 82% of always takers and 68% of never takers, which is what ours does on the field. So out of every 100 individuals, the researchers would get 38 always takers, 53 never takers and 9 compliers. After applying the deniers randomization procedure, only 7 (38*0.18) always takers, 17 (53*.32) never takers and 9 compliers would enter the experiment. Thus, the researchers would have to focus on 33/100 visitors only and have a first stage power of 27% within their experimental sample. A big improvement over applying visitors randomization and working with a 9% first stage effect.

**Figure 3:** Recruitment Summary and Timeline

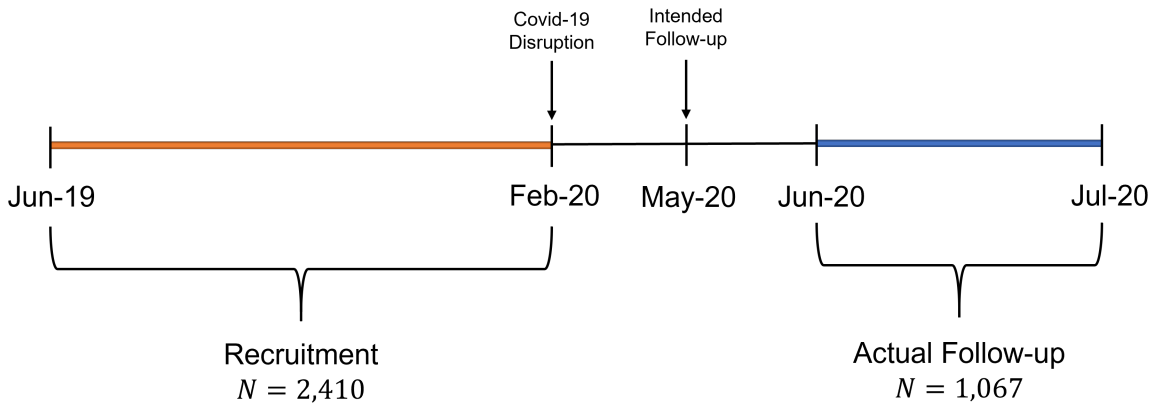
**(a)** Recruitment Summary




| Actual T | Mean |
|---|---|
| HbA1C | 9.50 |
| Weight | 80.3 |
| Age | 52.3 |
| **% Use Cda** | **49%** |

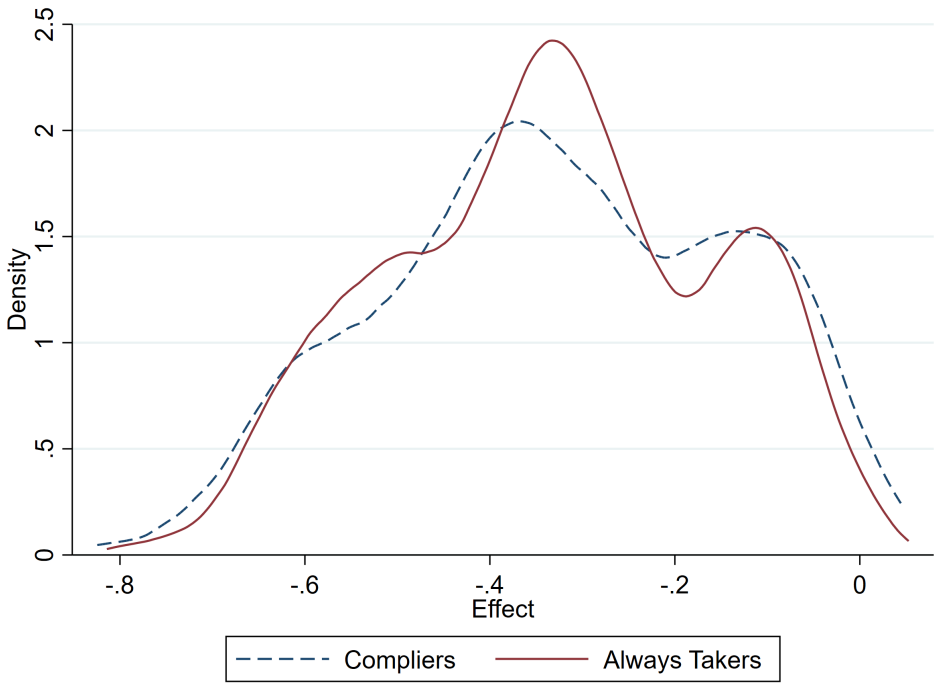
| Actual C | Mean |
|---|---|
| HbA1C | 9.52 |
| Weight | 80.2 |
| Age | 52.0 |
| **Use Cda** | **21%** |

**(b)** Experiment Timeline



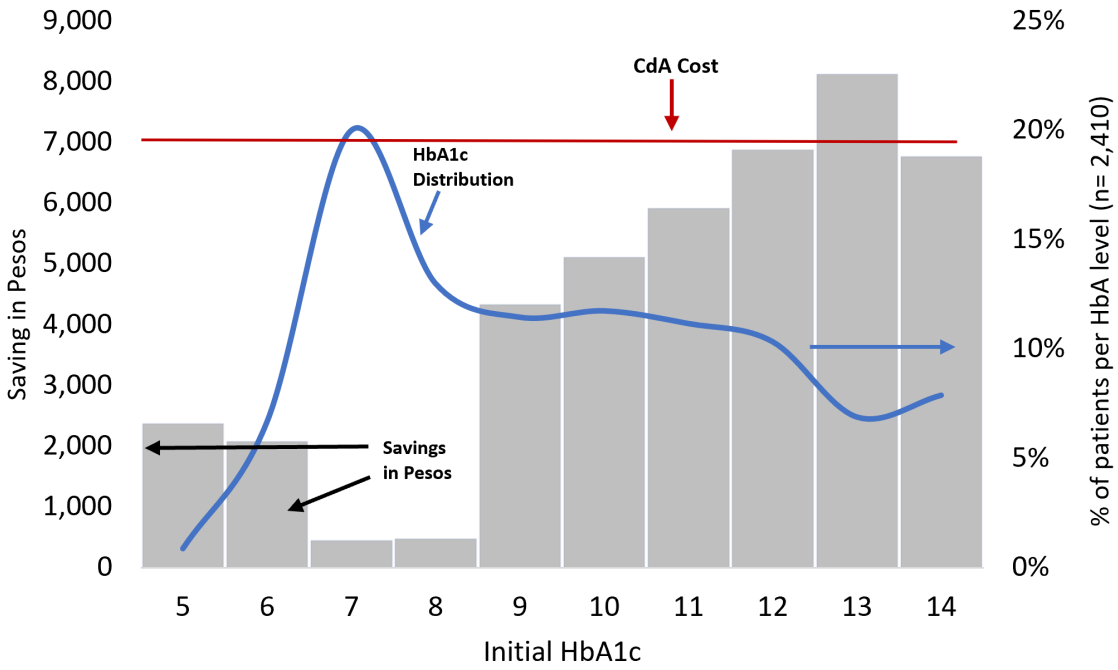

**Notes:** This figure summarizes our 8 month recruitment results as well as the timeline. We can see that there were nearly 8,000 diabetes patients that inquired about CdA, that our randomization was don evenly among treatment and control groups and that 94% of the patients were willing to answer our baseline survey. Moreover, we can see that through our deniers randomization design we were able to screen out 67% of the sample as always takers or never takers, which significantly increased our power for this experiment.

**Figure 4:** CATE: Compliers vs. Always Takers




**Notes:** This figure compares the distribution of the conditional average treatment effect we estimate based on Athey and Wager (2019) for our sample and how such an effect looks if we extrapolate to the rest of patients that showed interest in CdA. We can see both distributions look quite similar. Note that here we only utilize the reduced form estimates from being assigned to control or treatment and not the IV since we cannot know the endogenous choice that a regular patient would have made on whether or not to take up the offer by CdA. All the persons out of the experiment where assigned a zero en the treatment variable. We include as covariables: gender, HbA1c, BMI, age, social insurance and the clinic where the appointment took place. We omit the education variable because of missing values.

**Figure 5:** Savings from Hospitalizations Averted per Patient/Year (comparing T vs C)




**Notes:** This figure shows the savings we would observe from averted hospitalizations based on the reductions in HbA1c we causally estimate along with what the medical literature estimates and public spending data from the government. Specifically, we follow six steps. n six steps, recognizing the important heterogeneity in impacts by baseline blood sugar levels. First, we classify our sample into HbA1c baseline value bins, using the smaller nearest integer. Second, for those in the control group, in each bin we averaged their HbA1c observed at follow up ($HbA1C_{C_j}^{followup}$). Third, to go from HbA1c levels to health complications, we use the complications incidence tables by level of HbA1c from the widely cited UKPDS 35 study (King et al., 2001). This gives us the estimated complications for the control group by bin. Fourth, we apply an analogous method for the treatment group. We start from their baseline level and add in the conditional local average treatment effect (CLATE) of the respective bin, where the estimate is done as above but separately for each bin. That is for each bin j we calculate the number $HbA1C_{T_j}^{followup} = HbA1C_{T_j}^{baseline} + CLATE_j$, and map these to health complications. Fifth, we define averted hospitalizations as the difference expected hospitalizations between treatment and control group for each bin. Finally, to estimate the savings to the system from reduced hospitalizations, we multiply what each complication costs by the averted hospitalizations using the cost data from Barraza-Lloréns et al. (2015) for 2013 in Mexico updated for inflation.. Overall, we see that on average 55% of the costs would be recuperated by averted hospitalizations and that CdA is a potential savings mechanism for the government for most complicated patients. It is important to note that this figure is not considering substitution away from public sector services nor externalities from emptier clinics.

# References


AGUILAR, A., E. GUTIERREZ, AND E. SEIRA (2021): "The Effectiveness of Sin Food Taxes: Evidence from Mexico," *Forthcoming Journal of Health Economics*. [6]

AJALA, O., P. ENGLISH, AND J. PINKNEY (2013): "Systematic review and meta-analysis of different dietary approaches to the management of type 2 diabetes," *The American Journal of Clinical Nutrition*, 97. [4, 14]

ALMOMENTO.MX (2015): "¿Inicia privatización del IMSS?" Last accessed 9 March 2021, http://fmdiabetes.org/inicia-privatizacion-del-imss/. [1]

ANGRIST, J. D., G. W. IMBENS, AND D. B. RUBIN. (1996): "Identification of causal effects using instrumental variables." *Journal of the American statistical Association*, 434. [10]

ARNOLD, L. W. AND Z. WANG (2014): "The HbA1c and all-cause mortality relationship in patients with type 2 diabetes is J-shaped: a meta-analysis of observational studies," *Society for Biomedical Diabetes Research*, 11. [14]

ATHEY, S. AND S. WAGER (2019): "Estimating Treatment Effects with Causal Forests: An Application," *arXiv: Methodology*. [15, 17, 36]

BARRAZA-LLORÉNS, M., V. J. G. BARRÓN, C. H. VIVEROS, F. J. P. GUZMÁN, AND E. CRABLE (2015): *Carga Económica de la Diabetes Mellitus en México, 2013*, Fundación Mexicana para la Salud, A.C. [2, 16, 17, 37]

BASU, S., J. ANDREWS, S. KISHORE, PANJABI, AND D. STUCKLER (2012): "Comparative Performance of Private and Public Healthcare Systems in Low- and Middle-Income Countries: A Systematic Review," *PLOS MEDICINE*, 6. [4]

CHAN, D., D. CARD, AND L. TAYLOR (2020): "Is There a VA Advantage? Evidence from Dually Eligible Veterans," *Stanford University mimeo*. [4]

CHANDRA, A., J. GRUBER, AND R. MCKNIGHT (2010): "Patient cost-sharing and hospitalization offsets in the elderly," *American Economic Review*, 100. [4]

CLINIC, M. (2020): "Type 2 diabetes," Last accessed 9 March 2021, https://www.mayoclinic.org/diseases-conditions/type-2-diabetes/diagnosis-treatment/drc-20351199#:~:text=Metformin%20(Fortamet%2C%20Glumetza%2C%20others,body%20uses%20insulin%20more%20effectively. [5]

COLCHERO, M. A., J. A. RIVERA, B. M. POPKIN, AND S. WEN (2017): "Sustained consumer response: evidence from two-years after implementing the sugar sweetened beverage tax in Mexico," *Health Affairs*, 36. [6]

DAS, J., A. HOLLA, V. DAS, M. MOHANAN, D. TABAK, AND B. CHAN (2012): "In Urban And Rural India, A Standardized Patient Study Showed Low Levels Of Provider Training And Huge Quality Gaps," *Health Affairs*, 31. [4]

DAS, J., A. HOLLA, A. MOHPAL, AND K. MURALIDHARAN (2016): "Quality and Accountability in Health Care Delivery: Audit-Study Evidence from Primary Care in India," *American Economic Review*, 106. [4]

FRAKES, M., J. GRUBER, AND T. JUSTICZ (2021): "Public and Private Options in Practice: The Military Health System," *National Bureau of Economic Research*, No. w28256. [1, 4]

HEALD, A. H., M. STEDMAN, M. DAVIES, M. LIVINGSTON, R. ALSHAMES, M. LUNT, G. RAYMAN, AND R. GADSBY (2020): "Estimating life years lost to diabetes: outcomes from analysis of National Diabetes Audit and Office of National Statistics data," *Cardiovascular Endocrinology Metabolism*, 9. [14, 18]

IDF (2019): *IDF Diabetes Atlas 9th Edition*, International Diabetes Federation. [5]

IMSS (2013): "Modelo de prevención en diabetes mellitus (PREVENIMSS)," 5ta Conferencia Científica Anual sobre Síndrome Metabólico, https://www.slideshare.net/sindromemetabolico/modelo-de-prevencin-en-diabetes-mellitus-prevenimss. [19]

——— (2018): "Informe al Ejecutivo Federal y al Congreso de la Unión 2017-2018," . [5]

——— (2020): "Informe al Ejecutivo Federal y al Congreso de la Unión sobre la Situación Financiera y los Riestos del Instituto Mexicano del Seguro Social 2019-2020," . [1]

INEGI (2017): "Encuesta Nacional de Empleo y Seguridad Social: Principales Resultados," Last accessed 10 March 2021, https://www.inegi.org.mx/contenidos/programas/eness/2017/doc/presentacion_eness_2017.pdf. [6]

——— (2018): "Encuesta Nacional de Ingresos y Gastos de los Hogares," . [2]

——— (2019a): "Cuenta satélite del Sector Salud de México, 2019," Last accessed 10 March 2021, https://www.inegi.org.mx/temas/saludsat/. [6]

——— (2019b): "¿De qué mueren los mexicanos?" Last accessed 9 March 2021, http://cuentame.inegi.org.mx/poblacion/defunciones.aspx?tema=P. [2]

INSP (2012): "Encuesta Nacional de Salud y Nutrición 2016," . [6]

——— (2018): "Encuesta Nacional de Salud y Nutrición 2018," . [1, 2, 6]

IQEHC (2007): "What are disease management programs (DMP's)?" Last accessed 10 March 2021, https://www.ncbi.nlm.nih.gov/books/NBK279412/#:~:text=Disease%20management%20programs%20(DMPs)%20are,called%20%E2%80%9Cstructured%20treatment%20programs.%E2%80%9D. [7]

ISMAIL, K., K. WINKLEY, AND S. RABE-HESKETH (2004): "Systematic review and meta-analysis of randomised controlled trials of psychological interventions to improve glycaemic control in patients with type 2 diabetes," *The Lancet*, 363. [4, 14]

JAGER, W. (2003): *Breaking 'bad habits': a dynamical perspective on habit formation and change*, University of Groningen. [15]

KING, P., I. PEACOCK, AND R. DONNELLY (2001): "The UK prospective diabetes study (UKPDS): clinical and therapeutic implications for type 2 diabetes," *British Pharmacological Society Journals*. [4, 14, 17, 37]

KRAPEK, K., K. KING, S. S. WARREN, K. G. GEORGE, D. A. CAPUTO, K. MIHELICH, AND E. M. HOLST (2004): "Medication adherence and associated hemoglobin A1c in type 2 diabetes," *Annals of Pharmacotherapy*, 38. [14]

KRASS, I., P. SCHIEBACK, AND T. DHIPPAYOM (2015): "Adherence to diabetes medication: a systematic review," *Diabetic Medicine*, 32. [14]

LAW (1983): "Ley General de Salud," http://dof.gob.mx/nota_detalle.php?codigo=4652777&fecha=07/02/1984. [6]

——— (1995): "Ley del Seguro Social," http://www.imss.gob.mx/sites/all/statics/pdf/leyes/LSS.pdf. [5]

LEE, C., G. CHERTOW, AND S. ZENIOS (2009): "n Empiric Estimate of the Value of Life: Updating the Renal Dialysis Cost-Effectiveness Standard," *American Economic Review*, 12. [18]

NATHAN, D., J. KUENEN, R. BORG, H. ZHENG, D. SCHOENFELD, R. J. HEINE, AND A.-D. A. G. S. GROUP (2008): "Translating the A1C assay into estimated average glucose values," *American Diabetes Association*. [22]

NIDDK (2018): "The A1C Test & Diabetes," Last accessed 9 March 202, https://www.niddk.nih.gov/health-information/diagnostic-tests/a1c-test#:~:text=The%20A1C%20test%20measures%20the,level%20is%20below%205.7%20percent.https://medlineplus.gov/lab-tests/hemoglobin-a1c-hba1c-test/#:~:text=A%20hemoglobin%20A1c%20(HbA1c)%20test%20measures%20the%20amount%20of%20blood,the%20rest%20of%20your%20body. [5]

OECD (2019): "Diabetes Care," Last accessed 10 March 2021, https://www.oecd-ilibrary.org/sites/402e2e10-en/index.html?itemId=/content/component/402e2e10-en. [2, 6]

OJEDA, R. (2019): "Diabetes en México: así se comporta la segunda causa de muerte en el país," Last accessed 10 March 2021, https://www.milenio.com/ciencia-y-salud/diabetes-en-mexico-2020-estadisticas-y-porcentaje. [6]

PIMOUGUET, C., M. L. GOFF, R. THIÉBAUT, J. F. DARTIGUES, AND C. HELMER (2011a): "Effectiveness of disease-management programs for improving diabetes care: a meta-analysis," *CMAJ-JAMC*, 183. [3]

——— (2011b): "Systematic review and meta-analysis of different dietary approaches to the management of type 2 diabetes," *CMAJ*, 183. [4, 14]

SAEEDI, P., I. PETERSOHN, P. SALPEA, B. MALANDA, AND S. KARURANGA (2019): "Global and regional diabetes prevalence estimates for 2019 and projections for 2030 and 2045: Results from the International Diabetes Federation Diabetes Atlas, 9th edition," *Diabetes Research and Clinical Practice*, 157. [2, 5]

SSP (2018): "Afiliación histórica del Sistema de Protección Social en Salud," Last accessed 10 March 2021, https://www.gob.mx/salud/seguropopular/documentos/

afiliacion-historica-del-sistema-de-proteccion-social-en-salud#:~:text=A%20diciembre%20de%202018%2C%2053,la%20poblaci%C3%B3n%20total%20del%20pa%C3%ADs. [6]

——— (2021): "Catálogo de Clave Única de Establecimientos de Salud," Last accessed 10 March 2021, http://www.dgis.salud.gob.mx/contenidos/intercambio/clues_gobmx.html. [1, 6]

UMPIERRE, D., P. A. B. RIBEIRO, C. K. KRAMER, C. B. LEITÃO, A. T. N. ZUCATTI, M. J. AZEVEDO, J. L. GROSS, J. P. RIBEIRO, AND B. D. SCHAAN (2011): "Physical activity advice only or structured exercise training and association with HbA1c levels in type 2 diabetes: a systematic review and meta-analysis," *JAMA*, 305. [4, 14]

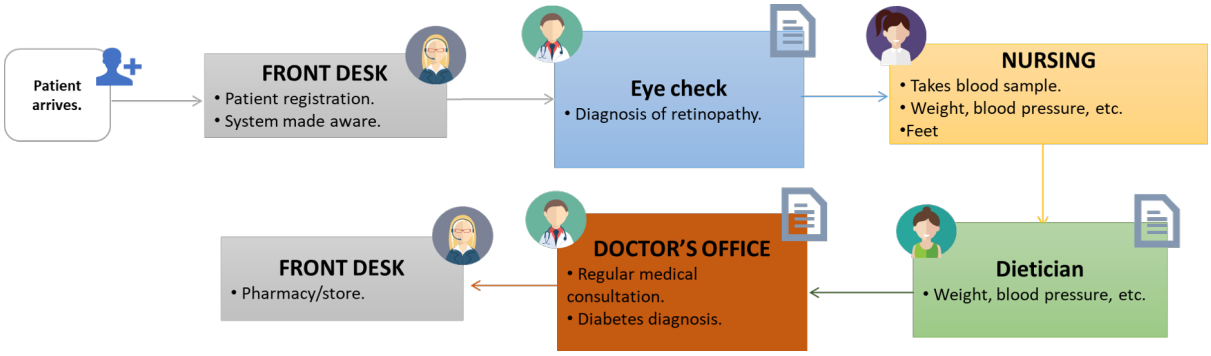
WHO (2021): "Mean fasting blood glucose," Last accessed 9 March 2021, https://www.who.int/data/gho/indicator-metadata-registry/imr-details/2380#:~:text=The%20expected%20values%20for%20normal,and%20monitoring%20glycemia%20are%20recommended. [5]

ÁNGEL RIVERA DOMMARCO, J., A. COLCHERO, M. L. FUENTES, T. G. DE COSÍO MARTÍNEZ, C. A. A. SALINAS, G. H. LICONA, AND S. BARQUERA (2018): *La Obesidad en Mexico Estado de la política públicay recomendaciones-para su prevención y control*, INSP. [6]

# A Description of CdA

This appendix presents additional context on the service that CdA provides. CdA is based on a one-stop-shop model for diabetes care that provides every service that a patient needs to monitor their health in one visit. The patient usually comes to the clinic and checks in with a receptionist at the front desk. The patient is then directed to visit 4 different stations in which she is reviewed for eye complications, then takes a blood test, gets vitals measured and is reviewed for diabetic foot, then visits the dietician and finally meets with the doctor to overview their condition. Finally the patient checks out at the front desk where she can acquire any medicine that she needs as well as a diverse set of diabetes related merchandise such as sugar-free chocolate and diabetes specialized shoes. Figure A.A.-1 summarizes this process.

**Figure A.A.-1:** CdA Service




**Notes:** This figure summarizes the stations that a patient goes to while visiting CdA

CdA has been expanding quickly over the last few years and aims to be accessible across Mexico. Their intention is to allow every patient to show up at the clinic that best suits them, whenever it best suits them. So for example, a patient could choose to get their first follow-up next to their workplace at clinic A during their lunch break and the next follow-up on a weekend at a clinic that is closer to their home. Figure A.A.-3 presents a map of CdA branches along with IMSS clinics as of December 2019 in the Monterrey area. Moreover, Table A.A.-1 compares the services provided at CdA to those provided by IMSS

**Figure A.A.-2:** CdA Branches in Monterrey


**Notes:** This figure summarizes branch locations for CdA. We include IMSS lcoations to highlight its widespread access

**Table A.A.-1:** IMSS vs CdA

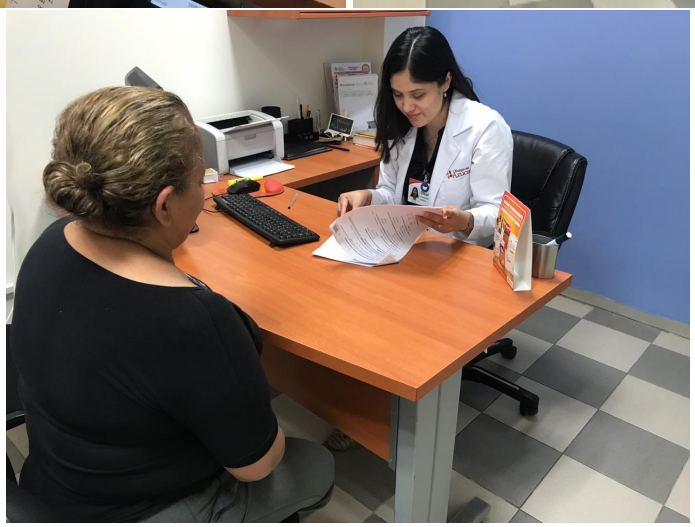
| Service | CdA | IMSS | Comparisson | IMSS costs |
|---|---|---|---|---|
| Doctor | 5 visits if control. 12 if not | 5 visits if control. 12 if not | = | 3990 |
| Nutrition | Personalized plan | General advice | CdA>IMSS | NA |
| Psychology | 2 visits | 0 visits | CdA>IMSS | 2582 |
| Dentist | 0 visits | 1 visit | IMSS>CdA | |
| General blood | 1/year | 1/year | = | 103 |
| HbA1c | Quarterly | Quarterly | = | 515 |
| Lipids | 2/year | 2/year | = | 206 |
| Microalbumina | 1/year | 1/year | = | 103 |
| Foot | Every 3 months | Every 3 months | = | 0 |
| Eye | once a year | Once a year | = | 1416 |
| Total Cost (pesos) | 7500 | | | 8915 |

**Notes:** This table compares services provided by CdA and IMSS as well as the spending that IMSS would incur in case it would provide the same services as CdA does, according to their public per-unit cost report.

Lastly, we present some pictures of how a clinic looks on the outside and inside below.

**Figure A.A.-3:** CdA Branch pictures



**Notes:** This figure exemplifies how a typical CdA branch looks

# B Deniers Randomization

In this appendix we revisit the argument behind the deniers randomization and explain where our assumptions play a key role. We develop the proof in a context where the entry randomization would be feasible and provide a causal LATE on compliers. That is, $Y_i$ captures the outcome for individual $i$, $D_i$ captures the enrollment decision by individual $i$ and there is a binary instrument $Z$ where $Z_i$ captures the instrument realization (random allocation into treatment or control) for individual $i$. Moreover, assume that there is a screening process $S$ that marks $S_i = 1$ if the individual is screened in and 0 otherwise. The outcome of individual $i$ is a function of decision to enroll, instrument and screening: $Y_i(Z,D,S)$. Similarly, the decision to enroll is a function of the instrument and the screening: $D_i(Z,S)$. For simplicity of exposition, we will denote $y_i(D_i)$ as the outcome when an individual $i$ gets enrolled or not and $D_i(Z_i)$ as the enrollment decision when the patient $i$ gets the instrument or not.

**Theorem 1** (Deniers LATE Equivalence)**.** *Let the following assumptions hold:*

1. *Stable Unit Treatment Value Assumption (SUTVA): for $Z,Z',D,D'$, if $Z_i = Z_i'$, then $D_i(Z,S) = D_i(Z',S)$ and if $Z_i = Z_i'$ , $D_i = D_i'$ then $Y_i(Z,D,S) = Y_i(Z',D',S)$ for all S.*
2. *Exclusion restriction: $Y(Z,D,S) = Y(Z',D,S)$ for all Z, $Z'$ and for all $D,S$.*
3. *Relevance (nonzero average causal effect of Z on D): $E[D_i(1,S) - D_i(0,S)] \neq 0$.*
4. *Monotonocity (no defiers): $D_i(1,S) \geq D_i(0,S)$ for all S.*

*On top of these assumptions, that are sufficient for estimating a causal LATE under a randomized instrument in visitors , we assume that:*

5. *$Pr(D_i(1) - D_i(0)|S_i = 0) = 0$. That is, no compliers are screened out by the screening process.*
6. *Screening does not affect enrollment decisions: $D(Z,S) = D(Z,S')$ for all S, $S'$.*
7. *Screening does not affect outcomes: $Y(Z,D,S) = Y(Z,D,S')$ for all S, $S'$ and for all D.*

*Under the previous assumptions,*

$$\beta_{IV}^{Deniers} = \frac{\mathbb{E}[y_i|Z_i = 1, S_i = 1] - \mathbb{E}[y_i|Z_i = 0, S_i = 1]}{\mathbb{E}[D_i|Z_i = 1, S_1 = 1] - \mathbb{E}[D_i|Z_i = 0, S_i = 1]} = \mathbb{E}[y_i(1) - y_i(0)|D_i(1) - D_i(0) = 1] = \beta_{IV}^{Visitors}.$$

*Proof.* Writing the potential outcome under being assigned to the instrument or not in terms of being screened in or screened out we get

$$\begin{aligned}\mathbb{E}[y_i|Z_i = 1] - \mathbb{E}[y_i|Z_i = 0] =& \Pr(S_i = 1)(\mathbb{E}[y_i|Z_i = 1, S_i = 1] - \mathbb{E}[y_i|Z_i = 0, S_i = 1]) \\ &+ \Pr(S_i = 0)(\mathbb{E}[y_i|Z_i = 1, S_i = 0] - \mathbb{E}[y_i|Z_i = 0, S_i = 0]).\end{aligned}$$

Expressing each one in terms of enrolling ($D$), we get

$$\Pr(S_i=1)(\mathbb{E}[D_i(1)y_i(1)+(1-D_i(1))y_i(0)|Z_i=1,S_i=1]-\mathbb{E}[D_i(0)y_i(1)+(1-D_i(0))y_i(0)|Z_i=0,S_i=1])$$

$$+\Pr(S_i=0)(\mathbb{E}[D_i(1)y_i(1)+(1-D_i(1))y_i(0)|Z_i=1,S_i=0]-\mathbb{E}[D_i(0)y_i(1)+(1-D_i(0))y_i(0)|Z_i=0,S_i=0]).$$

By independence of outcomes with respect to $Z_i$, we obtain

$$\Pr(S_i=1)\mathbb{E}[(D_i(1)-D_i(0))(y_i(1)-y_i(0))|S_i=1]+\Pr(S_i=0)\mathbb{E}[(D_i(1)-D_i(0))(y_i(1)-y_i(0))|S_i=0].$$

Let us define the set of always takers $AT$ as the subset of individuals $i$ such that $D_i(1)=D_i(0)=1$, the set of never takers $NT$ as the subset of individuals $i$ such that $D_i(1)=D_i(0)=0$, the set of compliers $C$ as the subset of individuals $i$ such that $D_i(1)-D_i(0)=1$ and the set of defiers $\delta$ as the subset of individuals $i$ such that $D_i(1)-D_i(0)=-1$. Noting that the set of defiers $\delta=\emptyset$ because of monotonicity, and that $D(1)-D(0)=0$ for $i\in AT\cup NT$, we get the following expressions:

$\Pr(S_i=1)\mathbb{E}[(D_i(1)-D_i(0))(y_i(1)-y_i(0))|S_i=1]$ =$\Pr(S_i=1)\mathbb{E}[y_i(1)-y_i(0)|i\in C,S_i=1]\Pr(i\in C|S_i=1)$ and
$\Pr(S_i=0)\mathbb{E}[(D_i(1)-D_i(0))(y_i(1)-y_i(0))|S_i=0]$ =$\Pr(S_i=0)\mathbb{E}[y_i(1)-y_i(0)|i\in C,S_i=0]\Pr(i\in C|S_i=0)$

Since $\Pr(i\in C|S_i=0)=0$, the second expression is 0. Thus,

$$\Pr(S_i=1)(\mathbb{E}[y_i|Z_i=1,S_i=1]-\mathbb{E}[y_i|Z_i=0,S_i=1])=\Pr(S_i=1)\mathbb{E}[y_i(1)-y_i(0)|i\in C,S_i=1]\Pr(i\in C|S_i=1)$$

Therefore, as $\Pr(i\in C|S_i=1)=\mathbb{E}[D_i|Z_i=1,S_1=1]-\mathbb{E}[D_i|Z_i=0,S_i=1]$

$$\beta_{IV}^{Deniers}=\frac{\mathbb{E}[y_i|Z_i=1,S_i=1]-\mathbb{E}[y_i|Z_i=0,S_i=1]}{\mathbb{E}[D_i|Z_i=1,S_1=1]-\mathbb{E}[D_i|Z_i=0,S_i=1]}=\mathbb{E}[y_i(1)-y_i(0)|(D_i(1)-D_i(0)=1,S_i=1]$$

Now, given that screening does not affect outcomes nor enrollment decisions and that if an individual is a complier she is screened in, the conditionality on the right hand side in terms of S is redundant. So,

$$\beta_{IV}^{\text{Deniers}}=\frac{\mathbb{E}[y_i|Z_i=1,S_i=1]-\mathbb{E}[y_i|Z_i=0,S_i=1]}{\mathbb{E}[D_i|Z_i=1,S_1=1]-\mathbb{E}[D_i|Z_i=0,S_i=1]}=\mathbb{E}[y_i(1)-y_i(0)|(D_i(1)-D_i(0)=1]=\beta_{IV}^{\text{Visitors}}$$

□

# C Other Tables and Figures

**Table A.C.-1:** CdA vs Average in ENSANUT

| Variable | CdA N | CdA Mean/SE | ENSANUT N | ENSANUT Mean/SE | Difference C-T |
|---|---|---|---|---|---|
| ***Panel A: Demographics*** | | | | | |
| Age | 1067 | 52.74 | 2120 | 56.79 | -4.05*** |
| | | [0.36] | | [0.30] | |
| Male | 1067 | 0.33 | 2120 | 0.38 | -0.05*** |
| | | [0.01] | | [0.01] | |
| ***Panel B: Health*** | | | | | |
| HbA1c | 1023 | 9.43 | 2120 | 7.51 | 1.92*** |
| | | [0.08] | | [0.05] | |
| BMI | 1024 | 31.03 | 1199 | 31.03 | 0.00 |
| | | [0.19] | | [0.18] | |
| Hypertension | 1067 | 0.38 | 2120 | 0.43 | -0.05** |
| | | [0.01] | | [0.01] | |
| ***Panel C: Health care and expenses*** | | | | | |
| Takes pills | 1025 | 0.61 | 2120 | 0.58 | 0.03* |
| | | [0.02] | | [0.01] | |
| Takes insuline | 1025 | 0.23 | 2120 | 0.14 | 0.09*** |
| | | [0.01] | | [0.01] | |
| I(Diet) | 1025 | 0.33 | 2120 | 0.27 | 0.06*** |
| | | [0.01] | | [0.01] | |
| I(Exercise) | 1025 | 0.28 | 2120 | 0.06 | 0.22*** |
| | | [0.01] | | [0.01] | |
| ***Panel D: Diabetic complications*** | | | | | |
| Ulcers | 1025 | 0.04 | 2120 | 0.06 | -0.01 |
| | | [0.01] | | [0.01] | |
| Heart attack | 1025 | 0.01 | 2120 | 0.01 | 0.00 |
| | | [0.00] | | [0.00] | |
| ***Panel E: Treatment at IMSS*** | | | | | |
| IMSS | 1025 | 0.57 | 2120 | 0.23 | 0.34*** |
| | | [0.02] | | [0.01] | |

**Notes:** This table presents Descriptive statistics of CdA patients that are in our experiment and the average diabetic surveyed in the 2018 health and nutrition survey (ENSANUT) * $p < 0.10$, ** $p < 0.05$. We can see that patients in our experimental sample are younger and more likely to be female. In terms of health, patients in our sample are in a worse condition in terms of HbA1c, but less likely to have hypertension. Both samples have similar BMI. Moreover, we can see that our sample is more likely to take pills and insuline, and also more likely to diet and exercise. They look similar in terms of complications, even though our sample is twice as likely to utilize IMSS services.

**Table A.C.-2:** Effect on Trust and Waiting Time

| | Dif Trust in Improving (1) | Dif Trust in Diagnosis (2) | Waiting time (2) |
|---|---|---|---|
| Use Cda | 1.92*** | 1.48** | -30.33*** |
| | (0.59) | (0.59) | (9.02) |
| Branch FE | Yes | Yes | Yes |
| Enrollment Month Fe | Yes | Yes | Yes |
| Basic Controls | Yes | Yes | Yes |
| Observations | 682 | 676 | 739 |
| R-squared | 0.02 | 0.02 | -0.05 |
| F | 82.97 | 92.88 | 82.49 |
| First coeff | 0.330 | 0.348 | 0.325 |
| Mean dep. var | 2.107 | 2.292 | 43.46 |

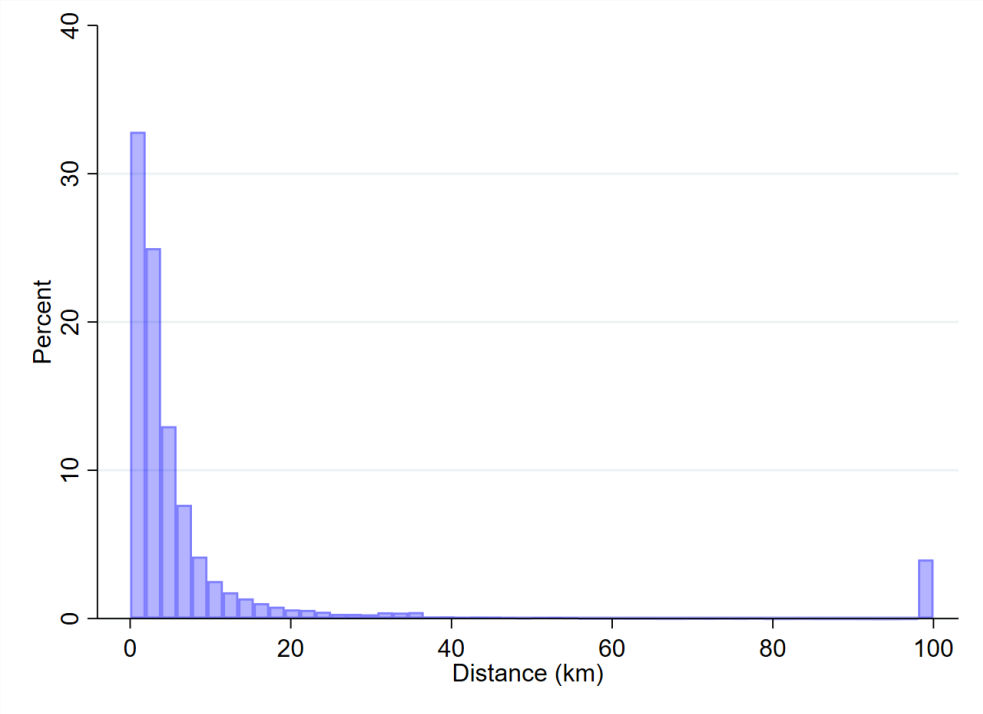
**Notes:** This table presents the results from the effect of getting CdA treatment on trust and waiting time. The first column captures the local average treatment effect (LATE) on the difference in the self-reported trust that the patient will improve their health through CdA, as opposed to through their health provider at baseline (the difference in the rows shown in Table 2), we also control by the baseline measure of the variable which is define as the difference from the baseline survey on the same questions. The second column carries out the same exercise for patients' trust in diagnoses that come from CdA as opposed to their health provider at baseline, we also control by the baseline measure of the variable which is define as the difference from the baseline survey on the same questions. The third column reports our results on waiting time. On our follow-up survey we asked how long do patients usually wait at the doctor they usually go to get diabetes care and estimated the difference based on our experiment, instrumenting going to CdA rather than to other providers by being in the treatment group. We winzorized at 5% the upper tail to omit implausible waiting times. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Table A.C.-3:** Comparing Patients With and Without Follow-up

| Variable | Patient without FU N | Patient without FU Mean/SE | Patient with FU N | Patient with FU Mean/SE | Difference C-T |
|---|---|---|---|---|---|
| ***Panel A: Demographics*** | | | | | |
| Age | 1215 | 51.54 | 1067 | 52.74 | -1.20** |
| | | [0.34] | | [0.36] | |
| Male | 1343 | 0.48 | 1067 | 0.33 | 0.15*** |
| | | [0.01] | | [0.01] | |
| % Incomer higher than C | 1343 | 0.67 | 1067 | 0.56 | 0.11*** |
| | | [0.01] | | [0.02] | |
| % High School or more | 1206 | 0.47 | 942 | 0.38 | 0.09*** |
| | | [0.01] | | [0.02] | |
| ***Panel B: Health and Health services*** | | | | | |
| HbA1c | 1339 | 9.59 | 1064 | 9.43 | 0.16 |
| | | [0.07] | | [0.08] | |
| BMI | 1342 | 30.76 | 1066 | 30.94 | -0.19 |
| | | [0.17] | | [0.18] | |
| Has IMSS, ISSSTE or Seguro Popular | 1343 | 0.77 | 1067 | 0.79 | -0.01 |
| | | [0.01] | | [0.01] | |
| Percentage that use IMSS, ISSSTE or Seguro Popular | 1343 | 0.74 | 1067 | 0.74 | 0.00 |
| | | [0.01] | | [0.01] | |
| HbA1c out of control (HbA1c>10) | 1339 | 0.41 | 1064 | 0.40 | 0.02 |
| | | [0.01] | | [0.01] | |
| ***Panel C: Beliefs*** | | | | | |
| Trust to improve following CdA recomendations (0-10) | 1168 | 9.26 | 913 | 9.35 | -0.09 |
| | | [0.05] | | [0.05] | |
| Trust to improve following current health provider recomendations (0-10) | 1089 | 7.94 | 861 | 7.90 | 0.04 |
| | | [0.08] | | [0.09] | |
| Trust in CdA diagnosis (0-10) | 1157 | 9.37 | 899 | 9.49 | -0.12* |
| | | [0.05] | | [0.05] | |
| Trust in current health provider diagnosis (0-10) | 1119 | 7.88 | 874 | 7.91 | -0.03 |
| | | [0.08] | | [0.09] | |

**Notes:** This table presents a comparisson between the patients who did answer our follow-up and the ones who did not. *(1) % High income*: Is an indicator variable equal to 1 if the person lived in zip codes that the AMAI classifies as above middle income. This variable is an administrative variable and is based on the place of residency of each potential client at CdA. *(2) Has IMSS, ISSSTE or Seguro Popular*: Is an indicator variable equal to 1 if the person declared to be affiliated to a public service to the medical staff at CdA. *(3) Percentage that use IMSS, ISSSTE or Seguro Popular*: Is an indicator variable equal to 1 if the person declared to go for medical attention to the public services in our baseline survey. *(4) Trust to improve following CdA recommendations (0-10)*: Is variable that measures the trust in improving with CdA. *(5) Trust to improve following current health provider recommendations (0-10)*: Is a variable that measures the trust in improving with their current health provider. Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$.

**Figure A.C.-1:** Distance to Clinic Distribution




**Notes:** This figure shows the distribution in distance form home to the clinic for the IMSS patients we utilize in our regressions. We winzorized the distances higher than 100 km.