UNIVERSITY OF SOUTHAMPTON

Faculty of Social Sciences
School of Economic, Social and Political Sciences

Three Essays in Health Economics

by

Rhys Llewellyn Thomas
BScSoc Economics (University of Southampton)
MSc Economics (University of Southampton)
ORCiD: 0000-0003-0124-0391

A thesis for the degree of
Doctor of Philosophy

March 2022

University of Southampton

Abstract

Faculty of Social Sciences
School of Economic, Social and Political Sciences

Doctor of Philosophy

Three Essays in Health Economics

by Rhys Llewellyn Thomas

This thesis presents three essays on policy-related topics in Health Economics. The specific policy topics this thesis explores are health insurance and inequality, spillovers or peer effects of health behaviours, and the impact of human resources in healthcare.

The first essay in this thesis, I analyse the redistributive effects of a publicly financed healthcare expansion. Using data from the Oregon Health Insurance Experiment (OHIE) we analyse the redistributive impacts of a publicly financed health insurance expansion. We use a residual inclusion methodology combined with quantile regression to estimate the heterogeneity in the effects of health insurance across the income distribution. We find that there are indeed redistributive impacts, even in the small income window we have access to, which would otherwise be concealed in a linear regression. Those at the lower end of the income distribution tend to have a substantial increase in their disposable income as a result of health insurance coverage, while those at the upper end see no change in income. We additionally estimate that increased employment in at risk households is driving this effect.

In next essay I analyse the spillover effects of a diabetes diagnosis. Diabetes is a unique condition, in that a positive change in lifestyle and behaviour, is both the first line treatment and the recommended method of preventing the disease. It is theoretically possible that by jointly partaking in diabetes treatment, partners of people with diabetes would substantially benefit from their partners' diabetes diagnosis. Using blood data from the Health Survey for England, and a fuzzy regression kink design, we causally estimate the effect of a diabetes diagnosis on health-related behaviours of the individual with diabetes, as well as, their partners. We find that a diagnosis of diabetes results in a significant increase in the probability of exercising and a decrease in the probability of currently being a smoker both for the diabetic individual and their partner. However, we find limited evidence of other lifestyle changes. From a public health perspective, our results are especially important for the evaluation of diabetes related policies, while positive spillovers, particularly within households, should be taken into account in the evaluation process.

In the final essay of this thesis, I analyse the impact of primary care physicians on health outcomes. Worldwide there is a growing concern that there are insufficient primary care physicians to meet demand. There is, however, mixed evidence on how effective primary care is in influencing population health outcomes. I estimate the effect of an increasing in primary care physicians using the *Programa Mais Médicos*. Although previous studies have used the *Programa Mais Médicos* to analyse the impact of a primary care physician supply, I exploit the variation in physicians allocated to each municipality and use only treated municipalities to identify the impact of primary care physicians on hospitalisations and mortality. I estimate the impact of primary care physicians using a generalised synthetic control estimator and find limited evidence of primary care physicians impacting health outcomes. These results question the notion that primary care physicians are a cost-effective means of improving population health.

Contents

List of Figures

List of Tables

Declaration of Authorship

I declare that this thesis and the work presented in it is my own and has been generated by me as the result of my own original research.

I confirm that:

1. This work was done wholly or mainly while in candidature for a research degree at this University;
2. Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated;
3. Where I have consulted the published work of others, this is always clearly attributed;
4. Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work;
5. I have acknowledged all main sources of help;
6. Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself;
7. None of this work has been published before submission

Signed:..... Date:.....

Co-Authorship Statement

Chapter 2 (*Redistributive effects of a publicly financed healthcare expansion*) and Chapter 3 (*The direct and spillover effects of diabetes diagnosis on lifestyle behaviours*) were co-written with Dr. Emmanouil Mentzakis (University of Southampton). My contribution to the production of these research works is outlined below:

- Identification of research question – Shared responsibility with co-author.
- Empirical strategy – Shared responsibility with co-author.
- Data analyses – My sole responsibility.
- Manuscript preparation – Shared responsibility with co-author.

Acknowledgements

As I reach the end of this PhD program, there are several people that I would like to thank for their substantial contribution to this thesis. I will begin these acknowledgements with those that have had the most salient influence on this project.

Firstly, I would like to thank Dr. Emmanouil Mentzakis. As a supervisor, he has always been on hand to discuss new ideas, offer advice, give feedback, and has invested a substantial amount of his own time on every part of this thesis. As a colleague, he has improved the papers in this thesis substantially, and I look forward to many more future projects together. More importantly, however, as a friend, he has always provided humour, encouragement, support and always made sure that I am in the right frame of mind to do my work. Dr. Thomas Gall deserves acknowledgment for providing valuable advice and always making me feel able to ask questions on whatever I have needed advice on. He has always provided words of wisdom when asked and has been a continued voice of reason throughout my time at Southampton.

I would like to thank the Department of Economics at the University of Southampton more generally. Throughout, the students and staff have played a vital role in my academic training. I would also like to acknowledge the South Coast DTP, not only for their financial support but for providing ample development opportunities.

Dr. Brendon McConnell has also provided valuable critique and feedback throughout, and this thesis has been improved dramatically thanks to his contribution. Dr. Thomas Hone and Prof. Christopher Millet also played a substantial role in the formulation of chapter four of this thesis, and their contribution is immensely appreciated.

I have been fortunate enough to be surrounded by countless supportive friends and family throughout my studies, all of whom have provided support and offered a welcome distraction when it was needed. I realise that I have been incredibly lucky to be surrounded by a very special group of people, and I'm extremely thankful for that. All of these people have played a much larger part in the completion of this thesis than they would realise and I want to express my gratitude to them.

Finally, I would like to thank my brother and my parents. They have played no small part in this thesis, and I will forever be grateful for everything they have done and continue to do for me. The completion of this thesis would not have been at all possible without them, and for that, they deserve a special thank you.

I Mam, Dad ac Elis, diolch am bopeth chi wedi rhoi a 'neud i fi.
Fyddau byth yn gallu dangos fy ngwerthfawrogiad digon.

Chapter 1

Introduction

Ever since the beginning of the century global life expectancy at birth has risen from 67 to 73 years, and adult mortality rate per 1000 population has fallen from 181 in 2000 to 142.3 in 2016 (World Health Organization, 2021b). These improvements in global health have also come with large changes in causes of death, with heart disease, stroke and chronic obstructive pulmonary diseases being the three leading causes of death globally, and only one communicable disease in the top ten causes of deaths in high-income countries (World Health Organization, 2021b). Longer life expectancy and changing trends in mortality have, however, brought new challenges to modern healthcare systems. The rise in non-communicable diseases requires the strengthening of preventative and screening interventions (World Health Organization, 2021a). The disparity in morality causes and life expectancy between high-income countries and developing nations also shows the need to tackle health inequalities, and these inequalities are not only present between countries but also within them. Tackling inequalities and improving the prevention and treatment of non-communicable diseases is also a priority of the Sustainable Development Goals (World Health Organization, 2019). These challenges require the substantial involvement of policymakers, and public policy will play a key role in addressing these issues.

Public policy, and the efficient allocation of healthcare resources is vital for ensuring that healthcare systems are able to handle these challenges. Throughout the COVID-19 pandemic the importance of public policy in relation to health has become clear. Countries have relied on well-designed and efficient healthcare services and evidence-based policies to slow the spread of the virus and to deal with this particular challenge, and public policy will continue to play a central role in future health challenges. The COVID-19 pandemic has also highlighted the substantial inequalities across the globe. It is necessary that sufficient research is conducted such that policymakers are well informed regarding the impacts of policies seeking to address these challenges.

Over the past several decades health economists have contributed substantially to our understanding of the impacts of public policy. They have contributed the discovery of econometric tools for empirically evaluating public policies, providing theoretical models for understanding behaviours and expected consequences of policies, and also contributing a framework for evaluating the cost effectiveness of interventions. These contributions have allowed for the improved efficiency and equity of healthcare systems, and also improvements in health outcomes of people worldwide. Health economics has been vital for dealing with the challenges that healthcare systems have faced and continues to be important for dealing with future challenges and delivering improved care and health outcomes.

This thesis seeks to add to the previous contributions of health economists by conducting necessary investigations into policy-related questions in health, and relies heavily on economists' methodological contributions to do so. The importance of empirical evidence for tackling these modern challenges, of the rise of non-communicable diseases and increasing health inequalities, cannot be understated. Using state of the art microeconometric techniques I analyse three distinct policy-related questions, and in doing so, I contribute empirical evidence necessary for tackling these challenges in a evidence driven way. In addition to the main aim of contributing to public policy discussions, this thesis also provides empirical evidence to support economic theory. The findings in this thesis are therefore relevant for policymakers and academics alike, and the results in this thesis allow governments to better design policies based on the evidence in this thesis, while also improving our understanding of healthcare systems and individual behaviours. The results in this thesis presents answers to policy-relevant questions, while also unveiling further questions and avenues of future research, that would, when answered, allow us to better understand how to tackle these modern challenges and future challenges in health.

The specific policy topics this thesis explores are health insurance and inequality, spillovers or peer effects of health behaviours, and the impact of human resources in healthcare. Specifically, the first essay analyses the income redistribution associated with an expansion of publicly financed health insurance, and further understanding the mechanisms behind this redistribution. The second essay of this thesis investigates the behavioural impacts of a diabetes diagnosis on the diagnosed individual, as well as their partners. The final chapter focuses on human resources in healthcare, and specifically the impact primary care physicians have on health outcomes. The findings in this thesis contribute to understanding of health-related public policy. My work also indirectly contributes to better understanding the wider benefits or lack thereof of various policies, and in doing so, providing useful evidence to ensure efficient allocation of healthcare resources and thereby allowing policymakers to tackle these modern healthcare challenges.

The remainder of this chapter will outline the background and specific contributions of each essay in this thesis.

1.1 Redistributive effects of a publicly financed healthcare expansion

In the first essay of this thesis, I analyse the impact an expansion of publicly financed health insurance has on income inequality.

There has been a wide body of evidence finding inequalities in health outcomes by income (Marmot et al., 1991; Kaplan et al., 1996; van Doorslaer et al., 1997; Humphries and van Doorslaer, 2000; Doorslaer et al., 2006), and universal healthcare coverage has been widely promoted as a means of reducing inequalities. Indeed, given this extensive body of evidence claiming that there are significant inequalities in health-related outcomes, it would seem implausible to claim that universal healthcare coverage, or more broadly publicly financed healthcare expansion, would have the same impact on the entire income distribution. In addition, claims by the WHO (2019b,a) imply that there is a potentially beneficial impact of publicly financed healthcare in favour of low-income individuals. Although there is a theoretical justification of a redistributive impact, there is limited empirical evidence to support the theory.

Much of the theoretical justification of publicly financed health coverage reducing inequality comes from the analysis of public good provision more broadly. A theoretical framework for the public provision of private goods was proposed by Besley and Coate (1991). Their paper explores the redistributive power of public financing of these goods. The basic idea is that all individuals contribute to a public good, but some high-income individuals do not use the good, therefore redistributing income away from the highest income individuals, and as a result reducing inequality. Glomm and Ravikumar (1992) provides further theoretical justification of an income redistribution from publicly provided private goods, this time in the context of education. Glomm and Ravikumar estimate the effect of public versus private financing of education on inequality. The authors theorise that in a private educational system children's endowment depend on parental income, while their endowment in a public education system depends only on the average income, therefore the public education system reduces inequality at a faster rate than the private system. Education and Health care share many characteristics, it is not uncommon for both goods to be provided publicly and privately, and both contribute to human capital. Given the parallels that can be drawn, it may be expected that the effects of their financing on inequality would be similar. Therefore, this paper potentially provides some evidence that publicly financing institutions that contribute to human capital

can have an impact on inequality. The theoretical findings of this paper are supported empirically by Sylwester (2002). Using data for 50 countries and analysing the change in the Gini coefficient from 1970 to 1990, Sylwester (2002) finds evidence in support of Glomm and Ravikumar (1992)'s theoretical findings.

Literature relating specifically to Health Insurance Financing and Inequality is sparse. However, one study that introduces analyses of the effects of public and private shares of health expenditures on income inequality was conducted by Bhattacharjee et al.. The study proposes a theoretical model and then verifies the findings by using empirical analysis. The model used in this paper is similar to that of Glomm and Ravikumar (1992) however instead of education expenditure the paper analyses health care expenditure. Human capital of the child is a function of parental health and human capital. Therefore, investment in health increases longevity of life but also child's human capital. Childhood health is well documented to have both a significant effect on schooling and labour but is also highly correlated with parental health. The authors use the prediction of their theoretical model as a hypothesis to test using Indian vaccine and expenditure data. The theoretical model's prediction was that "higher share of private to public health expenditures leads to higher income inequality" (Bhattacharjee et al., 2017). In addition to their theoretical contribution, they provide empirical evidence of the impact of share of private to public health expenditures on Gini coefficient. These empirical findings concur with the theoretical findings and the authors conclude that the empirical analysis supports the hypothesis that increased share of private healthcare spending increases inequality.

In this essay, I hypothesise that publicly financed healthcare can cause an income redistribution through two separate pathways. An expansion of publicly financed health insurance has already been shown to reduce healthcare expenditures. Finkelstein et al. (2012)'s analysis of the Oregon health insurance experiment already found there to be "a statistically significantly higher health care utilization (including primary and preventive care as well as hospitalizations), lower out-of-pocket medical expenditures and medical debt (including fewer bills sent to collection)" as a result of Medicaid coverage, compared to the uninsured control group. Although this analysis did not estimate heterogeneous treatment effects of Medicaid coverage, there is theoretical support of a differential impact by income groups. Previous research has found that the price elasticity of healthcare expenditures ranges from -0.04 to -1.5, being approximately -0.3 on average (Ringel et al., 2002). When this is combined with the finding that out-of-pocket payments are relatively largest for the lowest income group (Pannarunothai and Mills, 1997), it might be reasonable to expect that the left tail of the income distribution would have a disproportionately larger effect. In addition, Wagstaff et al. (2009) analysed a heavily subsidised voluntary health insurance scheme in China in 2003 and found that out of pocket spending saw a smaller increase for the poor relative to the richer group. Therefore, in this essay I

consider whether an expansion of publicly financed healthcare does heterogeneously impact the out-of-pocket costs of lower income individuals more so than those with higher income.

The alternative pathway in which Universal Healthcare Coverage can theoretically increase disposable income is through higher income due to better health outcomes. There is strong theoretical evidence in the literature of a link between health and income. Since Grossman (1972)'s model of health production, there has been extensive evidence of health having a positive causal impact on income (Luft, 1975; Bartel and Taubman, 1979; Chirikos and Nestel, 1985; Contoyannis and Rice, 2001). Given that Finkelstein et al. (2012) found that Medicaid coverage resulted in "better self-reported physical and mental health than the control group", it would be reasonable to expect an increase in income for those that received Medicaid coverage. Once again, heterogeneity is also theoretically justified. Previous work on income related health inequalities has shown there to a be a significant inequality gradient of better health for the well off and lower health for low income individuals (Wilkinson, 1997). These inequalities have been found to be particularly evident in the United States and the United Kingdom (van Doorslaer et al., 1997). A characteristic of the Grossman model is that as health stock increases the marginal benefit of health decreases, therefore the lowest income group are expected to have the largest marginal benefit of healthcare. This claim is somewhat supported by Manning et al. (1987), who found there to be some suggestive evidence of an increase in health of those that are both low health and low income, however there was little evidence to suggest that either low income or low health groups separately saw an increase in health as a result of publicly financed health insurance coverage. The heterogeneous effect of UHC on health outcomes by income groups can therefore theoretically translate to a heterogeneous increase in earnings in favour of the lowest income groups. Although Baicker et al. (2014) found there to be no effect on overall individual earnings as a result of Medicaid coverage, the theoretical pathway presented here suggests that heterogeneous effects of Medicaid on earnings is possible. Therefore, I also seek to analyse whether there is heterogeneity in the impact of publicly financed health insurance on earned income.

The aim of the first essay of this thesis is therefore to provide theoretical evidence of heterogeneity in the impact of publicly financed healthcare on household disposable income, and subsequently I make a claim regarding the redistributive power of public provision of healthcare. The findings of this essay provide evidence of a secondary redistributive effect, specific to healthcare provision, which works in parallel to findings of previous research on the public provision of private goods (Besley and Coate, 1991; Glomm and Ravikumar, 1992; Bhattacharjee et al., 2017). In this essay, I focus my attention on redistribution through household finances and not through public finances and taxation. Specifically, this essay uses the experimental setting of the Oregon Health Insurance Experiment (OHIE) to causally assess the heterogeneity

in the effect of publicly financed healthcare on disposable income. In addition, the essay separately estimates the effects on household earned income and healthcare expenditure. The essay also assesses the causal channels in which these effect work through.

1.2 The direct and spillover effects of diabetes diagnosis on lifestyle behaviours

The third chapter of this thesis has more of a focus on individual decisions related to health; however, the motivation and conclusions are still profoundly important for policymakers. This essay evaluates whether spillovers in behaviours exist between partners as a result of an update in health status.

When considering the cost-effectiveness of various public policies, typically health economists do not consider the effect of the policy or intervention on the health outcomes of anyone else, other than the targeted individual or population. However, in doing so, health economists may be substantially under-estimating the true impact of such policies and interventions, if spillover effects exists, and consequently many services may be under-provided.

The idea that policies impact a wider network of individuals and that this should be considered in the cost-evaluation framework is not a recent advancement, and the Panel on cost-effectiveness in Health and Medicine stated that "all health effects and costs that flow from it [the policy or intervention] are counted, regardless of who would experience them" (Russell et al., 1996; Shepard, 1999; Basu and Meltzer, 2005). However, even though health economists are aware that indirect effects or spillovers may exist, it is typically the case that the change in quality adjusted life years of the targeted population is the means of comparing those interventions (Garber and Phelps, 1997). However, the policies themselves may have substantial indirect impact on the health outcomes of individuals that are not specifically targeted by the programme. These kind of spillovers are theoretically and empirically justified.

At the household level, which is the focus of this essay, there are two theoretical pathways in which it might be expected that an intervention would impact other members of the household. The New Household Economic theory (Lancaster, 1966; Becker, 1981) proposes that individuals within the household produce and consume some shared output, and therefore any intervention which, in some way, impacts the production or consumption decisions of one productive household member, would also impact the production or consumption decisions of other productive household members. Based on this framework, it is reasonable to expect behaviours of other members living in the same household to change if a policy aims to change the

behaviours of some targeted individual. A further discussion of the theoretical justification of an indirect effect, or as framed in the essay, a spillover effect, is presented in chapter three.

Empirically, externalities, spillovers, peer effects, or indirect effects have been found to be present in a number of settings, one of the most prominent being in education (Hoxby, 2000; Boozer and Cacciola, 2001; Hanushek et al., 2003; Angrist and Lang, 2004; Hoxby and Weingarth, 2005; Whitmore, 2005; Fuest, 2007; Ammermueller and Pischke, 2009; Lavy and Schlosser, 2011; Sacerdote, 2011; Carman and Zhang, 2012; Burke and Sass, 2013). However, there is also a growing literature on the spillovers in health settings.

Work by Persson et al. (2021) focuses on the spillovers of a marginal ADHD diagnosis on younger siblings. They estimate that an older sibling being diagnosed with ADHD causes an increase in the probability of a diagnosis in the younger sibling. They additionally find that the probability of the sibling taking ADHD medication increases as a result. Gathmann et al. (2020) analyse the within household impacts of job displacements on health outcomes. Specifically, they analyse how a job displacement in one partner impact mortality and hospitalisations. They find that male job displacement has a substantial impact on the mortality risk of the displaced male and their spouse. Gathmann et al. clearly show that there is an interdependence between spouses, and that spillovers onto partners indeed exist and should be considered in the evaluation process. The work most related to this essay is Fadlon and Nielsen (2019)'s paper on the impact of an update of health knowledge on various health-related behaviours. Their works analyses the impact of fatal and non-fatal heart attacks on behaviours such as: statin consumption, hospital visits, consumption of opioids and smoking medication. They find that heart attacks cause significant changes in behaviours of close family members. They also find a significant impact on a wider network of co-workers of the heart attack sufferer.

As well as this literature in economics, there is also an extensive literature on spousal correlation in health-related behaviours. Much of the work thus far has focused on the correlation between spouses in terms of smoking behaviour (Barrett-Connor et al., 1982; Venters et al., 1984; Graham and Braun, 1999; Franks et al., 2002; Bloch et al., 2003; Clark and Etilé, 2006; Stimpson et al., 2006; Christakis and Fowler, 2008; Falba and Sindelar, 2008; Cobb et al., 2014) and alcohol consumption (Kolonel and Lee, 1981; Graham and Braun, 1999; Leadley et al., 2000; Leonard and Mudar, 2003; Stimpson et al., 2006; Falba and Sindelar, 2008). However Farrell and Shields (2002) and Falba and Sindelar (2008) also analyses physical activity, and find that there is a strong positive correlation of physical activity between household constituents. Kolonel and Lee (1981), Barrett-Connor et al. (1982), Macario and Sorensen (1998), Bove et al. (2003) and Lyu et al. (2004) all estimate the correlation between spouses' diets, and consistent with the other studies on concordance of behaviour, find that spouses' diets show a

strong correlation. However, these correlations extend beyond behaviours alone. There has been work documenting spousal correlation in mental health and physical health[1].

This work clearly shows that there is an interconnected-ness between spouses, relatives and even co-workers. This growing literature provides empirical evidence of these spillovers beyond what may usually be considered by policy makers; however, individuals' interconnected-ness means that we should expect these indirect impacts. The essay in chapter three contributes to this literature, focusing on the indirect or spillover effect of a diabetes diagnosis on behaviours known to be risk factors for health.

The overall aim of the essay in chapter three is to explore whether changes in the behaviours of one partner induces changes in behaviours of the other. Specifically, I analyse the impact of a diabetes diagnosis on the diagnosed individual, as well as the impact of the diagnosis on their partners. Diabetes provides a unique opportunity to analyse these spillovers, in that a positive change in lifestyle and behaviour, is both the first line treatment and the recommended method of preventing the disease. If individuals diagnosed with diabetes are compliant to treatment and change these lifestyle behaviours, then we may theoretically expect other household members to do the same. This essay empirically evaluates whether health-related spillovers do indeed exist, which is vital knowledge for policymakers and health economists seeking to evaluate the cost-effectiveness of various policies.

However, although documenting the existence of spillovers is vital for informing policy makers, understanding the causal mechanisms behind these indirect effects contributes to the economic literature more broadly. Alongside the main contribution of documenting causal spillovers, an additional contribution is the analysis of the causal channels in which these spillovers work, and whether it is information transfer or joint participation which is driving these effects.

1.3 More Doctors, better health? Considering doctor numbers in the Mais Medicos Programme

The final essay of this thesis returns to a topic which is more directly relevant to policy makers, by analysing the impact of human resources in healthcare. Specifically, I analyse the impact of primary care physicians on population health outcomes. There is growing concern that there are insufficient primary care physicians to provide adequate care to populations worldwide. However, given the mixed evidence of the impact of primary care physicians on health outcomes, it is not clear whether this will

[1] See systematic review by Meyler et al. (2007) and references therein for full discussion

eventually be problematic for population health. This essay examines whether primary care physicians do indeed impact the health outcomes of the population.

The role of primary care physicians in many healthcare systems is to treat common diseases and illnesses, provide preventative care and to be a link between patients and the rest of the healthcare system. Primary care physicians are particularly important for attempting to tackle the rise in non-communicable diseases. Non-communicable diseases, as mentioned, are of growing concern, and are a substantial burden on modern healthcare system. Given that preventative treatment is the most influential means of dealing with this burden, primary care is reasonably thought of as a vital component of the strategy for tackling these challenges. The ability to treat common illnesses also alleviates strain on the rest of the healthcare system too, because they are able to treat patients without the need to use scares healthcare resources.
Well-functioning healthcare systems also rely on primary care physicians to refer patients to the relevant specialities. It is widely accepted that primary care plays a vital role in improving population health through these services, and there is significant literature findings strong associates between strong primary healthcare system and health outcomes (Starfield, 1998; Starfield et al., 2005; Guanais and Macinko, 2009; Macinko et al., 2010; Caley, 2013; Rao and Pilot, 2014; Rasella et al., 2014; Bitton et al., 2019). This is especially the case in areas that typically lack a reliable healthcare system and, in some communities, can be the only healthcare available (World Health Organization, 2021c). Previous work has made the claim that primary care also contributes to the efficient functioning of the healthcare system as a whole (Starfield et al., 2005).

Given the role that primary care physicians play within a wider healthcare system it is clear to see, theoretically, how increasing the number of primary care physicians would improve population health directly. Increasing primary care physician supply has the potential of increasing the probability of early detection of various conditions, through increasing the capacity of regular check-ups. It also allows for an increase in the amount of preventative treatments that would be given, and provides further opportunity for patients to receive care for common conditions that would have otherwise not been performed if there was an under supply of physicians. But equally as important, increased provision of preventative services and the management of common illness has the potential to reduce the strain on other healthcare services. Early detection increases probability of survival with most conditions and reduces the risk of hospitalisations. Preventative services also decrease risk of hospitalisations by reducing the risk of developing some diseases. Therefore, an increase in primary care physician numbers can both impact health outcomes directly, but also indirectly through reduced strain on the rest of the healthcare system. The direct and indirect effect on population health could potentially make increasing primary care physician numbers a cost-effective means of improving population health. These theoretical

benefits of increasing physician supply, do however, rely on the assumption that there is excess demand within the healthcare system for these services. Indeed, there is worldwide concern that there are insufficient primary care physicians to meet the demand (World Health Organisation, 2006; Gladu, 2007; Gorman and Brooks, 2009; Hoyler et al., 2014; Petterson et al., 2012; Truglio et al., 2012; Frisch, 2013; Islam, 2014; Majeed, 2015; The Kaiser Family Foundation, 2020).

Theoretically the benefits of primary care physicians are clear, and the concern regarding insufficient physician numbers to provide this care is of immediate concern to policymakers. However, the empirical evidence is somewhat more ambiguous and previous studies do not all align in their findings, an extensive discussion of these findings is presented at the beginning of chapter four. It is not clear whether increasing physician numbers should be a priority for policymakers seeking to improve population health. If primary care physicians can significantly contribute to improving population health, then reducing the primary care physician deficit should be a priority for policymakers. Whereas if primary care physicians are found to have marginal or limited impact on health outcomes, then tackling the physician deficit may not be a cost-effective means of improving health outcomes, and resources may be more efficiently allocated to improve population health and ease the burden on the healthcare system. The mixed evidence is problematic for policymakers seeking to allocate healthcare resources, and without clear evidence it is impossible for policymakers to be able to efficiently allocate those resources. It is vital that the benefits of physicians are well understood. This is especially important given that human resources in healthcare is one of the largest contributors to health budgets in many countries, with an average of 42.2% of all government health expenditure being spent on wages and salaries of healthcare workers (World Health Organisation, 2006). This essay therefore seeks to contribute concrete evidence of the effectiveness of primary care physicians in improving health outcomes of the population, and in doing so, I seek to contribute reliable empirical evidence to inform future policy.

The essay in this chapter investigates the impact of physicians on health outcomes, specifically assessing whether increasing primary care physician's density in a region has a beneficial effect on mortality and hospitalisations. This essay uses a unique setting in which an increase in physician supply was driven by importing foreign doctors to Brazil. The results in this chapter contribute an empirical estimate of the impact of primary care physicians on health outcomes; therefore, providing accurate estimates for a cost-effectiveness analysis to be conducted, and further allowing for an efficient allocation of healthcare resources.

Finally, to conclude, the fifth chapter provides an overview of the findings and contributions of this thesis, and the implications of those results on future policy.

Chapter 2

Redistributive effects of a publicly financed healthcare expansion

2.1 Introduction

Health equity and inequalities in the provision, utilization and financing of healthcare draw significant interest from researchers and policy makers alike. A large body of literature has consistently provided evidence of a socio-economic gradient in health inequalities (Marmot et al., 1991; Kaplan et al., 1996; van Doorslaer et al., 1997; Humphries and van Doorslaer, 2000; Doorslaer et al., 2006), where the lowest parts of the income distribution experience worse health outcomes compared to their higher income counterparts. In addition, there is evidence of strong associations between income and healthcare utilization with the lowest income groups utilizing less healthcare services, despite them experiencing worse health outcomes (O'Donnell and Propper, 1991; van Doorslaer and Wagstaff, 1992; Valdivia, 2002; Doorslaer and Koolman, 2004; Gundgaard, 2006; Dixon et al., 2007), as well as low-income groups being less likely to access healthcare services (Puffer, 1986; van Doorslaer and Wagstaff, 1992; van Doorslaer et al., 2000; Waters, 2000). Furthermore, healthcare expenditures are highly regressive and therefore low-income groups bear disproportionately larger costs of care (Wagstaff et al., 1989; Holahan and Zedlewski, 1992; Rasell et al., 1994; Pannarunothai and Mills, 1997; Galbraith et al., 2005; Ketsche et al., 2011). These disproportionately large health care costs or access constraints can offer explanations for such inequalities, whereby low-income (often uninsured or partially-insured) individuals cannot afford or access needed but expensive heath care and subsequently experience lower health which itself translates to lower income creating a vicious cycle (Deaton, 2003; Auerbach and Kellermann, 2011; Ketsche et al., 2011; Christopher et al., 2018).

The introduction of universal health coverage, or more generally publicly financed healthcare coverage has often been proposed as a way to break this cycle and promote healthy lives (World Health Organization, 2010; U.N. General Assembly, 2013, 2015). Nevertheless, if such policy expansion is to parallelly improve the position of everyone in the system, inequalities will not be mitigated even if there is an increase in absolute population health. If, on the contrary, heterogeneous effects are experienced by different parts of the income distribution, such policies will not only improve healthcare access and health but might also aid in the reduction of income inequalities. Claims by the World Health Organization (WHO, 2019a,b) suggest beneficial impacts on income in favour of low-income individuals, yet much of the evidence is based on population average effects or overall policy impacts of publicly financed healthcare provision.

In principle, public healthcare provision can increase disposable income through, at least, two channels, directly through reduced health care expenditure and indirectly through improved health. Indeed, Card et al. (2008), Finkelstein and McKnight (2008) and Finkelstein et al. (2012) found the expansion of publicly financed health insurance lead to an increase health care utilization (including primary and preventive care as well as hospitalizations), reduce out-of-pocket medical expenditures and medical debt (including fewer bills sent to collection), with similar findings also reported by the RAND health experiment (Manning et al., 1987). At the same time, since Grossman's model of health production, the positive effect of health on income has been established in numerous empirical studies (Luft, 1975; Bartel and Taubman, 1979; Chirikos and Nestel, 1985; Contoyannis and Rice, 2001).

Nevertheless, empirical evidence of heterogeneous effects of publicly financed insurance across the income distribution is limited. The fact that out-of-pocket payments are larger for the lowest income group (Wagstaff et al., 1989; Holahan and Zedlewski, 1992; Rasell et al., 1994; Pannarunothai and Mills, 1997; Galbraith et al., 2005; Ketsche et al., 2011) would suggest that the left tail of the income distribution would experience a disproportionately larger effect. Similarly, the significant socio-economic gradient of health would again suggest a larger relative benefit to be obtained by those less well-off (Wilkinson, 1997), especially in countries with pronounced inequalities such as the U.S. and the U.K. (Wagstaff and van Doorslaer, 1992; Wagstaff et al., 1999; van Doorslaer et al., 1997).

From a theoretical perspective, as health stock increases the marginal benefit of health decreases, where the lowest income groups are expected to have the largest marginal benefit from healthcare (Grossman, 1972). The left-hand side of the income distribution have worse health than those on the right hand side of the distribution (Marmot et al., 1991; Kaplan et al., 1996; van Doorslaer et al., 1997; Humphries and van Doorslaer, 2000; Doorslaer et al., 2006). Theoretically therefore, an expansion of publicly financed health insurance has the potential to disproportionately increase the

income of those on the left-hand side of the distribution, if health limits the ability to work in these groups. The increase in health stock promotes an increase in hours worked in these individuals, which is one of the foundational conclusions of the Grossman model. We might not observe increases in earned income for those on the right-hand side of the distribution because they already have high health which allows them to supply more labour than those with worse health. Such a position is somewhat supported by Manning et al. (1987) who reported an increase in health of those that are both of low health and low income, while there was little evidence to suggest that either low income or low health groups saw an increase in health as a result of insurance experiment. In our theoretical setting, we expect those individuals that benefited from the RAND experiment, in terms of improved health, to have increased their labour supply as a result (Manning et al., 1987). The heterogeneous effect of coverage expansion on health outcomes by income groups can therefore theoretically translate to an heterogeneous increase in earnings in favour of the lowest income groups. That said, Baicker et al. (2014) found no evidence of an effect of Medicaid coverage on overall individual earnings, although the average treatment effects estimated in their study might be masking heterogeneous treatment effects.

The significant literature on inequalities in the financing of healthcare also provides evidence that an expansion of publicly financed health insurance would reduce inequalities. Previous analysis has shown that healthcare costs are highly regressive (Wagstaff et al., 1989; Wagstaff and van Doorslaer, 1992; Rasell et al., 1994; Wagstaff et al., 1999), and this suggests that reducing the out-of-pocket medical costs would have disproportionate benefits in their favour. In the United States context Wagstaff et al. found that those at the lower end of the income distribution pay more for healthcare, relative to their wealth, than those at the higher end of the distribution. They found that this unequal payment for healthcare was greater than both the UK and the Netherlands. Compared to a number of other developed countries, the US has the most regressive system for financing healthcare in terms of total payments, and this seems to be mostly driven by private expenditures, more specifically out-of-pocket expenditures (Wagstaff and van Doorslaer, 1992; Wagstaff et al., 1999). Analysis by van Doorslaer et al. (1999) found that, in terms of healthcare expenditures, overall the US system redistributes away from low-income individuals to higher-income individuals, and does so much more severely than any other of the 12 OECD countries they analyse. Further, although spending on healthcare for the highest income decile in the US is 60% higher than the lowest decile (which is potentially problematic in itself, given that low-income groups have a higher need for healthcare (Valdivia, 2002; Gundgaard, 2006; Dixon et al., 2007)), the lowest decile pay about 20% of disposable income to finance their healthcare, whereas this figure is only 8% for the highest (Holahan and Zedlewski, 1992). This large difference can be explained by the ways in which low-income individuals usually pay for their healthcare, which is typically out-of-pocket, rather than employer insurance

contributions that usally pay the health expenditure of those at the higher end of the distribution. Rasell et al. (1994) showed that out-of-pocket spending was 8.5 times larger for low-income families than for high-income families, and more recent work by Galbraith et al. (2005) found families under the federal poverty line (FPL) spend about $ 120 per $1,000 of income on healthcare, whereas those with income above 400% of the FDL spent just $ 38 per $1,000 of income. Ketsche et al. (2011) find that in the lowest income quintile, 20% of family income was spent on healthcare expenditures (either paying privately or through tax contributions), whereas this figure was at most 16% in the other income categories. This burden has also been shown to have increased across the period of 2001 to 2009, with the annual average growth being 2.7% for premiums and out-of-pocket expenditures over that period (Blumberg et al., 2014). In addition, the share of health insurance units with a healthcare expenses being greater than 30% of their income rose from 6.2% in 2001 to 8.2% in 2009. These disparities suggest that a publicly financed health insurance coverage would reduce the burden of healthcare financing disproportionately for low-income groups. Indeed, Galbraith et al. (2005) analyse the impact of a public healthcare coverage on financial burden of low-income families, and estimated a decrease in the financial burden of 785 percent when compared to private coverage over 1 year. This beneficial impact on financial strain from publicly financed health insurance has also been estimated by Finkelstein et al..

This mechanism likely extends beyond the individual-level, with high healthcare costs also reducing income at the household-level. Prior to the health insurance expansion, low-income households without the means to send unhealthy household members to receive healthcare may have substituted formal for informal care to reduce out-of-pocket costs (Golberstein et al., 2009; Weaver and Weaver, 2014). This then reduces household income because households reallocate resources away from labour supply to provide informal care. If these households were covered by publicly financed health insurance, the reduction in formal healthcare costs would induce household to utilise, otherwise costly, formal care, rather than providing this care informally.

For this theoretical channel to be observed in practice, it would need to be the case that informal care and formal care are substitutes. There is a substantial literature documenting this substitution between informal and formal care (Clark, 2002; Bonsang, 2009; Bolin et al., 2008; Pickard, 2012; Mommaerts, 2018), and the evidence suggests that the availability of informal care reduces length of hospitalisations, specifically for long term hospital stays (Weaver and Weaver, 2014). The presence of a caregiver at home has also been found to "decreased service utilization, emergency room use, hospital utilization, home health, inpatient expenditures, inpatient and short-term nursing home stays (Carmichael and Charles, 2003; Coe et al., 2019; Carmichael et al., 2010; Condelius et al., 2010; Dorin et al., 2014; Van Houtven and

Norton, 2004; Yoo et al., 2004), and decreased minutes of publicly funded care (Hayward et al., 2004)" (Friedman et al., 2019). The increased cost of healthcare coverage causing an increase in informal care is intuitive, and this literature on the substitution between formal and informal care suggest such an effect, however, there is also empirical evidence to support such a result. Golberstein et al. (2009) analyse the Medicare home health care programme and find that payment caps increased the cost of formal care, therefore increasing the amount of care provided informally, specifically in low-income households. Of course, this disproportional effect is not unexpected, as low-income individuals are the most likely to go without healthcare (Valdivia, 2002; Gundgaard, 2006; Dixon et al., 2007) due to the disproportionately higher relative costs they face (Wagstaff et al., 1989; Holahan and Zedlewski, 1992; Rasell et al., 1994; Pannarunothai and Mills, 1997; Galbraith et al., 2005; Ketsche et al., 2011) and therefore, at a household level, a health insurance expansion would likely have a greater benefit for low-income households. Finally, our proposed theoretical channel also assumes that a decrease in informal care would result in changes to labour supply, and there is also a wide literature which finds a negative relationship between informal care provision and labour supply (albeit mostly for older adults) (Mentzakis et al., 2009; Michaud et al., 2010; Van Houtven et al., 2013; Nguyen and Connelly, 2014; Carmichael and Ercolani, 2016; Schmitz and Westphal, 2017; Houtven et al., 2019; Mazzotta et al., 2020; Mozhaeva, 2021; Hollingsworth et al., 2021). Most relevantly however, Kim and Lim (2015) analyse a formal care subsidy programme in South Korea, they find that the substitution from informal care is heterogeneous by physical function level, with the highest substitution to formal care for the least physically able. They claim that for those carers that care for the least able, a formal care subsidy "policy may lead to increased labor supply of individuals caring for this population."

The informal care theoretical channel may also be further compounded as individuals that receive formal care, instead of informal care, see improvements in the quality of their care, subsequently improving their health and reducing the need for further care. There is some literature in support of this theory, several papers estimate that household labour supply decreases as a result of ill-health in one household member (García-Gómez et al., 2013; Jeon and Pohl, 2017; Riekhoff and Vaalavuo, 2021). However, many of these papers find that this decrease in household labour supply is driven by female spouse decreasing labour supply as a result of their male partner becoming ill (Coile, 2004; Siegel, 2006). This decrease in labour supply is likely due to household member now needing to provide care to the sick household member, which is related to the channel we discuss above. It would seem reasonable, given these results and the literature on formal and informal care, that lowering the costs of healthcare would result in less dramatic decreases in household labour supply. Indeed, Fadlon and Nielsen (2021) analyse non-fatal health shocks in Denmark, where health insurance coverage is relatively comprehensive, and conclude that a non-fatal

heart attack or stroke has no significant impact on spouse's labour supply. Given the generous publicly funded healthcare coverage in Denmark, one may conclude that spouses' do not need to reduce labour supply to provide informal care, as this care is provided publicly at a low cost to the household. Furthermore, Braakmann (2014) analyse the impact a disability has on own and spouses labour supply using data from Germany, and again find no evidence that spouses labour supply changes as a result of a disability. Once again, the context is important as Germany has a universal healthcare system, which may mean there is no requirement for households to trade costly formal care for informal care. Indeed, Stabile et al. (2006) find that a more generous public home care programme was associated with a decrease in informal care provided and also an increase in self-reported health, although the focus of this study was on older adults. These papers provide some evidence in favour of publicly financed health insurance increasing labour supply at the household-level. Given that low income households are more likely to experience this ill-health, it is possible that an expansion of publicly financed health insurance would disproportionately impact those at the lower end of the income distribution. We, therefore, expect that analysing the impact of a health insurance expansion at the household-level will yield substantial redistributive effects, which magnify the individual-level effects, as a result of the household resource allocation pathway.

This paper estimates the heterogeneous effects of a publicly financed healthcare expansion on income. In particular, we revisit the Oregon Health Insurance Experiment (OHIE) and exploit its experimental design to estimate its impact on household disposable income. We separately examine the effects on household earned income and healthcare expenditure to offer insight into possible causal pathways.[1]

Several papers have previously analysed the OHIE and have investigated a wide range of outcomes, including health care utilisation, medical debt, emergency department use, dental care, prescription use, various medical outcomes, labour supply outcomes, and voting participation (Finkelstein et al., 2012; Baicker et al., 2013, 2014; Taubman et al., 2014; Finkelstein et al., 2016; Baicker et al., 2017, 2018a,b; Baicker and Finkelstein, 2019; Sacarny et al., 2020).

The two most closely related papers to ours are Finkelstein et al. (2012) who analyse, among other outcomes, medical expenditure, and Baicker et al. (2014) who analyse labour supply. Both papers exploit the experimental design of the OHIE, where individuals were given the opportunity to apply for Medicaid if they were selected by the lottery. Not all lottery selected individuals eventually received Medicaid coverage and therefore lottery selection is used as an instrument for Medicaid coverage to estimate the impact on individual outcomes. Our paper follows a similar approach to

[1] Our contribution to the discussion of redistributive effects of healthcare provision is specifically channelled through household finances that exist in parallel to effects of public provision of private goods (Besley and Coate, 1991; Glomm and Ravikumar, 1992; Bhattacharjee et al., 2017) and general equilibrium effects are outside the scope of this paper.

these previous papers, and further details of the analysis are presented in Sections 2.2 and 2.3. These papers present causal impacts of Medicaid coverage using this approach, but additionally include intention-to-treat estimates, which use the lottery selection in a reduced form specification, which the authors claim to be akin to policy effect estimates. Finkelstein et al. analyse a wide range of outcomes, including healthcare utilisation, self-reported health, preventative treatment, and importantly for this paper, financial strain and healthcare expenditures. Finkelstein et al. estimates a significant fall in whether the individual owes money for medical expenses (18%), whether borrowed money to pay for medical bills (15.4%), and whether has been refused medical treatment because of medical debt (3.6%), all within the past six months. They also estimate there to be a significant fall of 20% in the probability of any medical expenses in the previous six months as a result of receiving Medicaid coverage. Finkelstein et al. also estimate a reduced form quantile regression on out-of-pocket medical expenditures, but do not estimate causal quantile estimates in their paper. However, they find that the reduced form specification shows that the magnitude of impact of the policy on out-of-pocket expenditures was increasing across the distribution, which suggests an unequal impact in favour of low-income groups.

Baicker et al. focus on the impact of Medicaid coverage on labour supply outcomes and benefits received. Baicker et al. estimate there to be no significant change in any earnings, amount of earnings, and earnings above the federal poverty line, in both their intention-to-treat and instrumental variable estimates. They do, however, find an increase in amount of food stamps received, but no evidence of an increase in benefits from any other source. Baicker et al. additionally do not estimate any heterogeneity in the effect on earnings from Medicaid coverage.

This paper adds to the contributions of these papers in two important ways. Firstly, this papers main aim is to analyse the redistributive impacts of Medicaid coverage, to do so we estimate IV-quantile regressions on disposable income (income net of medical expenditure), as well as separately for income and out-of-pocket medical expenditure. The previous papers analysing the OHIE have not estimated the heterogenous effects on income, and although Finkelstein et al. previously estimated quantile regressions on out-of-pocket payments, we estimate the equivalent outcome with IV-quantile regressions to provide causal estimates of Medicaid coverage. In addition, we analyse the impact of Medicaid at the household level rather than the individual level. We believe that household-level reallocation of resources, that we discuss above, is an important pathway when analysing the redistributive effects. Given that the OHIE extended coverage to all adults in the household if they won the lottery, we are able to conduct this analysis at the household level, and make a claim regarding the impacts at a higher level of aggregation. When analysed at the

household-level, we may, in fact, find that the income effects are even larger at the household level, given the informal care channel discussed above.

Briefly, we find significant heterogeneous effects on household disposable income with estimates offering tentative evidence of a secondary redistributive pathway specific to healthcare provision. We acknowledge that results are based on a low-income sub-sample, which nevertheless contains a large portion of the income distribution in Oregon, suggesting that findings offer lower bound estimates for the presented effects and the redistributive power of the coverage expansion mechanism.

The rest of the paper is organised as follows: Section 2.2 presents the experimental set-up for the Oregon experiment. Section 2.3.1 discusses the data and the features that guide some of our econometric specifications, Section 2.3.2 develops the empirical strategy and econometric models to be estimated, Section 2.4 presents the results and sensitivity checks, while Section 2.5 discuss the findings and concludes.

2.2 Background to the Oregon Health Insurance Experiment

In 2008, the Oregon Health Plan (OHP) consisted of two programs, namely OHP Standard and OHP Plus. The OHP Plus provided coverage for those eligible for Medicaid, i.e. children and pregnant women, the disabled, and families who received TANF (Temporary Assistance to Needy Families). OHP Standard was a state expansion of the traditional Medicaid program that covered (with no enrollee cost-sharing) low-income adults not categorically eligible for OHP Plus that met certain criteria, namely: adults aged 19-64 not otherwise eligible for public insurance who are Oregon residents, are U.S. citizens or legal immigrants, have been without health insurance for six months, have income below the federal poverty level (FPL), and have assets below $2,000 (Finkelstein et al., 2012). Due to budget constraints OHP Standard had been closed to new enrollees since 2004 but natural attrition of enrollees had meant that in 2008 the state had sufficient funds to reopen OHP Standard to new entrants. As the state budget would be unlikely able to cover the predicted demand for new OHP enrolments, a lottery would take place to decide who would receive OHP Standard coverage.

The lottery was opened for five weeks in total between the end of January and the end of February of 2008. The signup process was designed to be simple so that the barrier to signing up for the lottery was small (only some demographic information was requested at sign-up), and there was a large publicly funded campaign to ensure that the population of Oregon were aware of the lottery. Although entrance in the lottery was done individually, success in the lottery meant that the winner, as well as every member of their household, could apply for OHP Standard. There was no eligibility criteria for applying to the lottery, however being selected in the lottery allowed the

household to apply for OHP Standard, and at this stage the households eligibility for OHP Standard would be checked. As a result, approximately 30% of those selected by the lottery did not eventually receive OHP Standard as they were not eligible (Finkelstein et al., 2012). Given this eligibility criteria, the OHIE data, discussed in more detail in Section 2.3, is not representative of Oregon as a whole, but rather has a disproportionately high number of low-income households, and indeed was designed to be so.

Originally there were 100,600 applicants for the lottery, however, once duplicates, deceased, and ineligible applicants were removed there were 74,992 individuals in the sample. Between March and September of 2008, eight random lotteries were drawn to select 29,589 individuals (from 24,912 households) that had been successfully chosen in the lottery. Once individuals received notification of lottery selection there was a short period in which they were able to apply for OHP Standard. Hence, success in the lottery did not guarantee that a household would be enrolled on OHP Standard. Only about 30% of the individuals selected in the lottery successfully enrolled onto OHP Standard. Of those that did not successfully enrol, about half did not send the application back in time, and the other half did not meet the eligibility criteria at the time of enrolment (mostly income criteria). Once enrolled, individuals remained on OHP Standard indefinitely, unless they failed to resubmit the required paperwork or moved outside of the eligibility criteria (Finkelstein et al., 2012).

Figure 2.1 shows the household income distribution for Oregon households in 2008 and the income eligibility requirement thresholds. The broken lines show the federal poverty lines for 2008 and different household sizes, with green (dotted) line for one person households ($10,400), blue (short dashed) line for four person households ($21,200), and finally the red (long dashed) line for six person households ($28,400) (Allen et al., 2010; Department of Health and Human Services, 2015). Although coverage expansion and its effects were focused on low-income individuals, the histogram potentially suggests a significant mass for which this policy would be relevant.

2.3 Data and estimation

2.3.1 Data

The analysis uses the OHIE public data[2]. The OHIE data was collected specifically for the purpose of analysing the impacts of Medicaid coverage on a variety of outcomes, and there is a substantial body of previous work using this data (Finkelstein et al., 2012; Baicker et al., 2013, 2014; Taubman et al., 2014; Finkelstein et al., 2016; Baicker

[2]Full detail of the public use datasets is available at `http://www.nber.org/oregon/data.html`.

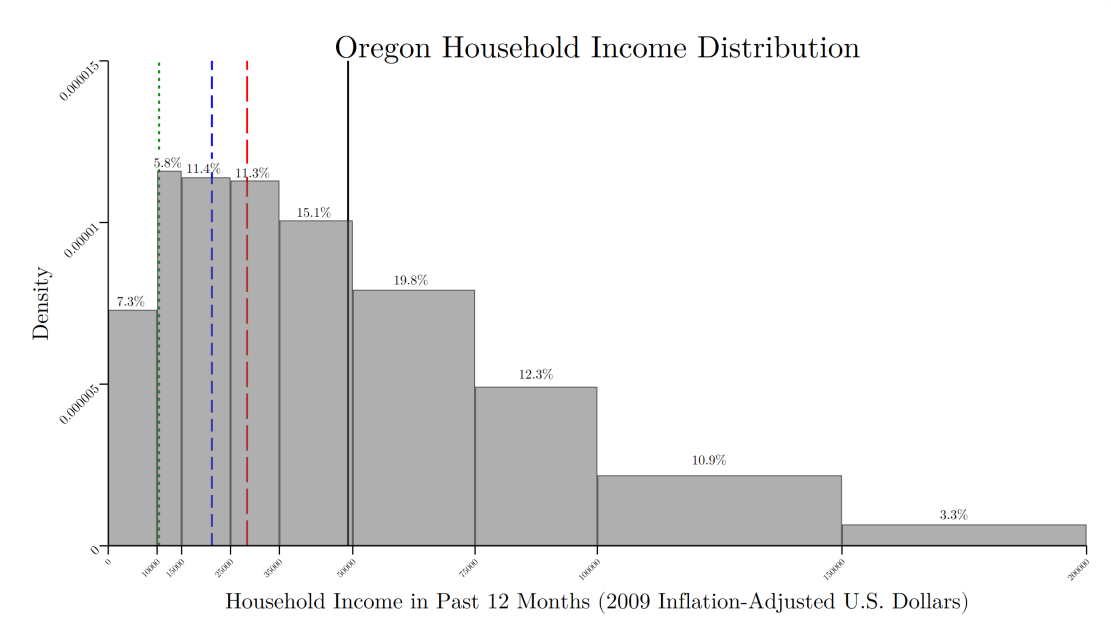


FIGURE 2.1: Distribution of household income for Oregon in 2009

Note: Income categories shown in the histogram are: less than $10,000, $10,000 to $14,999, $15,000 to $24,999, $25,000 to $34,999, $35,000 to $49,999, $50,000 to $74,999, $75,000 to $99,999, $100,000 to $149,999 and finally $150,000 to $199,999. There is an omitted category which is $200,000 or more, the mass for this category is 2.9%. Green (dotted) line: 2008 federal poverty line for households of a single person. Blue (short dashed) line: 2008 federal poverty line for households of four people. Red (long dashed) line: 2008 federal poverty line for households of six people. Black (solid) line: Median income. Data presented in this figure are estimates, margins of error are not presented but are available from data source. This data is sourced from the United States Census Bureau: American Community Survey 5-Year Estimates 2005-2009 and Department of Health and Human Services.

et al., 2017, 2018a,b; Baicker and Finkelstein, 2019; Sacarny et al., 2020). The OHIE data is comprised of 74,922 individuals (66,385 households) that applied to the lottery, initially there were 100,600 lottery applicants, however the data collectors removed 25,608 records from the analysis sample because of duplicate entries, ineligibility for the lottery, having died pre-lottery, and because they would not have been eligible for OHP Standard if they were selected in the lottery.

The OHIE data was collected from a variety of sources. A list of individuals that signed up to the lottery and the data they provided upon signup, Medicaid enrolment and application, mortality, and benefits data were provided by state departments[3]. The data also includes hospital discharge data from the Oregon Association of Hospitals and Health Systems, and credit report data from the TransUnion's Consumer Credit Database.

[3]These include: Oregon's Department of Human Services Division of Medical Assistance Programs, Oregon's Office of Health Policy and Research and Oregon's Department of Human Services Children, Adults and Families Division, Oregon's Center of Health Statistics.

However, the main part of the OHIE data, and the part which much of the analysis in this paper uses, is the mail survey. Three mail surveys were conducted, an initial mail survey and a 12 month survey which were sent to the same sample of 58,405 (29,589 treated, 28,816 controls) individuals, of which there were 26,423 individual responses for the initial mail sample, and 23,741 in the 12 month survey. There was additionally a six month survey which was sent to a sub-sample of 11,756 individuals, of which there were 5,411 responses. Once we aggregate to the household-level (discussed further below) and remove missing values we have an estimating sample of 18,653 households, of which we have 9,094 treated households, and 9,559 control households.

The mail survey was designed to be thorough in eliciting responses. The mail surveys initially consisted of a screener postcard, and the three mail surveys. If the mail surveys were un-deliverable, then there were several attempts to find a more up-to-date address from the post office, LexisNexis people search, and Cascade Direct change of address database, or by phoning the participants. Responders were given $5 for returning the survey, and entered into a lottery to win $200. In addition, a more intensive procedure was conducted on a sub-sample of 30% of those that did not respond initially. This more intensive follow-up included several phone calls, and were mailed an additional two times. If surveys were un-deliverable in this sub-sample, then attempts were made to find an up-to-date address using Google, whitepages.com, MySpace, Facebook, commercial databases, court documents and marriage licenses. This process of eliciting responses was reasonably successful in achieving a high proportion of responses, but selection into response could be a source of bias in our estimates. A discussion of the balance between treatment and control groups below provides evidence that disproportionate selection into response not being being a major concern in this data.

Notification of treatment status was done prior to the data collection and therefore the mail survey provides information on post-treatment impacts. A variety of variables are collected in the surveys, including health and healthcare utilization, as well as, demographic and socio-economic characteristics. Because eligibility was, partly, determined by household income, income data in the OHP Standard was collected through the question "What was your gross household income (before taxes and deductions are taken out) for the last year? Please include any cash assistance or unemployment you may have received."[4] Responses are given in 22 discrete categories with the first being $0 and from $1 onwards, in intervals of $2,500, up until $50,000. Conversion to a continuous measure for the analysis takes the midpoint of

[4]Given the eligibility criteria, and income being self-reported, there may be some concern that income might be strategically misreported. If income was misreported, then this would invalidate your identification strategy. We explore the potential that income was misreported in the appendix Section 2.A.1. We don't find any evidence of strategic misreporting based on this analysis.

each discrete category to be household income [5]. Health care expenditure are elicited for the last six-months through the question "In the last 6 months, how much money did you spend on medical care for yourself? Include anything you paid for your health care", with responses on a an open-ended format of dollar amounts.

Given the level of aggregation of income and the objective to estimate the policy effects of OHP standard coverage on disposable income, the analysis is carried out at the household level. To this effect, individual-level continuous variables are aggregated to the household level, while for categorical variables the share of household members in each variable category is computed. Disposable income is constructed by dividing household income by two to get its six-month equivalent, summing up individual healthcare expenditure across all household members to match the aggregation level of household income and finally subtracting household expenditure from household income.

Finkelstein et al. (2012) provide extensive results on the balance between treatment and control units based on observable characteristics, and, on the whole, show that the treatment and control groups are well balanced. We additionally show our own evidence of treatment control balance, at the household level, in Table 2.2. However, one can see from our Table 2.2, and Finkelstein et al. also show that there are disproportionately more larger households in the treatment group, and the groups are not balanced along this dimension. This is due to large households being able to increase their probability of being treated. Larger households could increase their probability because although randomisation happened at the individual level, if selected all household members were given the chance to apply for OHP Standard, therefore by ensuring that other members of their household also sign up there was higher chance of receiving OHP Standard. Although this was expected *a priori* and controls were purposefully over-sampled for this reason, the take-up was lower than expected and therefore when randomising the OHIE "ran-out" of larger households to use as controls. For this reason, we follow Finkelstein et al. and control for multiple lottery sign-ups in the household, as well as number of household members throughout our analysis to handle this issue.

Descriptive statistics of the estimation sample are presented in table 2.1. Table 2.3 presents means and 95% confidence intervals for households selected in the lottery and those not selected in the lottery. Table 2.4 shows the proportion of household with positive expenditure, and the mean expenditure for those with positive expenditures, for a number of healthcare services. Figure 2.2 shows the income distribution of earned income in our sample.

[5] Under the assumption that individual characteristics are not correlated with the precise position of an individual within the interval, the linearisation of income does not introduce any bias in the models. Nevertheless, the additional noise will likely inflate standard errors compared to the case of a precisely measured income but such variable is not available in the published dataset.

A notable feature of Figure 2.2 is that there is a very high proportion of the sample who report zero income, or very low income. Firstly, as mentioned, the OHIE was designed specifically such that those eligible for Medicaid coverage were from low-income households, therefore we would expect there to be a high proportion of low-income individuals in our sample. The high proportion of households reporting zero income may be particularly surprising however. Previous analysis of the OHIE have also reported a high proportion of zero-income individuals in their data, with the proportion of individuals reporting "any earnings" being just 0.55, and therefore this mass of zero reported income is not unique to our data (Finkelstein et al., 2012; Baicker et al., 2014). Further, the American Community Survey 1-year Estimates (albeit for the year 2010) (United States Census Bureau, 2010) show that approximately 76% of Oregon household have any earnings. Therefore this high proportion of zero-income households is a feature of the income distribution of Oregon, and the United States more generally, not unique to our data. The high proportion of households that report no earnings in our sample is therefore not particularly unexpected, and if anything, we would expect that zero-income household are more disproportionately represented in our data given the low-income criteria of the programme.

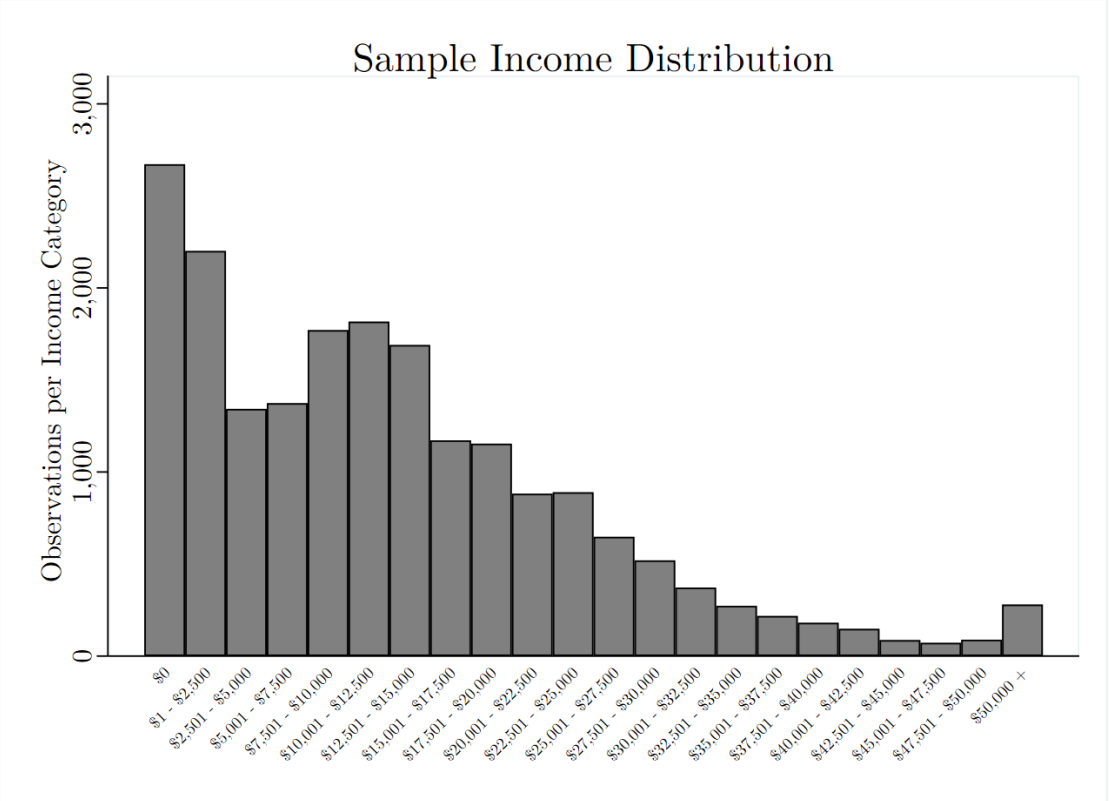


FIGURE 2.2: Distribution of earned income in sample

Note: The left most category are households with zero income, the second is $1 onwards to $2,500 and then in intervals of $2,500, up until $50,000. The final category is greater than $50,000.

Variable	Mean	Standard Deviation
Lottery Winner	0.49	0.50
Medicaid Recipient	0.17	0.37
Sex		
Proportion of Females in Household	0.61	0.45
Age		
19-24 year olds	0.07	0.26
25-34 year olds	0.21	0.39
35-44 year olds	0.19	0.38
45-54 year olds	0.28	0.43
55-65 year olds	0.21	0.40
Race		
Proportion of White household members	0.84	0.36
Proportion of Black household members	0.03	0.18
Proportion of Hispanic household members	0.11	0.31
Proportion of Other ethnic origin household members	0.18	0.38
Household Composition		
Number of Adults in each household	2.10	1.10
Number of Children (under 19 years old) in each household	0.83	1.24
Earned Income		
Proportion of Households with positive income	0.87	0.33
Income of households with a positive earned income	13,265.75	11,718.98
Employment		
Employed	0.49	0.50
Working under 20 hours	0.10	0.30
Working between 20 and 30 hours	0.11	0.32
Working over 30 hours	0.29	0.45
Observations	18,653	

TABLE 2.1: Descriptive Statistics of the Estimating Sample

2.3.2 Estimation

The analysis exploits the unique design of the OHIE to examine heterogeneity of effects of OHP Standard coverage in both intent-to-treat and instrumental variable estimations for the extensive and intensive margin on three outcomes measured at the household level, namely, household disposable income (i.e. household income net of health expenditures), and subsequently on its constituent parts, i.e. healthcare expenditure and household earned income.

Through the lottery mechanism, lottery selection can be interpreted as the intent-to-treat effect, describing the impact of the policy change on the overall population including those that did not eventually receive OHP Standard after being selected by the lottery. At the same time, given that those not receiving coverage, despite being selected by the lottery, are likely to share common unobservable

Variable	Selected by lottery	Not selected by lottery	T-test p-values
Sex			
Proportion of Females in Household	0.608	0.617	0.1566
Age			
19-24 year olds	0.074	0.075	0.8080
25-34 year olds	0.208	0.209	0.8386
35-44 year olds	0.196	0.192	0.4867
45-54 year olds	0.269	0.283	0.0216
55-65 year olds	0.210	0.208	0.7562
Race			
Proportion of White household members	0.836	0.845	0.0828
Proportion of Black household members	0.030	0.036	0.0234
Proportion of Hispanic household members	0.110	0.103	0.1102
Proportion of Other ethnic origin household members	0.186	0.182	0.3815
Household Composition			
Number of Adults in each household	2.120	2.089	0.0511
Number of Children (under 19 years old) in each household	0.864	0.803	0.0007
Earned Income			
Proportion of Households with positive income	0.880	0.870	0.0589
Income of households with a positive earned income	13,658.228	12,894.873	0.0000
Employment			
Employed	0.509	0.479	0.0000
Working under 20 hours	0.106	0.099	0.0989
Working between 20 and 30 hours	0.114	0.115	0.9025
Working over 30 hours	0.307	0.278	0.0000
Observations	9,094	9,559	

TABLE 2.2: T-test of the difference in means of lottery winners and those that were not selected in the lottery in the Estimating Sample

characteristics, lottery selection can be used as an instrument for OHP Standard coverage (Finkelstein et al., 2012; Baicker et al., 2014).[6]

2.3.3 Intent-to-treat

The extensive margin for the intent-to-treat (ITT) specification is described as a binary probit model:

$$P(Y_{j,i} > 0) = \Phi\left(\beta_{1,j} Lottery\ Selection_i + \beta_{2,j} X_i\right) \tag{2.1}$$

where $j = 1, .., 3$ denotes the three possible outcomes in the analysis, i.e. disposable income, healthcare expenditure or earned income, for household i. $Lottery\ Selection_i$ is a binary variable equal to 1 if the household was selected in the lottery to receive OHP Standard. X_i denotes a range of household level characteristics that the regression controls for, including average household age, proportion of females in household, proportion of white household members, proportion of black household members, proportion of Hispanic household members.

[6] Although a pre-treatment survey is available for lottery sign ups, information on household income and health expenditures was not collected. Instead, lottery sign-up specific variables are included, such as date of sign-up, number of people in the household on the lottery list and whether application was approved. A difference-in-differences estimation strategy for causal estimation is therefore not possible.

Variable	Selected by lottery	Not selected by lottery
Sex		
Proportion of Females in Household	0.61	0.62
	[0.5988,0.6171]	[0.6082,0.6264]
Age		
19-24 year olds	0.07	0.07
	[0.0688,0.0793]	[0.0697,0.0801]
25-34 year olds	0.21	0.21
	[0.2002,0.2162]	[0.2014,0.2173]
35-44 year olds	0.20	0.19
	[0.1884,0.2040]	[0.1848,0.2000]
45-54 year olds	0.27	0.28
	[0.2599,0.2774]	[0.2744,0.2920]
55-65 year olds	0.21	0.21
	[0.2018,0.2181]	[0.2001,0.2161]
Race		
Proportion of White household members	0.84	0.85
	[0.8288,0.8438]	[0.8383,0.8526]
Proportion of Black household members	0.03	0.04
	[0.0269,0.0339]	[0.0326,0.0400]
Proportion of Hispanic household members	0.11	0.10
	[0.1039,0.1167]	[0.0971,0.1092]
Proportion of Other ethnic origin household members	0.19	0.18
	[0.1786,0.1944]	[0.1740,0.1892]
Household Composition		
Number of Adults in each household	2.12	2.09
	[2.0972,2.1432]	[2.0672,2.1104]
Number of Children (under 19 years old) in each household	0.86	0.80
	[0.8382,0.8896]	[0.7783,0.8273]
Earned Income		
Proportion of Households with positive income	0.88	0.87
	[0.8727,0.8864]	[0.8633,0.8770]
Income of households with a positive earned income	13,658.23	12,894.87
	[13408.7673,13907.6884]	[12658.4441,13131.3014]
Employment		
Employed	0.51	0.48
	[0.4990,0.5198]	[0.4690,0.4893]
Working under 20 hours	0.11	0.10
	[0.0996,0.1123]	[0.0926,0.1046]
Working between 20 and 30 hours	0.11	0.11
	[0.1077,0.1208]	[0.1084,0.1212]
Working over 30 hours	0.31	0.28
	[0.2975,0.3165]	[0.2687,0.2867]
Observations	9,094	9,559

TABLE 2.3: Comparison of characteristics of lottery winners and those that were not selected in the lottery (with 95% confidence intervals in parentheses) in the Estimating Sample

We additionally control for household composition, because throughout the analysis we use household-level income which has not been equivalised (i.e. not adjusted for household size). One concern is that when estimating the heterogeneous impacts of Medicaid using household-level outcomes that have not been equivalised, we may be inadvertently estimating heterogeneous impacts by household size indirectly. In other words, we are estimating the heterogeneous effects by household size, because larger households will be disproportionately found in higher income groups by virtue of having more individuals able to earn income. We therefore include household

Variable	Mean	Standard Deviation
Doctor Visits Expenditures		
Proportion of Households with positive expenditure	0.35	0.48
Expenditure of households with a positive health expenditure	730.22	2,620.61
Emergency Room Expenditures		
Proportion of Households with positive expenditure	0.08	0.27
Expenditure of households with a positive health expenditure	2,796.60	6,895.11
Drug Prescriptions Expenditures		
Proportion of Households with positive expenditure	0.39	0.49
Expenditure of households with a positive health expenditure	450.74	1,514.40
Other Healthcare Expenditures		
Proportion of Households with positive expenditure	0.12	0.33
Expenditure of households with a positive health expenditure	914.85	2,469.41
Total Healthcare Expenditure		
Proportion of Households with positive expenditure	0.49	0.50
Expenditure of households with a positive health expenditure	1,547.49	4,738.68
Observations	18,653	

TABLE 2.4: Healthcare Expenditure Statistics of the Estimating Sample

composition (# of adults in the household, # of children (under 18) in the household) as additional control variables to ensure we are estimating the heterogeneity by income, not household size indirectly [7].

Although sign up to the lottery was done at the individual level the entire household was treated with the opportunity to receive OHP Standard. This means that household members can increase their probability of receiving treatment by ensuring that other members of their household also sign up. As such, probability of treatment, is therefore, conditional on the number of household members that applied. To account for endogeneity, in addition to household size, X_i also includes a control for multiple lottery sign-ups.

Conditional on a positive outcome, the intensive margin for ITT is specified as a quantile regression, where quantile regression coefficients are estimated by minimizing the following objective function:

$$Q_\tau(Y_{j,i}|Y_{j,i} > 0) = \sum_{i:Y_{j,i} \geq W_i\beta_{j,\tau}}^{N} q[Y_{j,i} - W_i\beta_{j,\tau}] + \sum_{i:Y_{j,i} < W_i\beta_{j,\tau}}^{N} (1-q)[Y_{j,i} - W_i\beta_{j,\tau}] \quad (2.2)$$

where τ is the quantile to be estimated and $W_i = (Lottery\ Selection_i, X_i)$.

[7] An alternative approach may be to equivalise household income, however given that we have a categorical variable, doing so would induce further noise into our outcome measure and inflating standard errors, which we note above, and we would prefer to avoid doing this. In appendix Section 2.A.5 we estimate our models using a per-capita measure of household income to show that our results are robust to method of handling this issue.

2.3.4 Instrumental Variable

To estimate the causal impact of OHP Standard coverage expansion, instrumental variables are estimated with lottery selection as the exogenous change (i.e instrument) in Medicaid coverage. The ITT estimates are not considered to be causal estimates of the impact of Medicaid coverage because although lottery selection was randomly assigned, not all of those that won the lottery eventually received OHP Standard. As discussed above, approximately 30% of those that won the lottery did not receive Medicaid coverage, about half of those being due to them falling outside the eligibility criteria, and half because they did not submit the relevant paperwork in time. Therefore, the ITT estimates will likely underestimate the true impact.

Using Medicaid coverage directly in a reduced form estimator (i.e. without an instrument) would also likely lead to bias estimates. Given that many individuals did not receive Medicaid coverage after winning the lottery because of them falling outside of the eligibility criteria, there is essentially a selection into treatment by eligibility status. Indeed this is particularly salient in our context, as we are interested in outcomes which are specified explicitly in the eligibility criteria. Our primary outcome of interest is disposable income, and part of the eligibility criteria was having income below a given threshold. Therefore using Medicaid coverage directly would downward bias our estimates as we would be comparing those that did receive Medicaid coverage, and therefore had income below the income threshold, with a group containing households that were rejected from receiving OHP standard because their income was too high. In addition, it may be the case that those that did not submit their paperwork in time share some unobservable characteristics which would also bias the estimates. As a result, we require the use of lottery selection as an instrument for Medicaid coverage, and this is what we do for our IV estimates.

The identifying assumption is that lottery selection had no direct impact on outcomes, but rather only impacted outcomes through Medicaid coverage. This assumption seems plausible in our setting, as the lottery selection was, by construction, random, and there are no other obvious channels in which lottery selection may impact our outcomes, aside from by impacting Medicaid coverage. The only way in which lottery selection may impact our outcomes is through a placebo effect, however we don't anticipate these effects to be substantial, if they exist at all. Therefore, we believe that the IV estimates will provide a causal interpretation of Medicaid coverage on disposable income.

For the extensive margin a consistent two-stage instrumental variable probit estimator is used (Roodman, 2011), while for the intensive margin a two stage residual inclusion (2SRI) methodology is applied [8] as it gives numerically identical estimates as the

[8] To ensure the standard errors produced in the second stage are correct the entire system of equations is bootstrapped and iterated 250 times.

typical two stage least squares approach, but has also been shown to provide consistent estimates for a broad class of non-linear models, like those estimated here (Terza et al., 2008a) and is an approach previously used by Fang et al. (2009)[9]. In both extensive and intensive margins, the first stage is given by:

$$OHP\ Standard_i = \delta_1 Lottery\ Selection_i + \delta_2 X_i + u_i \tag{2.3}$$

where $OHP\ Standard_i$ is a binary variable denoting 1 if the household received OHP standard and 0 if they did not.

The second stage of the extensive margin is subsequently specified as a probit model

$$P(Y_{j,i} > 0) = \Phi\left(\beta_{1,j}\widehat{OHP\ Standard}_i + \beta_{2,j}X_i\right) \tag{2.4}$$

with the second stage of the intensive margin given by

$$Q_\tau(Y_{j,i}|Y_{j,i} > 0) = \sum_{i:Y_{j,i} \geq V_i\beta_{j,\tau}}^{N} q[Y_{j,i} - V_i\beta_{j,\tau}] + \sum_{i:Y_{j,i} < V_i\beta_{j,\tau}}^{N} (1-q)[Y_{j,i} - V_i\beta_{j,\tau}] \tag{2.5}$$

where $V_i = (OHP\ Standard_i, X_i, \widehat{u}_i)$, with $\widehat{u}_i$ being the first stage residual that is included as an additional independent variable.

2.3.5 Identifying effects pathways

Given the objective of healthcare coverage expansion, the reduction in household healthcare expenditure is an expected corollary, even if the heterogeneity of effect across the expenditures distribution is a new insight into how the effect works. The spillover effect, however, of the OHP expansion on earned household income is not as obvious and exploring some of the possible channels that the effect might work through is of interest.

Direct evaluation of such channels is not feasible and as such we rely on indirect tests indicative as the underlying relations. First, we examine whether increased income could be the result of an increase in employment or in the hours worked. Using the lottery mechanism, we causally estimate the effect of the health plan expansion on the probability of household employment, employment type and hours of work. Second, we check whether changes in employment patterns are the result of improvements in household health. Under known evidence, which is also empirically supported in our data, low-income individuals have lower health than the right-hand side of the income distribution. In our sample, the left hand side of the income distribution also have lower proportion of households with an employed household member. An

[9] In appendix Section 2.A.2 we present estimates from more typical estimation procedures for the intensive margin and show that our findings are robust to estimation method.

increase in employment due to OHP Standard coverage would be driven by individuals not working pre-coverage, which in our case are more concentrated in the lower-income categories. Given that those on the left hand side of the distribution are less likely to work and have the highest marginal benefit to healthcare, *a priori* we expect increases in the health of this group to lead to larger changes in employment. Under the assumption that the lottery mechanism has no correlation to health other than through healthcare coverage expansion, we directly instrument health with the lottery mechanism and obtain causal effects of changes in heath (brought about by the expanded health plan) on employment type and hours of work. Our measure of health is a binary measure of whether, on average, the household reported 'Good, Very Good or Excellent' self-assessed health, or whether they, on average, considered their health to be 'Fair or Poor'. This measure is clearly not a precise measure of health, and we loose a lot of nuance from simply using self-reported health and by aggregating to the household level. There may be better ways of assessing overall household health, however, it is still useful as means of measuring general household health, and self-assessed health has been previously been found to have increased due to the Medicaid coverage (Finkelstein et al., 2012) and therefore has been found to pass the IV relevance condition we require. It is also one of the only measures of health which does not have any healthcare utilisation or access components embedded. Therefore, although we acknowledge its limitations as a measure of health, we believe it is still useful as a means of assessing whether health is a mediator in the causal channel between Medicaid coverage and employment.

Finally, we examine, whether changes in earned income are the result of potential household resources being freed-up as a result of coverage expansion. To this effect, we identify households at increased risk of needing intra-household care as those where at least half the household members indicated that "went without prescription medication" because they could not physically get the prescription or could not get to pharmacy. These households are disproportionately represented in the low-income categories, and therefore if freeing up household resources is a means by which income increases, we would expect low-income categories to benefit more from the health insurance expansion. We repeat the earned household income estimations for this restricted sub-sample to analyse whether this is indeed the case.

2.4 Results

2.4.1 ITT estimations

ITT estimates are presented in Tables 2.5 and 2.6. Results show that intent-to-treat had no statistically significant effect (extensive or intensive margin) on disposable or earned income. Yet, significant effects on the probability of positive out-of-pocket

expenditure were found for Doctors, ER visits, medical prescriptions, other expenditure and total expenditure with intention-to-treat decreasing expenditure by 6.8, 1.3, 6.0, 0.1 and 6.5 percentage points (p.p.), respectively. Significant effects on the size of the expenditure (conditional on being positive) was observed only for the 25 th percentile of prescription expenditure, and total expenditure with \$21.9, \$54.7, \$49.4 and \$63.2 drops for the 25^{th}, 40^{th}, 50^{th} and 60^{th} percentiles.

	Household Disposable Income						
	Extensive Margin	Intensive Margin					
		Linear	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
Lottery Selection	0.0057	188.0	199.4	60	251.3	267.3	129.9
	(0.0049)	(172.2)	(173.6)	(103.2)	(254.9)	(245.6)	(241.7)
Observations	18,653	15,375					

Notes: Estimations control for multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. Extensive margins are reported as marginal effects. Robust standard errors in parentheses. *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.5: Intention-to-treat estimates of the effect of lottery selection on household disposable income

2.4.2 IV estimations

Disposable Income

The first row of Table 2.7 shows the causal effect of Medicaid coverage expansion on disposable income, with the first column presenting the findings for the extensive margin, then the linear estimates of the intensive margin and subsequent columns presenting effects for the 25^{th}, 40^{th}, 50^{th}, 60^{th} and 75^{th} percentiles. For the extensive margin, those that received Medicaid coverage as a result of winning the lottery, on average, saw a statistically significant increase in the probability of having a positive disposable income of approximately 4.5 percentage points. Among those with positive disposable income, quantile regressions shows a clear gradient of the effect of Medicaid with an almost monotonic drop of the effect as one moves from the left to the right hand side of the distribution. Those on the far left hand side of the disposable income distribution (i.e. 25^{th} percentile) saw a statistically significant increase in disposable income in the region of \$1,667 and a \$1,243 for the 60^{th} percentile, while those on the right hand side of the distribution (i.e. 50^{th}, 60^{th} and 75^{th} percentile) saw no statistically significant change in disposable income.

Healthcare Expenditure

Panel (a) in Table 2.8 presents the results of the probit, linear and quantile estimates of the effect on out-of-pocket (OOP) healthcare expenditures. For the extensive margins,

(a)	Out-of-Pocket Expenditures						
	Extensive Margin	Intensive Margin					
		Linear	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
Doctors							
Lottery Selection	-0.0683***	85.61	0.0	0.0	-8.362	0.0	0.0
	(0.0069)	(64.49)	(2.550)	(3.069)	(12.02)	(7.798)	(20.15)
Observations	18,653	6,552					
ER							
Lottery Selection	-0.0130***	240.7	0.0	-32.68	17.17	69.02	113.2
	(0.0039)	(376.2)	(40.40)	(57.94)	(76.77)	(128.4)	(242.4)
Observations	18,653	1,431					
RX							
Lottery Selection	-0.0600***	18.28	-9.385***	-4.615	-5.538	-4.631	0.430
	(0.0070)	(38.50)	(3.112)	(4.634)	(6.872)	(9.552)	(17.40)
Observations	18,653	7,269					
Other							
Lottery Selection	-0.0098**	-115.4	-6.012	-10.69	-22.40	-16.78	-2.930
	(0.00477)	(107.2)	(11.49)	(12.11)	(24.02)	(26.34)	(63.05)
Observations	18,653	2,248					
Total							
Lottery Selection	-0.0652***	39.06	-21.93***	-54.66***	-49.29***	-63.19**	-81.53
	(0.0072)	(100.2)	(7.658)	(14.93)	(18.66)	(27.00)	(54.62)
Observations	18,653	9,124					

(b)	Earned Household Income						
	Extensive Margin	Intensive Margin					
		OLS	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
Lottery Selection	0.0006	173.1	79.68	0.0	0.0	310.7	0.0
	(0.00487)	(173.4)	(162.4)	(65.97)	(160.5)	(292.5)	(174.7)
Observations	17,886	15,645					

Notes: Estimations control for multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. Extensive margins are reported as marginal effects. Robust standard errors in parentheses. *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.6: Intention-to-treat estimates of the effect of lottery selection on household out-of-pocket expenditures and household income

IV causal estimates are statistically significant, with Medicaid reducing the probability of any health expenditures by approximately 22 p.p. for Doctor visits, 5.2 p.p. for ER visits, 18.2 p.p. for drug prescriptions, 3.4p.p. for other expenditure and 18.7 p.p. for total OOP expenditure. For the intensive margins result show similar patterns to those found for disposable income above, although significance is present only for total OOP expenditure. There is a clear and significant gradient in the estimated effects of Medicaid coverage suggesting significant heterogeneity in the impact of OHP standard coverage across the expenditure distribution. Those with higher healthcare

	Household Disposable Income						
	Extensive Margin	Intensive Margin					
		Linear	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
OHP Standard	0.0453***	822.6	1,665.5**	1,243.3**	618.2	912.1	-1,141.0
	(0.0147)	(557.4)	(655.0)	(496.8)	(712.5)	(759.0)	(771.3)
1^{st} stage F-test	4,729.40	3,487.04					
Observations	18,433	15,184					

Notes: Estimations control for multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. Extensive margins are reported as marginal effects. Robust standard errors in parentheses. *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.7: Instrumental variable estimates of the effect of the Oregon health plan expansion on household disposable income

expenditures seeing larger decreases in their expenditures as a result of OHP standard coverage. Specifically, those at the 25^{th} percentile have a causal drop in expenditures of $65.8 with a progressively increase in the magnitude as one moves to higher percentiles, namely a drop of $143.5, $215.3 and $263.7 for the 40^{th}, 50^{th} and 60^{th} percentiles, respectively. The linear intensive margin are insignificant for each type of health expenditure.

Household Earned Income

Results for earned income are given in Panel (b) of Table 2.8, where findings very much resemble those for disposable income. As a result of OHP Standard expansion, the probability of reporting positive household income greater increases by approximately 3.8 p.p., while for those with positive income there is again significant heterogeneity in the effect across the income distribution. Moving from left to right of the distribution we find a gradual decrease of the effect of heath care coverage on income, starting at $2,925 for the 25^{th} percentile, dropping to $1,365 for the 50^{th} percentile and turning negative and insignificant (i.e. -$118) for the 75^{th} percentile.

Identifying earned income pathways

Labour Supply

Panel (a) in Table 2.9 shows that OHP Standard coverage causes an increase in the probability that a household member is employed by approximately 4 p.p., an effect that is wholly driven by changes in the probability that a household member becomes an employee. Turning to number of hours worked, we find little significance with an increase in the probability of working more than 30 hours by 3.7 p.p. being significant at the 10% level. As might be expected, in our sample the proportion of households with at least one employed member increases through the income distribution, so it

(a)	Out-of-Pocket Expenditures						
	Extensive Margin	Intensive Margin					
		Linear	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
Doctors							
OHP Standard	-0.218***	171.6	-9.126	-18.54	-60.18	-100.9**	-197.3*
	(0.0194)	(236.6)	(34.92)	(25.14)	(62.70)	(49.05)	(117.4)
1^{st} stage F-test	4,729.40	623.03					
Observations	18,433	6,481					
ER							
OHP Standard	-0.0518***	620.8	21.40	-96.42	58.00	280.0	61.45
	(0.0116)	(1077.8)	(236.1)	(368.0)	(529.7)	(843.8)	(1530.9)
1^{st} stage F-test	4,729.40	119.18					
Observations	18,433	1,413					
RX							
OHP Standard	-0.182***	19.27	-29.69**	-31.21	-36.49	-33.51	-31.59
	(0.0199)	(105.1)	(14.88)	(20.44)	(30.79)	(42.70)	(82.75)
1^{st} stage F-test	4,729.40	907.28					
Observations	18,433	7,187					
Other							
OHP Standard	-0.0342**	-378.8	-19.71	-24.55	-69.78	-108.8	-5.411
	(0.0138)	(296.8)	(46.40)	(45.33)	(75.06)	(106.1)	(285.9)
1^{st} stage F-test	4,729.40	354.70					
Observations	18,433	2,212					
Total							
OHP Standard	-0.187**	-72.55	-65.78**	-143.5**	-215.3***	-263.7**	-393.0
	(0.0202)	(292.1)	(31.89)	(63.11)	(81.06)	(115.8)	(243.2)
1^{st} stage F-test	4,729.40	1261.03					
Observations	18,433	9,020					

(b)	Earned Household Income						
	Extensive Margin	Intensive Margin					
		Linear	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
OHP Standard	0.0379***	805.9	2,925.3***	2,031.4***	1,364.7**	1,233.0	-117.7
	(0.0146)	(562.7)	(639.9)	(534.9)	(646.1)	(938.5)	(850.8)
1^{st} stage F-test	4,729.40	3,346.92					
Observations	18,433	15,449					

Notes: Estimations control for multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. Extensive margins are reported as marginal effects. Robust standard errors in parentheses. *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.8: Instrumental variable estimates of the effect of the Oregon health plan expansion on household out-of-pocket expenditures and household income

might be reasonable to expect that there is a differential impact on employment probability by income category. Therefore, these results do, at least partially, explain the heterogeneous effect in terms of earned income.

These substantial changes in labour supply may be considered to be unexpected over the relatively short time frame we analyse (12 months post coverage). Therefore, we present further analysis of the effect of Medcaid coverage on labour supply in appendix Section 2.A.3. We use the initial survey zero month and the 6 month survey

to show that these effects on labour supply take until the 12 month survey to take effect.

(a)	Effect of OHP Standard on Employment					
	Any Employment	Employee	Self-Employed	Under 30 hours	Over 20 hours	Over 30 hours
OHP Standard	0.0406*	0.0467**	-0.000374	0.0141	-0.0112	0.0368**
	(0.0208)	(0.0202)	(0.0139)	(0.0129)	(0.0135)	(0.0187)
1^{st} stage F-test	4,626.01	4,626.01	4,626.01	4,696.02	4,696.02	4696.02
Observations	18,055	18,055	18,055	18,055	18,055	18,055
(b)	Effect of Good Household Health on Employment					
	Any Employment	Employee	Self-Employed	Under 30 hours	Over 20 hours	Over 30 hours
Good Health	0.293*	0.366**	-0.0386	0.105	-0.0946	0.250
	(0.168)	(0.168)	(0.114)	(0.107)	(0.113)	(0.155)
1^{st} stage F-test	34.54	34.54	34.54	34.21	34.21	34.21
Observations	18,057	18,057	18,057	18,057	18,057	18,057

Notes: In Panel (b) Good Health is a recoded self-assessed health question which takes the value of one if all household members report good/very good/excellent health and the value of zero if at lest one member report fair/poor health. Estimation proceeds by directly instrumenting health through the lottery mechanism of the Oregon health plan expansion. Estimations control for multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. Robust standard errors in parentheses *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.9: Instrumental variable estimates of the effect of the Oregon health plan expansion and of good household health on the probability of household employment

Exploring further what potentially drives such increase in labour supply, Panel (b) of Table 2.9 suggests that some of this effect is channelled through improvements in household health. In other words, improvements in household health, as a result of OHP Standard expansion, almost fully capture the increase a household member being an Employee. It is worth noting that the measure of health is at the household-level, as is employment, therefore these results may not be explained by an individual seeing improvements in their health and therefore increasing their individual labour supply. Indeed, over a short 12-month time span, it is unlikely that if a household is unemployed due to ill-health their health will improve enough to seek employment due to Medicaid coverage. Rather, it is more likely that one household members health improves enough so that the burden of informal care is reduced, but not necessarily eliminated, therefore increasing the labour supply of other household members [10].

As previously, heterogeneity in earned income is possibly the result of heterogeneity in the health distribution across income percentiles, whereby worse health is more pronounced among the less well-off. A χ^2 test confirms significant differences in the earned income distribution between households with and without good health (χ^2 test p.value $= 0.00$).

[10] Work documenting the change in labour supply as a result of changes in household-level health is discussed in detail in the introduction (Coile, 2004; Siegel, 2006; Stabile et al., 2006; García-Gómez et al., 2013; Braakmann, 2014; Jeon and Pohl, 2017; Fadlon and Nielsen, 2021; Riekhoff and Vaalavuo, 2021).

Finally, Table 2.10 focuses on small sub-sample of households identified as being at risk of needing intra-household care. We find that coverage expansion significantly boosts income both at the extensive margin with 18 p.p. more likely to report positive income and the intensive margins with an almost uniform effect across all members of that group irrespective of their level of income. Such effects are also consistent with increased household supply, whereby household members, as a result of OHP can take up employment and increase their household income.

	Extensive Margin	Intensive Margin					
		Linear	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
OHP Standard	0.175**	10,259.1***	12,848.3***	12,889.5***	13,766.0***	10,482.8**	16,928.7***
	(0.0822)	(3,053.331)	(4,204.1)	(3,900.7)	(4,036.6)	(4,912.6)	(4,863.1)
1^{st} stage F-test	183.95	122.65					
Observations	1,317	1,070					

Notes: Households at increased risk of need of intra-household care are identified as those where at least half the household members indicated that "went without prescription medication" because they could not physically get the prescription or could not get to pharmacy. Estimations control for multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. Extensive margins are reported as marginal effects. Robust standard errors in parentheses. *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.10: Instrumental variable estimates of the effect of the Oregon health plan expansion on household disposable income for households at increased risk of need of intra-household care.

2.5 Discussion

Using the Oregon Health Insurance Experiment this paper estimates novel heterogeneous treatment effects for disposable income, earned income and healthcare expenditure. Quantile estimates of the effect of Medicaid coverage on disposable income have a clear gradient across the income distribution, benefiting those at the left hand side of the income distribution. Similarly, we find that those with higher health expenditure also benefit more from coverage, while coverage expansion also leads to increases in household earned income for those less well-off through increases in household employment and improvements in household health. Overall, our findings offer tentative evidence of a secondary redistributive pathway, specific to healthcare provision and policies expanding publicly financed health coverage, which works in parallel to findings of previous research on the public provision of private goods (Besley and Coate, 1991; Glomm and Ravikumar, 1992; Bhattacharjee et al., 2017).

Grossman (1972)'s theory of health and human capital offers the basic theory whereby an expansion of healthcare coverage increases labour supply as health improves due to the new healthcare access. Indeed our results suggest significant positive effect of Medicaid coverage on the probability of at least one household member being employed and offer evidence that such effect is likely channelled though improvements in health. Yet, given that households in the lower end of the income

distribution are less likely to have any employed household members and feature worse health, heterogeneous effects of Medicaid coverage arise across the income distribution. On the other hand the drop in households' healthcare expenditures due to the policy is intuitive and an intended consequence. To this effect, our estimates are in line with expectations and findings in previous studies (Finkelstein et al., 2012; Baicker et al., 2013). However, this paper goes a step further, and identifies a clear gradient in the impact across the expenditure distribution. Once again, the effect for those with the highest healthcare expenditures is of a larger magnitude than those with the lowest healthcare expenditures.

At face value the estimates presented here seem at odds with Baicker et al. (2014). However, difference in the level of analysis (i.e. household rather than individual data) could be driving these differences. Unfortunately, we do not have access to the individual level dataset to replicate the original analysis. Nevertheless, as an approximate test, we split the sample to single- and multi-individual households and estimate a model with the same outcome as Baicker et al., that being "any earnings", and this analysis is included in appendix Section 2.A.4. We confirm that for for single individual households, OHP Standard has no significant effects on employment whereas there are significant employment effects for multi-adult households. Indeed, this also provides additional evidence in favour of our conclusion that Medicaid expansion frees up household resources, rather than individual resources, so that households can increase labour supply.

Tentatively, we offer suggestive evidence of publicly financed health insurance reducing income inequality, which is empirical evidence in line with previous findings of Bhattacharjee et al. (2017). However, as the OHIE relates to a limited low-income sub-sample of the population, we caution against generalizations, albeit if anything effects should be stronger when considering the whole income distribution. In addition, findings are constrained by the local average treatment (LATE) interpretation of the estimated coefficients, which concentrates on the impact of receiving Medicaid for those that the lottery selection enabled them to receive Medicaid coverage. Not only does this interpretation mean we are only able to make a claim regarding those within the eligible range, but in addition we are not able to make a claim regarding those that did not eventually receive coverage after winning the lottery.

When discussing our results in the context of income inequality, it is worth noting that we do not consider general equilibrium effects. Specifically, redistribution arising from the financing of public health coverage is not analysed in our paper and in this context we have no means of doing so. The OHIE was financed through state budget which had become substantial enough to cover a small proportion of the population, and was not financed through an increase in tax revenues or similar. It has been previously found that public provision of private goods has redistributive effects through taxations (Besley and Coate, 1991). This paper serves as evidence of a

secondary redistribution effect of publicly financed health coverage, alongside the tax redistribution.

In conclusion, using the Oregon Health Insurance Experiment we to estimate the heterogeneous treatment effects on disposable income, healthcare expenditures and income. Findings tentatively suggest that publicly financed health insurance has redistributive power by heterogeneously reducing healthcare costs and increasing household income in favour of low income households, and this causal pathway is separate to the previously studied redistribution through public financing and taxation.

2.A Appendix

The appendix of this paper includes several robustness checks, and additional results. The first section considers the possibility that income is strategically misreported by households to ensure that they remain within the eligibility criteria, and as a result keep their Medicaid coverage. Then we present our main estimates, this time using more standard estimation techniques, to show our findings are robust to estimation method. Next we present employment results, but instead using the Zero and Six month surveys to assess the temporal dynamics. The following section reports estimates comparing the labour supply effects of single person households and multiple person households are presented and discussed, so that we are able to compare our results to those of Baicker et al. (2014). Finally, we present, once again, our main results, this time using a simple transformation to change total household-level outcomes to equalised measures.

2.A.1 Strategic Misreporting of Income

Given that in our setting there was an eligibility criteria, not only to receive Medicaid coverage in the first instance but also to keep coverage, there may be a concern that households strategically misreport their income to keep Medicaid coverage. We seek to assess whether strategic misreporting is a concern in our setting, and whether this invalidates our analysis.

Although we have no means of directly assessing misreporting, we would expect that if strategic misreport was an issue in our setting we would observe large mass points just below the eligibility cut-off. For the Oregon Health Insurance experiment the income criteria for eligibility was that households had income below the federal poverty line (Finkelstein et al., 2012), and therefore for one person households the cutoff in 2009 was $10,830, for two persons was $14,570, three persons being $18,310 and for four person was $22,050 (Department of Health and Human Services, 2009). We, therefore, present the distribution of income separately for household sizes of one to four people, as well as the eligibility cut-off for the corresponding household size. We show the income distribution of those that did not receive Medicaid coverage in figure 2.A1 and the sample of individuals that received Medicaid coverage in figure 2.A2. The blue line represents the eligibility cutoff for each household size, and the red line is a kernel density plot for each sample. It is also worth noting that those treated with Medicaid coverage may have an incentive to misreport income, but those that were not treated would have no incentive to misreport their income, as they could not get coverage once the lotteries had been drawn.

If we compare figures 2.A1 and 2.A2 the first noticeable difference between the two is that in the sample of those that received Medicaid there are far fewer households to

the right hand side of the cut-off than in the sample of those that did not receive Medicaid. Overall, there is also a decreasing number of individuals in each category across the distribution. This is, of course, to be expected, given that it was a requirement that those initially receiving Medicaid coverage met the eligibility criteria. This in itself is not a concern, as we would expect the distribution of the Medicaid coverage sample to have lower income than the non-selected sample precisely because of the eligibility criteria. There are also large mass points at very low income levels, and then again at the $10,000 to $17,500 range for all household sizes, and this is present in both Medicaid recipient and non-recipient samples.

Further inspecting the distribution of Medicaid recipients, there are not any substantially large mass-points just to the left of the eligibility cut-off, and no points that do not conform to the overall trend. There are higher numbers of individuals to the left of the cut-off, but this on the whole follows the trend of the data, and therefore there does not appear to be substantial evidence of misreporting. The only case where there is some evidence of a large mass-point just to the left of the cut-off is for single person households. However if this is compared to the non-recipient distribution the same mass exists for that sample too, and therefore we don't believe there is substantial evidence in favour of a hypothesis of strategic misreporting. We believe that because we do not observe any large mass-points to the left of the eligibility cutoffs, strategic misreporting of income is not a major concern, and therefore supports the reliability of our results.

2.A.2 Alternative Estimation Method Robustness

Given that our choice of estimation technique may not be considered the standard approach, in this section we estimate our main results using alternative estimation techniques to ensure that our main results are robust to choice of estimator. In addition, we show that the findings using a unconditional quantile regression (UQR) are similar to the conditional quantile regression (CQR) estimates we present in the main text.

2.A.2.1 Disposable Income and Out-of-pocket Medical Expenditures

Throughout the paper we use the two-stage Residual Inclusion Method (2SRI) for estimating the instrumental variable quantile regressions. Although this method has previously been used in combination with the quantile regression (Fang et al., 2009) and provides provide consistent estimates for a broad class of non-linear models, like those estimated in this paper (Terza et al., 2008b), it is not a standard approach to estimating quantile regressions. Therefore, we estimate Quantile Treatment Effects (IVQTE) estimates using Powell (2020b)'s generalized quantile regression (GQR).

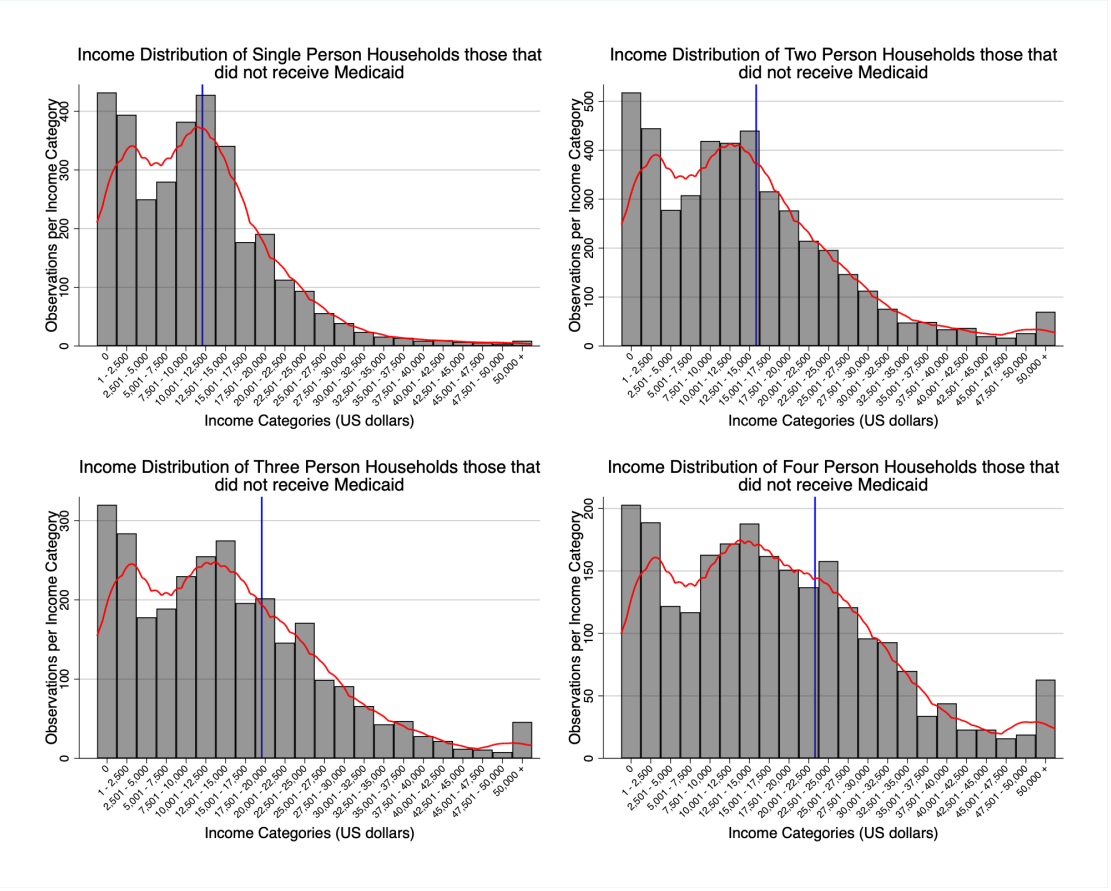


FIGURE 2.A1: Earned Income Distribution of Households that did not receive Medicaid coverage by Household Size

Note: The left most category are households with zero income, the second is $1 onwards to $2,500 and then in intervals of $2,500, up until $50,000. The final category is greater than $50,000. Red line is the kernel density estimate. Blue line indicates the 2009 Federal Poverty Line for the corresponding household sizes (Department of Health and Human Services).

Aside from being a more typical approach to estimation, using Powell's GQR to estimate the IVQTE has an additional benefit. In this paper we are primarily interested in the impact of receiving publicly financed health insurance on inequality. In other words, we are asking whether receiving publicly financed health insurance impacts different parts of the distribution differently. If there is an increase in income at the lower end of the distribution and no effect at the top end of the distribution we have found evidence of publicly financed health insurance reducing inequality. However, a CQR estimator, like the 2SRI-quantile regression we use in the main text, does not always provide this interpretation. Instead, the CQR provides the interpretation, in this example, of the Medicaid expansion increasing the income of the lower end of the *conditional* income distribution. Importantly the 20th percentile of the income distribution does not necessarily correspond to the 20th percentile of the conditional

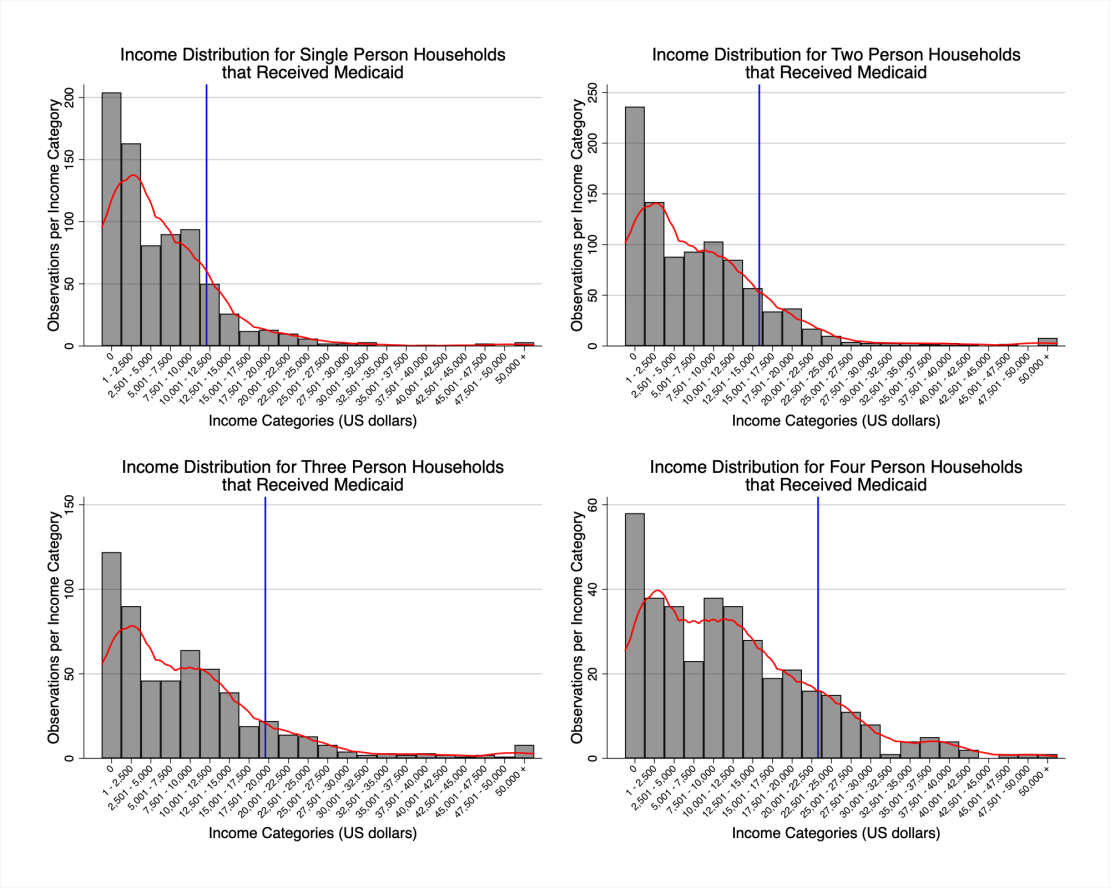


FIGURE 2.A2: Distribution of earned income of Households that Received Medicaid Coverage by Household Size

Note: The left most category are households with zero income, the second is $1 onwards to $2,500 and then in intervals of $2,500, up until $50,000. The final category is greater than $50,000. Red line is the kernel density estimate. Blue line indicates the 2009 Federal Poverty Line for the corresponding household sizes (Department of Health and Human Services).

income distribution, which makes the interpretation more difficult. However, if the data generating process is close to a parallel location shift (i.e. distribution is maintained but is shifted equally for each percentile) for the additional covariates, then the estimated treatment effect of the CQR is a consistent estimator for the UQR (Borah and Basu, 2013).

When estimating the CQR using the 2SRI method, we find that our control variables are relatively stable across quantiles, and therefore we believe the data generating process is close to a parallel location shift. However, given that our estimate does not explicitly estimate a UQR, to ensure that our interpretation is correct it seems reasonable to use a methodology that does so. Therefore, our alternative estimation method will be a UQR, and we will compare those results to the CQR estimates we present in the main paper, to ensure we are correct in our interpretation.

In this section, we therefore present estimates using Powell (2020b)'s GQR. The GQR estimates unconditional quantile estimates, in the presence of an endogenous variable, an instrument, and other covariates, as we have in this case. We therefore estimate the GQR where receiving OHP Standard is our treatment variable, lottery selection as our instrument, and we include the following proneness variables (variables used for ensuring exogeneity while still estimating a UQR): lottery signups, household level age, sex, ethnicity, number of children and number of adults. The estimates from the GQR with disposable income as the outcome variable are presented in Table 2.A1 and those with total healthcare expenditures are presented in Table 2.A2.

For disposable income, the overall pattern is the same, however estimates are substantially larger than for the 2SRI method. We, once again, have large significant effects for the 25th and 40th quantile estimates, however in this case the 40th quantile estimates are larger than for the 25th quantile. There is also evidence of an increase in disposable income at the 60th percentile, which we did not find in the 2SRI method, however the magnitude is substantially smaller than the estimates to the left of the median.

Comparing the out-of-pocket expenditures GQR estimates to the 2SRI estimates, the trend is broadly the same. The GQR estimates seem to be less precise than the 2SRI estimates, however the magnitudes are similar. The only substantial deviations are the 25th and 50th quantiles, which are both much smaller in magnitude in the GQR and are estimates less precisely. Overall, the GQR estimates show that the conclusions we draw regarding the distributive power of Medicaid are robust to the choice of specification, and show that our 2SRI estimates are close to the unconditional quantile estimates.

	Household Disposable Income				
	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
OHP Standard	2795.8***	4651.8**	790.6	1682.3**	1444.8
	(81.5)	(2344.3)	(468.5)	(831.7)	(1345.8)
Observations	16670				

Notes: Table presents Estimates of the Unconditional IV Quantile Treatment Effects (IVQTE), using Generalized Quantile Regression (GQR), of the effect of the Oregon health plan expansion on household disposable income, using the OHIE lottery as the instrument. Standard errors in parentheses. The proneness variables are multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. The numerical optimisation is conducted in two-steps: Grid-search method is used where the starting values are the 2SRI-quantile estimates, with a limit of 5,000 either side of those values. Then an adaptive MCMC optimisation procedure was used, with a Metropolis-within-Gibbs sampler, 1,000,000 draws, a burn rate of 0.5, an acceptance rate of 0.44, and keeping every 10^{th} draw. *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.A1: Estimates of the Unconditional IV Quantile Treatment Effects (IVQTE), using Generalized Quantile Regression (GQR), of the effect of the Oregon health plan expansion on household disposable income

	Total Out-of-Pocket Medical Expenditures				
	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
OHP Standard	1.94*	-134.7***	-82.4	-178.6*	-294.3
	(1.15)	(23.4)	(96.7)	(98.0)	(167.1)
Observations	9908				

Notes: Table presents Estimates of the Unconditional IV Quantile Treatment Effects (IVQTE), using Generalized Quantile Regression (GQR), of the effect of the Oregon health plan expansion on household out-of-pocket expenditures, using the OHIE lottery as the instrument. Standard errors in parentheses. The proneness variables are multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. The numerical optimisation is conducted in two-steps: Grid-search method is used where the starting values are the 2SRI-quantile estimates, with a limit of 5,000 either side of those values. Then an adaptive MCMC optimisation procedure was used, with a Metropolis-within-Gibbs sampler, 1,000,000 draws, a burn rate of 0.5, an acceptance rate of 0.44, and keeping every 10^{th} draw. *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.A2: Estimates of the Unconditional IV Quantile Treatment Effects (IVQTE), using Generalized Quantile Regression (GQR), of the effect of the Oregon health plan expansion on total household out-of-pocket medical expenditures

2.A.2.2 Earned Income

For consistency, in the main text we also use the 2SRI quantile regression to estimate the distributional impacts on earned income, however, again, this may not be considered the most conventional approach. Given that we have categorical data for household income, a more conventional approach to estimating the distributional effects may be to estimate ordered probit models, therefore in this section we estimate those to ensure our findings are robust to the choice of estimator.

The IV ordered probit estimator we use is described as follows. The first stage:

$$OHP\ Standard_i = \delta_1 Lottery\ Selection_i + \delta_2 X_i + u_i \tag{2.6}$$

where $Lottery\ Seclection_i$ is a binary variable denoting 1 if the household was selected in the lottery and 0 if they did not, and $OHP\ Standard_i$ is a binary variable denoting 1 if the household received OHP standard and 0 if they did not. X_i denotes the same set of control used in the main text.

The second stage is subsequently specified as an ordered probit model:

$$P(Y_i > j) = \Phi\left(\beta_{1,j}\widehat{OHP\ Standard}_i + \beta_{2,j}X_i\right) \tag{2.7}$$

Where Y_i denotes earned income, and j denotes the income categories, which takes one of 21 values, \$0, \$1 and then in intervals of \$2,500, up until \$48,751.
$\widehat{OHP\ Standard}_i$ is the predicted probability of receiving OHP Standard, predicted using the first stage in Equation 2.6.

The marginal effects of this estimator are presented in Figure 2.A3. The black dot represents the marginal effect of being in an income category greater than the corresponding value on the x-axis (j from Equation 2.7), with the bars representing the 95% confidence intervals.

These results broadly follow the results from the 2SRI quantile estimator in the main text. Receiving Medicaid coverage increases the probability of being in income categories greater than \$0 through to \$5001. However, the IV ordered probit estimates also show a decreased probability of being in higher income categories for j values \$ 17,501 to \$25,001. This deviates somewhat from the 2SRI Quantiles estimates in that we don't find this decrease in income. However, these results still show that there is a redistributive effect in terms of earned income. Lower income categories show an increase in their income, whereas the higher income categories do not have an increase in their income. Overall, this shows that our results are robust to choice of estimator.

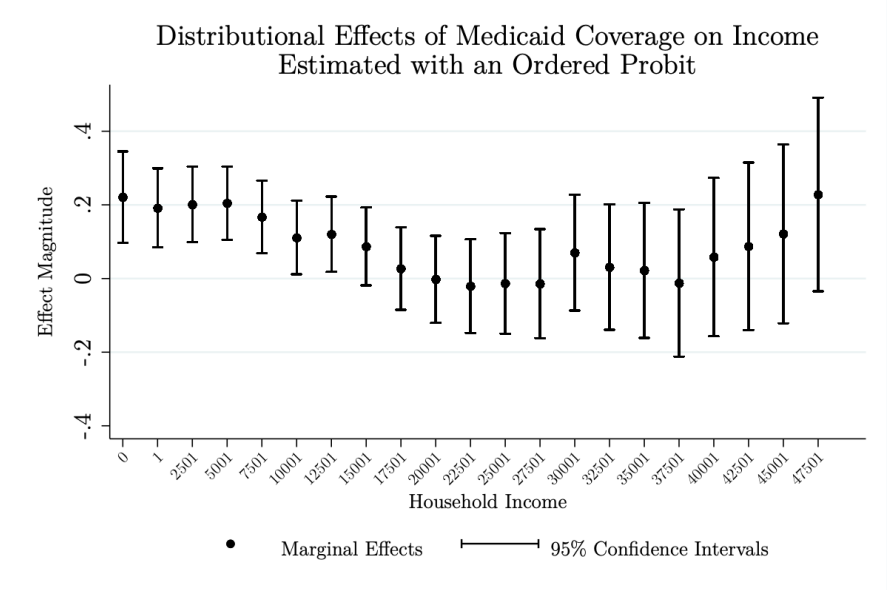


FIGURE 2.A3: IV Ordered Probit Model, estimating the distributional impacts of OHP Standard on earned income

Note: IV Ordered Probit estimates of the impact of OHP Standard on earned income. Lottery selection is used as the instrument for the treatment variable, whether received OHP Standard. Additional control variables being: multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. Dots are the marginal effects, and denote the probability of being in an income group greater than the corresponding value on the x-axis. Bars represent the 95% confidence interval of the marginal effects.

2.A.3 Zero and Six Month Employment Estimates

In the main text we find significant and large estimates of the impact of Medicaid coverage on labour supply. Findings such large and significant effects after just 12 month of receiving Medicaid may be surprising. Although the effect is theoretically justified, the impact on labour supply may be expected to materialise over a longer time-frame. Therefore, we present estimates using the zero and six month surveys to provide insight into the temporal dynamics of the labour supply effects.

Our hypothesis is that when using the zero month survey we would find null effects for labour supply. It should not be the case that labour supply changes are found immediately post Medicaid coverage, and therefore analysing the zero-month survey is akin to a falsification test. In relation to the six month survey, we would expect the magnitude of the effects to be smaller than we find for the 12 month survey, as changes may begin to materialise over that time-frame, but we would expect the full effect to materialise over the longer term.

(a)	Zero Month (Initial) Survey					
	Any Employment	Employee	Self-Employed	Under 30 hours	Over 20 hours	Over 30 hours
OHP Standard	0.000380	0.0227	-0.0226*	-0.00206	0.00133	0.00424
	(0.0181)	(0.0178)	(0.0120)	(0.0115)	(0.0122)	(0.0165)
1^{st} stage F-test	5,903.67	5,901.55	5,901.55	5,906.42	5,906.42	5,906.42
Observations	20,339	20,327	20,327	20,474	20,474	20,474

(b)	Six Month Survey					
	Any Employment	Employee	Self-Employed	Under 30 hours	Over 20 hours	Over 30 hours
OHP Standard	0.0648	0.0240	0.0407	0.0503*	-0.0220	0.0681*
	(0.0432)	(0.0421)	(0.0299)	(0.0294)	(0.0289)	(0.0389)
1^{st} stage F-test	912.14	912.14	912.14	917.38	917.38	917.38
Observations	4,822	4,822	4,822	4,852	4,852	4,852

Notes: Estimations control for multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. Extensive margins are reported as marginal effects. Robust standard errors in parentheses. *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.A3: Instrumental variable estimates of the effect of the Oregon health plan expansion on household income using Zero and Six Month Surveys

Using the same specification as in the main text, we present the impacts of Medicaid coverage on labour supply. Table 2.A3 presents the estimates using the zero and six month survey. As expected, and reassuringly, we find no significant impact on labour supply in the zero month survey. However, we also find no effect in the six month survey. This suggests that it takes between six and twelve months for these labour supply effects to take effect and are not immediate.

2.A.4 Single Person and Multiple Person Household Labour Supply Effects

As discussed in the main text, on initial inspection our results may seem at odds with the results of Baicker et al. (2014), who also use the Oregon Health Insurance experiment to analyse the impact of Medicaid coverage on labour supply. Specifically, they analyse the impact of Medicaid coverage on whether individual has any earnings, the amount of earnings and earnings above the federal poverty line. They find no evidence of Medicaid impacting any of these outcomes, whereas in our analysis we find evidence suggesting that labour supply does increase. As mentioned in the main text, the main difference in our analysis which explains these results is that our analysis is conducted at the household-level, rather than the individual-level which is the case for Baicker et al.'s analysis.

Given that we do not have access to the individual-level income data, we are unable to exactly replicate Baicker et al.'s analysis, however as an as an approximate test, we split the sample to single- and multi-adult households and estimate the impact on whether the household as any earnings, as Baicker et al. do in their paper. The aim is to, as closely as we can, replicate their findings showing a null effect of Medicaid coverage on labour supply, which implies that there is no individual-level effect.

If our results are consistent with the findings of Baicker et al. then we expect there will be a significant impact of any earnings when we analyse multi-person households which will not be present when we analyse single person households. This would suggest that there is an important household reallocation of resources which increases household-level labour supply, which would not be found at the individual-level.

We estimate the following specification separately for single-person households and for multi-person households separately, where the first stage is described by:

$$OHP\ Standard_i = \delta_1 Lottery\ Selection_i + \delta_2 X_i + u_i \tag{2.8}$$

where $OHP\ Standard_i$ is a binary variable denoting 1 if the household received OHP standard and 0 if they did not.

The second stage is subsequently specified as a probit model:

$$P(Y_i > 0) = \Phi\left(\beta_1 \widehat{OHP\ Standard_i} + \beta_2 X_i\right) \tag{2.9}$$

where Y_i denotes household earnings, for household i. We use positive income as our measure of labour supply for this analysis, as this is what is used by Baicker et al. (2014). $Lottery\ Selection_i$ is a binary variable equal to 1 if the household was selected in the lottery to receive OHP Standard. X_i denotes a range of household level

characteristics that the regression controls for, including average household age, proportion of females in household, proportion of white household members, proportion of black household members, proportion of Hispanic household members, # of adults in the household, # of children (under 18) in the household, and whether there were multiple lottery sign ups.

	Positive Household Income	
	Single Adult Households	Multi-Adult Households
OHP Standard	0.0417	0.0354**
	(0.0256)	(0.0177)
1^{st} stage F-test	1490.50	3245.39
Observations	5,140	13,293

Notes: Estimations control for multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. Extensive margins are reported as marginal effects. Robust standard errors in parentheses. *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.A4: Instrumental variable estimates of the effect of the Oregon health plan expansion on household disposable income

Table 2.A4 presents the results for single- and multi-adult households. We find that there is no significant impact of Medicaid coverage on any earnings for single person households, whereas we do find a significant impact for multi-adult households. These results show that our findings are not at odds with those of Baicker et al., but rather, suggests that there is a household-level reallocation which leads to an increase in employment and income, which is not present for adults living alone.

2.A.5 Per-capita Equivalisation Estimates

Throughout the paper we analyse the distributional effects at the household-level, and although we control for household size throughout our analysis, we may wish to instead analyse an equivalised measure of household-level. One concern is that when estimating the heterogeneous impacts of Medicaid using household-level outcomes that has not been equivalised, we may be inadvertently estimating heterogeneous impacts by household size indirectly. Explaining this using an example, if (plausibly) smaller households are disproportionately found at the lower end of the income distribution, and it is only small households that are impacted by receiving Medicaid we may observe similar quantile estimate as we present in the main text i.e. increase in disposable income for the lower end of the distribution. We don't believe this is a major concern, and (as mentioned) we do control for household composition throughout our analysis, however in this section we estimate the same results in the

main text but using a simple equivalised measure, we do this to show that our findings are robust to the way that household-size is handled in the analysis.

Rather than estimating the impact of Medicaid coverage on total household income, in this section we transform the outcome variables, namely disposable income, medical expenditure, and income, into per individual values. We use a simple transformation of dividing the household income by the number of adults in the household. Although this is not an ideal approach, as it clearly induces further measurement error into our outcomes, which are already measured with error, it does provide a simple interpretation of the impact of Medicaid coverage on per capita outcomes.

The estimates from the per-capita equivalisation are presented in Table 2.A5 and Table 2.A6. Once again, the pattern is the same as the results we present in the main text, and they show that there is a gradient in the effect of Medicaid coverage which favours the lower end of the distribution. As one would expect the magnitude of the effects are smaller for the per capita measure than the household equivalents in the main text. One deviation worth mentioning is that the 40^{th} quantile estimate for disposable income and earned income has a larger impact than the 25^{th} percentile in both cases. However, both the 25^{th} and the 40^{th} quantile estimates are larger than any of the quantile estimates to their right and therefore the overall trend is preserved.

	Household Disposable Income						
	Extensive Margin	Intensive Margin					
		Linear	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
Lottery Selection	0.0453***	8.838	740.9***	902.4***	280.5	-267.2	-1263.6***
	(0.0147)	(293.3)	(279.6)	(309.4)	(324.5)	(358.6)	(432.0)
1^{st} stage F-test	4,729.40	3,487.04					
Observations	18,433	15,184					

Notes: Outcome measure are a simple transformation of dividing the original variable with the number of adults in the household. Estimations control for multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. Extensive margins are reported as marginal effects. Robust standard errors in parentheses. *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.A5: Intention-to-treat estimates of the effect of lottery selection on per capita household disposable income

(a)	Out-of-Pocket Expenditures						
	Extensive Margin	Intensive Margin					
		Linear	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
Doctors							
OHP Standard	-0.218***	46.30	-9.563	-22.52	-0.236	-9.506	-60.92
	(0.0194)	(107.7)	(16.49)	(28.38)	(34.70)	(43.45)	(87.15)
1^{st} stage F-test	4,729.40	623.03					
Observations	18,433	6,481					
ER							
OHP Standard	-0.0518***	435.7	18.34	49.91	21.60	62.26	252.8
	(0.0116)	(604.8)	(102.6)	(181.7)	(284.6)	(448.1)	(820.1)
1^{st} stage F-test	4,729.40	119.18					
Observations	18,433	1,413					
RX							
OHP Standard	-0.182***	-1.399	-12.70*	-17.85	-18.25	-16.13	-12.57
	(0.0199)	(85.90)	(7.707)	(11.35)	(17.82)	(24.21)	(38.28)
1^{st} stage F-test	4,729.40	907.28					
Observations	18,433	7,187					
Other							
OHP Standard	-0.0342**	-305.2*	-8.237	-9.076	-13.37	-32.53	-112.2
	(0.0138)	(183.4)	(20.19)	(36.38)	(40.37)	(69.21)	(160.8)
1^{st} stage F-test	4,729.40	354.70					
Observations	18,433	2,212					
Total							
OHP Standard	-0.187***	-103.0	-37.21*	-81.80***	-100.0**	-129.4**	-196.9
	(0.0202)	(167.6)	(19.36)	(30.41)	(42.26)	(62.54)	(129.5)
1^{st} stage F-test	4,729.40	1,261.03					
Observations	18,433	9,020					

(b)	Earned Household Income						
	Extensive Margin	Intensive Margin					
		Linear	25^{th}	40^{th}	50^{th}	60^{th}	75^{th}
OHP Standard	0.0379***	334.0	1,267.2***	1,560.0***	868.2**	-311.3	-619.2
	(0.0146)	(333.3)	(265.3)	(408.8)	(431.1)	(422.0)	(435.8)
1^{st} stage F-test	4,729.40	3,346.92					
Observations	18,433	15,449					

Notes: Outcome measure are a simple transformation of dividing the original variable with the number of adults in the household. Estimations control for multiple lottery signups, household level age, sex, ethnicity, number of children and number of adults. Extensive margins are reported as marginal effects. Robust standard errors in parentheses. *** denotes p-value ≤ 0.01, ** denotes p-value ≤ 0.05, * denotes p-value ≤ 0.10

TABLE 2.A6: Instrumental variable estimates of the effect of the Oregon health plan expansion on per capita household out-of-pocket expenditures and household income

Chapter 3

The direct and spillover effects of diabetes diagnosis on lifestyle behaviours

3.1 Introduction

There is substantial literature documenting a positive correlation in spousal behaviours with much of the work focusing on smoking behaviour and alcohol consumption (Christakis and Fowler, 2008; Falba and Sindelar, 2008). Similar strong positive correlation of behaviour between spouses has also been reported for physical activity Farrell and Shields (2002); Falba and Sindelar (2008) and diet (Macario and Sorensen, 1998; Bove et al., 2003). However, such correlations extend beyond behaviours alone with previous work reporting spousal correlation in mental and physical health (Meyler et al., 2007; Di Castelnuovo et al., 2009). Three theories have been put forward to understand the causal pathways of these strong empirical correlations, namely: assortative matching, shared environment, and joint household decision making (Clark and Etilé, 2006; Cutler and Glaeser, 2010; Chiappori et al., 2012).

Assortative matching views partners' characteristics and preferences as complements which drive individuals to match with partners they share preferences and characteristics with (Becker, 1973). In a shared environment partners make decisions individually based on their preferences, but are constrained by shared resources and exposed to common shocks, which give rise to observed correlated behaviours. An epidemiological dimension is implicit whereby partners who share a common environment are also exposed to common health risks factors. An additional channel under this pathway relates to shared information. Partners not only share resources, but also share information sets, by transferring information between each other, Clark

and Etilé (2006) call this social learning. Common information sets mean that partners also have similar expectations of future uncertainty and risk, and as a result make similar behavioural choices (Khwaja et al., 2006). Finally, joint household production leans on the theory of New Home Economics where households jointly produce goods which enter individuals' utility functions (Lancaster, 1966; Becker, 1981). Individuals within the household bargain and as a result produce and consume some shared output, implying a correlation both in behaviour and health. Payoffs from producing and subsequently consuming a particular good is a function of own private payoffs, and an externality from their partner consuming the same good. As with assortative matching, if behaviours or specific consumption goods are complements, then partners may choose to jointly produce and consume them, which results in empirical correlations in consumption and behaviour.

The latter two of these theories suggest that if an individual was to have health knowledge that would lead to curative or require preventative changes in behaviour, then such changes would likely have a beneficial spillover onto their partner. Only a handful of studies have explored such externalities in the context of health. Fadlon and Nielsen (2019) analyse the spillover effects on an extended network of individuals as a result of fatal and non-fatal heart attacks. They find significant and persistent increases in statin consumption of spouses, children and co-workers of individuals who had a non-fatal heart attack, and offer evidence in support of both learning new health information, and salience explaining the estimated effect. Fletcher and Marksteiner (2017) use experimental data to estimate spillover effects of smoking cessation therapy program and alcoholism treatments. They find significant impact in both behaviours and their experimental design can reasonably preclude a matching in the marriage market explanation. However, their results are at odds with the conclusions by Clark and Etilé (2006) who show that social learning and household decision making play a minor role in explaining raw correlations between partners. Once controlling for individual random effects smoking behaviours are statistically independent between partners, suggesting that all spousal correlation in smoking behaviour is the result of correlations in the individuals' effects, which Clark and Etilé interpret as evidence of assortative matching. Finally, Janssen and Parslow (2021) examine the presence of spillover effects within a household when looking at the impact of pregnancy on alcohol consumption. Pregnancy persistently reduces household alcohol consumption with reductions in purchasing of both beer and wine. Given that males are the prominent beer drinkers in the United States, the authors interpret this as evidence in favour of a spillover effect from females onto males in the household.

In this paper we investigate the effect of diabetes on individual and partner' lifestyle behaviours, namely physical activity, diet, alcohol and smoking consumption. These lifestyle behaviours are well established risk factors of non-communicable diseases

(Willi et al., 2007; Ezzati and Riboli, 2012, 2013) and constitute the first line of treatment of diabetes (WHO, 2016). Using blood sample data from the Health Survey for England (HSE) dataset we exploit a seemingly arbitrary cut-off of diabetes risk and through a fuzzy regression kink design we causally estimate the impact of own diabetes on own behaviour, as well as, the effects of own diabetes status on partners' behaviour. The identification strategy allows us to exclude assortative matching as a causal pathway, while through mediation analysis we decompose the spillover effect into its shared environment and joint household production contributions. Finally, we present falsification tests over multiple health outcomes that would not be expected to be impacted by diabetes status. Further, in the appendix of this paper, we explore three sources of heterogeneity over observables. First, we test whether own behaviour changes as a function of living with a spouse or not. Second, we use time since diabetes diagnosis to examine differential impact on own and partner lifestyle outcomes, which, in the absence of panel data, approximates long-term effects or recidivism to pre-diagnosis behaviours. Third, we assess whether there are observable heterogeneities by individual education.

Briefly, we find significant effects of diabetes diagnosis on own physical activity and smoking, while partners' of individuals with a diabetes diagnosis also increase their physical activity and decrease their probability of currently being a smoker. Spillover effects are mostly driven by partner's behaviour and less so by the partner's diabetic status. We find almost no evidence of heterogeneity of the effect of own or partner diabetes on behaviour by presence of partner in the household, time since diagnosis or education. All of our falsification tests support our identification strategy and provide evidence towards the robustness of the results.

We contribute to the literature in a number of ways. First, we provide evidence within the household economics literature that observed correlated partners' health behaviours are not limited to assortative matching, but that social learning and joint household decision making are important components of the observed correlation. Indeed, we contribute to the household economics literature, and find empirical evidence of the joint household decision making theory in a health context. Second, we contribute to the existing literature on diabetes, by causally estimating the behavioural responses of a diabetes diagnosis (Hut and Oster, 2018; Oster, 2018; Kim et al., 2019). This is related to how these behaviours are determined and influenced, as well as to individuals' compliance with first line treatments for diabetes. Our results suggest that individuals with diabetes comply with some treatments and that this behavioural change is persistent over time. Our results are of particular importance to health policy makers, as the evidence for substantial positive spillover effects from diabetes diagnoses potentially suggests additional health benefits that are currently not accounted for in the evaluation of health care policies in this area. Finally, we contribute to a new and growing literature on health-related spillover effects by

analysing the effects of a health shock on lifestyle behaviours commonly acknowledged as important risk factors of non-communicable diseases.

This paper is organised as follows, first we offer background for the context and premise of the paper, specifically, we discuss diabetes in detail, noting the institutional setting as well as previous literature in this area. Second, we present the theory and literature on spousal correlation and how such theories fit in our setting. Third, we present the data and move onto our identification and estimation strategy. Then, we present our results and validate the identifying assumptions. Finally, we discuss our findings, and place them within a wider context.

3.2 Background

3.2.1 Diabetes

The World Health Organization (WHO) defines diabetes as "a chronic, metabolic disease characterized by elevated levels of blood glucose (or blood sugar), which leads over time to serious damage to the heart, blood vessels, eyes, kidneys, and nerves" (WHO). Diabetes is classified into two types, type 1 and type 2. Of the 4.7 million people with diabetes in the UK, approximately 8% have type 1, which occurs when insulin production in the body is limited (Diabetes UK, 2019). Although there is limited understanding on its causes, diet or lifestyle are not known to have any impact on the probability of having or developing type 1 diabetes. Type 2 diabetes affects approximately 90% of those with diabetes, and occurs when the body becomes resistant to insulin and is usually found to be a result of poor diet and lifestyle (Helmrich et al., 1991; Hu et al., 2001).

Glycated haemoglobin (HbA1c) refers to the amount of haemoglobin (i.e. protein within red blood cells) which has been "glycated". This occurs when the body processes sugar, and glucose in the blood then attaches to haemoglobin proteins. The red blood cells which contain the haemoglobin proteins usually survive for between 8 and 12 weeks, and therefore HbA1c is considered to be an average blood sugar level over the previous three months. HbA1c is considered a useful measure in the diagnosis of diabetes, in that it provides an indication of blood sugar level over a longer duration.[1]

The World Health Organisation recommends an HbA1c of 6.5% as the cut-off point for diagnosing diabetes, while stating that values below 6.5% do not exclude a diabetes diagnosis (WHO, 2011). Levels below 6% are considered normal blood sugar levels

[1] An alternative measure, blood glucose level, is the concentration of sugar in the blood at a single point in time and is highly variable within individuals, and more dependent on very recent consumption than persistent behaviour.

and therefore low-risk, while levels between 6% and 6.5% are considered at high risk of becoming diabetic, also called pre-diabetes. However, while the link between HbA1c and the probability to develop diabetes is well-established, the choice of specific cut-off for diabetes and pre-diabetes are relatively arbitrary.[2] Nevertheless, although pre-diabetes usually has no symptoms, NICE[3] recommends that "for people at high risk (a high risk score and fasting plasma glucose of 5.5 - 6.9 mmol/l, or HbA1c of 42 - 47 mmol/mol [6.0 - 6.4%]), offer a blood test at least once a year (preferably using the same type of test). Also offer to assess their weight or BMI." NICE (2012).

Therefore, individuals who have been found to be pre-diabetic and at high risk of type 2 diabetes have a significantly higher probability of being diagnosed with diabetes simply as a result of being subject to annual assessment of their HbA1c level. On the other hand, individuals just below the threshold of 6.0%, while having similar probability of actually having diabetes as those just above the threshold, have a much lower probability of being diagnosed as a result of them not being annually tested, as per the NICE guidelines.

Our analysis focuses on the impact of a diabetes diagnosis on risk-factors commonly associated with non-communicable diseases. Clinical recommendations regarding such risk-factors are clear and well-known to the general population, rendering a priori expectation of the effects straightforward. Namely increasing physical activity and vegetables consumption and decreasing tobacco and alcohol consumption mitigate the risk of developing diabetes and are important first-line treatments of the disease (WHO). On the contrary, while the health benefits of fruit are well established, recommendations on fruit consumption for diabetic patients is somewhat ambiguous and possibly misunderstood by the general population [4] making a priori expectations unclear.

[2] Yudkin and Montori (2014) state that "glycaemia are continuous, with no inflections to provide obvious cut-off points. Cut-offs for the diagnosis of diabetes are based on thresholds for risk of retinopathy. Lesser degrees of hyperglycaemia increase the risk of developing diabetes and maybe arterial disease. But in both cases the risk is graded, making any choice of cut-off point purely arbitrary." This claim is also supported by NICE (2011, 2012)

[3] The National Institute for Health and Care Excellence (NICE) is an executive non-departmental public body of the Department of Health which publishes guidelines for clinical practice and the use of healthcare technologies in the National Health Service.

[4] On one hand, experts encourage fruit consumption due to their low energy density, and high content of vitamins, minerals, phytochemicals and dietary fibre. On the other hand, others argue that fruit should be limited due to the high carbohydrate content which raises blood sugar, which is problematic in those with diabetes (Forouhi et al., 2018). NHS advice states that those with diabetes should "eat a wide range of foods - including fruit", the advice also states that individuals should "keep sugar, fat and salt to a minimum" (NHS, 2018), which can potentially cause confusion due to the high sugar content of fruit. Indeed, there are a number of ongoing campaigns to resolve understanding of the guidelines (Diabetes UK). However, confusion is present both among healthcare professionals and patients with 25% and 57%, respectively, stating that "fresh fruit can be eaten freely with little effect on blood glucose levels" (Speight and Bradley, 2001). Forouhi et al. (2018) state that "consumption of fruits should be guided within the overall dietary pattern of an individual, their taste and other preferences and by their glycaemic control and need for antidiabetic medication, supported by healthcare professionals".

3.2.2 Spousal Correlation

As discussed in Section 3.1, there is theoretical justification for the presence of a spillover effect from one of the partners being diagnosed with diabetes. Firstly, a diabetes diagnosis transfers health information to the patient both in relation to their own health state (i.e. diagnosis of the disease) and to the disease itself (i.e. causes and consequences of diabetes). Social learning implies that this knowledge would be passed on from patient to partner and having the same information set each partner updates their expectations of future risk and uncertainties. Whether this new information promotes behavioural changes is dependent on idiosyncratic preferences, structural determinants of health and their information set pre-diagnosis (Orphanides and Zervos, 1995). However, if an individual has a preference for health but they were not, previously, fully informed of the risks of diabetes, we would expect the newly acquired information to result in a reduction in the probability or level of engaging in risky health behaviours.

For the health information causal channel, the effect on partners' behaviours is independent of the observed behaviours of the diabetic individual post-diagnosis. The partner privately re-evaluates and makes new utility maximising decisions based on their new information set that was transferred to them by their partners (Cutler and Glaeser, 2010), but based on their own idiosyncratic preferences. Although the information set would be shared between partners, their preferences are not identical, and therefore realised behaviours are not perfectly correlated. The magnitude of this effect is moderated by the information set pre-diagnosis. Partners in possession of realistic expectations of the risks of diabetes pre-diagnosis would not substantially change their expectations and would require smaller adjustments to their behaviour as a result of the new information. The claim here being that individuals' preferences remain stable, but the expectation of uncertain events is updated.

Secondly, if a diabetes diagnosis changes the optimal consumption of health-related activities of the diabetic individual, through the updated information channel discussed above, we can also expect it to impact the production and consumption decisions of the other productive household members (i.e. partners) through joint household decision making (Becker, 1973, 1981). For instance, post-diagnosis, physical activity may have higher expected payoff for the diabetic partner. A non-diabetic partner with strong preference for joint time consumption (Jenkins and Osberg, 2004) may choose to participate in physical activity even if they gain relatively less utility from physical activity *per se* compared to other household production activities (Cutler and Glaeser, 2010). However, a positive spillover is not necessarily always the case[5]

[5]Presence of a non-compliant to treatment diabetic partner or a stronger dislike for physical activity than preference for joint time consumption for the non-diabetic parter could also explain explain minimal behavioural change for the non-diabetic partner.

making the effect of a diabetes diagnosis through this causal channel, while still possible, somewhat more ambiguous.

Finally, assortative matching on diabetes diagnosis would imply that individuals actively seek partners with diabetes (even if they themselves are not diabetic) and would also require diagnosis to happen pre-match. Hence, it is less likely that assortative matching is the driving force behind our findings. What is possible, however, is that individuals match based on behaviours which may impact the cause of diabetes. For instance, individuals sharing a dislike for physical activity or preference for smoking match in the marriage market, these individuals are more likely to be diagnosed with diabetes precisely as a result of the shared preferences. In such case, partners' diabetes status would be endogenous. However, this is not the causal effect we estimate in the present paper and our identification strategy minimizes the possibility that our estimates are the result of assortative matching.

3.3 Data

The paper uses data from the Health Survey for England (HSE) for years 2003 to 2015. HSE is an annual cross sectional dataset aiming to monitor trends in national health. More than 9,000 addresses are sampled over the course of the calendar year. Within each household, all individuals are eligible for survey inclusion, however children under 15 years old are asked to complete a different survey. In addition to the individual questionnaire, all respondents are eligible for a nurse visit, in which individuals' physical measurements and a blood sample are taken. Once taken, the blood sample is sent to a specialist laboratory to measure among others, glycated haemoglobin (HbA1c). Although 82.4% of individuals (across all years) agreed to be contacted for a nurse visit, only 34.7% of the full sample had blood samples taken for analysis. Of the 56,245 individuals who had blood taken in the survey, 53,450 individuals had valid HbA1c measurements [6].

Our selection of outcomes analysed (i.e. physical activity, diet, tobacco and alcohol) focus on behaviours that have all been shown to cause diabetes, and have been outlined as a first line treatment for managing and treating diabetes (WHO, 2016). Physical exercise is taken as the response to "any exercise done in the last four weeks". Information relating to diet in the HSE is limited, however we use two relevant variables, "whether consumed any vegetables yesterday" and "whether consumed any fruit yesterday", while smoking and drinking behaviour are captured by "whether currently a smoker" and "whether currently a drinker" excluding those that

[6] A change in calibration of the equipment used for analysis HbA1c was made in 19th of September 2013, which resulted in a slight change in result for equivalent blood samples. Throughout the analysis we use "valid HbA1c result", as recommend in the Health Survey for England documentation, which adjusts the results post-2013 to be equivalent to pre-2013 results for the same blood samples.

are never smokers or drinkers, respectively. As measures of behaviours, these measures are limited and crude, neverthless they should provide insight into the response of individuals to a diabetes diagnosis, the spillover effect, and the channels in which this effect works.

Table 3.1 provides descriptive statistics of the data used in the analysis. The first column provides means and standard deviations of a number of observable characteristics and stated health-related behaviours for the entire HSE sample, including those that did not have blood measurements taken. In subsequent columns we give summary statistics of the sub-sample of individuals who did have blood taken for analysis and whose data is used in our estimations. We break descriptive statistics into those with measured HbA1c levels below and above the 6.0% cut-off. The right-most columns in the table are descriptive statistics of the sub-sample of individuals who have HbA1c results in the data and additionally have partners living in their household with HbA1c results in the data. These are also separately broken down into HbA1c levels below and above 6.0%.

	HSE Adult Sample	Blood Sample			Blood and Partner Sample		
		All	Below Kink	Above Kink	All	Below Kink	Above Kink
Observable Characteristics							
Age†	49.52	51.53	49.11	63.91	51.95	50.04	62.19
	(18.72)	*(17.63)*	*(17.29)*	*(13.66)*	*(15.19)*	*(14.84)*	*(12.75)*
Males	0.45	0.46	0.45	0.49	0.48	0.47	0.56
	(0.50)	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*
Any Qualifications	0.74	0.76	0.8	0.58	0.78	0.81	0.62
	(0.44)	*(0.43)*	*(0.40)*	*(0.49)*	*(0.42)*	*(0.39)*	*(0.48)*
Degree level education	0.20	0.22	0.23	0.13	0.24	0.25	0.15
	(0.40)	*(0.41)*	*(0.42)*	*(0.34)*	*(0.42)*	*(0.43)*	*(0.35)*
Partner living in household	0.64	0.67	0.67	0.64	–	–	–
	(0.48)	*(0.47)*	*(0.47)*	*(0.48)*			
Household Size†	2.69	2.59	2.68	2.16	2.90	2.96	2.57
	(1.39)	*(1.32)*	*(1.34)*	*(1.15)*	*(1.17)*	*(1.18)*	*(1.05)*
Employed	0.60	0.61	0.65	0.37	0.67	0.71	0.43
	(0.49)	*(0.49)*	*(0.48)*	*(0.48)*	*(0.47)*	*(0.45)*	*(0.50)*
Equivalised Income†	30,457.38	31,732.89	32,834.61	25,894.29	33,227.54	34,392.52	26,659.59
	(27,527.94)	*(27,879.07)*	*(28,212.34)*	*(25,253.62)*	*(26,157.03)*	*(26,439.76)*	*(23,445.57)*
Self-assessed general health (1 = Very Good, 5 = Very Poor)	2.04	1.98	1.89	2.43	1.93	1.85	2.36
	(0.95)	*(0.91)*	*(0.87)*	*(1.00)*	*(0.87)*	*(0.83)*	*(0.97*
Glycated Hemoglobin (HbA1c)	–	5.61	5.39	6.73	5.60	5.39	6.72
		(0.75)	*(0.33)*	*(1.17)*	*(0.73)*	*(0.32)*	*(1.16)*
Stated Behaviours							
Physical Activity †	0.44	0.46	0.5	0.26	0.46	0.48	0.27
	(0.50)	*(0.50)*	*(0.50)*	*(0.44)*	*(0.50)*	*(0.50)*	*(0.44)*
Vegetable Consumption	0.53	0.53	0.53	0.54	0.54	0.54	0.55
	(0.50)	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*	*(0.50)*
Fruit Consumption	0.61	0.62	0.61	0.67	0.63	0.63	0.67
	(0.49)	*(0.48)*	*(0.49)*	*(0.47)*	*(0.48)*	*(0.48)*	*(0.47)*
Currently a drinker	0.85	0.89	0.90	0.82	0.90	0.91	0.84
	(0.36)	*(0.32)*	*(0.30)*	*(0.38)*	*(0.30)*	*(0.29)*	*(0.37)*
Currently a smoker	0.21	0.19	0.19	0.18	0.16	0.16	0.16
	(0.40)	*(0.39)*	*(0.39)*	*(0.38)*	*(0.37)*	*(0.36)*	*(0.37)*
Ever a drinker	0.90	0.93	0.94	0.90	0.94	0.94	0.91
	(0.30)	*(0.25)*	*(0.24)*	*(0.29)*	*(0.24)*	*(0.23)*	*(0.29)*
Ever a smoker	0.58	0.59	0.58	0.63	0.58	0.57	0.62
	(0.49)	*(0.49)*	*(0.49)*	*(0.48)*	*(0.49)*	*(0.50)*	*(0.49)*
Number of Observations	121,849	53,146	44,448	8,698	32,910	27,740	5,170

Table shows the mean and, in parentheses, the standard deviation of observable characteristics and stated behaviours. The HSE adult sample column shows the descriptive statistics for the entire Health Survey for England sample whom have a full set of non-missing observations for our control variables, including those that did not have valid HbA1c measurements. The blood sample column shows only the sub-sample of individuals whom we have valid HbA1c measurements for. Blood and Partner sample represents the sub-sample of individuals who had both valid HbA1c measurements and that we were able to identify partners in the Health Survey for England. Below kink columns represent the sub-sample of individuals with HbA1c levels below 6.0%, and above kink columns represent the sun-sample of individuals with HbA1c levels above 6.0%.
† denotes variables which were not available to us for all years of the survey, and therefore the true number of observations used to calculate them are less than the number of observations denoted at the bottom of the table.

TABLE 3.1: Descriptive Statistics

The Blood and Partners sample is substantially smaller than the Blood Sample. Not all individuals included in the blood sample have partners, and not all partners that responded had valid HbA1c measurements, therefore we would expect and indeed

observe fewer observations for this sample. Variables marked with a † in Table 3.1, denote variables that they were not asked in every year of the survey, and therefore the number of observations for these variables are smaller than the total number of observations given at the bottom of the table. One example is physical activity, which was not surveyed in all years but only in 2003, 2004, 2006, 2008, 2012. This is also true for household size and equivalized income, but for different years.

It is worth noting that in our sample, individuals who have ever been diagnosed as diabetic were, on average, diagnosed 10.06 years ago (standard deviation of 10.46). Therefore our results are not interpreted as the immediate effect of a diabetes diagnosis, unlike previous studies that observe behavioural responses in a short-time frame post-diagnosis (Hut and Oster, 2018; Oster, 2018; Kim et al., 2019). These studies use a panel data structure and observe the pre-diagnosis period, and a short time frame post diagnosis, up to four years in Kim et al.'s setting. Because on average we observe individuals who were diagnosed in the distant past, our Marginal Treatment Effect (MTE) is more akin to the long-term effect of a diabetes diagnosis. This additionally allows us to investigate the temporal effects over a longer time-frame than previous studies, and indeed we do analyse these temporal effects. We note, however that our identification strategy is not invalidated by such data structure and we present it in detail in the following section.

3.4 Identification Strategy

The aim of this paper is to estimate the causal impact of own or partner's diabetes diagnosis on a variety of health related lifestyle behaviours, specifically, tobacco and alcohol consumption, physical activity and diet. This relationship can be described by the following equation:

$$Y_i = \theta_0 + \theta_1 EverD_i + \theta_2 EverD_j + e_i \tag{3.1}$$

where Y_i denotes the health related lifestyle behaviour of interest and $EverD_i$ denotes whether individual i has ever been diagnosed with diabetes, and $EverD_j$ denotes whether the partner of individual i, person j, has ever been diagnosed with diabetes. A naive OLS of this form, using survey data, would most likely provide biased estimates of both θ_1 and θ_2.

The first and possibly most salient source of bias is simultaneity. It is possible that individuals with diabetes may display behaviour damaging to their health compared to those without diabetes. Such correlation, however, ignores that these individuals would have been diagnosed as having diabetes precisely because they behaved in this damaging way. Indeed, the causes of type 2 diabetes are poor lifestyle factors (Helmrich et al., 1991; Hu et al., 2001). However, the fact that lifestyle is often

determined by environmental and socio-economic factors, this channel further incorporates omitted determinants of behaviours. A second source of endogeneity that would bias least squares estimation of θ_2 in equation (3.1) is matching in the marriage market (Dupuy and Galichon, 2014). Individuals selectively marry along similar traits and therefore ignoring this channel through a naive estimation will again bias estimates of the spillover effect.

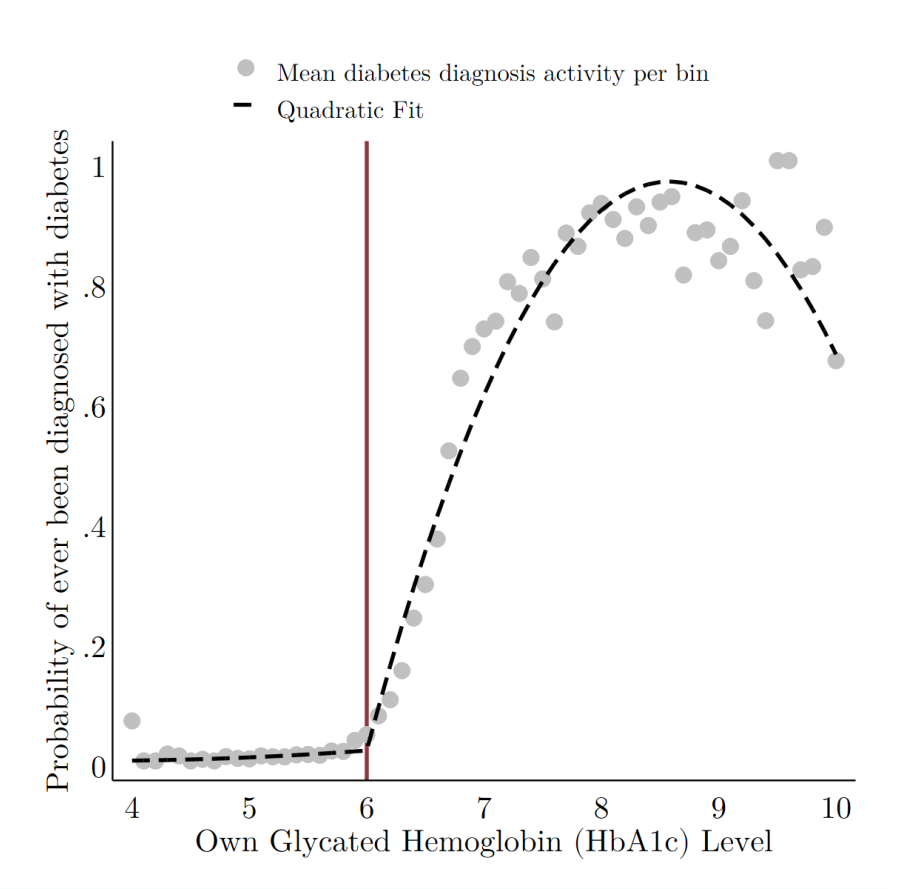
In the following section we will present our approach to estimating the impact of a diabetes diagnosis on own behaviours, handling the simultaneity bias. Then, in section 3.4.2, we will present our identification strategy for estimating the unbiased spillover effect of own diabetes diagnosis on partner's behaviour.

3.4.1 Regression Kink Design

To identify the causal effect of diabetes diagnosis on health-related behaviours, we utilise a regression kink design (RKD), where the kink is a slope change in the treatment probability of a binary treatment variable. Figure (3.1) motivates the use of the RKD within this setting. As shown, there is an increasing but consistently low probability of ever being diagnosed with diabetes when plotted against HbA1c, until the kink point of 6%[7], at which point there is a dramatic increase in the slope of the probability of being diagnosed. As discussed in Section 3.2.1, NHS recommends that individuals with a glycated hemoglobin (HbA1c) level above 6% are offered annual blood tests to monitor their blood sugar levels, and to diagnose diabetes as early as possible. The initial test could be for a variety of reasons, sometimes as part of a regular check up offered by the NHS, or if an individual shows symptoms that warrant a blood test. It is worth emphasising that such precise kink in the probability of a diabetes diagnosis is not supported in the medical sense as Yudkin and Montori (2014) explicitly explain that an inflection point of diabetes risk does not indeed exist, meaning that the assignment of diabetes risk is arbitrary. We will use this arbitrary threshold of 6% as an exogenous threshold to identify the effect of diabetes diagnosis on behaviour. The intuition is that individuals just below the 6% threshold are virtually identical in terms of actual diabetes risk as those just above the 6% threshold. However, despite a very similar baseline risk, those just above the 6% threshold are increasingly likely to have been diagnosed with diabetes given the NHS recommendation.

Dong (2011) provides the theoretical framework for identification in our setting, whereby the RKD identifies the causal effect of a binary treatment when there is no discontinuity in the probability of treatment but rather a kink. When the policy rule is implemented with some error (i.e. the kink is not deterministic) a fuzzy RKD design

[7] We explore the possibility of an alternative jump or kink point in the appendix section 3.A.5, by testing the fit of alternative kink/jump-points and specifications.



NOTE: Mean of the probability of ever being diagnosed with diabetes per bin. Bin width of 0.1 for glycated hemoglobin levels between 4 and 10. Quadratic fit is separately estimated for the left and right hand sides of the kink. Red line represents the kink point, where glycated hemoglobin is a value of 6.0.

FIGURE 3.1: Probability of Diabetes Diagnosis by HbA1c Level

can be implemented (Card et al., 2015). A fuzzy RKD combines the RKD with a two-stage least squares (2SLS) specification. The first stage identifies the effect of the kink on the probability of treatment:

$$EverD_i = \gamma_0 + \gamma_1(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \nu_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \nu_p^+(x_i - k)^p D_i\right] + \xi_i \quad (3.2)$$

where $EverD_i$ is a binary variable taking the value of one for individual i if they have ever been diagnosed with diabetes, and zero otherwise. x_i denotes the running variable, which is HbA1c level in this case, and k is the kink point of 6%. $D_i = \mathbb{1}(x_i \geq k)$, is an indicator variable, taking the value of one if the individual's level of HbA1c is above the kink point, and where $(x_i - k)D_i$ is the excluded instrument for the fuzzy RKD. p^* denotes the highest order of polynomial used in the regressions, ν_p^- and ν_p^+ are the estimates of the polynomial function below and above the kink point, respectively.

We then estimate the following second stage regression where the the kink is used as an instrument for the binary treatment, whether ever diagnosed with diabetes:

$$Y_i = \beta_0 + \beta_1 \widehat{EverD_i} + \left[\sum_{p=1}^{p^*} \alpha_p^- (x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \alpha_p^+ (x_i - k)^p D_i\right] + \epsilon_i \tag{3.3}$$

where Y_i denotes the health related behavioural outcome of interest. $\widehat{EverD_i}$ is the predicted probability, from the first stage, of ever being diagnosed with diabetes, while again the terms in the square brackets denote the polynomial function below and above the kink point. In line with Gelman and Imbens (2019), the main analysis uses quadratic polynomial specifications to estimate effects, while linear specifications are also reported in sensitivity tests. Under the assumptions outlined by Dong (2011) and Card et al. (2015) (see section 3.5 for details), the coefficient β_1 can be interpreted as the unbiased Marginal Treatment Effect (MTE) of ever having been diagnosed with diabetes.

Identification comes from the exogenous variation that the kink provides in the probability of diabetes diagnosis. This relies on the assumption that those just to the left of the kink are almost identical to those just to the right of the kink and it was random variation that resulted in them falling either side of the kink-point. Given that in the dataset diabetes diagnosis is predetermined (i.e. past diagnosis), yet HbA1c is contemporaneous, this potentially creates confusion over identification but does not invalidate it.

The fact that individuals to the left of the 6% cut-off may have received a past diabetes diagnosis and others to the right of 6% may not have had a diagnosis [8] suggests that we do not have a strictly deterministic function of diabetes diagnosis by HbA1c level but a kinked function (i.e. the change in the probability of diagnosis around the cut-off) driven by a policy rule. There is no medical reason for this kink in the diabetes probability, and most individuals are not even aware of their HbA1c level.

[8] There are individuals to the right of the kink-point that are not diagnosed with diabetes, and indeed being to the right does not strictly increase their probability of being diagnosed with diabetes. These individuals can be thought of as never-takers. Being to the right of the kink-point does not increase their probability of being diagnosed. Various explanations could be offered for this, the most salient being individuals who never engage with the healthcare system, regardless of their health outcomes, or who refuse blood tests. Correspondingly, there are individuals who are to the left of the kink-point and yet have been diagnosed with diabetes. Firstly, being to the left of the cut-off does not strictly eliminate the probability that an individual is diagnosed with diabetes, these individuals also face a small probability that they were diagnosed with diabetes. These individuals can be thought of as always takers and not defiers. Although these individuals have a diabetes diagnosis, being to the left does not make them defiers, and they do not per se violate the monotonicity assumption. It is implausible that we have defiers in our setting. A defier in our setting would need to have a decreasing probability of ever being diagnosed with diabetes due to being to the right of the cut-off, which does not seem reasonable. In other words, being to the left of the kink-point leads to a higher probability of ever being diagnosed with diabetes than being to the right of the kink point, which seems implausible. An individual may have a positive diabetes diagnosis, however, being to the right of the kink would always increase this probability.

It is this exogeneous kink that identification rests upon, and not HbA1c per se, or an individual's place in the HbA1c distribution. Past and present lifestyle behaviours can both be correlated and impacting HbA1c (this would certainly be expected as a result of diabetes treatment) but are all unable to precisely affect HbA1c location around the kink-point (Dong, 2011).

Hence, exogeneity would require that kinks around the cut-off would not be expected for lifestyle behaviours and, by implication, any kink in behaviours would be driven by the kinked probability of diabetes diagnosis. Such assumption (i.e. kinks not being present in the structural outcome equation) is, by and large, innocuous as there is no reason why the kink in HbA1c should directly impact behaviour. On the contrary, the running variable HbA1c may be reasonably included in the structural outcome equation, however inclusion of the kink itself is hard to justify intuitively. Rather, the kink has a predictive effect on diabetes diagnosis, hence its relevant as an instrument. The kink can only plausibly impact behaviours through its effect on probability of diabetes diagnosis.

As with regression discontinuity designs (RDD) there is a bias-variance trade-off to be made when selecting the estimation sample. A narrow bandwidth around the kink point will reduce the chances of misspecification error, given that around the kink-point the functional form is likely to be closer to linear. However smaller samples will not have sufficient power to reject a false null hypothesis because of the larger variance in the estimates. Large samples will improve precision of the estimates but will also increase the chances that the functional form is misspecified, therefore increasing the risk of bias (Cattaneo et al., 2020). In our data we observe HbA1c measurements to one decimal place, and therefore we have data which looks more discrete in nature around the cut-off. For this reason, we limit our polynomial specification to a quadratic, to ensure we are not over-fitting to our data. In addition, we choose a bandwidth that is relatively large so that we have sufficient power to reject a false null hypothesis.

However, to ensure that our results are robust, we transparently present a number of alternative specifications and bandwidths in sensitivity tests. Given the few observations of individuals who have been diagnosed as having diabetes on the right hand side of the kink-point, we increase that bandwidth and keep the left-hand side bandwidth much narrower where small sample size is less of a problem (i.e. asymmetric bandwidths). Our main set of results, uses a bandwidth of 4.0% on the right hand side of the cut-off and 2.0% on the left hand side (i.e. HbA1c values of 4% to 10% are included in the estimation sample).

To improve precision and reduce bias of our estimates (Imbens and Lemieux, 2008) we additionally include the following covariates in our estimating equation: a gender dummy, a continuous age variable, we also include a binary indicator of whether

individual has degree level education, and a binary indicator denoting whether a partner lives in the household.

3.4.2 Partner's Diabetes Status

To handle the endogeneity in the effect of partner's diabetes diagnosis on own behaviour, we adapt the previous setup by using the partner's kink as an instrument for partner's probability of being diagnosed with diabetes. The first stage of the 2SLS is specified as

$$EverD_j = \lambda_0 + \lambda_1(x_j - k)D_j + \left[\sum_{p=1}^{p^*} \rho_p^-(x_j - k)^p\right] + \left[\sum_{p=2}^{p^*} \rho_p^+(x_j - k)^p D_j\right] + u_i \quad (3.4)$$

where j denotes the partner, $EverD_j$ is whether partner has ever been diagnosed with diabetes, and x_j denotes the partners HbA1c level. The second stage estimating the causal relationship is

$$Y_i = \delta_0 + \delta_1\widehat{EverD_j} + \left[\sum_{p=1}^{p^*} \tau_p^-(x_j - k)^p\right] + \left[\sum_{p=2}^{p^*} \tau_p^+(x_j - k)^p D_j\right] + \varepsilon_i \quad (3.5)$$

Once again, Y_i denotes the health related behavioural outcome of interest. $\widehat{EverD_j}$ is the predicted probability, from the first stage, of partner ever being diagnosed with diabetes, while again the terms in the square brackets denote the polynomial function below and above the kink point. As discussed previously, causal identification requires reasonable bandwidths either side of the kink-point. Using the same bandwidths for partners as for own, the estimation sample is reduced as it is restricted to those who have partners, and those partners have HbA1c levels within the bandwidths. As previously, the same set of covariates for both i and j (excluding whether partner lives in the household) are included in the regression.

We interpret these results to be a spillover effect, and exclude the possibility that our estimates are the result of assortative matching. To exclude assortative matching, we require that matching does not happen based on being either side of the kink-point. It is certainly possible to assume that individuals match based on their relative position in the HbA1c distribution, or some unobservable variable correlated with HbA1c, and indeed, doing so does not violate the identifying assumption, however it seems less plausible that individuals would specifically match based on being just either side of the kink-point. For matching to explain our estimates, it would require individuals to be aware enough of their own HbA1c level at the time of matching, and to selectively match based on being either side of the arbitrary kink-point. Given that most individuals are not aware of their own HbA1c for this to be possible, and there

appears to be no underlying incentive to match based on this arbitrary threshold, it seems implausible that assortative matching would be affecting our estimates.

3.5 Validity of identifying assumptions

For RKD estimates to be considered the MTE of diabetes diagnosis, two observable implications must hold (Card et al., 2015). The first relates to the smooth density of the assignment variable and empirically tests the assumption of no deterministic sorting. The second relates to the lack of discountinuity or kinks in the pre-determined covariates and tests the assumption that the marginal effect of the assignment variable on the outcome is smooth.

3.5.1 Smooth density of the assignment variable

The smooth density of the assignment variable implies no discontinuity in its density (an assumption similar to that required for RDD settings) but additionally for the RKD case, requires the lack of a kink in its density. While one's position in the distribution can be coarsely influenced by changes in diet and other health behaviours, the value of HbA1c is not able to be manipulated precisely as would be required for it to exhibit a kink or discontinuity at the threshold given Yudkin and Montori (2014). However, this observable implication of the RKD assumptions is testable, and therefore we do so to ensure that this assumption does hold in our context.

McCrary (2008) provides a test for deterministic sorting for continuous assignment variables but ignores the stronger version of the assumption requiring no kink. There are two important considerations for testing this assumption in our setting. The first issue that we face is that the McCrary test is designed with continuous assignment variables in mind, however in our data HbA1c levels are rounded to the nearest 0.1. The discrete nature of our assignment variable can lead to both size and power issues if we were to use the McCrary test. Therefore, instead we use the Frandsen (2017) test for manipulation when the assignment variable is discrete.

The second consideration is that the tests proposed by both McCrary (2008) and Frandsen (2017) do not claim to explicitly test the stronger assumption of no jump *or* kink in the density of the assignment variable, required for the RKD. However, the Frandsen (2017) test allows the user to choose a degree of departure from linearity which is tolerated, by choosing the value of the bound coefficient k. A choice of $k = 0$ implies a null hypothesis of linearity and an alternative hypothesis of non-linearity around the threshold (i.e. jump or kink), which would mean that our assumption of smooth density fails. As a result, we set the bound coefficient to equal zero and report the p-value of this test.

Figure 3.2 presents graphically the density of the assignment variable by HbA1c. The density is neither uniform nor entirely smooth across the entire range of HbA1c levels, however it is clear that there is no graphical evidence of either a jump or a kink in the density at the kink point of 6% (red vertical line). The graph also shows the p-value from the Frandsen (2017) test, which is unable to reject the null of linearity across the threshold suggesting that the first identifying principle for our RKD holds. Such findings are not particularly surprising, given that by nature HbA1c is extremely difficult to exactly manipulate and influence around the threshold.

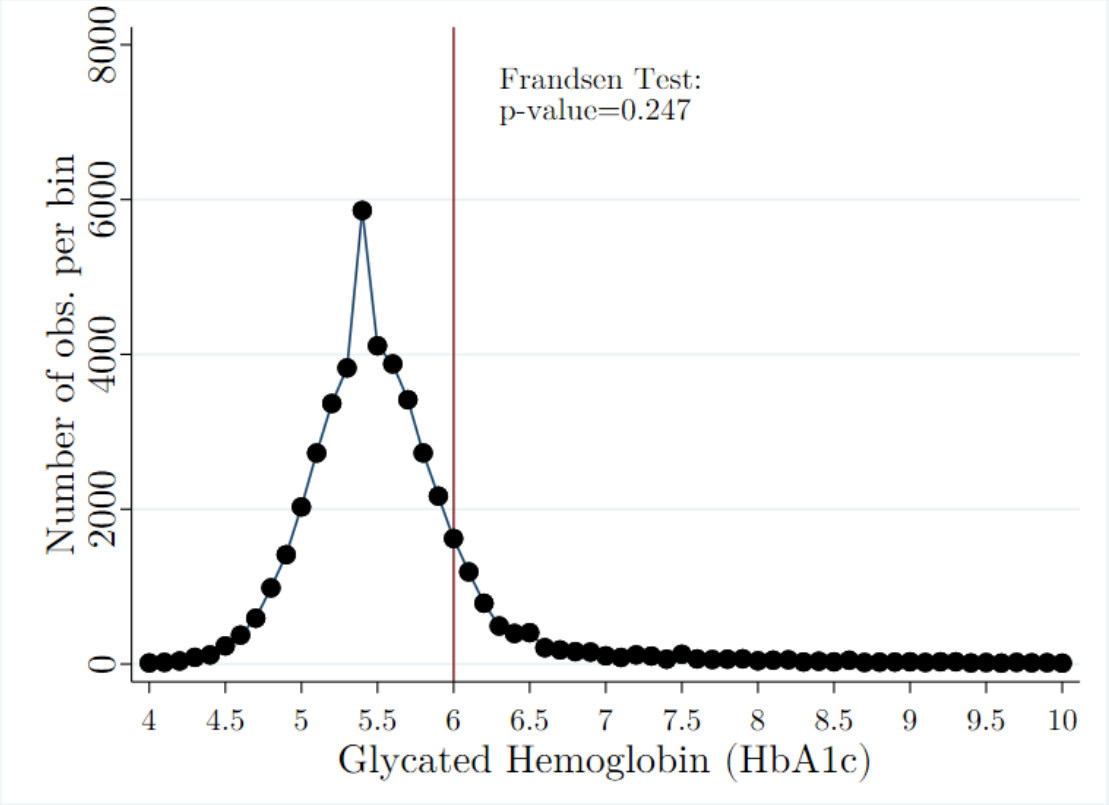


FIGURE 3.2: Smooth Density of the Assignment Variable

NOTE: Number of observations per bin. Bin width of 0.1 for glycated hemoglobin levels between 4 and 10. Graph also shows Frandsen (2017) discontinuity statistic.

3.5.2 Predetermined Variables

This assumption is similar to the "test of random assignment" commonly required in randomized control trials. As above, this observable implication is more restrictive than the equivalent RDD implication as in addition to the lack of any discontinuity it also requires the lack of any kink in the pre-determined variables. We assess whether the observable assumption holds in our setting by visual inspection and graphically

present the mean values per bin by the assignment variable for a number of predetermined variables.

Card et al. (2015) make clear this observable implication relies on the existence of a set of variables which, by definition, are not determined by the treatment. As such, we are somewhat limited in terms of the variables available at our disposal for testing. HSE is a cross-sectional study and most survey questions refer to specific points in time without eliciting information about the past, and in the cases where they do, it is unknown if such information relates to periods prior or post treatment. However, we examine a number of relevant variables, namely age, gender, self-reported health, whether individual has degree level education, whether the individual has any educational qualifications [9], whether a partner lives in the household, whether ever a smoker and whether ever a drinker.

Graphical results are given in Figure 3.3. There is no evidence of clear discontinuities or kinks at the kink point for any of the variables presented here, validating our second necessary assumption and suggesting that interpretation of the results of the RKD as MTEs is valid.

3.6 Main estimation results

3.6.1 Effect of own diagnosis

Table 3.2 presents estimates of the effect of own diabetes diagnosis on own behaviour. The relevance of the kink as an instrument for ever being diagnosed with diabetes is given in the first stage coefficients available in appendix table 3.A1 with results suggesting a highly statistically positive significant effect of the kink on probability of being diagnosed with diabetes. The first row of Table 3.2 gives the coefficient β_1 from equation (3.3). We find that being diagnosed with diabetes significantly increases the probability of having done some physical activity in the last four weeks and significantly reduces the probability of currently being a smoker. We find no evidence to suggest an impact on consumption of fruit or vegetable, and there is no evidence to suggest that diabetes diagnosis changes drinking behaviour. [10]

[9] Any qualification corresponds to a long list of education qualifications surveyed in the HSE, which include (but not limited to) degree education, high school and professional qualifications (i.e. teaching, nursing, vocational).

[10] In addition to the estimates presented, Figure 3.A21 in the Appendix shows the reduced form quadratic prediction graphically imposed over the mean outcomes per bin for HbAqc levels, where the reduced form estimates are from $Y_i = \chi_0 + \chi_1(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \psi_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \psi_p^+(x_i - k)^p D_i\right] + \mu_i$. The graphs show similar results to the 2SLS estimated with physical activity having the clearest slope change around the kink point, whereas fruit, smoking and alcohol consumption show a far more subtle changes in slope.

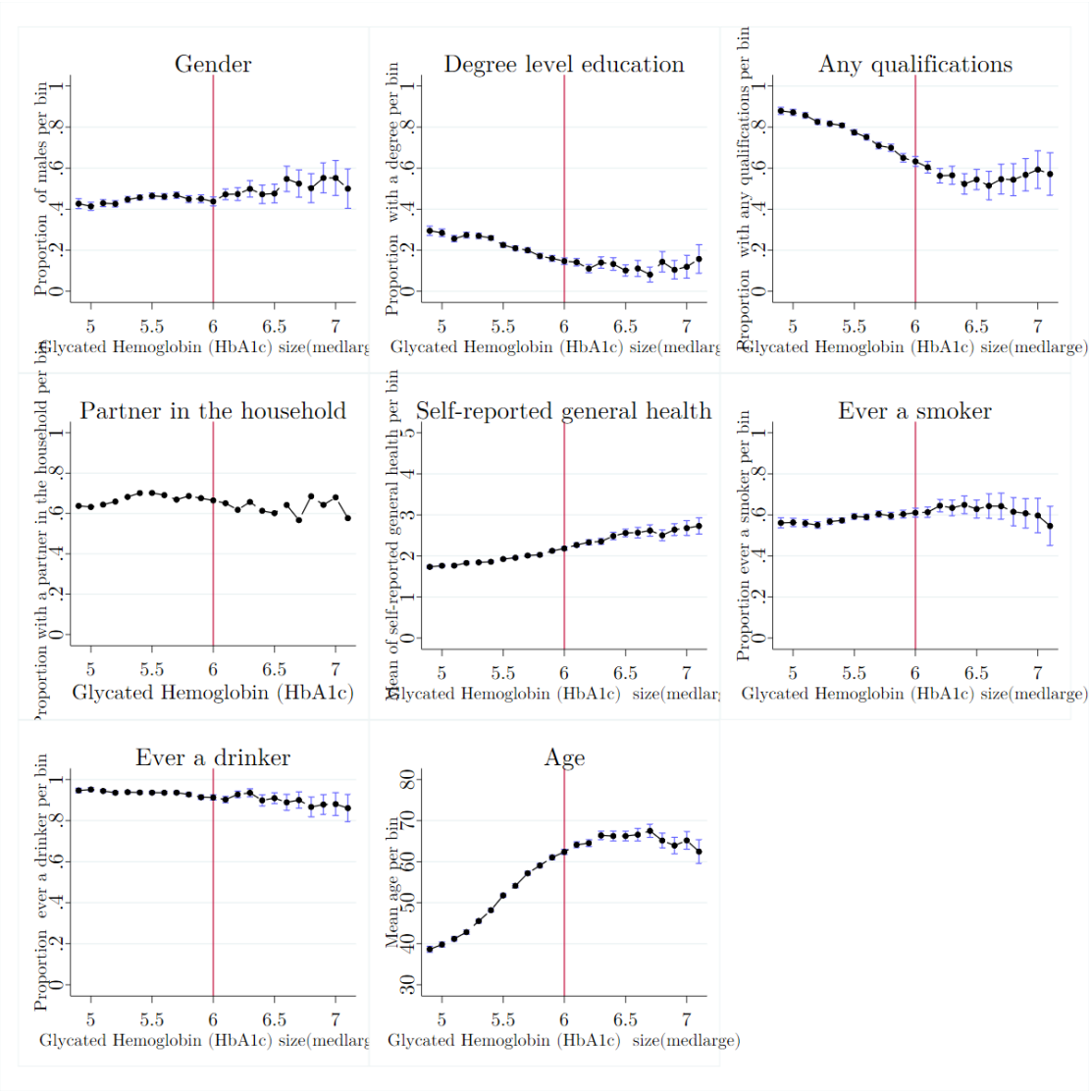


FIGURE 3.3: Predetermined variables

NOTE: Graphical representation of the mean of each predetermined variable by glycated hemoglobin (HbA1c) level. Each graph shows the mean of the predetermined variable per bin, with a bin width of 0.1. 95% confidence intervals are represented by the blue lines. Predetermined variables included are gender, ethnicity, degree level education, any qualifications, whether a partner lives in the household, whether ever a smoker, whether ever a drinker and age. Red line represents the kink point of 6.0 %.

3.6.1.1 Sensitivity to alternative bandwidths and polynomials

To assess the sensitivity of results to alternative specifications and bandwidths we explore a series of robustness graphs in appendix figures 3.A1 to 3.A5. Graphs show the point estimate, β_1, and the corresponding 90% and 95% confidence interval, from equation 3.3, estimated using 2SLS for each Y_i outcome of the main analysis. Specifications vary by polynomial order (i.e. linear or quadratic) and the selected

	Exercise	Vegetable Consumption	Fruit Consumption	Currently a Smoker	Currently a drinker
OLS Estimates					
Effect of Own Diabetes	-0.0978***	-0.0009	0.0405***	-0.0403***	-0.0671***
	(0.0090)	(0.0073)	(0.0069)	(0.0074)	(0.0048)
Obs.	42,407	82,070	82,130	47,427	82,828
RKD Estimates					
Effect of Own Diabetes	0.203***	0.0376	0.0650	-0.414***	0.00843
	(0.0688)	(0.0480)	(0.0454)	(0.0562)	(0.0248)
First Stage $F-Statistic$	562.06	1505.81	1505.49	932.48	1546.82
Obs.	20641	39666	23432	44828	41686

Notes: RKD coefficients are estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 4.0 on the right hand tail, and 2.0 on the left hand side for panels. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification includes the following controls: Age, and dummies for whether individual i is male, a partner lives in the household, and has degree level education. OLS coefficients estimated using equation $Y_i = \theta_0 + \theta_1 EverD_i + \theta_2 W_i + e_i$, they include all observations in sample, and the same controls W_i as the RKD estimates. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.2: Fuzzy RKD estimates of change in own behaviour as a result own diabetes diagnosis

bandwidths for above and below the cutoff (bounds of the estimation sample). The upper bound describes the relative bandwidth above the kink point with the lower bound being the corresponding bandwidth below the kink point, (i.e. a lower bound of 2 corresponds to a HbA1c value of 4%, a bandwidth of 2% below the kink-point of 6%. An upper bound of 3 corresponds to a HbA1c value of 9%, a bandwidth of 3% above the kink-point).

Inspecting Figure 3.A1, for physical activity, across all specifications point estimates are above zero and in almost all cases confidence intervals exclude zero. Overall, results seem robust with physical activity estimates not being overly sensitive to specification chosen.

Vegetable consumption and fruit consumption estimates in Figures 3.A2 and 3.A3, respectively, follow a similar pattern to one another. For quadratic specifications the estimates are both close to zero in magnitude, and have a relatively tight confidence interval which includes zero in almost every case. However, for both fruit and vegetable the linear specifications seem to have a positive and significant effect. We are cautious in claiming that an effect exists for either outcome, given that our main specification, a quadratic polynomial, supports a null effect, and that significance of these estimates are clearly specification dependent. We therefore conservatively claim lack of evidence of an effect of diabetes on vegetable or fruit consumption.

Findings for smoking behaviour, Figure 3.A4, are similar to those of physical activity with point estimates varying little across specifications and all specifications featuring tight confidence intervals excluding zero. Estimates from a quadratic specification

appear to be very robust and all sitting within a small interval around -0.3 also with tight confidence intervals.

Finally, alternative specifications for the effect of diabetes diagnosis on alcohol consumption are presented in Figure 3.A5. Almost all specifications have confidence intervals which include zero and are also tightly bounded around zero, especially for our preferred specifications with a quadratic polynomial.

3.6.2 Spillover effect

The spillover estimates as a result of partners' diabetes diagnosis, i.e. parameter δ_1 in eq. 3.5, are presented in Table 3.3. In this case, partner's kink is used as an instrument for partner diabetes diagnosis and its relevance is given in the first stage estimates implying very good identification properties. 2SLS estimates are presented in first row, with findings suggesting very similar patterns to those of own diabetes diagnosis[11]. Specifically, we find significant positive effects for exercising in the past four weeks and significant negative effects for currently being smoker, in the former the magnitude is similar to that of the effect of own diagnosis and about half as large for the latter. There is some suggestive evidence of a change in fruit consumption, however these results are not robust when we look at the sensitivity to alternative specifications.

3.6.2.1 Sensitivity to alternative bandwidths and polynomials

We additionally assess the sensitivity of our spillover estimates in appendix figures 3.A6 to 3.A10. Broadly these figures follow similar patterns to those for the effect of own diabetes diagnosis. One point of difference is that confidence intervals for spillover effects are substantially larger than those for own behaviour. This is to be expected given differences in the estimation sample sizes between spillover and own effects. Indeed, we find that large confidence intervals are especially present in specifications with narrow bandwidths or higher order polynomials, and therefore power might be of concern in these cases. Nevertheless, the pattern for figures 3.A6 to 3.A10 follow a similar pattern to the effect on own, and indeed the results for physical activity, and smoking do not appear to be sensitive to specification and the majority of specifications are significantly different from zero.

[11] Reduced form RKD estimates from $Y_i = \chi_0 + \chi_1(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \psi_p^-(x_i-k)^p\right] + \left[\sum_{p=2}^{p^*} \psi_p^+(x_i-k)^p D_i\right] + \mu_i$ are plotted in Table 3.A22 in the Appendix. Physical activity once again exhibits the most prominent slope change, with little evidence of a slope change elsewhere.

	Exercise	Vegetable Consumption	Fruit Consumption	Currently a Smoker	Currently a drinker
OLS Estimates					
Partner's Diabetes	-0.116***	-0.0154	-0.0140	0.0155	-0.0260***
	(0.0122)	(0.0103)	(0.00978)	(0.0110)	(0.00564)
Observations	19589	37789	37800	21064	38165
RKD Estimates					
Partner's Diabetes	0.235**	0.0166	-0.0907	-0.227***	0.0372
	(0.0967)	(0.0666)	(0.0626)	(0.0738)	(0.0315)
First Stage $F-Statistic$	281.56	758.24	758.54	433.05	771.09
Obs.	10581	20013	20015	11313	20941

Notes: Coefficients are estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 4.0 on the right hand tail, and 2.0 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification includes the following controls: Age, and dummies for whether individual i is male, and has degree level education. Additionally these estimates include the same set of controls for individual j. OLS coefficients estimated using equation $Y_i = \theta_0 + \theta_1 EverD_j + \theta_2 W_i + \theta_3 W_j + e_i$, they include all observations in sample, and the same controls (W_i and W_j) as the RKD estimates. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.3: Fuzzy RKD estimates of change in own behaviour as a result of partner's diabetes diagnosis

3.7 Robustness checks

3.7.1 Simultaneous Own and Partner's diabetes status

Having obtained evidence for the consistency of the RKD estimations in our setting we pursue sensitivity issues and examine the robustness of the effect of own and partner diabetes diagnoses on own behaviour when both effects are simultaneously identified and estimated. In this specification own and partners' kinks are used as instruments for own and partners' probability of being diagnosed diabetic. Two separate first stage estimations are required, one equation for own, $z = i$, and one for partner, $z = j$.

$$EverD_z = \eta_0 + \eta_1(x_i - k)D_i + \eta_2(x_j - k)D_j + \left[\sum_{p=1}^{p^*} \chi_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \chi_p^+(x_i - k)^p D_i\right] + \left[\sum_{p=1}^{p^*} \zeta_p^-(x_j - k)^p\right] + \left[\sum_{p=2}^{p^*} \zeta_p^+(x_j - k)^p D_j\right] + q_z \quad (3.6)$$

Obtaining predicted probabilities for both equations, the second stage is correspondingly defined as

$$Y_i = \kappa_0 + \kappa_1 \widehat{EverD}_i + \kappa_2 \widehat{EverD}_j + \left[\sum_{p=1}^{p^*} \pi_p^- (x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \pi_p^+ (x_i - k)^p D_i\right] + \left[\sum_{p=1}^{p^*} \phi_p^- (x_j - k)^p\right] + \left[\sum_{p=2}^{p^*} \phi_p^+ (x_j - k)^p D_j\right] + r_i \quad (3.7)$$

Results are given in Table 3.4. Own diabetic diagnosis, increases the probability of exercise, increases the probability of fruit consumption and decreases the probability of currently smoking. Partner's diagnosis also increases the probability of exercise, however the effect for smoking behaviour is lost in these specifications.

Overall, findings confirm the main analysis albeit for some specifications significance is reduced substantially. This is the result of smaller sample sizes and reduced estimation power. We note that given the set-up, the relevant estimation sample only includes those who have HbA1c levels within the bandwidths, have partners, and those partners also have HbA1c levels within the bandwidths. Further, in support of power issues as the reason behind lower significance levels, we note that comparisons of the corresponding 2SLS estimates between tables 3.2, 3.3 and 3.4 suggest that for the vast majority of models, coefficients magnitudes are comparable, and indeed are almost identical for physical activity and smoking behaviour, but effects in table 3.4 are estimated with less precision and hence much higher standard errors.

3.7.2 Falsification Tests

As additional robustness checks, we present falsification tests, where we use the identification strategy presented in section 3.4.1 to estimate the effect on several outcomes which we *a priori* expect to be zero. Estimating a null effect in outcomes which we do not expect to be effected by a diabetes diagnosis provides further evidence that our identification strategy is valid and our estimated effects are not spurious.

We analyse the effect on a set of three other medical outcomes, namely whether individuals take: antibiotics, anti-depressants or statins. We, further, include one other pre-determined variable, whether ever been in paid employment, to extend the falsification checks beyond only medical outcomes. In addition, to check the robustness of identification for disentangling own and spillover effects, we examine the own and spillover effect of diabetes diagnosis on whether currently taking anti-diabetic medication. In this case, we would expect to find a strong own effect, but

	Exercise	Vegetable Consumption	Fruit Consumption	Currently a Smoker	Currently a drinker
OLS Estimates					
Own Diabetes	-0.112***	0.00708	0.0364***	-0.0312***	-0.0625***
	(0.0118)	(0.00981)	(0.00939)	(0.00976)	(0.00632)
Partner's Diabetes	-0.107***	-0.0157	-0.0163*	0.0170	-0.0225***
	(0.0120)	(0.00983)	(0.00957)	(0.0109)	(0.00550)
Obs.	19,456	37,497	37,508	20,903	37,898
RKD Estimates					
Own Diabetes	0.214*	0.103	0.187**	-0.358***	0.0753*
	(0.116)	(0.0783)	(0.0753)	(0.0879)	(0.0430)
Partner's Diabetes	0.244**	0.0508	-0.121	-0.201**	0.0453
	(0.121)	(0.0782)	(0.0745)	(0.0928)	(0.0379)
First Stage $F-Statistic$	41.01	156.99	156.99	144.60	168.56
Obs.	8064	15055	15055	8408	15871

Notes: Coefficients are estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 6.0 on the right hand tail, and 3.0 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification includes the following controls: Age, and dummies for whether individual *i* is male, and has degree level education. Additionally these estimates include the same set of controls for individual *j*. OLS coefficients estimated using equation 3.1, they include all observations in sample, and the same controls as the RKD estimates. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.4: Fuzzy RKD estimates of change in own behaviour as a result of own and partner's diabetes diagnosis

no evidence of a spillover effect onto partners as partner's diabetes diagnosis should not *per se* increase probability of own receiving anti-diabetic medication.

Estimates of the effects on these outcomes are presented in Table 3.5, and, we additionally present the same robustness graphs for our main estimates in appendix figures 3.A11 - 3.A20. Firstly, as expected, we find clear evidence of an increase in probability of taking anti-diabetic medication for own diabetes diagnosis but no evidence of a spillover effect. In terms of our other estimates, reassuringly we find no evidence of an effect on any of the outcomes used in the falsification tests. The robustness graphs also support the results presented in Table 3.5, in almost all specifications we estimate we find null effects, aside from the own effect on antidiabetes medication, where there is clear significant effects across all specifications. All in all, testing strongly supports our identification strategy.

3.8 Causal Pathways

As discussed in detail in section 3.2.2, the correlation between spouses can theoretically be attributed to assortative matching, shared environment and joint

	Whether taking				Whether even been in paid employment
	Anti-diabetic medication	Antibiotic medication	Anti-depressant medication	Statins	
(a)					
Effect of Own Diabetes	0.883***	0.000726	-0.00786	-0.0207*	0.0374
	(0.0298)	(0.0271)	(0.0553)	(0.0109)	(0.0310)
First Stage $F-Statistic$	567.89	567.89	567.89	1432.29	965.46
Obs.	12138	12138	12138	34638	19546
(b)					
Partner's Diabetes	0.00881	0.0281	-0.0265	-0.000439	0.0694
	(0.0717)	(0.0292)	(0.0691)	(0.0138)	(0.0431)
First Stage $F-Statistic$	292.62	292.62	292.62	708.54	468.66
Obs.	5604	5604	5604	16528	9833

Notes: Coefficients are estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 4.0 on the right hand tail, and 2.0 on the left hand side for panels (a) and (b). Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification includes the following controls: Age, and dummies for whether individual *i* is male, a partner lives in the household, and has degree level education. Panel (b) additionally include the same set of controls for individual *j*, but excluding whether partner lives in the household. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.5: Fuzzy RKD estimates of change in own behaviour as a result own and partner's

household decision making. Our identification strategy allows us to plausibly exclude attributing spillover effects to assortative matching, which leaves us with two possible channels. The results we present in Table 3.3 from estimating equation 3.5 are the combined effect of these two pathways. In this section we seek to decompose spillover effects, and assess the contribution of shared environment and joint household production to the overall spillover effect. To do so we conduct a mediation analysis, where we separately identify changes in own behaviour that are the result of partner's diagnosis (i.e. direct effect of diagnosis), and changes in own behaviour that are the result of the induced change in partner's behaviours (i.e. indirect effect), see Figure 3.4 for illustration.

To assess the direct and indirect effects in this setting a typical a mediation analysis would estimate the following equations:

$$Y_i = \varphi_M^T \, EverD_i + e^M \tag{3.8}$$

$$Y_j = \varphi_Y^M \, Y_i + \varphi_Y^T \, EverD_i + e^Y \tag{3.9}$$

Where $\varphi_M^T \times \varphi_Y^M$ represents the indirect effect, or the effect of $EverD_i$ through the mediator Y_i, and φ_Y^T is the direct effect of the diagnosis on partner's behaviour. In this setting both Y_i and $EverD_i$ are endogenous, but we only have a single instrument, that being the kink in the fuzzy RKD framework. However, it is still possible to conduct a mediation analysis given this restriction. Dippel et al. (2020) outline a framework for

doing mediation analysis for cases with only one instrument. To do mediation analysis in this setting we require just one additional assumption, which we believe is reasonable in our setting, that being: "the confounding variable that jointly affects $EverD_i$ and Y_i is independent of the confounding variable that jointly causes Y_i and Y_j" (Dippel et al., 2020). We present and discuss such assumption and its implications below.

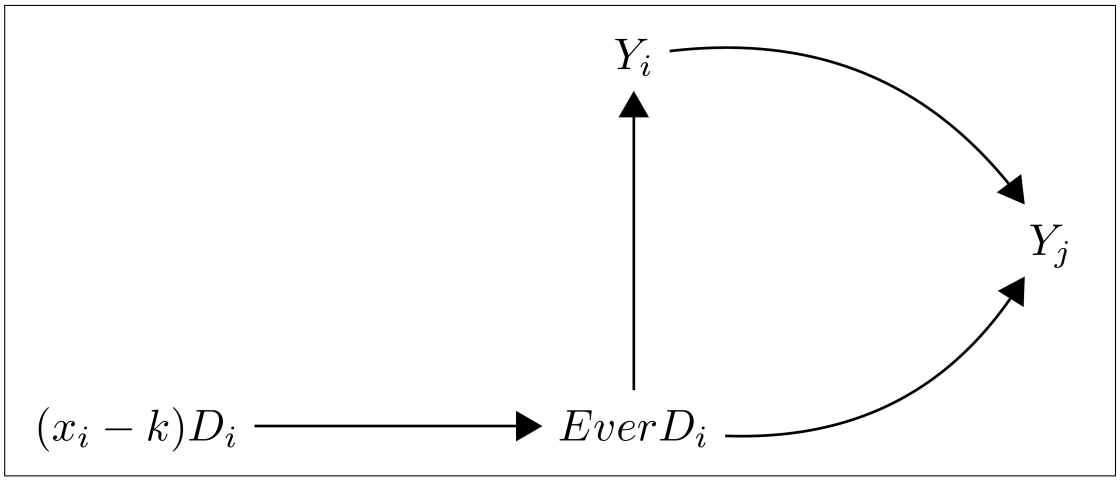


FIGURE 3.4: Causal Pathway of the spillover effect

NOTE: $(x_i - k)D_i$ denotes the kink, which we use as the instrument in the fuzzy RKD specification. $EverD_i$ is the diabetes status of individual i. Y_i is the health-related behaviour of individual i, and Y_j is the health-related behaviour of individual j. The pathway $EverD_i \rightarrow Y_j$ is considered to be the direct effect of individual i's diabetes diagnosis on the behaviours of individual j. The pathway $EverD_i \rightarrow Y_i \rightarrow Y_j$ is the indirect effect, where the diagnosis of i causes a change in j's behaviours which is the result of the induced change in i's behaviours. In other words, the effect of the diagnosis $EverD_i$ on Y_j, through the mediator Y_i.

To estimate the direct and indirect effects using a mediation model with one instrument, we follow the approach outlined by (Dippel et al., 2020). We estimate the following four equations:

$$EverD_i = \beta_T^Z (x_i - k)D_i + f(x_i - k) + \epsilon^T \tag{3.10}$$

$$Y_i = \beta_M^T \widehat{EverD_i} + f(x_i - k) + \epsilon^M \tag{3.11}$$

$$Y_i = \gamma_M^Z EverD_i + \gamma_M^T (x_i - k)D_i + f(x_i - k) + \xi^M \tag{3.12}$$

$$Y_j = \beta_Y^M \widehat{Y}_i + \beta_Y^T EverD_i + f(x_i - k) + \epsilon^Y \tag{3.13}$$

where $EverD_i$ is whether individual i has ever been diagnosed with diabetes, and $\widehat{EverD_i}$ is the predicted probability from eq. 3.10. Y_i denotes the health related behavioural outcome of interest, and $\widehat{Y}_i$ is the predicted equivalent from 3.12. x_i denotes the running variable (HbA1c level), and k is the kink point of 6%. $D_i = \mathbb{1}(x_i \geq k)$ is an indicator variable, taking the value of one if the individual's level

of HbA1c is above the kink point. $f(x)$ represents the polynomial function used throughout the analysis in this paper: $(\left[\sum_{p=1}^{p^*} \nu_p^- (x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \nu_p^+ (x_i - k)^p D_i\right])$. β_Y^T is the direct effect, and the indirect effect is $\beta_M^T \times \beta_Y^M$. Equations (3.10) and (3.11) are the same specifications as eqs. (3.2) and (3.3).

In addition to the usual exclusion restrictions for the instrument (i.e. the kink $(x_i - k)D_i$) in the Y_i and Y_j outcome equations (see section 3.4.1), causal estimation of direct and indirect effects, additionally requires that the confounder in Y_i and Y_j outcome equations be independent. More formally, we require that $\epsilon^T \perp\!\!\!\perp \epsilon^Y$, which is akin to stating that the confounding variable that jointly affects $EverD_i$ and Y_i is independent of the confounding variable that jointly causes Y_i and Y_j (Dippel et al., 2020). The implication of this assumption is that an additional exclusion restriction is required, such that our instrument can be used as an instrument for the mediator Y_i when conditioned on $EverD_i$ for the Y_j outcome equation $((x_i - k)D_i \perp\!\!\!\perp Y_j(Y_i) \mid EverD_i)$. It is important to note that this assumption does not assume away the endogeneity of $EverD_i$ in the Y_i outcome equation.

This identifying assumption is reasonable in our setting, as the unobserved confounder which causes bias in the Y_i outcome equation when estimating the impact of $EverD_i$, is different to the one that that causes the bias in Y_i in the Y_j outcome equation. As discussed in Section 3.4, when estimating the effect of $EverD_i$ on Y_i we are concerned with bias arising from simultaneity, where those that behave in a more damaging way for their health are more likely to receive a diabetes diagnosis. Whereas when estimating the impact of Y_i on Y_j the source of bias is assortative matching. However, one way in which this assumption may be violated is if own diabetes diagnosis impacts partner's behaviour through increasing the probability of partner being diagnosed with diabetes. In other words, if own diabetes status impacts partner's diabetes status directly (not through any other channel) and it is this that induces the changes in partner's behaviour. If the spillover effect worked through this channel we would expect the magnitude of the spillover effect to fall when controlling for own and partner's diabetes status in the same regression. However, as we show in Section 3.7.1 the magnitude of the spillover effect is nearly identical making such causal channel unlikely.

Testing the additional requirement that the instrument is relevant for the mediator Y_i when conditioned on $EverD_i$, we present F-statistics for eq. (3.12) in table 3.6. F-statistics values, as expected, are much smaller than for eq. 3.10, however, for the two outcomes in which we find evidence of a spillover effect (physical activity and tobacco consumption), they suggest our instrument is valid.

β_Y^T is an estimate of the effect of change in partner j's behaviour that is a result of partner i's diagnosis itself. We attribute this pathway to the health information causal channel. In this case, the diabetes diagnosis of partner i has a "direct" effect on partner

j's behaviours. As a result of the diagnosis, partner i receives new health information, possibly from a physician, about their diagnosed condition which they then share with the non-diagnosed partner j. The transfer of information from partner i to j therefore provides j with a new information set which they use to privately re-evaluate their optimal behaviour. This informational transfer may induce a change in partner j's behaviour if the new health information changes expected future payoffs. However, the magnitude of the effect is dependent on the pre-diagnosis information set, as well as idiosyncratic preferences.

The indirect effect $\beta_M^T \times \beta_Y^M$ captures the change in own behaviour that is caused by the induced change in partner's behaviours. This effect is attributed to the joint household decision making causal pathway. If jointly participating in these activities are complements, that is behaviours co-move independent of diabetes status or new health information, because individuals' gain utility from jointly participating in these behaviours, then it is reasonable to attribute the spillover to joint household decision making. The complementarity of these behaviours induce a change in partner j's behaviours as a result of i's diagnosis-induced behavioural change. This is clear in the case of smoking behaviour, as we would expect that quitting tobacco would be more difficult if another household member continued consuming tobacco. In terms of physical activity, individual j may get utility or dis-utility from exercising, however joint time with their partner may provide sufficient utility to render exercising a utility increasing choice.

Table 3.6 provides estimates of the direct and indirect effects from the mediation analysis. For physical activity and tobacco consumption we find that the spillover effect is driven by partner's behaviour Y_i, and we find limited evidence that the diagnosis itself is causing a change in behaviours of j. Results suggest that the estimated spillover effect we find is the result of joint household production rather than information sharing. For the remaining outcomes of vegetable consumption, fruit consumption and currently being a drinker, similarly to the absence of total effects, we find no evidence of direct or indirect effects.

3.9 Conclusion

Diabetes is a unique condition, in that a positive change in lifestyle and behaviour, is both the first line treatment and the recommended method of preventing the disease. By jointly partaking in diabetes treatment, partners of people with diabetes could substantially benefit from their partners' diabetes diagnosis. In this paper we estimate the causal effect of own or partner's diabetes status on own lifestyle behaviours, namely exercising, eating habits, smoking and drinking. Exploiting national guidelines around the levels of sugar in the blood and recommendation for annual

	Exercise	Vegetable Consumption	Fruit Consumption	Currently a Smoker	Currently a drinker
Total Effect	0.235**	0.017	-0.0901	-0.353***	0.032
	(0.0967)	(0.0667)	(0.0626)	(0.1052)	(0.0307)
Direct Effect					
Partner's Diagnosis ($EverD_i$)	-0.040	-0.017	-0.017	-0.005	0.009
	(0.0269)	(0.0171)	(0.0268)	(0.0341)	(0.0175)
Indirect Effect					
Partner's Behaviour (Y_i)	0.275**	0.034	-0.073	-0.345**	0.023
	(0.1278)	(0.0648)	(0.0953)	(0.1409)	(0.0329)
First Stage $F-Statistic$ Eq. 3.10: $(x_i-k)D_i$ on $EverD_i$	281.56	758.28	758.58	432.79	771.12
First Stage $F-Statistic$ Eq. 3.12: $(x_i-k)D_i$ on $Y_i\|EverD_i$	12.28	0.63	4.64	14.14	2.39
Obs.	10,581	20,011	20,013	7,004	20,286

Notes: The total effect corresponds to the coefficient δ_1 from equation 3.5, albeit for a slightly smaller sample in some cases. The direct effect corresponds to β_Y^T in equation 3.13, and the indirect effect corresponds to $\beta_M^T \times \beta_Y^M$ in equations 3.11 and 3.13. Each stage is estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 4.0 on the right hand tail, and 2.0 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. All stages include the following controls: Age, and dummies for whether individual i is male, a partner lives in the household, and has degree level education, as well as the same set of controls for individual j. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.6: Total, Direct and Indirect Effect Estimates from Mediation Analysis

testing for those above a specific threshold, a fuzzy kink regression design is implemented using data on blood samples and individual behaviours from the Health Survey for England (HSE) dataset.

Findings show that individuals who have ever been diagnosed with diabetes significantly increase their physical activity and reduce probability of currently being a smoker, suggesting compliance with first line treatment guidelines for diabetes. In analysis included in the appendix of this paper, we additionally find evidence of persistence over time in the effect, given that we observe individuals, on average 10 years post their initial diabetes diagnosis, and find no evidence of a heterogeneous on behaviours by time since diagnosis over time. Most importantly, we uncover substantial spillover effects from diabetes diagnosis in the form of an increase in physical activity and reduction in the probability of smoking for the partners of those diagnosed with diabetes. Through our identification strategy such effects are likely to be a combination of joint household decision making and health-related information transfer between partners.

Comparing our results of the own effect to those of previous studies, our estimated impact on diet differ to those of Hut and Oster (2018) and Oster (2018), and are

somewhat at odds with the impact on physical activity estimated by Kim et al. (2019). Hut and Oster estimated there to be significant and positive changes in diet post-diagnosis, and found that increased fruit purchases was the fourth largest contributor to these dietary changes. However, their results somewhat suggest that the improvements in diet begin to fade over time. They also find that single-person households do not significantly change their diet as a result of a diabetes diagnosis. Finally, they find that individuals with college education or higher improve their diet marginally more than the average as a result of a diagnosis. The findings of Oster (2018) follow a similar pattern to the results of Hut and Oster, in that calories purchased of fruit and vegetables both increase in the month post-diagnosis, however once again, the effect appears to decrease over time, and between months 2-12 post-diagnosis there is no significant increase in calories purchased of fruit and vegetables. Although our results do not directly confirm these studies we once again note the difference in time-since-diagnosis between studies and suggest that our findings largely follow the temporal pattern of those studies. Given that the average time since diagnosis in our sample is over 10 years, and that Hut and Oster and Oster both find decreasing effects over time, it might be expected that the effects reduce to zero in the long-run. However, when we analyse the temporal effects for diet, we again find no evidence that there are changes over time. Kim et al. finds there to be no significant increase in physical activity as a result of a diabetes diagnosis in either the short-run (1 or 2 years) or the long-run (3 or 4 years), whereas we find there to be both a significant and persistent change in physical activity as a result of a diabetes diagnosis.

Unfortunately, there are no studies to directly compare our estimated spillover effects onto partners to, albeit our broader conclusions do concur with previous studies, with the exception of Clark and Etilé (2006). Clark and Etilé found that the correlation between partners' smoking behaviour was driven mainly by matching in the marriage market, whereas our findings, as well as those of Fletcher and Marksteiner (2017), find there to be significant spillover effects in terms of smoking behaviour. In terms of alcohol consumption, comparisons with Fletcher and Marksteiner are harder, given that they investigate the own and spillover effects of alcoholism treatment, rather than a diabetes diagnosis. Unlike Janssen and Parslow (2021), we find no evidence in favour of a change in alcohol consumption as a result of the diabetes diagnosis, however our results concur with theirs in that both studies find evidence of persistent effects and evidence of a spillover effect in behaviours, albeit for different behaviours. Finally, although again we cannot directly compare our results to Fadlon and Nielsen (2019), both studies find significant health-related behavioural spillovers.

From a public health perspective, confirmation of long-term compliance of diabetics to first line treatments and necessary lifestyle changes is reassuring, at least in relation to physical activity and smoking. However, further work is required on how to induce

behavioural changes in terms of diet and alcohol consumption in diabetic patients. From a policy perspective, our findings suggest that benefit evaluation of diabetes interventions needs to be revisited in the presence of substantial spill-over effects, as their current benefit-cost ratio is likely to be substantially underestimated, especially in relation to physical activity and smoking of partners, from a diabetes diagnosis.

3.A Appendix

The appendices of this chapter present a number of robustness graphs and RKD plots, which are discussed in the main text. As well as estimates of the first-stage of the fuzzy RKD. We also present estimates of heterogenity in the impact of the own and spillover effects of a diabetes diagnosis by observable characteristics, and also an inspection of the kink-point location.

The robustness graphs show point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimate shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by: the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample. All possible combinations are presented, none of which are excluded from these figures. The white dot represents the main specification which we present in our tables in the main text. The upper bound describes the relative bandwidth above the kink point with the lower bound being the corresponding bandwidth below the kink point, (i.e. a lower bound of 2 corresponds to a HbA1c value of 4%, a bandwidth of 2% below the kink-point of 6%. An upper bound of 3 corresponds to a HbA1c value of 9%, a bandwidth of 3% above the kink-point).

Figures 3.A21 and 3.A22 show the reduced form quadratic prediction graphically imposed over the mean outcomes per bin for HbAqc levels, where the reduced form estimates are from
$Y_i = \chi_0 + \chi_1(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \psi_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \psi_p^+(x_i - k)^p D_i\right] + \mu_i.$

3.A.1 First-Stage Estimates

	Exercise	Vegetable Consumption	Fruit Consumption	Currently a Smoker	Currently a drinker
Second-Stage					
Effect of Own Diabetes	0.203***	0.0376	0.0650	-0.414***	0.00843
	(0.0688)	(0.0480)	(0.0454)	(0.0562)	(0.0248)
First-Stage					
$(x_i - k)D_i$	0.675***	0.738***	0.738***	0.700***	0.730***
	(0.0285)	(0.0190)	(0.0190)	(0.0229)	(0.0186)
First Stage $F - Statistic$	562.06	1505.81	1505.49	932.48	1546.82
Obs.	20641	39666	23432	44828	41686

Notes: RKD coefficients are estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 4.0 on the right hand tail, and 2.0 on the left hand side for panels. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification includes the following controls: Age, and dummies for whether individual i is male, a partner lives in the household, and has degree level education. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.A1: First and Second-Stage Estimates of the Fuzzy RKD estimates of change in own behaviour as a result of own diabetes diagnosis

	Exercise	Vegetable Consumption	Fruit Consumption	Currently a Smoker	Currently a drinker
Second-Stage					
Partner's Diabetes	0.235**	0.0166	-0.0907	-0.227***	0.0372
	(0.0967)	(0.0666)	(0.0626)	(0.0738)	(0.0315)
First-Stage					
$(x_j - k)D_j$	0.678***	0.743***	0.743***	0.735***	0.738***
	(0.0404)	(0.0270)	(0.0270)	(0.0353)	(0.0266)
First Stage $F - Statistic$	281.56	758.24	758.54	433.05	771.09
Obs.	10581	20013	20015	11313	20941

Notes: Coefficients are estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 4.0 on the right hand tail, and 2.0 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification includes the following controls: Age, and dummies for whether individual i is male, and has degree level education. Additionally these estimates include the same set of controls for individual j. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.A2: First and Second-Stage Estimates of the Fuzzy RKD estimates of change in own behaviour as a result own, partner's, own and partner's diabetes diagnosis

3.A.2 Robustness Graphs

	Exercise	Vegetable Consumption	Fruit Consumption	Currently a Smoker	Currently a drinker
Second-Stage					
Own Diabetes	0.214*	0.103	0.187**	-0.358***	0.0753*
	(0.116)	(0.0783)	(0.0753)	(0.0879)	(0.0430)
Partner's Diabetes	0.244**	0.0508	-0.121	-0.201**	0.0453
	(0.121)	(0.0782)	(0.0745)	(0.0928)	(0.0379)
First-Stage					
$(x_i - k)D_i$	0.525***	0.600***	0.600***	0.600***	0.595***
	(0.0402)	(0.0274)	(0.0274)	(0.0330)	(0.0263)
$(x_j - k)D_j$	0.525***	0.600***	0.600***	0.597***	0.605***
	(0.0402)	(0.0274)	(0.0274)	(0.0352)	(0.0261)
First Stage $F - Statistic$	41.01	156.99	156.99	144.60	168.56
Obs.	8064	15055	15055	8408	15871

Notes: Coefficients are estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 6.0 on the right hand tail, and 3.0 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification includes the following controls: Age, and dummies for whether individual i is male, and has degree level education. Additionally these estimates include the same set of controls for individual j. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.A3: First and Second-Stage Estimates of the Fuzzy RKD estimates of change in own behaviour as a result of own and partner's diabetes diagnosis

	Whether taking				Whether even been in paid employment
	Anti-diabetic medication	Antibiotic medication	Anti-depressant medication	Statins	
Own Effect (a)					
Second-Stage					
Effect of Own Diabetes	0.883***	0.000726	-0.00786	-0.0207*	0.0374
	(0.0298)	(0.0271)	(0.0553)	(0.0109)	(0.0310)
First-Stage					
$(x_i - k)D_i$	0.776***	0.776***	0.776***	0.753***	0.747***
	(0.0325)	(0.0325)	(0.0325)	(0.0199)	(0.0240)
First Stage $F - Statistic$	567.89	567.89	567.89	1432.29	965.46
Obs.	12138	12138	12138	34638	19546
Spillover Effect (b)					
Second-Stage					
Partner's Diabetes	0.00881	0.0281	-0.0265	-0.000439	0.0694
	(0.0717)	(0.0292)	(0.0691)	(0.0138)	(0.0431)
First-Stage					
$(x_j - k)D_j$	0.826***	0.826***	0.826***	0.774***	0.749***
	(0.0483)	(0.0483)	(0.0483)	(0.0291)	(0.0346)
First Stage $F - Statistic$	292.62	292.62	292.62	708.54	468.66
Obs.	5604	5604	5604	16528	9833

Notes: Coefficients are estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 4.0 on the right hand tail, and 2.0 on the left hand side for panels (a) and (b). Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification includes the following controls: Age, and dummies for whether individual *i* is male, a partner lives in the household, and has degree level education. Panel (b) additionally include the same set of controls for individual *j*, but excluding whether partner lives in the household. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.A4: First and Second-Stage Estimates of the Fuzzy RKD estimates of change in own behaviour as a result own and partner's

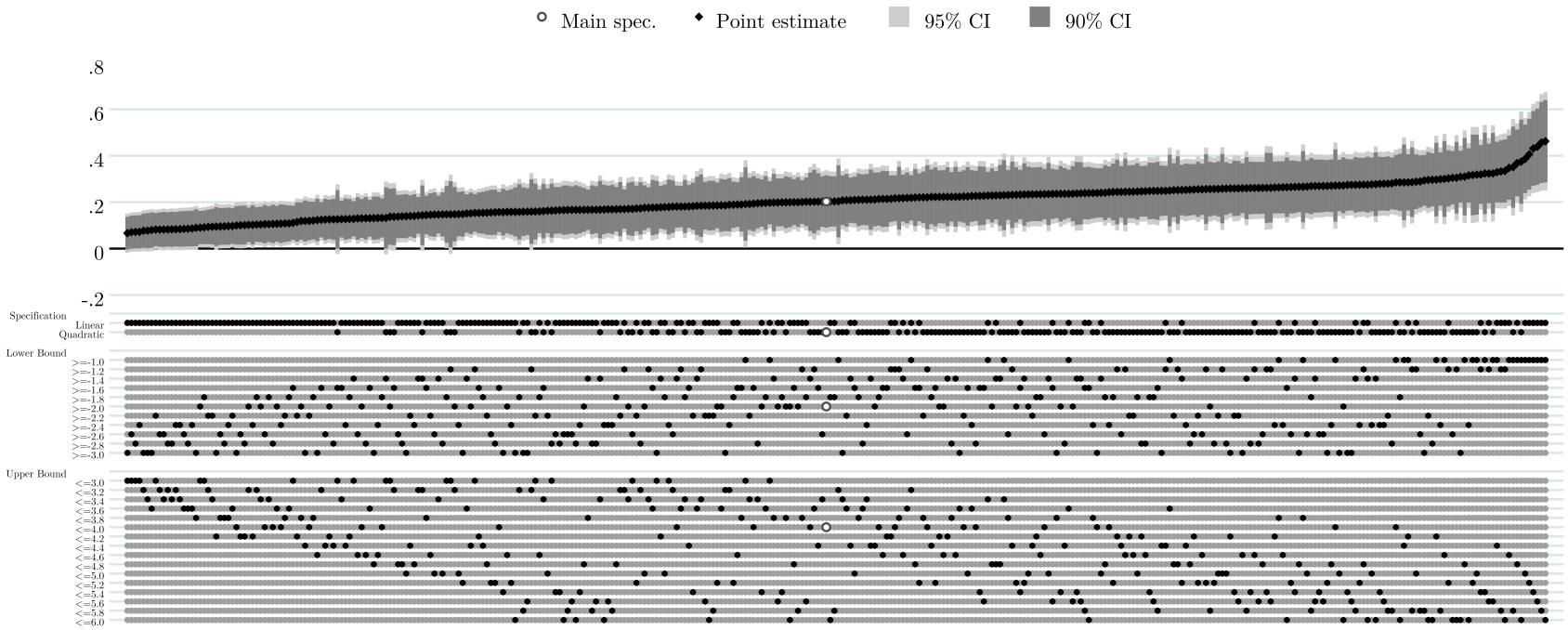


FIGURE 3.A1: Sensitivity to alternative bandwidths and polynomials - Physical Activity

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

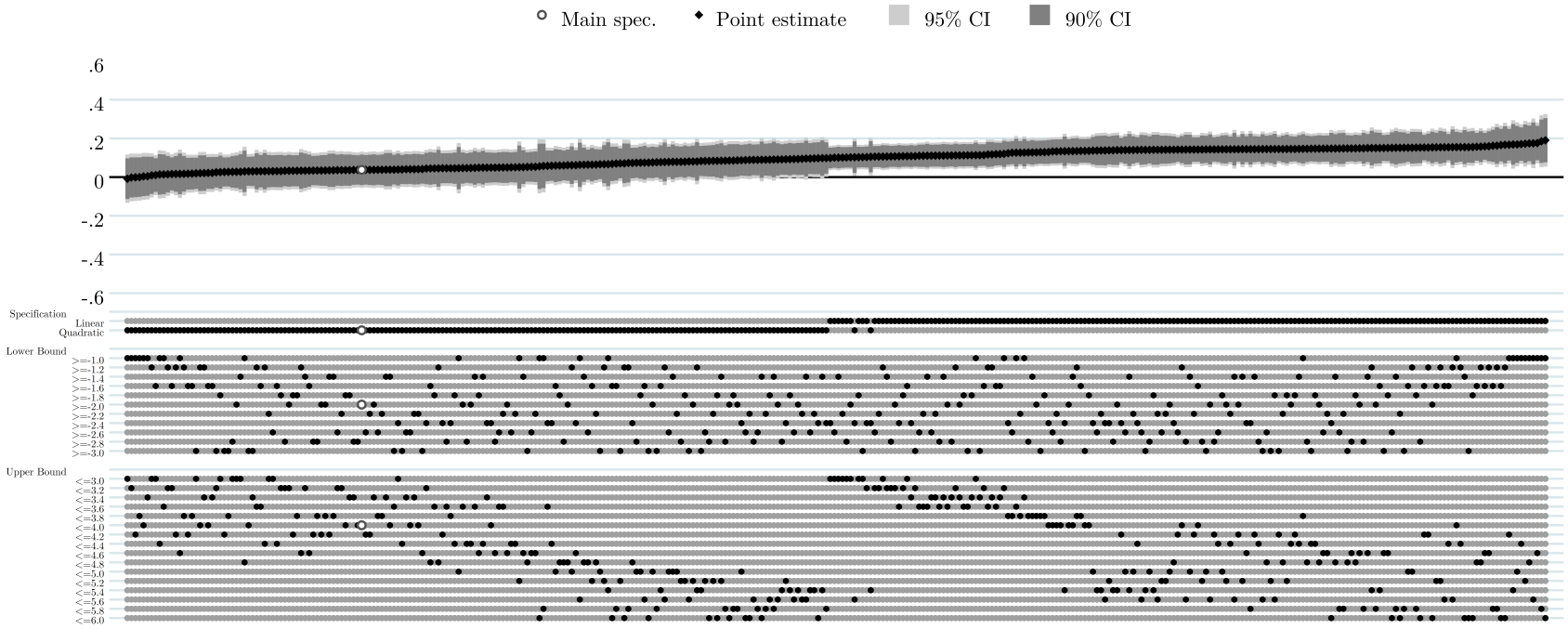


FIGURE 3.A2: Sensitivity to alternative bandwidths and polynomials - Vegetable Consumption

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

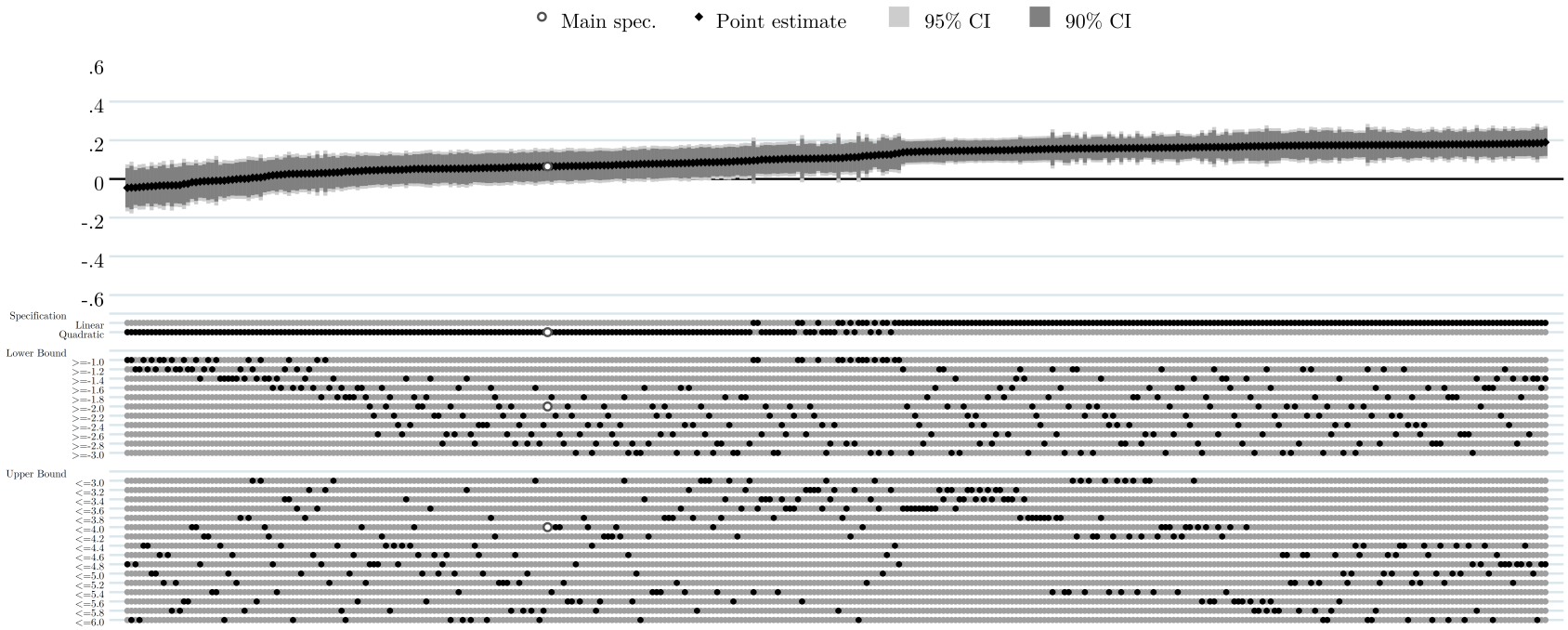


FIGURE 3.A3: Sensitivity to alternative bandwidths and polynomials - Fruit Consumption

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

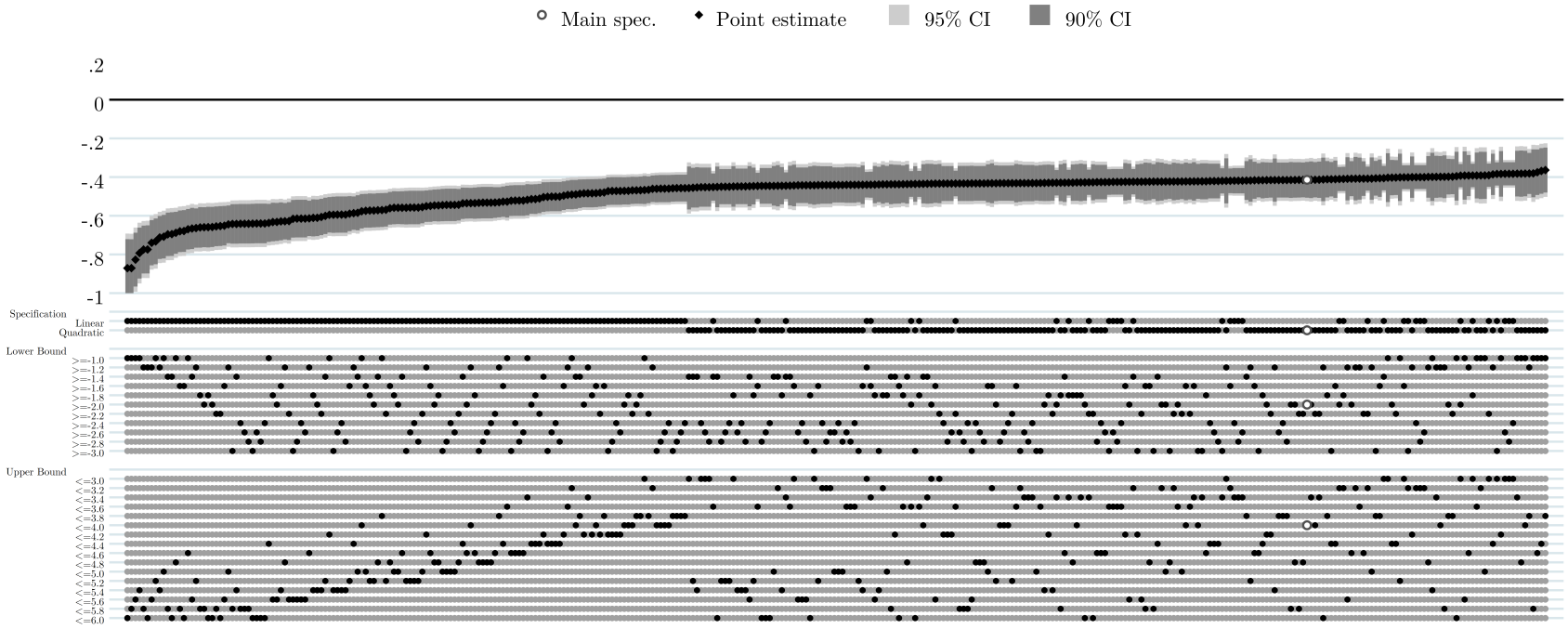


FIGURE 3.A4: Sensitivity to alternative bandwidths and polynomials - Smoking Behaviour

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

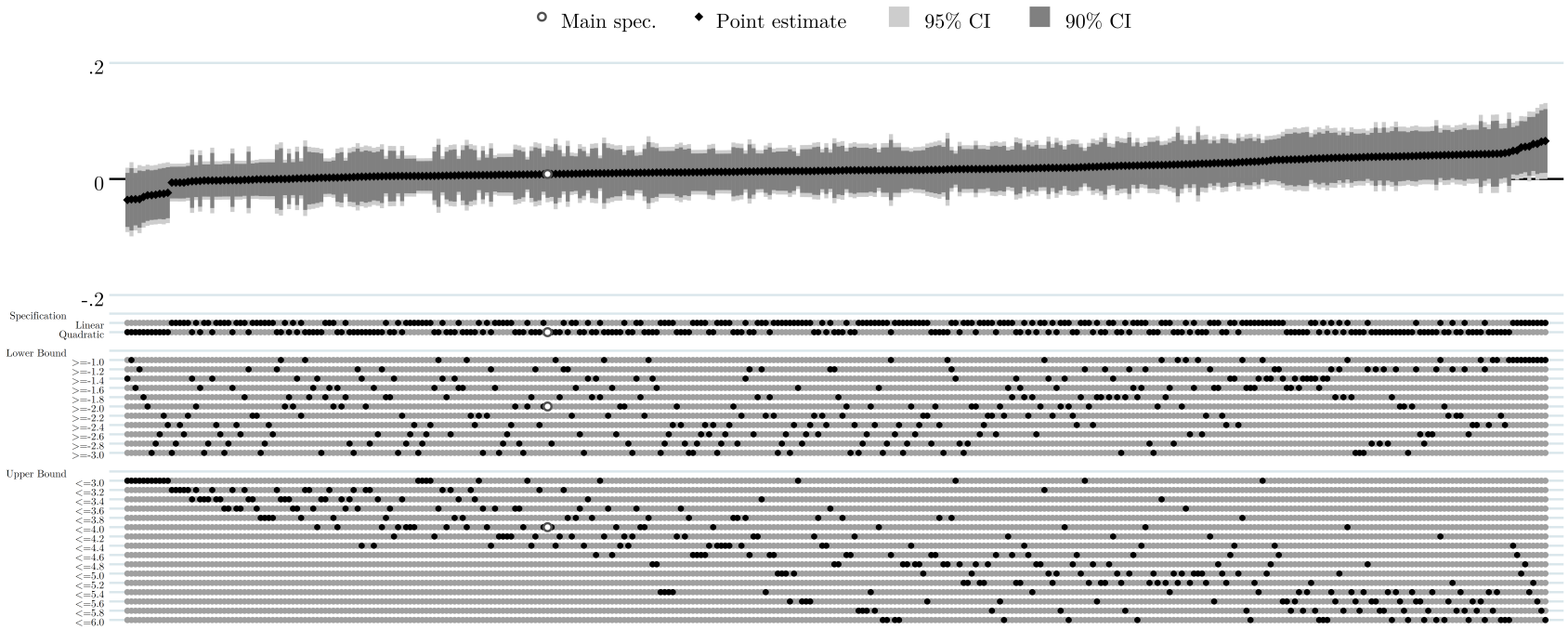


FIGURE 3.A5: Sensitivity to alternative bandwidths and polynomials - Alcohol Consumption

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

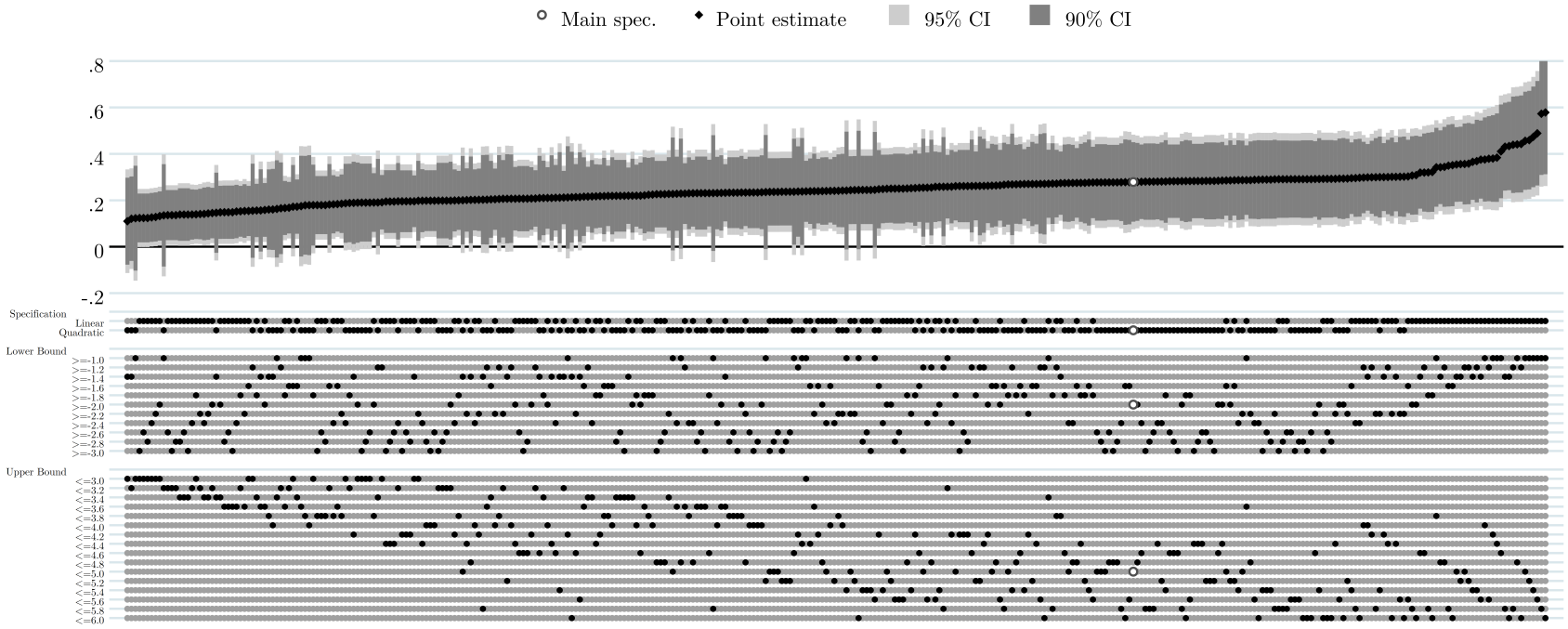


FIGURE 3.A6: Sensitivity to alternative bandwidths and polynomials - Partner Spillover Estimates of Physical Activity

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

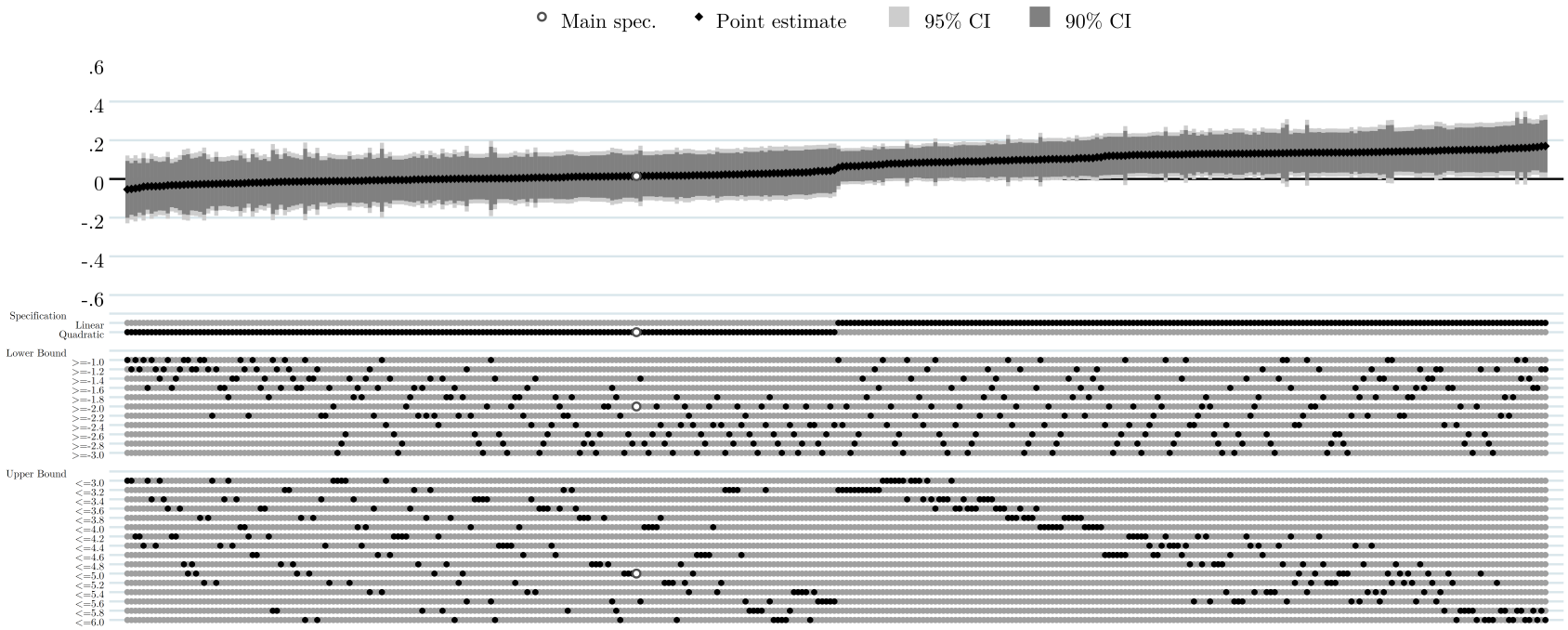


FIGURE 3.A7: Sensitivity to alternative bandwidths and polynomials - Partner Spillover Estimates of Vegetable Consumption

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

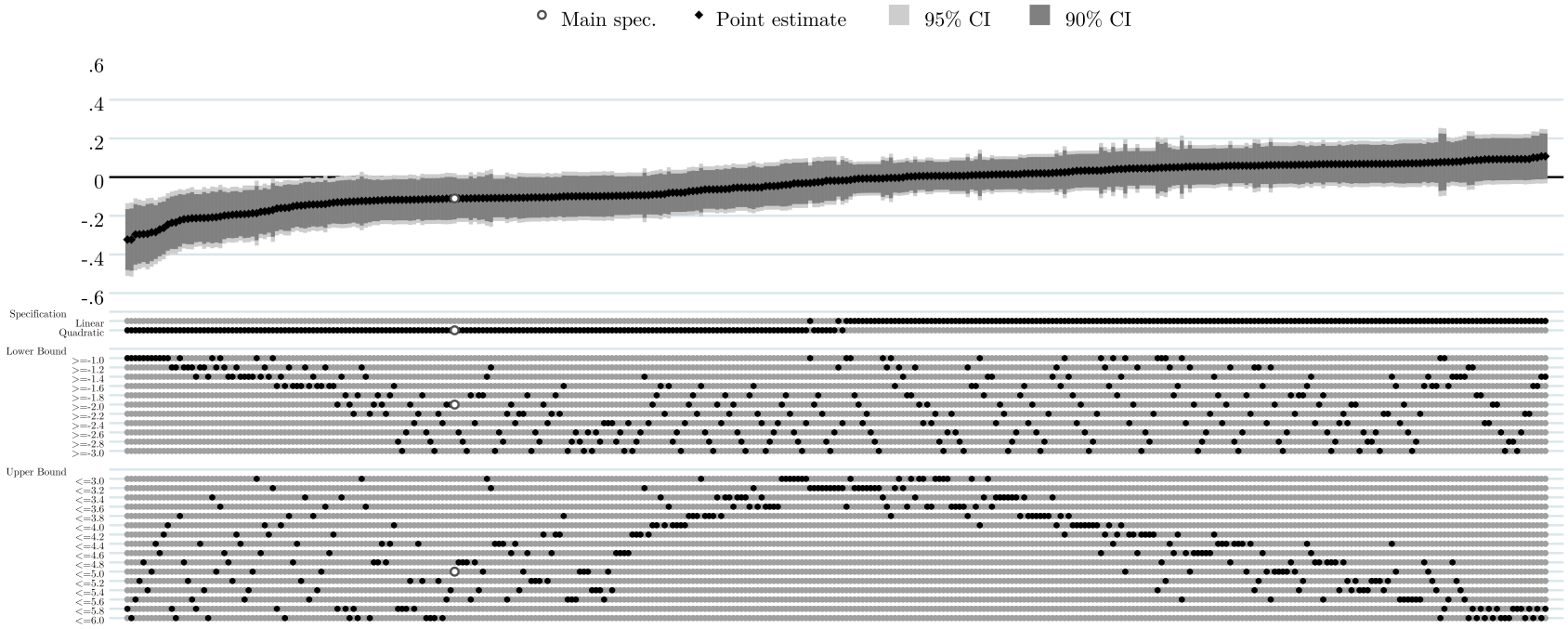


FIGURE 3.A8: Sensitivity to alternative bandwidths and polynomials - Partner Spillover Estimates of Fruit Consumption

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

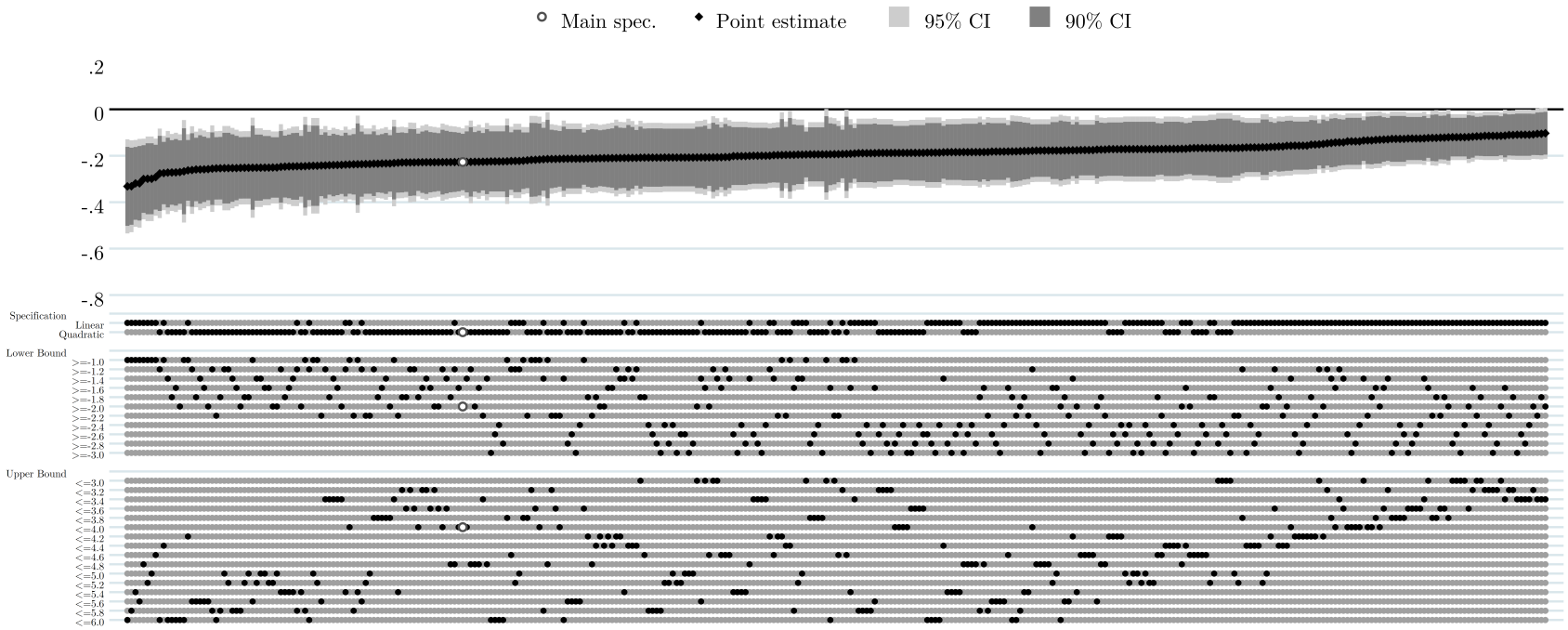


FIGURE 3.A9: Sensitivity to alternative bandwidths and polynomials - Partner Spillover Estimates of Smoking Behaviour

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

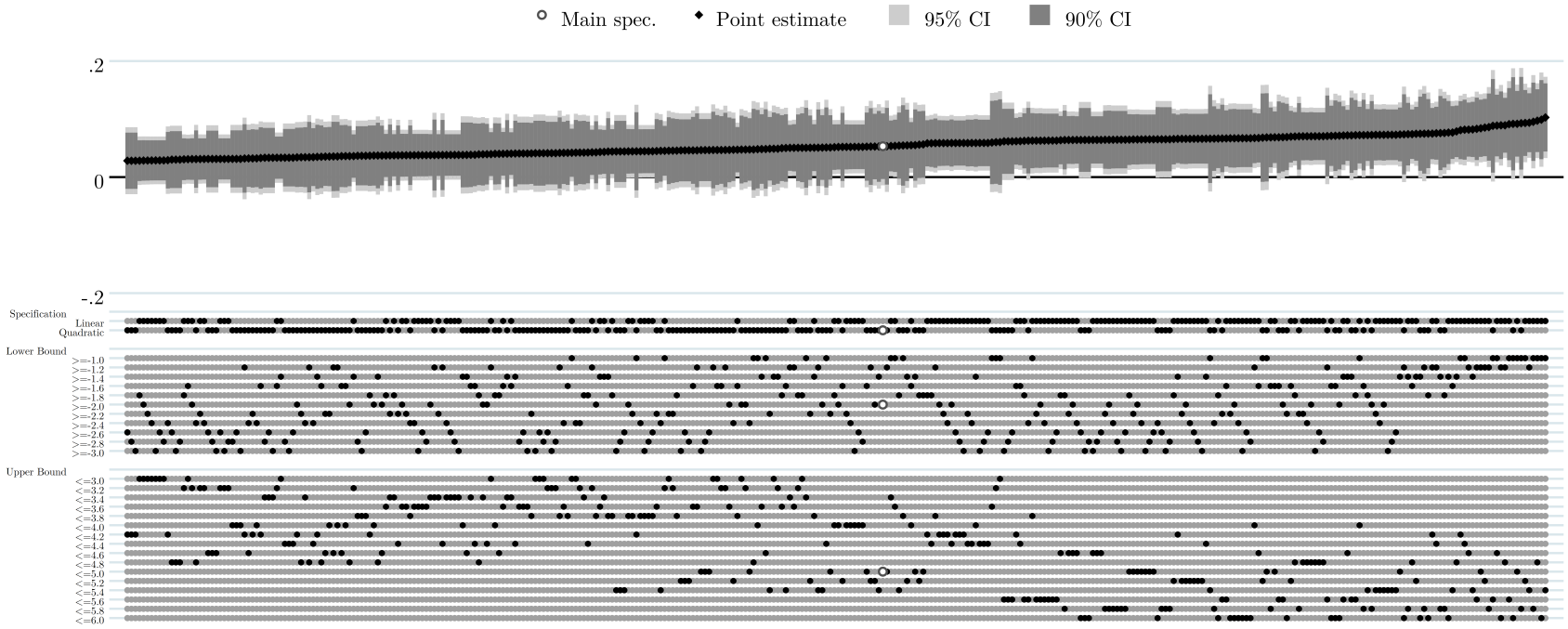


FIGURE 3.A10: Sensitivity to alternative bandwidths and polynomials - Partner Spillover Estimates of Alcohol Consumption

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

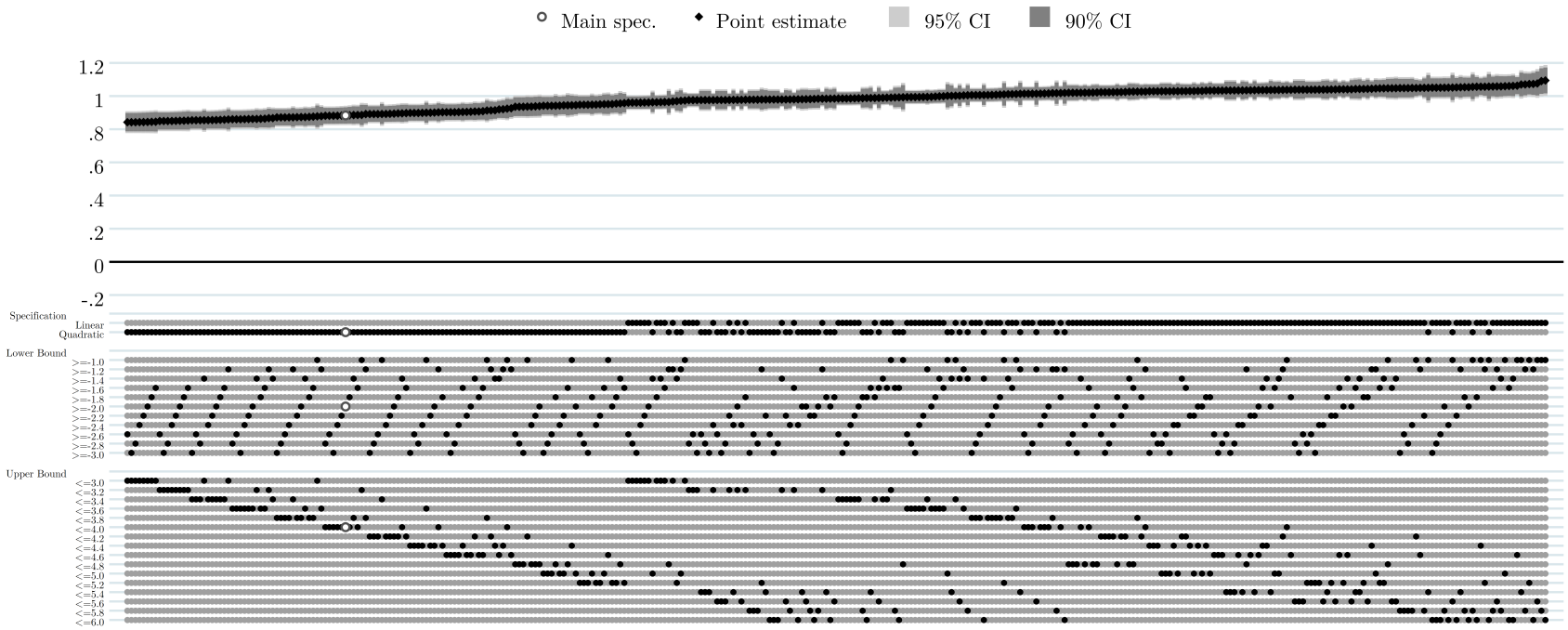


FIGURE 3.A11: Sensitivity to alternative bandwidths and polynomials - Whether taking Anti-diabetic medication

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

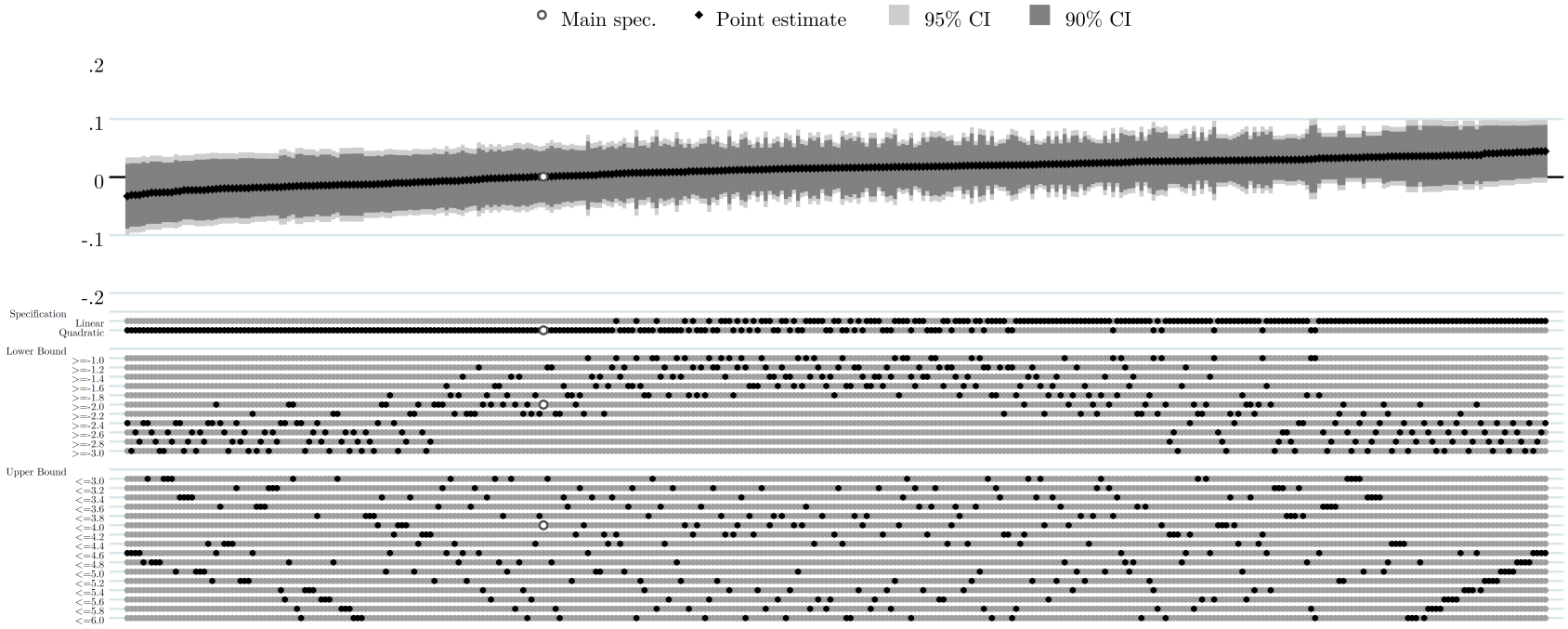


FIGURE 3.A12: Sensitivity to alternative bandwidths and polynomials - Whether taking Antibiotic medication

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

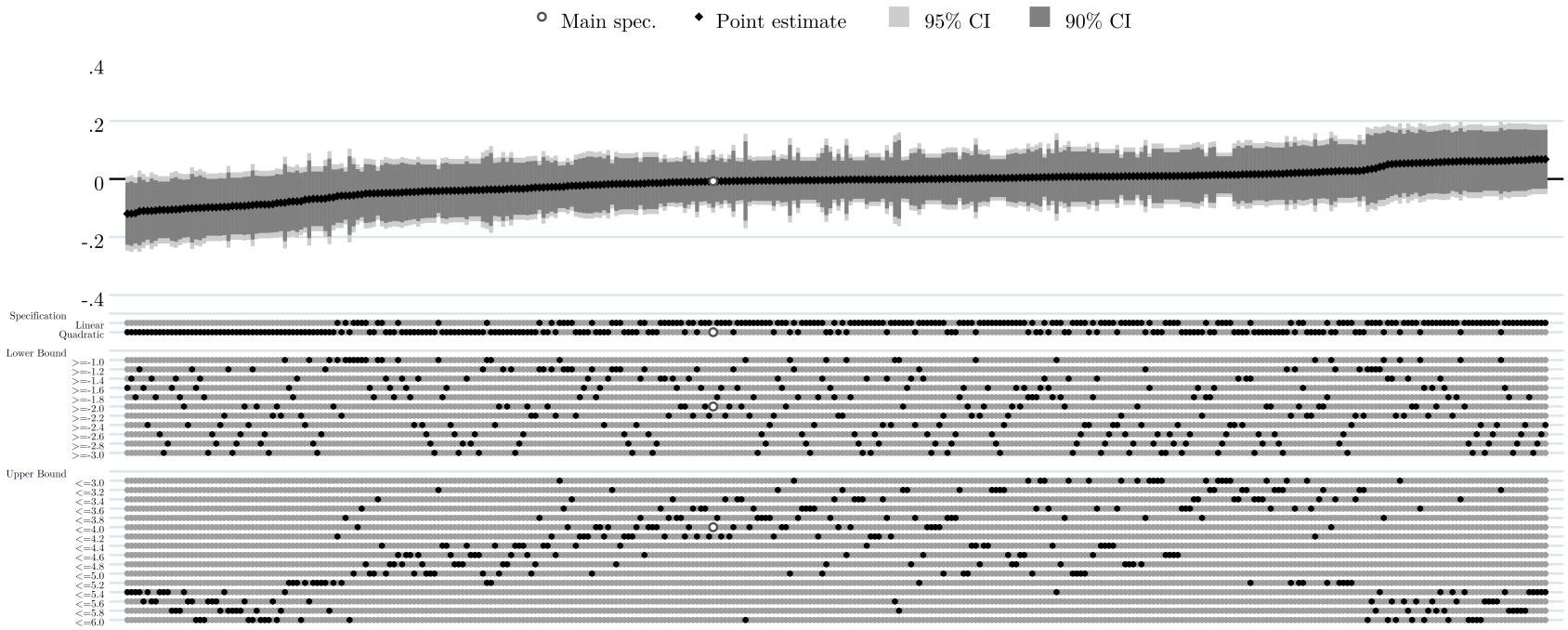


FIGURE 3.A13: Sensitivity to alternative bandwidths and polynomials - Whether taking Anti-depressant medication

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

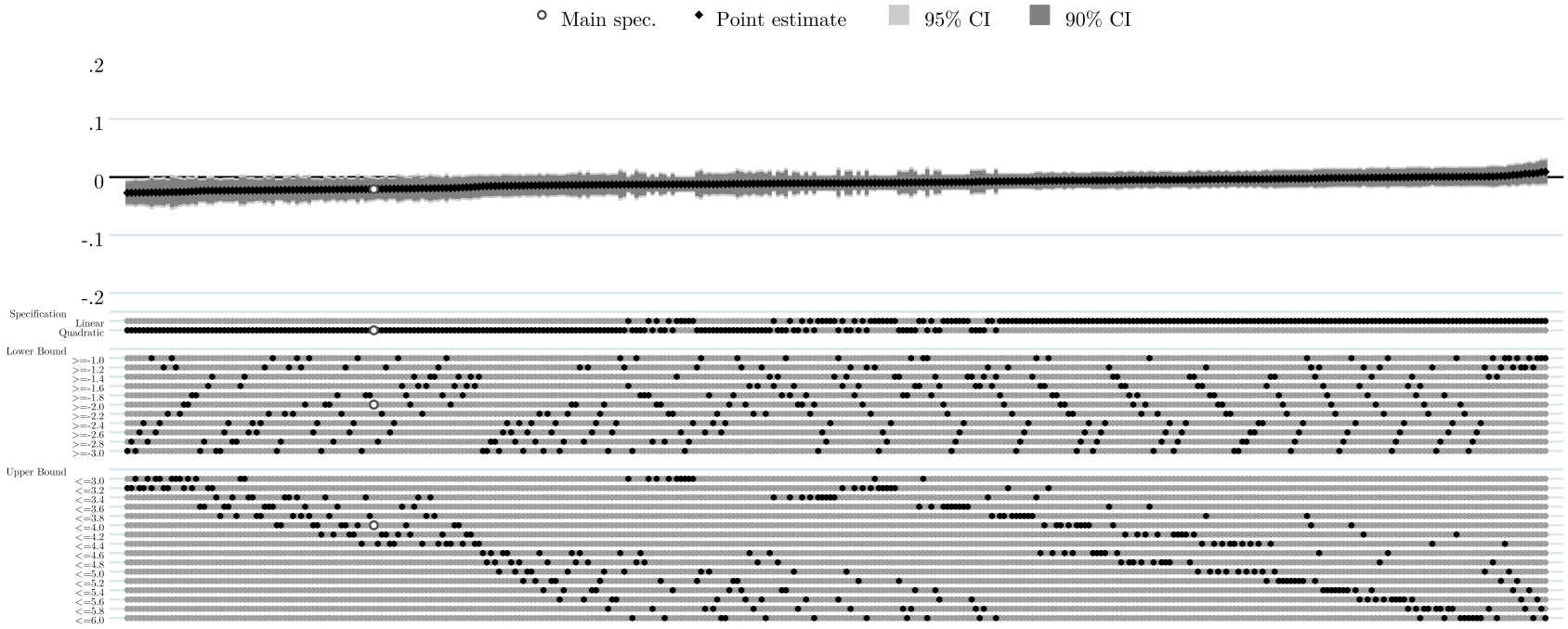


FIGURE 3.A14: Sensitivity to alternative bandwidths and polynomials - Whether taking Statins

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

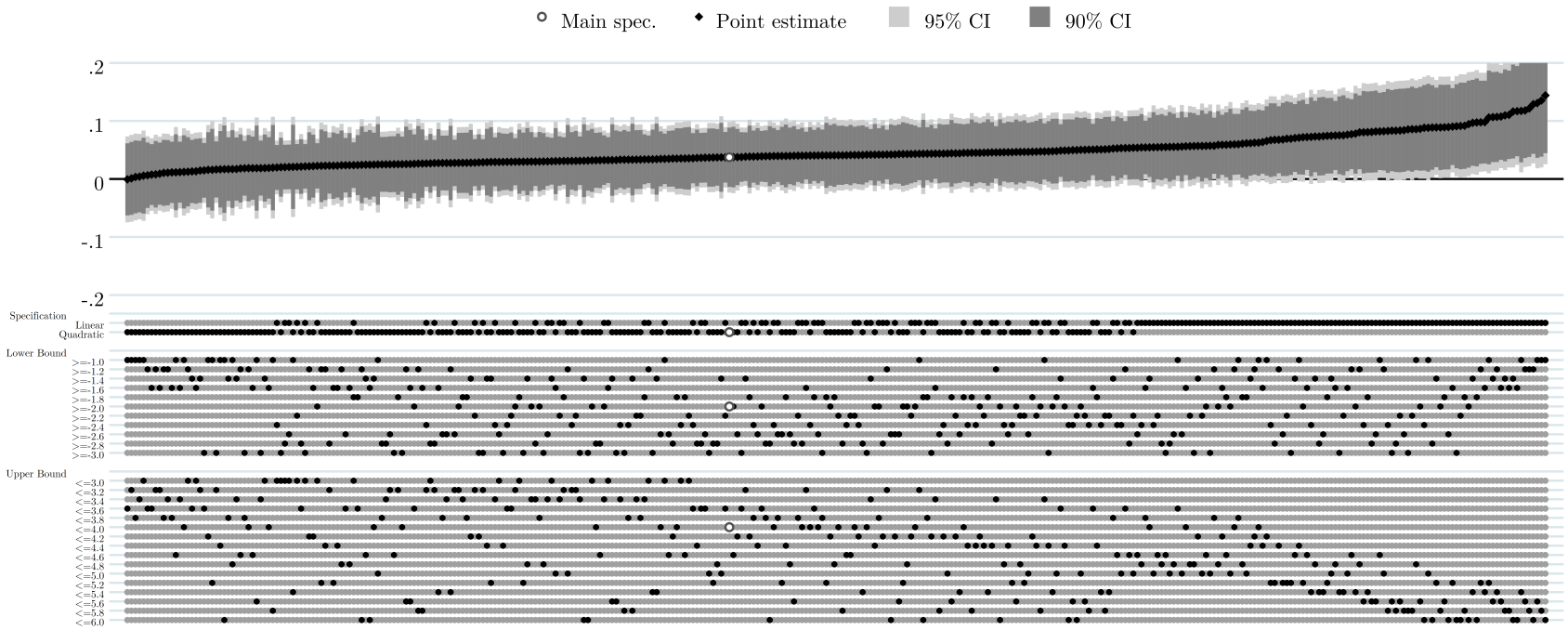


FIGURE 3.A15: Sensitivity to alternative bandwidths and polynomials - Whether ever had a job

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

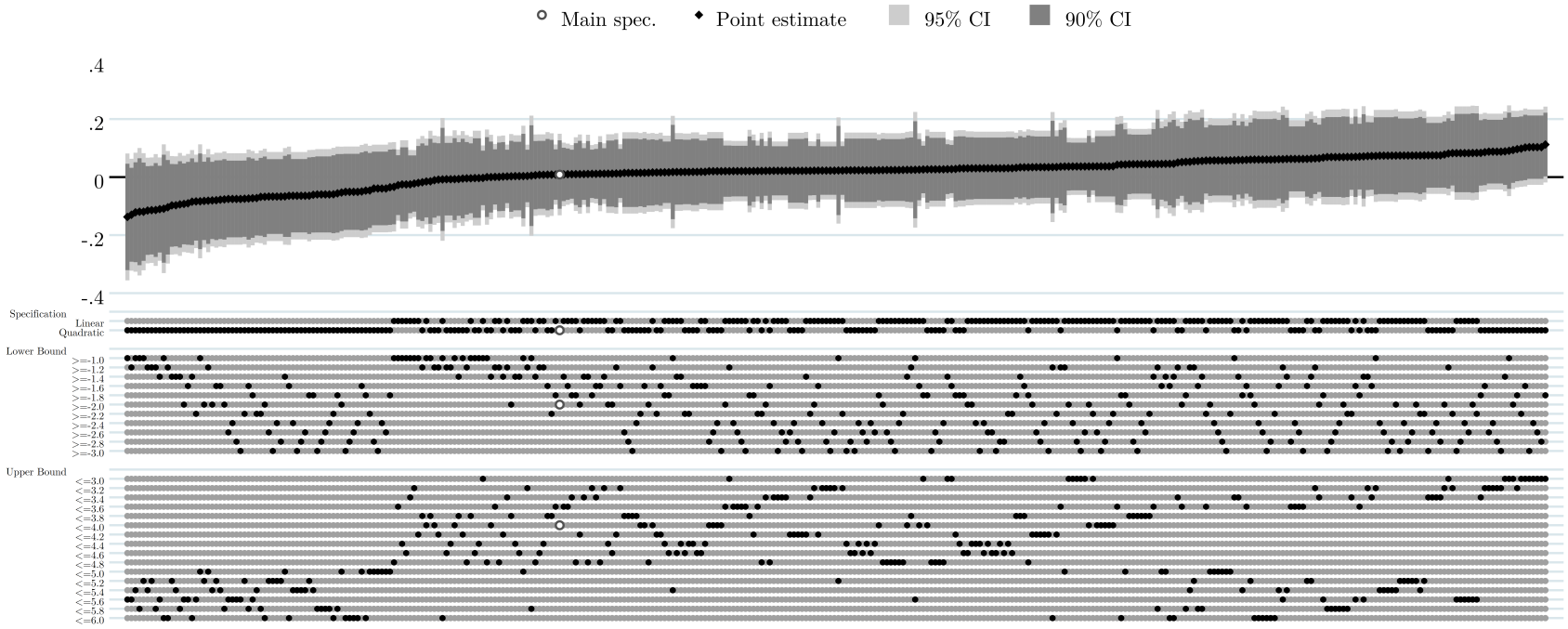


FIGURE 3.A16: Sensitivity to alternative bandwidths and polynomials - Spillover effect of whether taking Anti-diabetic medication

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

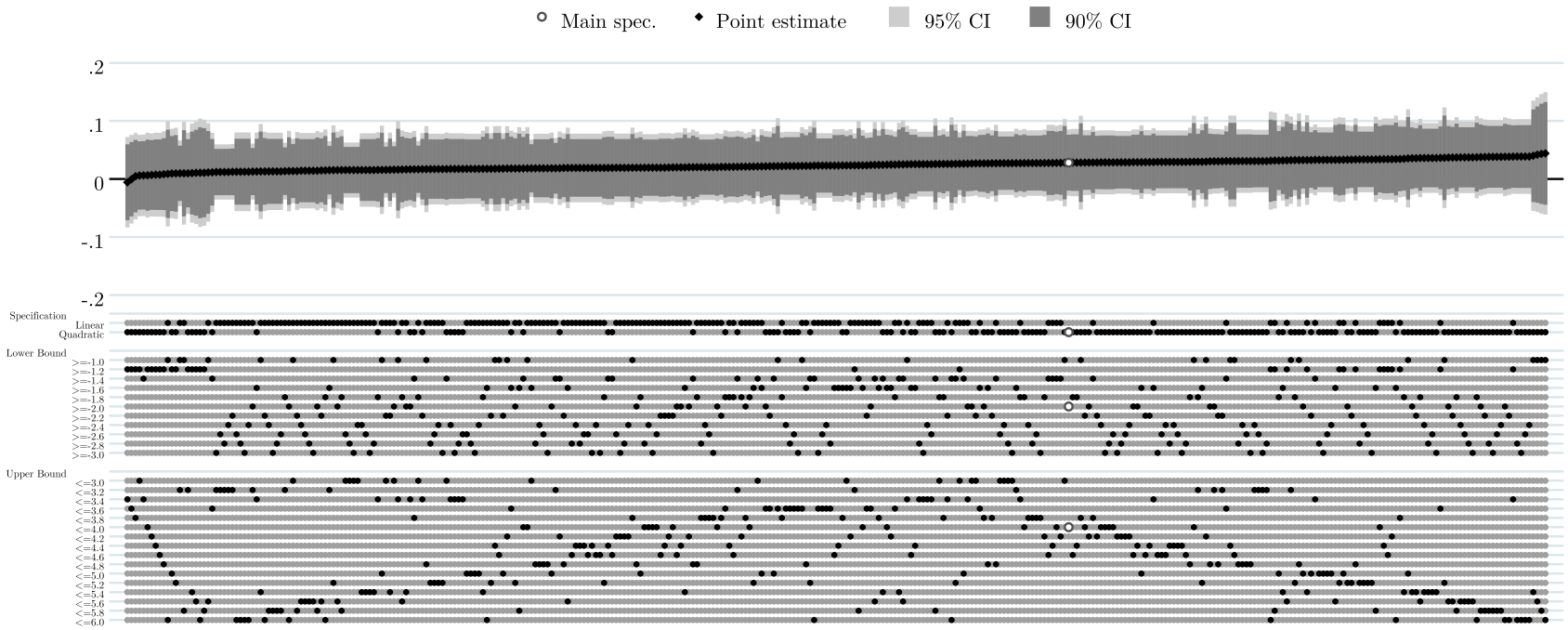


FIGURE 3.A17: Sensitivity to alternative bandwidths and polynomials - Spillover effect of whether taking Antibiotic medication

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

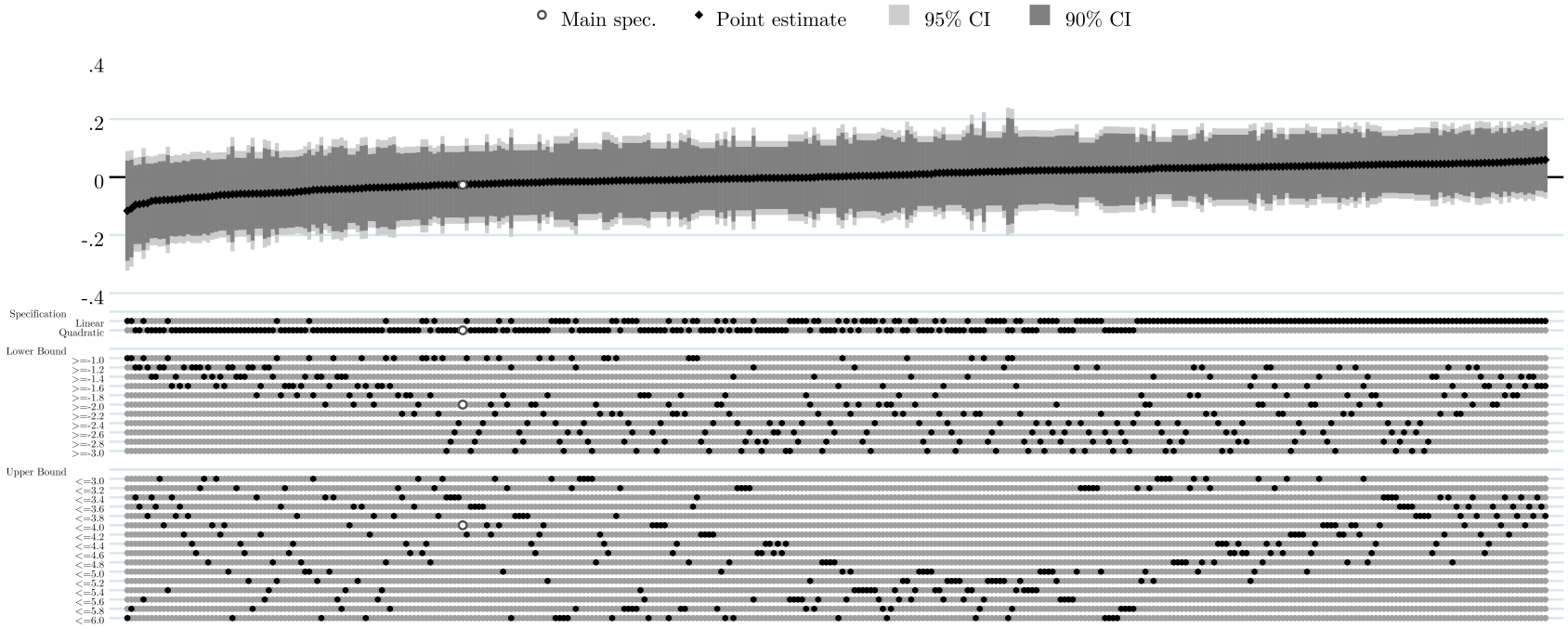


FIGURE 3.A18: Sensitivity to alternative bandwidths and polynomials - Spillover effect of whether taking Anti-depressant medication

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

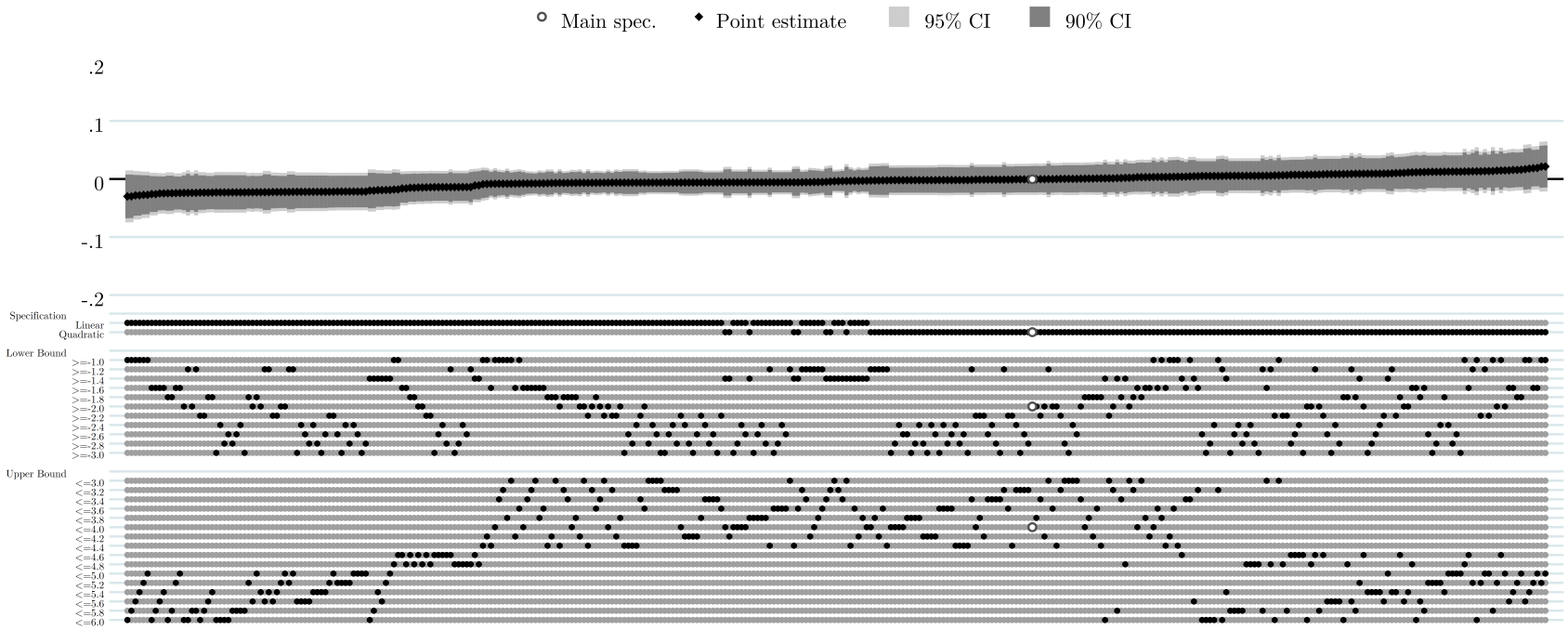


FIGURE 3.A19: Sensitivity to alternative bandwidths and polynomials - Spillover effect of whether taking Statins

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

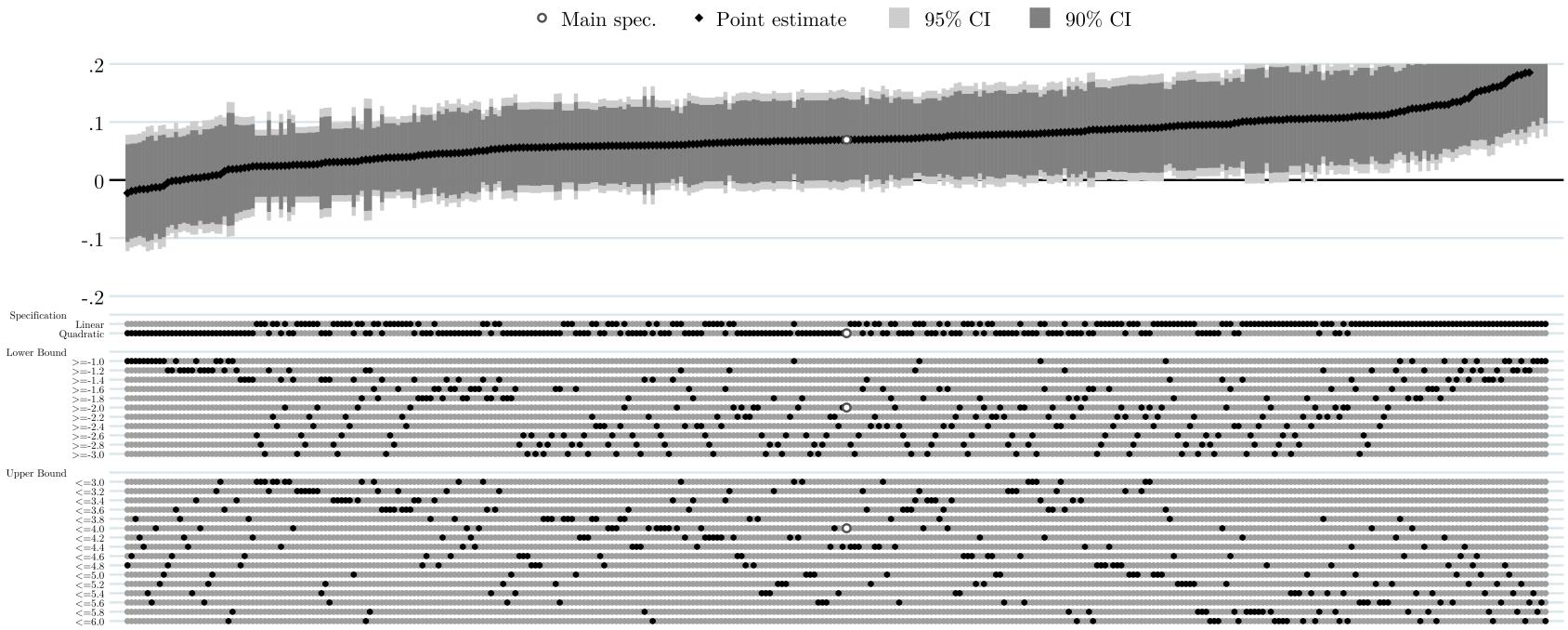


FIGURE 3.A20: Sensitivity to alternative bandwidths and polynomials - Spillover effect of ever having a job

NOTE: The figure shows point estimates, 90% and 95% confidence intervals across a variety of specifications all using the two-stage least squares estimation procedure outlined in equations (3.2) and (3.3). Each point estimates shows the corresponding specification underneath, where a black dot represents that it was estimated using that polynomial, lower bound and upper bound. We estimate alternative specifications by, the order of polynomial, the upper bound of the estimation sample, and the lower bound of the estimation sample, and present all possible combinations, none of which are excluded from this figure. The white dot represents the main specification which we present in our tables in the main text. We thank Peter Eibich, Uri Simonsohn and Hans H. Sievertsen for developing the idea and the code for this figure. Stata code is available at *https://github.com/hhsievertsen/speccurve*.

3.A.3 Reduced Form Regression Kink Design Graphs

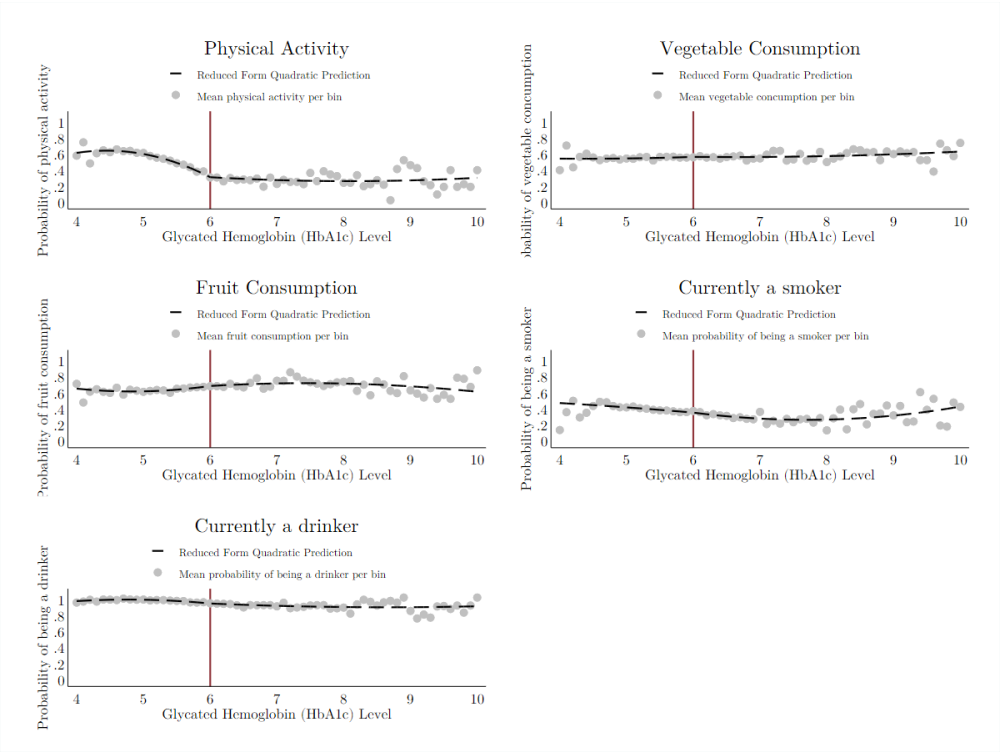
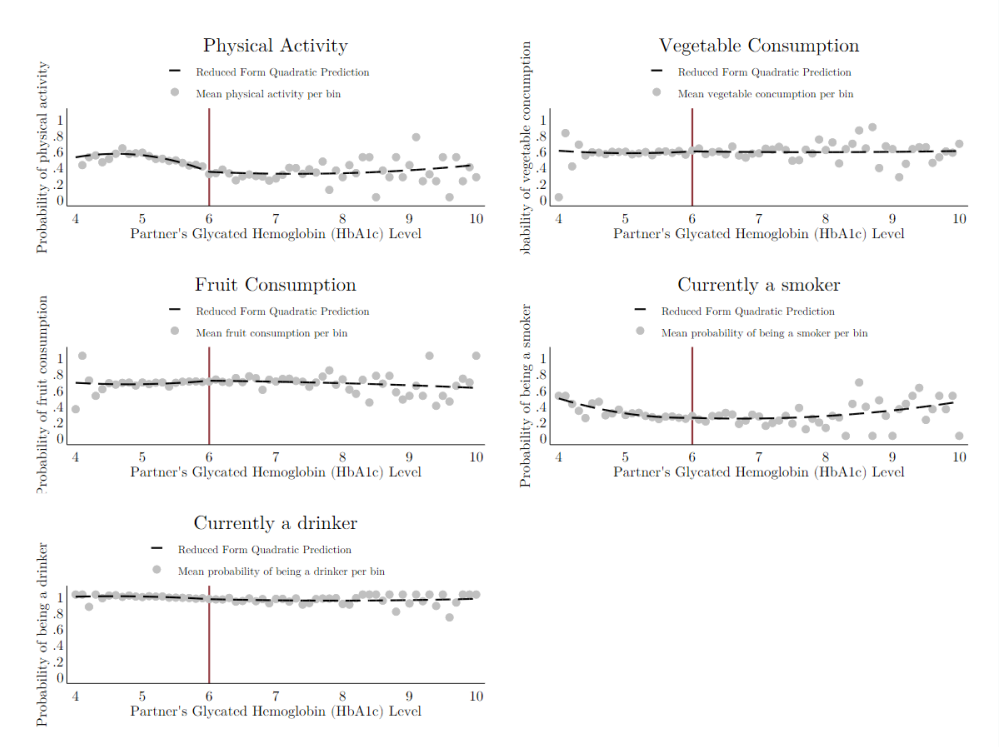


FIGURE 3.A21: Graphical Representation of Reduced Form RKD Results - Own Glycated Hemoglobin

NOTE: These figures are a graphical representation of the RKD. Figures show the mean outcomes per bin (grey points), where bin width is 0.1, between HbA1c levels between 4.0 and 10.0. Black dashed line represents the quadratic prediction from the reduced form a regression of the form: $Y_i = \chi_0 + \chi_1(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \psi_p^- (x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \psi_p^+ (x_i - k)^p D_i\right] + \mu_i$. The red lines represents the kink point where HbA1c is 6.0%. Precise estimates of the fuzzy RKD using equations (3.2) and (3.3) are available in table (3.2).



NOTE: These figures are a graphical representation of the partner RKD. Figures show the mean outcomes per bin, where bin width is 0.1 (grey points), between HbA1c levels between 4.0 and 10.0. Black dashed line represents the quadratic prediction from the reduced form a regression of the form: $Y_i = \sigma_0 + \sigma_1(x_j - k)D_j + \left[\sum_{p=1}^{p^*} \phi_p^-(x_j - k)^p\right] + \left[\sum_{p=2}^{p^*} \phi_p^+(x_j - k)^p D_j\right] + \zeta_i$. The red lines represents the kink point where HbA1c is 6.0%. Precise estimates of the fuzzy RKD using the first stage and second stage in equations (3.6) and (3.7) respectively are available in table (3.3).

FIGURE 3.A22: Graphical Representation of Reduced Form RKD Results - Partner's Glycated Hemoglobin

3.A.4 Observed heterogeneity in RKD estimates

Following the main analysis and identification properties, we assess whether effects of a diabetes diagnosis are heterogenous across observables both for own and partners' diagnoses. We explore three sources of heterogeneity. First, we test whether those that live with a spouse behave differently to those that do not. Second, in an attempt to estimate whether the impact of diagnosis on behavioural change varies over time we analyse whether those being diagnosed for longer behave differently to those recently diagnosed. In the absence of panel data, differential impact by time since diagnosis approximates long-term effects or recidivism to pre-diagnosis behaviours. Finally, we estimate whether there are observable heterogeneities by education.

For estimation, we derive the Hetergoenous Marginal Treatment Effect (HMTE) in a similar vein to Becker et al. (2013), by replacing the MTE of $EverD_i$ in equation 3.3 and manipulating it to allow for heterogeneous effects along the variable z_i. This is implemented by replacing coefficients with an interaction, the general case being $\gamma = \widehat{\gamma} + \widetilde{\gamma} z_i$ where z_i denotes the trait across which heterogeneity is examined. The first stage equation is re-written as:

$$EverD_i = \mu_0 + \mu_1 z_i + \mu_2 (x_i - k) D_i + \mu_3 (x_i - k) D_i z_i + \sum_{p=1}^{p^*} \left[\nu_p^- (x_i - k)^p + \nu_p^- (x_i - k)^p z_i \right] + \sum_{p=2}^{p^*} \left[\nu_p^+ (x_i - k)^p D_i + \nu_p^+ (x_i - k)^p D_i z_i \right] + w \quad (3.14)$$

The second stage of the 2SLS is then described by:

$$Y_i = \psi_0 + \psi_1 z_i + \psi_2 \widehat{EverD_i} + \psi_3 \widehat{EverD_i} z_i + \sum_{p=1}^{p^*} \left[v_p^- (x_i - k)^p + v_p^- (x_i - k)^p z_i \right] + \sum_{p=2}^{p^*} \left[v_p^+ (x_i - k)^p D_i + v_p^+ (x_i - k)^p D_i z_i \right] + m_i \quad (3.15)$$

The parameters of interest here are ψ_2 and ψ_3, where the estimate of ψ_3 describes the heterogeneity in the treatment effect over the trait under inspection z_i. The rest of the notation is as previously. The inclusion of an additional term to estimate, ψ_3, which is dependent on the endogenous variable $EverD_i$ requires an additional instrument for the 2SLS estimates to be correctly identified. We, therefore, estimate an auxiliary first

stage regression:

$$EverD_i z_i = \omega_0 + \omega_1 z_i + \omega_2 (x_i - k) D_i + \omega_3 (x_i - k) D_i z_i + \sum_{p=1}^{p^*} \left[\sigma_p^- (x_i - k)^p + \sigma_p^- (x_i - k)^p z_i \right] + \sum_{p=2}^{p^*} \left[\sigma_p^+ (x_i - k)^p D_i + \sigma_p^+ (x_i - k)^p D_i z_i \right] + z_i \quad (3.16)$$

The above framework refers to own behaviour in response to own diabetes diagnosis. We extend this approach to partners and estimate whether there is heterogeneity in own behaviour as a result of partner diagnosis and heterogeneity according to time since partner's diagnosis and their educational level.

Our HMTE estimation strategy closely follows that of Becker et al. (2013) with the key difference being that ours is implemented within an RKD instead of an RDD setting. For HMTE estimation we require that two additional assumptions hold in addition to those discussed in the previous section. First, that there is continuity of the interaction variables at the threshold vector. In our setting we require a stronger version of this, namely that there is neither a jump nor a kink in the interaction variables at the threshold. To check whether this assumptions holds, we plot the average per bin of the interaction variables against Glycated Hemoglobin (HbA1c). Figure 3.3 shows, amongst other variables, whether individual has degree level education. As discussed previously there is little evidence of either a jump or kink at the threshold HbA1c level of 6%.

Time since diagnosis cannot be handled in a similar fashion as is not observed (i.e. it does not exist) for those that have never been diagnosed. To make HMTE effects estimation possible, for those with missing observations, we follow Kleven et al. (2019) and assign placebo time-since-diagnosis values by randomly drawing values with replacement, from observed individuals who have a time-since diagnosis values. For the analysis, we demean the variable so that ψ_2 represents the effect for the average time since diagnosis. To ensure smooth density of time since diagnosis we present a similar graphic to those in figure 3.3 but for time since diagnosis in figure 3.A23. There is no clear evidence of a jump or a kink in time since diagnosis at the threshold, however it is worth keeping in mind that for those not diagnosed with diabetes, the time since diagnosis values are placebo values.

The second required assumption is the random assignment of the interaction variable conditional on covariates. In this setting, we require that z_i is not correlated with the error term in the estimating equation. To ensure that this is indeed the case, we include a number of observable individual level characteristics in the estimating equations, which we also include in our main estimates, namely a gender dummy, a continuous age variable, we also include a binary indicator of whether individual has

degree level education in the estimating equations where we are not directly estimating the heterogeneity along this dimension.

3.A.4.1 Partner in Household

Table 3.A5 presents the effect of own diabetes diagnosis by whether an individual lives with a partner or not. Having a partner in the household on its own, increases the probability of consuming vegetables, reduces the probability of smoking, while also increases the probability of drinking. Yet, there is little heterogeneity on the effect of own diabetes diagnosis on any of the own outcomes.

3.A.4.2 Time Since Diagnosis

Heterogeneity estimates across time-since-diagnosis are given in Table 3.A6 with Panel (a) showing the effect of own diabetes and Panel (b) the effect of partner's diabetes diagnosis. For both own and partner's diabetes main effects we find that diagnosis increases exercise and reduces smoking with no variation in any of the estimates by time since diagnosis. Such finding, supports a hypothesis of habit formation, whereby individuals make positive lifestyle changes that they consistently maintain going forward. This is somewhat contrary to Kim et al. (2019) who find that for their specific outcomes measures (i.e. outpatient visits, medicated days, basic exercise) there were no significant long-run effects.

It is also reassuring to note that time-since-diagnosis, as a main effect, is insignificant in almost all models, which is precisely what we would expect, given that time since diagnosis for individuals who have not had a diabetes diagnosis is a placebo time since diagnosis, or placebo time since partner's diagnosis.

3.A.4.3 Education

Finally, heterogeneity in terms of educational attainment is presented in Table 3.A7. On average, those with degree level education tend to make better lifestyle choices than those without degree level education. Those with degree level education are more likely to exercise, eat vegetables and fruit, and are less likely to smoke. However, they are also more likely to currently be a drinker, which is somewhat at odds with what we would expect. In terms of the interaction between diabetes diagnosis and education, we find limited evidence of a heterogeneous effect by education for most of our outcomes, the only exception being fruit consumption.

We find that degree educated individuals decrease their fruit consumption in response to a diagnosis, whereas those without a degree increase their consumption of fruit. At

first glance this may be somewhat perplexing, however as discussed in Section 3.2.1 and its footnotes, clinical guidelines state that fruit should not be eaten freely, and although its consumption is encouraged, the amount should be limited. Considering that degree educated eat more fruit than those without degree level education, a higher proportion of them are at the upper bound, or exceed the recommended fruit consumption prior to a diagnosis, and therefore the diagnosis induces them to reduce their fruit consumption. However, there is potential concern in that we find no evidence to suggest that these individuals offset their decrease in fruit consumption with an increase in vegetable consumption, which would be medically recommended.

	Exercise	Vegetable Consumption	Fruit Consumption	Currently a Smoker	Currently a drinker
Own Diabetes	0.141	0.0548	-0.0221	-0.325***	0.00394
	(0.120)	(0.0833)	(0.0793)	(0.100)	(0.0464)
Partner in HH	-0.00825	0.0403**	0.0195	-0.116***	0.0293***
	(0.0241)	(0.0161)	(0.0152)	(0.0180)	(0.00805)
Own Diabetes x Partner in HH	0.101	-0.0261	0.122	-0.119	0.0123
	(0.146)	(0.102)	(0.0971)	(0.121)	(0.0547)
Obs.	20641	39666	39690	23432	41686

Coefficients are estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 4.0 on the right hand tail, and 2.0 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification includes the following controls: Age, and dummies for whether individual *i* is male, a partner lives in the household, and has degree level education. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.A5: Heterogeneous fuzzy RKD estimates of change in own behaviour as a result own diabetes diagnosis by whether individual has a partner

	Exercise	Vegetable Consumption	Fruit Consumption	Currently a Smoker	Currently a drinker
(a)					
Own Diabetes	0.201***	0.0322	0.0694	-0.419***	0.0137
	(0.0693)	(0.0482)	(0.0456)	(0.0566)	(0.0250)
Time Since Own Diagnosis (TSoD)	0.000114	-0.000173	0.000523	-0.000485	0.000829***
	(0.00112)	(0.000727)	(0.000698)	(0.000764)	(0.000308)
Own Diabetes x TSoD	-0.00337	0.00285	-0.00363	0.00422	-0.00544**
	(0.00726)	(0.00466)	(0.00454)	(0.00511)	(0.00239)
Obs.	20641	39666	39690	23432	41686
(b)					
Partner Diabetes	0.237**	0.0108	-0.0761	-0.218***	0.0410
	(0.0988)	(0.0672)	(0.0631)	(0.0730)	(0.0312)
Time Since Partner Diagnosis (TSpD)	-0.0000979	-0.000407	0.000876	0.000737	-0.000129
	(0.00166)	(0.00106)	(0.00106)	(0.000998)	(0.000539)
Partner Diabetes x TSpD	0.00428	0.00510	-0.0119*	-0.00707	-0.00257
	(0.00960)	(0.00697)	(0.00714)	(0.00877)	(0.00401)
Obs.	10563	19983	19985	11296	20924

Coefficients are estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 4.0 on the right hand tail, and 2.0 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification includes the following controls: Age, and dummies for whether individual *i* is male, a partner lives in the household, and has degree level education. Panel (b) additionally include the same set of controls for individual *j*, but excluding whether partner lives in the household. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.A6: Heterogeneous fuzzy RKD estimates of change in own behaviour as a result own and partner's diabetes diagnosis by time-since-diagnosis

	Exercise	Vegetable Consumption	Fruit Consumption	Currently a Smoker	Currently a drinker
(a)					
Own Diabetes	0.266***	0.0507	0.109**	-0.397***	0.00176
	(0.0757)	(0.0535)	(0.0509)	(0.0620)	(0.0286)
Own College Degree (OCD)	0.249***	0.108***	0.136***	-0.134***	0.0176**
	(0.0303)	(0.0191)	(0.0176)	(0.0222)	(0.00782)
Own Diabetes x OCD	-0.368*	-0.0301	-0.234**	-0.0235	0.0562
	(0.192)	(0.124)	(0.114)	(0.150)	(0.0498)
Observations	20641	39666	39690	23432	41686
(b)					
Partner Diabetes	0.256**	0.0402	-0.0606	-0.215***	0.0427
	(0.104)	(0.0717)	(0.0677)	(0.0806)	(0.0339)
Own College Degree (OCD)	0.173***	0.0768***	0.108***	-0.119***	0.0239**
	(0.0432)	(0.0275)	(0.0256)	(0.0259)	(0.0111)
Partner Diabetes x OCD	0.0114	-0.158	-0.134	-0.110	0.0222
	(0.284)	(0.219)	(0.202)	(0.199)	(0.0945)
Observations	10581	20013	20015	11313	20941

Coefficients are estimated using a quadratic specification each side of the kink point, equally weighted observations, and with bounds of 4.0 on the right hand tail, and 2.0 on the left hand side. Parentheses includes cluster-robust standard errors, clustered at the household level. Each specification includes the following controls: Age, and dummies for whether individual *i* is male, a partner lives in the household, and has degree level education. Panel (b) additionally include the same set of controls for individual *j*, but excluding whether partner lives in the household. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 3.A7: Heterogeneous fuzzy RKD estimates of change in own behaviour as a result own and partner's diabetes diagnosis by educational level

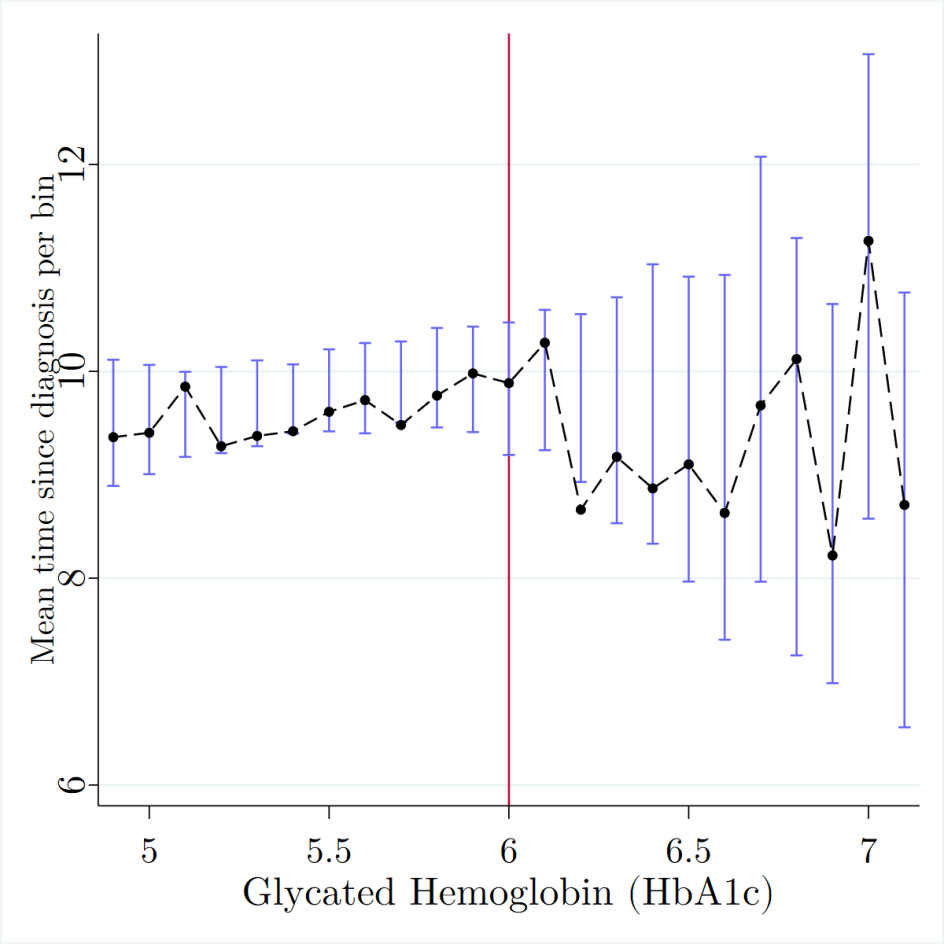


FIGURE 3.A23: Time Since Diagnosis including placebo values

NOTE: Graphical representation of the mean of time since diabetes diagnosis by glycated hemoglobin (HbA1c) level. Black dots are the mean of the time since diagnosis per bin, with a bin width of 0.1, including the placebo time since diagnosis used in the analysis for that do not have observed time since diagnosis. 95% confidence intervals are represented by the blue lines. This confidence interval is constructed by randomly re-assigning the placebo values, and then bootstrapping this sample. This two-step procedure is done 250 times to estimate the bootstrap confidence intervals. Red line represents the kink point of 6.0 %.

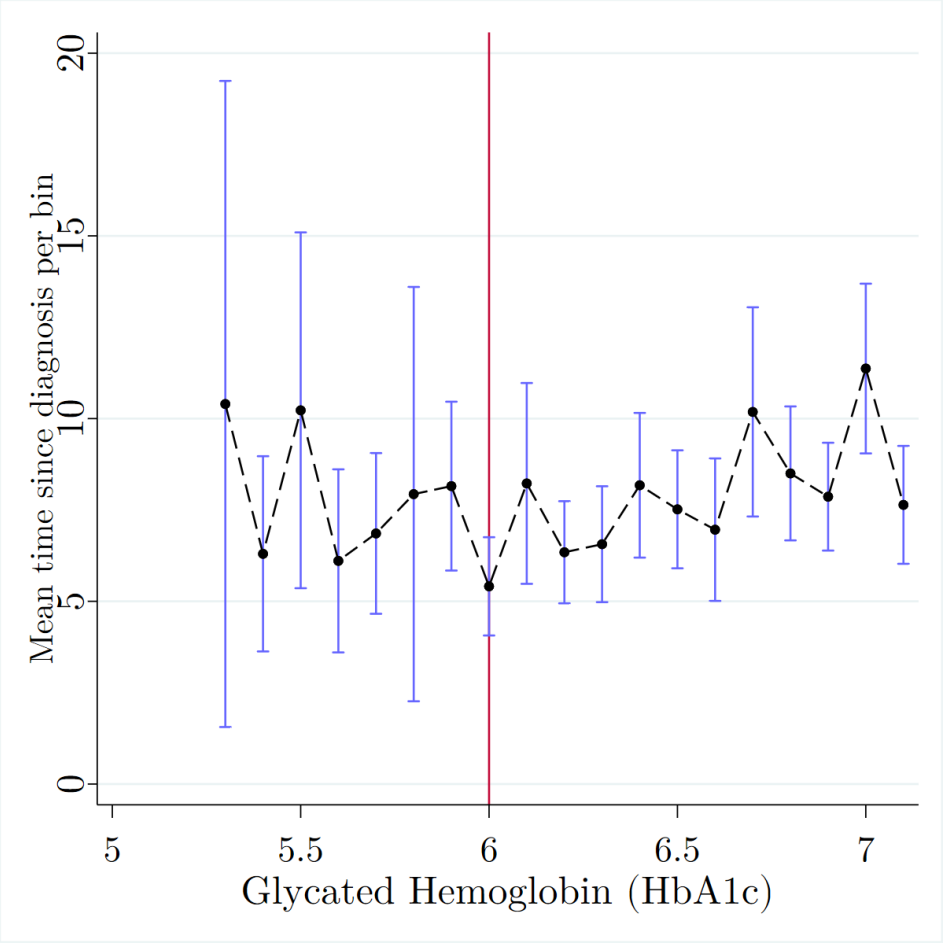


FIGURE 3.A24: Time Since Diagnosis without placebo values

NOTE: Graphical representation of the mean of time since diabetes diagnosis by glycated hemoglobin (HbA1c) level. Black dots are the mean of the time since diagnosis per bin, with a bin width of 0.1, of only those with valid time since diagnosis values. 95% confidence intervals are represented by the blue lines. Red line represents the kink point of 6.0 %.

3.A.5 Test for location of the Kink

In this section we seek to investigate alternative kink locations to ensure that the kink used in our analysis is correct. Although theoretically we expect a kink-point at a glycated hemoglobin (HbA1c) level of 6.0%, it is possible that the true data generating process is different, and that the kink point may be at some other location. Indeed, there is an another candidate jump or kink point, which has theoretical support; a HbA1c level of 6.5% is the threshold for receiving a diabetes diagnosis. Therefore, it is reasonable to investigate possible kink points further.

We follow a similar approach to Landais (2015), where we attempt to find the real location of the kink, if we did not have any theoretical guidance, and we were not aware of where the kink was. This approach estimates the first stage of the RKD specification for a number of virtual values of the kink point (k), and inspect to see which value of the kink point maximises the adjusted R-square. Given that Dong (2011)'s framework allows for a jump and a kink to be used as instruments for the endogenous variable, we estimate two different specifications; we estimate the first stage fuzzy RKD equation (equation 3.2), which is used to estimate the results we present in the main text. This is given by:

$$EverD_i = \gamma_0 + \gamma_1(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \nu_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \nu_p^+(x_i - k)^p D_i\right] + \xi_i \quad (3.17)$$

where $EverD_i$ is a binary variable taking the value of one for individual i if they have ever been diagnosed with diabetes, and zero otherwise. x_i denotes the running variable, which is HbA1c level in this case, and k is the kink point of 6%. $D_i = \mathbb{1}(x_i \geq k)$, is an indicator variable, taking the value of one if the individual's level of HbA1c is above the kink point, and where $(x_i - k)D_i$ is the excluded instrument for the fuzzy RKD. p^* denotes the highest order of polynomial used in the regressions, ν_p^- and ν_p^+ are the estimates of the polynomial function below and above the kink point, respectively.

In addition, we estimate a Regression Probability Jump and Kink (RPJK) first stage, where both the jump and kink are used as instruments. This first stage is described by:

$$EverD_i = \gamma_0 + \chi D_i + \gamma_1(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \nu_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \nu_p^+(x_i - k)^p D_i\right] + \xi_i \quad (3.18)$$

Note the additional term D_i and the additional parameter χ to be estimated. Once again $D_i = \mathbb{1}(x_i \geq k)$ is an indicator variable, taking the value of one if the individual's level of HbA1c is above the kink point. This term estimates the "jump" in probability of receiving a diabetes diagnosis above the threshold k. Indeed, as discussed above, there may be reasonable theoretical justification of a jump in the probability, either at the 6.0% or 6.5% threshold, and therefore we should explore the possibility that there is a jump in the probability.

We estimate both of these specification, bootstrapping 250 times, and report the mean adjusted R-Squared values of these replications in figure 3.A26. R-Squared values for equation (3.17) (without a jump) are presented in blue, and R-squared for equation (3.18) (which includes a jump) are presented in black. The red circle denotes the maximum value of the adjusted R-Square, which was achieved by specification (3.18) and for a kink point (k) of 6.1%.

The first thing to note is that the R-squared value is increasing through values of k initially, up to the value of 6.1%, with both specifications performing almost identically, with a difference in R-Squared of less than 0.00015 between the two. For values above 6.1% the performance of the specifications diverge. The performance of specification (3.17) decreases relatively quickly after reaching the maximum R-squared value, whereas the performance of specification (3.18) does decrease, but at a slower rate, with values of k between 6 and 6.5 being relatively comparable. It is also worth noting that R-squared values of specifications (3.17) and (3.18) are almost identical for k values of 6, 6.1 and 6.2. This suggests that there is no performance gain from adding a discontinuity term to the first stage in these cases.

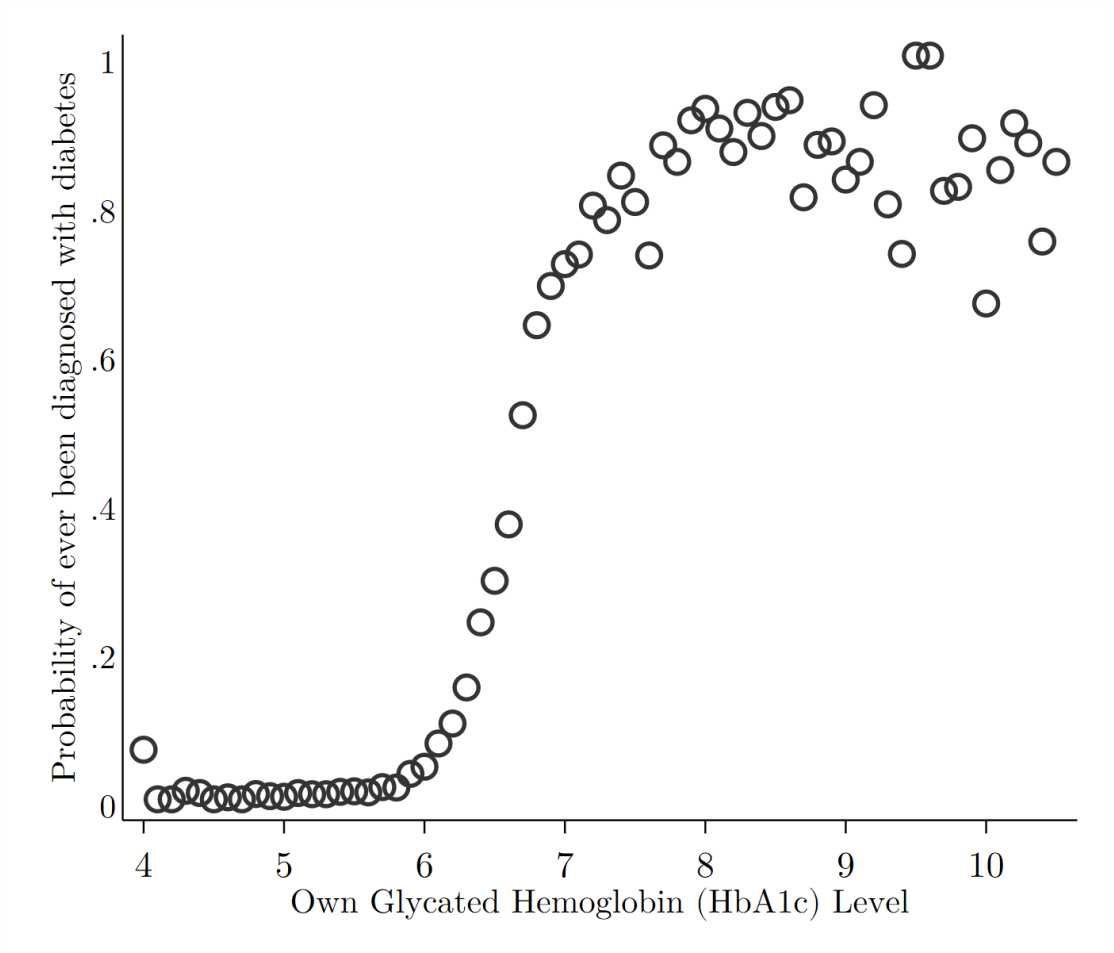
We believe this exercise provides evidence in support of your empirical strategy, in that the kink point we use in our analysis is very close to the highest performing value of k, and indeed the R-Squared of k values 6.0% and 6.1% are not statistically significantly different. In addition, these specifications also do not benefit from having a discontinuity term (i.e. jump).

The alternative candidate value of k which has theoretical support ($k = 6.5$) clearly fits the data less well. Both jump and kink specifications for $k = 6.5$ perform less well than the 6.0% threshold, and therefore we have no reason to believe that the true data generating process has a threshold point of 6.5%. Although we are confident this exercise is supportive of our kink point, we do one further step to ensure that the kink point is indeed at 6.0. We take the specification (3.18) and a threshold point of 6.5 (which is both the furthest from the used k value, and the point which has theoretical support) and illustrate the fitted values of this specification in figure (3.A27).

Figure 3.A27 shows the first stage used in the main text, with a kink point of 6.0% and no discontinuity term, in blue. The alternative candidate threshold of 6.5% which additionally includes a discontinuity term is displayed in red. The first point to note is

that both specifications estimate very similar functions, with the only divergent point being in the range of approximately 5.8% and 6.5, and even across this range the functions are not too dissimilar. The specification with a discontinuity term does not estimate a large jump in probability at the threshold. Across this range the specification without the discontinuity may in fact be preferred, because there is no clear and obvious jump in the probability at either threshold, rather the change in probability is more reasonably modelled by a slope change. In addition, the Kink and Jump Specification with a k value of 6.5 seems to fit values below 6 less well. Below 6.5 the Kink and Jump specification estimates the function to be convex, whereas the true data generating process appears to be more linear for these values.

We believe this inspection provides further evidence in favour of the kink point indeed being at 6.0, and that a discontinuity term provides no measurable benefit in the first stage. We therefore conclude that the data does support the threshold of 6.0% and there is no gain from including a discontinuity term in our first stage specification, which is supportive of our empirical strategy.



NOTE: Mean of the probability of ever being diagnosed with diabetes per bin. Bin width of 0.1 for glycated hemoglobin levels between 4 and 10.

FIGURE 3.A25: Probability of Diabetes Diagnosis by HbA1c Level

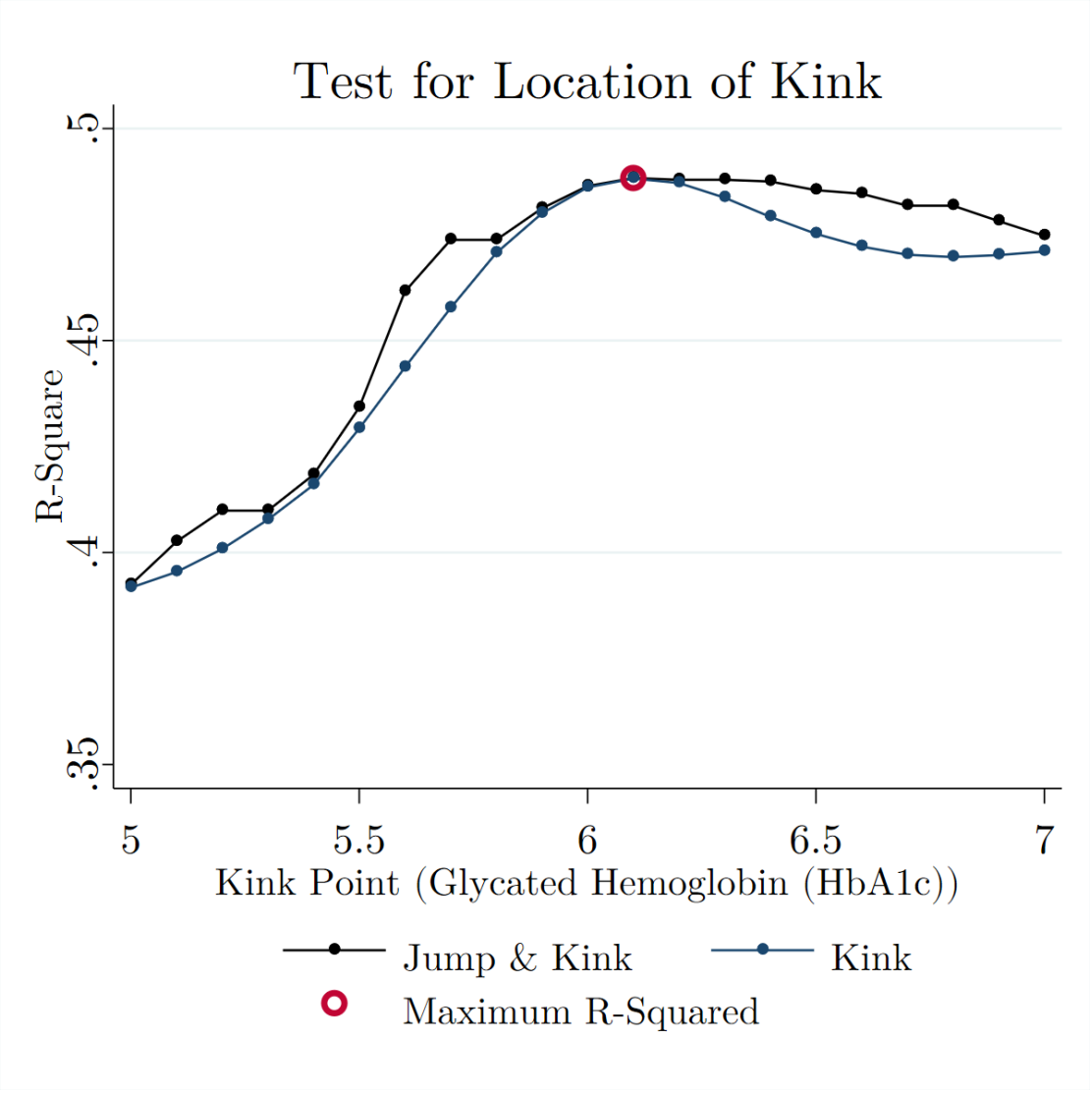
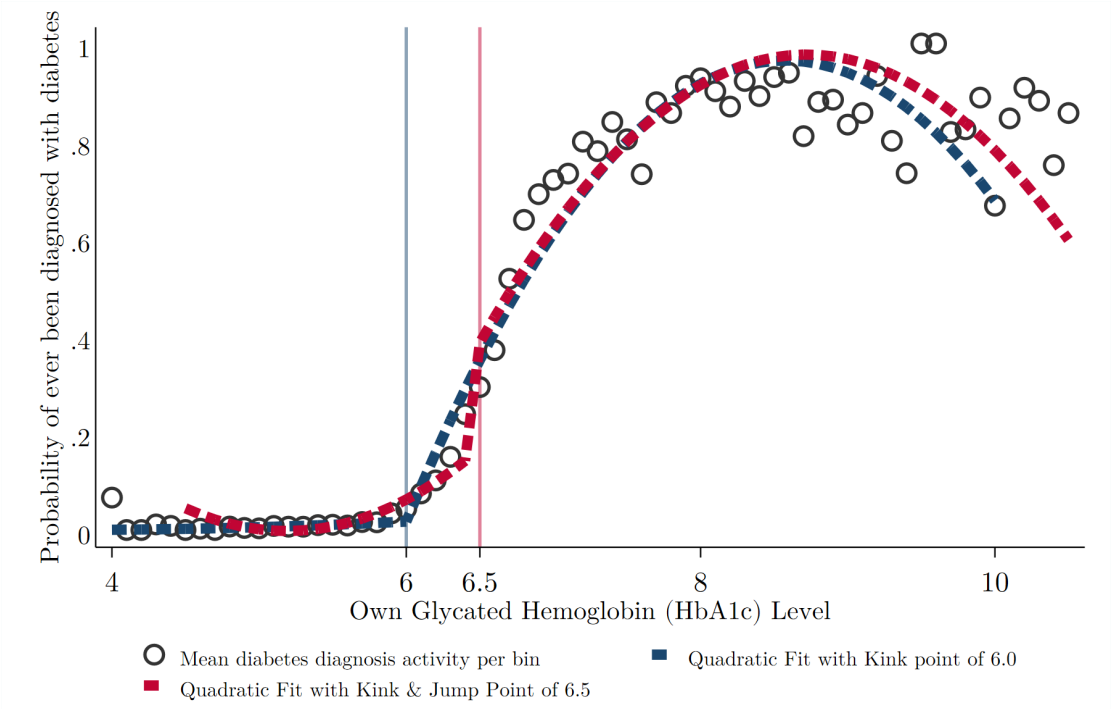


FIGURE 3.A26: Test for location of the kink

NOTE: Value of adjusted R-Square for first stage of the Regression Probability Kink (RPK) specifications (shown in blue) or Regression Probability Jump and Kink (RPJK) (shown in black), for different virtual kink points (k). All values of virtual kink point (k) are presented, with the real kink point at kink point $k = 6$. Red circle denotes the specification with the highest adjusted R-squared value, which was the jump and kink specification with a k of 6.1%. Specifications are bootstrapped 250 times and mean values of adjusted R-squared are shown. We use a relative bandwidth of -2 for the lower bound, and 4 for the upper bound, which is the same bandwidth used in the main specification presented in this paper. Regression Probability Kink is estimated by:
$EverD_i = \gamma_0 + \gamma_1(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \nu_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \nu_p^+(x_i - k)^p D_i\right] + \xi_i$ Regression Probability Jump and Kink is estimated by:
$EverD_i = \gamma_0 + \gamma_1 D_i + \gamma_2(x_i - k)D_i + \left[\sum_{p=1}^{p^*} \nu_p^-(x_i - k)^p\right] + \left[\sum_{p=2}^{p^*} \nu_p^+(x_i - k)^p D_i\right] + \xi_i$
Where $D_i = \mathbb{1}(x_i \geq k)$.



NOTE: Mean of the probability of ever being diagnosed with diabetes per bin. Bin width of 0.1 for glycated hemoglobin levels between 4 and 10. Two specifications are presented. In blue, is the Regression Probability Kink specification, with a kink point (k) of 6.0. In red, the Regression Probability Jump and Kink with a kink point (k) of 6.5. Quadratic fit is separately estimated for the left and right hand sides of the kink for both specifications, with the Regression Probability Jump and Kink also including a jump term of $D_i = \mathbb{1}(x_i \geq k)$. Blue and red vertical lines represents the two kink points under inspection, either a glycated hemoglobin value of 6.0 or 6.5.

FIGURE 3.A27: Probability of Diabetes Diagnosis by HbA1c Level

Chapter 4

More Doctors, better health? Considering doctor numbers in the Mais Medicos Programme

4.1 Introduction

According to the WHO "investments in Primary Healthcare improve equity and access, health care performance, accountability of health systems, and health outcomes" and primary care is considered to be the "front door of the health system" (WHO, 2021). In most healthcare systems worldwide, primary care is the initial contact with the healthcare system and referral to specialists or further treatment is usually initiated by primary care physicians. Primary care is also usually the setting in which preventative treatment and treatment for common illnesses are administered. A strong primary care system has also been linked to better health, lower costs, decrease of unnecessary hospitalisations (Starfield et al., 2005; Guanais and Macinko, 2009; Macinko et al., 2010; Rasella et al., 2014) and can alleviate the burden of poverty on health (Shi, 2012). Indeed, countries with strong primary healthcare systems have been found to "have fewer low birth weight infants, lower infant mortality, especially post-neonatal mortality, fewer years of life lost due to suicide, fewer years of life lost due to all except external causes and higher life expectancy at all ages except at 80 years" (Starfield and Shi, 2002; Rao and Pilot, 2014), and is also associated with decreased health inequality (Szwarcwald et al., 2010; Hone et al., 2017; OECD, 2020). Primary care physicians, therefore, play a key role in the distribution of healthcare resources as well as a direct impact on individuals' health outcomes.

However, by 2030 it has been estimated that 5 billion people will lack access to healthcare (World Bank, 2021b). The deficit in physicians is likely a key contributor to this and there is an estimated deficit of of 2.4 million doctors, nurses and midwives

worldwide (World Health Organisation, 2006). Insufficient primary care physician numbers is of increasing global concern and could be having a significant negative impact on health. In the US the workload for primary care physicians was estimated to rise by 29% from 2005 to 2015, but the increase in number of primary care physicians was only estimated to grow by between 2% and 7% (Truglio et al., 2012). It is estimated that over 14,000 primary care doctors are required reduce the deficit in the US (The Kaiser Family Foundation, 2020) and this number is expected to increase to 52,000 by 2025 (Petterson et al., 2012; Frisch, 2013). This deficit in primary care doctors is not limited to the US, and shortages in primary care physicians have been documented in Australia and New Zealand (Gorman and Brooks, 2009), Canada (Gladu, 2007; Islam, 2014), the United Kingdom (Majeed, 2015), and there is also a known shortage in many low and middle income countries (World Health Organisation, 2006; Hoyler et al., 2014).

Although the deficit in primary care physicians is well document, the literature quantifying the benefits of primary care physicians is mixed. There is reasonable theoretical justification for a benefit in healthcare outcomes from increasing primary care physician supply, however, the empirical evidence is less clear. It is not well established as to whether increasing primary care physician supply would indeed lead to the improvements in health, and importantly whether increasing primary care physician numbers is the most cost-effective way of improving public health. It is hypothesised that if there was indeed unmet need in primary care which was due to insufficient primary care physicians, then it would be expected that an increase in physician numbers would increase utilisation of primary care services. We would then expect more preventative treatments to be administered to the population, more treatment of common illnesses and conditions, as well as increased referrals to specialists. As a result, it would be expected that mortality would fall due to this increased use of primary care services and preventative treatment.

Basu et al. (2019) analysed the relationship between primary care physician supply in the US, and found that an increase in primary care physician density was associated with an increase in life expectancy and reduced mortality from cardiovascular, cancer, and respiratory complications. In support of these findings, Anand and Bärnighausen (2004) conducted a cross-country analysis and found that doctor density, not specifically primary care doctor density, was associated with a decrease in maternal, infant and under-five mortality rate, but found that there was no evidence that nurse density impacted the same set of outcomes. An analysis of general practitioner density in England was conducted by Gulliford (2002) and found that all cause mortality, avoidable mortality, hospital admissions with acute and chronic conditions were all negatively associated with doctor density. However, the author did acknowledge that location characteristics may be confounding the estimates. Conversely, Aakvik and Holmås (2006) found that the impact of general practitioners

on mortality was insignificant in Norway. Fadlon and Van Parys (2020) instead analyse the impacts of switching to a primary care doctor with a different style of care in the US, and find that there are indeed significant long-term implications as a result of the intensity of care a primary care physician provided. Namely, they find increases in number of diagnoses, chronic conditions, number of emergency department visits, number of avoidable hospitalisations, and increased probability of receiving diabetes care from switching to a higher intensity primary care physician.

It is therefore unclear whether primary care doctor density plays a role in determining mortality and hospitalisations. This paper seeks to provide empirical estimates of the marginal effect of an increase in primary physician supply on hospitalisations and mortality. We do so by using a unique setting ideal for this type of analysis. Specifically, in this paper, we use the Programa Mais Médicos (PMM) and a generalised synthetic control estimator (Powell, 2021) to estimate the effect of an increase in physician primary care physician supply on mortality and hospitalisations. However, we deviate from previous work analysing the PMM, in that we use only municipalities that received PMM doctors, to ensure that selection into treatment does not bias in our estimates.

The Programa Mais Médicos (PMM) aimed to improve the provision of primary care to underserved communities specifically through increasing primary care physician supply. A policy to significantly increase physician supply would normally occur due to a redistribution of physicians, rather than a unilateral increase in primary care numbers. Redistributions of physicians would usually either be due to physicians from other specialisms moving into primary care, and therefore a decrease in physicians working in other specialists, or due to a geographical redistribution which would move physicians from one region to another. In both cases, there would be a substantial spillover effect as a result of these redistributions, which would make estimating the effect of a change in primary care physician density on health outcomes more challenging. However, the increase in physician supply in the PMM was driven by importing doctors from Cuba, who were imported specifically to fill these PMM vacancies. Because the change in physician numbers was mostly not the result of redistribution of doctors in Brazil but rather importing from outside of Brazil, the spillover and general equilibrium effects are less of a concern in this setting, which makes it an ideal setting for this type of analysis.

There has already been work analysing the effectiveness of the PMM. Carrillo and Feres (2019) estimate that PMM doctors simply substituted for nurses in municipalities that received PMM doctors. They find minimal evidence that prenatal visits increase, but do estimate that prenatal visits by physicians increase, specifically the PMM resulted in 0.7 more prenatal visits by physicians per 1,000 residents each bi-month. They additionally found no evidence to suggest that this increase did resulted in better infant health outcomes. There was no statistically significant change

in prematurity, birth weight or infant mortality. These findings suggest that although these areas may have had less primary care doctors than demanded, the demanded care was not unmet but rather provided by nurses in these regions. The increase in physicians supply caused a substitution away from care delivered by nurses but had very little impact on health outcomes. These findings are not unexpected as there has been previous work concluding that nurses are able to deliver a comparable quality of care to doctors (Laurant et al., 2005). Tentatively, these results do suggest that primary care physicians may not have been the most cost-effective way to improve health outcomes in this setting.

A comprehensive analysis of PMM was conducted by Mattos and Mazetto (2019), in which they analyse the impact of PMM on healthcare utilisations and mortality. They estimate an average 9.4% increase in doctor consultations in PMM municipalities, with a larger increase in children consultations (11.1%) and marginally less for the elderly (7.1%). Number of appointments were also found to increase (6.7%), and was driven by preventative health appointments (5.9%). They additionally found an increase in referrals in the order of 12.5% and some less robust evidence of a decrease in overall hospitalizations. There was, however, robust evidence of a decrease in hospitalizations from infectious and parasitic diseases of approximately 6%. Finally, they analysed mortality, and precisely estimated null effects for overall, elderly, maternal and infant mortality indicators.

Fontes et al. (2018) focus their analysis on ambulatory care sensitive (ACS) hospitalisations, which are conditions considered to be preventable by primary care interventions (Purdy et al., 2009), and find a negative impact of the PMM on ACS hospitalisations. They additionally find that the effect magnitude increases in each of the three years after the introduction of PMM, with an average decrease of 8% over those three years. Which suggests that increase in utilisation from the PMM had beneficial impacts on health.

Previous work by Hone et al. (2020) found there to be an overall decrease of amendable mortality in the order of 1.06 mortalities per 100,000 population. However, they found there to be no statistically significant impact of PMM in non-priority municipalities on amendable mortality, however amendable mortality decreased by 1.26 in priority municipalities. This suggests that the PMM was beneficial to previously underserved communities. They also estimated a gradient in amendable mortality by quintiles of baseline primary care doctor density, where, in general, the municipalities with the lowest levels of doctors pre-PMM saw a decrease in amendable mortality, and those municipalities with higher number of doctors pre-PMM saw no significant impact from the PMM. This work, in particular, provides evidence to suggest that previously underserved municipalities disproportionally benefited from the PMM programme through improvements in health.

However, a concern of these previous studies is that municipalities that received treatment are systematically different to those that did not receive PMM doctors. Previous work has condensed the PMM into a binary classification for whether received any PMM doctors or not. These studies have therefore estimated the average treatment effect of receiving any PMM doctors, which could be considered an overall policy effect. A concern with these previous studies, however, is that they compare treated and untreated municipalities. This is a concern because the policy was explicitly designed to be non-random, and had a clear priority criteria. Also, municipalities were required to express interest in participating in the PMM, and therefore there may be a form of self-selection into the programme. This paper deviates from these previous analyses, in that we aim to utilise the variation in PMM doctors to identify the impact of PMM doctors. We exploit the variation in PMM doctors between municipalities by using a generalised synthetic control estimator (Powell, 2021), so that we are able to limit the sample to only municipalities that received PMM doctors. This eliminates the concern that we are comparing systematically different municipalities, and allows us to identify unbiased estimates of the impact of an increase in primary care physician supply.

Our paper aims to estimate the marginal effect of an increase PMM physician density on hospitalisations and mortality, which is also analogous to estimating the dose-response. We hypothesis that an increase in primary care physicians to these areas that were previously underserved would result in a decrease in both hospitalisations and mortality due to physicians servicing the unmet need in these areas. Alongside our estimates of the impact on total hospitalisations and total mortality rates, we also estimate the impact on hospitalisations and mortality rates for ambulatory sensitive conditions (ACS) and non-ambulatory sensitive conditions (non-ACS) separately. ACS conditions are ones which are considered to be preventable by primary care interventions (Purdy et al., 2009), and previous studies have claimed that there to be a 75% preventability of these conditions (Sundmacher et al., 2015). These conditions include: chronic diseases in which management of them can reduce risk of mortality and hospitalisations, acute diseases where early diagnosis and intervention can reduce risk of major complications, and finally vaccine-preventable conditions, which are ones in which a vaccine administered in primary care can almost fully eliminate the risk of developing those diseases. It is therefore expected that the primary care physician expansion would have the largest impact on ACS conditions, and we explore whether this is indeed the case. Indeed, we hypothesis that the decrease in hospitalisations and mortality due to the increase in physician supply would be driven by decreases in hospitalisations and mortality from ACS conditions, due to these being the conditions that primary care has the largest influence over. We expect there to be limited, if any, impact on non-ACS conditions.

In estimating these effect, we contribute to the literature in two key ways. Firstly, we

contribute to the literature evaluating the PMM by providing estimates of the effect of treatment intensity on mortality and hospitalisations, as well as providing estimates that are purged of selection biases because we use only treated municipalities. These estimates also contribute to the literature more broadly, by providing empirical estimates of the effect of an additional primary care physician on mortality and hospitalisations, in a setting in which spillover and general equilibrium effects are less of a concern. We also add to the literature by providing within country estimates of the impact of primary care doctor density for a developing country. Literature that has previously estimated the impact of doctor density on health outcomes have prominently been focused on high-income countries in which the healthcare systems are highly developed and where doctor densities are higher. Our results are useful estimates for policy makers in determining the most cost-effective use of healthcare resources, and indeed whether an increase in primary care physician numbers can be expected to improve health outcomes. We find no evidence that an increase in primary care physician density has any impact on mortality, whether it is total mortality, or mortality from ambulatory sensitive conditions (ACS) and non-ambulatory sensitive conditions (non-ACS). We do, however, find some limited evidence that total hospitalisations falls.

The paper is organised as follows: firstly, we provide background of the Programa Mais Médicos (PMM), before moving onto the data used in this paper. Then, we present the identification strategy including a description of the atypical inference procedure we use. Results are then presented before providing evidence that our identifying assumptions holds. Finally, we discuss our findings and place them within a wider context.

4.2 *Programa Mais Médicos* (PMM) Background

In July 2013 Programa Mais Médicos (PMM) was launched in Brazil with one of the main objectives being to reduce primary care physician deficits in underserved areas (Mattos and Mazetto, 2019). The programme comprised of three main components: to establish new medical schools, provide funding for construction and refurbishment of clinics, and, the most relevant component for this paper, was the emergency increase in the number of primary care doctors. This was achieved, firstly, by offering generous salaries, as well as housing and food benefits, to Brazilian physicians that relocated to work in areas with shortages of doctors. However, only 10% of the vacancies were filled in the first round (Pereira et al., 2016). To fill the remaining vacancies, an international cooperation was made between Brazil and Cuba, facilitated by Pan American Health Organization (PAHO), which allowed the participation of Cuban doctors in the Mais Medicos programme (Harris and Harris, 2016; Santos et al., 2017). PMM doctors from abroad were required to participate in additional training, which

included Portuguese language classes and background on Brazil and the healthcare system (Mattos and Mazetto, 2019).

The Ministry for Health in Brazil set a criteria for PMM prioritization. If one of the following criteria was satisfied, then a municipality would be considered a priority municipality for receiving PMM physicians: 20% or more of the population in extreme poverty; the 100 municipalities with the lowest income per capita and a population over 80,000; state capitals, metropolitan regions and municipalities with extreme poverty; and, municipalities with low human development indexes, or semi-arid and Quilombo communities (Oliveira et al., 2016; Carrillo and Feres, 2019; Hone et al., 2020).

For a municipality to receive PMM doctors they were required to express their interest in participating in the programme. However, the priority criteria was not strictly adhered to and many non-priority municipalities received doctors under the programme (Oliveira et al., 2016). By 2018, the programme was implemented in 4,524 of the 5,570 municipalities.

As a result of critical comments by Jair Bolsonaro, the Cuban government then withdrew Cuban doctors from Brazil in November 2018, which brought an end to Cuban doctors involvement in the programme (Santos et al., 2018)[1].

4.3 Data

To analyse the impact of PMM physicians we construct a panel of municipalities over the years 2007 to 2018. Our constructed panel includes bi-yearly observations of each municipality. Although data from the period before 2007 is available for some of our variables, we choose not to use this data due to substantial primary care changes prior to 2007. Specifically, in 2006 the Brazilian government initiated a national primary care policy in an attempt to improve the access and quality in primary care (Paim et al., 2011). We seek to avoid including structural breaks in our data to ensure that we are able to generate reasonable synthetic controls for each municipality.

Our main variables of interest are, total hospitalisations and mortality rates (per 1,000 population), and the number of full time equivalent (FTE) PMM and non-PMM primary care physicians rates (per 1,000 population), which came from the Brazilian

[1] Future work may consider using this event as a shock to asses a decrease in physician supply. Given that it is unlikely that Bolsonaro's critical comments and the Cuban government's reaction would have been foreseen or expected. The withdrawal of PMM doctors could reasonably considered to be an exogenous shock. The shock is very unlikely to be anticipated and would have impacted every treated municipality simultaneously. This shock, therefore, may allow for an unbiased estimate of the impact of a decrease in physician supply to be estimated. Indeed, it may provide a better setting to assess the impacts of the PMM programme than using its implementation, given the challenges that staggered roll-out, non-random assignment and anticipation effects cause in the setting investigated in this paper.

Ministry of Health. In our analysis we use hospitalisation rates, mortality rates and FTE doctor rates, rather than raw values. In addition to total hospitalisation and mortality rates, we also have hospitalisation and mortality rates from ambulatory care sensitive conditions (ACS) [2] and non-ambulatory care sensitive conditions (non-ACS) separately. ACS conditions are ones in which the need for emergency care is considered to be avoidable (Purdy et al., 2009; Sundmacher et al., 2015; Hodgson et al., 2019). It is therefore useful to consider ACS conditions separately from non-ACS as we may expect primary care physicians to be able to impact hospitalisations and mortality from these conditions more easily than non-ACS conditions, which primary care physicians may have limited influence over.

Our data also includes: total number of nurses, health facilities, and hospital beds within a municipality, which were also collected from the Brazilian Ministry of Health. In addition, we also have a number of time-invariant characteristics from the 2010 Brazilian Census, which includes a rich set of demographic and socio-economic characteristics.

4.3.1 Descriptive Statistics

Table 4.1 presents descriptive statistics of a range of municipality level variables in 2012, which was the year prior to the introduction of the PMM, and in 2014 the year after the introduction of the PMM.

From this table we show that both hospitalisations and mortality rates are lower for municipalities that would eventually receive PMM doctors prior to the policy introduction. This is also the case for each different type of hospitalisation and mortality. This pattern persists to the year after the introduction of PMM and differences are similar in magnitude.

In terms of doctor numbers, there is an expected lower number of primary care physicians in municipalities that eventually receive PMM physicians. This is also the case for the density of nurses, hospital beds, health facilities, percentage of private insurance and amount of health spending per person. Once again, this pattern persists post PMM introduction.

Municipalities that received PMM doctors are also, on average, poorer, with per capita GDP being lower for those municipalities, and also percentage of extremely poor population being higher. Infrastructure is on average worse in receiving municipalities, in relation to piped water and sanitation. Interestingly, there appears to be a relationship between whether the ruling political party of the municipality was

[2] Conditions that we consider to be ACS are based on the Brazilian list (ordinance number 221, 17 April 2008. Ministry of Health. Brazil), which are: nutritional diseases and anaemia, asthma and pulmonary diseases, hypertension, heart failure, cerebrovascular diseases, diabetes, epilepsy, gastric ulcer.

	Pre-PMM Introduction		Post-PMM Introduction	
	Never Received MM Doctors	*Received MM Doctor*	*Never Received MM Doctor*	*Received MM Doctor*
Health Outcomes				
Hosp. per 1,000	69.12	63.20	70.57	65.12
	(26.34)	(24.50)	(26.85)	(25.36)
ACSC Hosp. per 1,000	18.34	16.92	17.77	16.63
	(13.68)	(12.77)	(13.88)	(12.87)
non-ACSC Hosp. per 1,000	50.78	46.28	52.80	48.49
	(16.88)	(15.99)	(17.78)	(17.01)
Mortality per 1,000	6.54	6.03	6.73	6.21
	(1.94)	(1.72)	(1.84)	(1.69)
ACSC Mortality per 1,000	1.92	1.81	1.93	1.82
	(0.83)	(0.72)	(0.83)	(0.70)
non-ACSC Mortality per 1,000	4.62	4.21	4.80	4.39
	(1.51)	(1.26)	(1.45)	(1.26)
Health				
FTE MM physicians per 1,000	–	–	–	0.10
	(0.00)	(0.00)	(0.00)	(0.10)
FTE non-MM physicians per 1,000	0.57	0.44	0.60	0.41
	(0.26)	(0.21)	(0.27)	(0.23)
Nurses per 1,000	0.50	0.41	0.63	0.54
	(0.37)	(0.31)	(0.47)	(0.39)
Hospital Beds per 1,000	3.08	2.39	2.96	2.36
	(2.52)	(1.99)	(2.47)	(1.99)
Health Facilities per 1,000	3.30	2.56	3.23	2.54
	(2.42)	(1.95)	(2.38)	(1.95)
Health Spending per person	645.83	485.19	815.99	607.54
	(289.66)	(211.09)	(365.99)	(271.03)
% with Private Insurance	9.05	8.22	10.05	8.98
	(11.97)	(11.43)	(12.34)	(11.94)
Economic				
GDP per capita	18466.49	15175.34	21921.42	17920.77
	(30079.18)	(16538.24)	(33740.87)	(17257.61)
Income per capita	540.45	482.77	540.45	482.77
	(30079.18)	(16538.24)	(33740.87)	(17257.61)
% of extremely poor	7.60	12.21		
	(9.74)	(12.02)		
Political				
Ruling party in opposition to PMM	0.34	0.25		
	(0.47)	(0.43)		
Demographic				
Population	9212.23	40072.79		
	(13896.56)	(224880.49)		
% Rural Population	35.14	36.41		
	(21.43)	(22.18)		
% of the income of the poorest 20%	4.36	3.58		
	(1.50)	(1.52)		
% in households with piped water	89.19	84.77		
	(12.55)	(15.06)		
% with inadequate sanitation	4.85	10.21		
	(8.81)	(13.40)		
Geographic				
Semi-arid zone	0.15	0.24		
	(0.36)	(0.43)		
Legal Amazon	0.09	0.15		
	(0.29)	(0.36)		
Municipality Area	625.85	1736.46		
	(1120.45)	(6196.82)		
Municipality Altitude	490.40	394.63		
	(261.74)	(297.03)		
Municipality distance to capital	281.54	246.71		
	(147.20)	(166.56)		
Average winter rainfall	48.27	63.89		
	(49.52)	(57.48)		
Average summer rainfall	192.20	167.80		
	(63.66)	(76.41)		
Observations	1046	4524	1046	4524

† denotes variables which are time-invariant in our data.

TABLE 4.1: Descriptive Statistics for year prior to PMM introduction by whether municipality received PMM doctors

opposed to the introduction of the PMM. The primary care physician expansion was a "politically-contentious" (Hone et al., 2020) part of the PMM, and therefore this relationship may be expected.

In terms of demographic and geographic variables, populations of receiving municipalities were on average higher, with a marginally higher proportion living in

rural areas, and municipalities were much larger in land area. Receiving municipalities were also more likely to be semi-arid or legal Amazon areas.

There are clearly substantial differences between receiving municipalities and those that never received PMM physicians. Although the priority criteria was not strictly adhered to, the municipality averages do follow a general pattern of preference for those regions that were considered priority. Municipalities receiving PMM doctors were on average poorer, more rural, and more likely to be semi-arid, and these municipalities were the ones targeted by the programme.

The variation in PMM doctor numbers received by municipalities was also substantial. Figure 4.1 shows the distribution of PMM doctor numbers by municipality. As is shown, the distribution is heavily skewed and the majority received 1 or 2 PMM doctors, with 2 doctors being the 60th percentile. Indeed, the 95th percentile is 10 doctors but there are some extreme cases and in our data, one municipality received 230 full-time equivalent doctors.

Figure 4.2 shows the number of physicians across our study period. The figure shows numbers for total physicians in treated municipalities, as well as separately for PMM and non-PMM doctors. In addition, numbers of physicians in never treated municipalities are shown. Firstly upon initial programme roll-out it is clear to see the increase in PMM numbers, and this contributed to an increase in total number of physicians in these municipalities. However, it is also clear to see that at the same time there is a decrease in non-PMM doctors. It is therefore reasonable to presume, given the programme design, that some non-PMM doctors may have converted to PMM doctors, or that PMM doctors simply replaced non-PMM doctors. Although, it does not appear to be the case that physicians moved away from never-treated municipalities to work in treated ones, as PMM doctors, because although there is a small increase in physician numbers over the study period, there is no clear change in trend for never-treated municipalities.

Due to the fact that some doctors who were already working in priority municipalities became PMM doctors, the "pass-through rate" of PMM doctors is not equal to one. In other words, a one unit increase in PMM doctor does not directly correspond to a one unit increase in physicians overall. This is not surprising as Hone et al. already found such a result, however this requires consideration when interpreting our results.

4.4 Identification Strategy

The main aim of this paper is to estimate the effect of an increase in PMM physicians on both hospitalisations and mortality. As discussed, previous studies used a binary indicator of whether a municipality received any number of PMM doctors, and

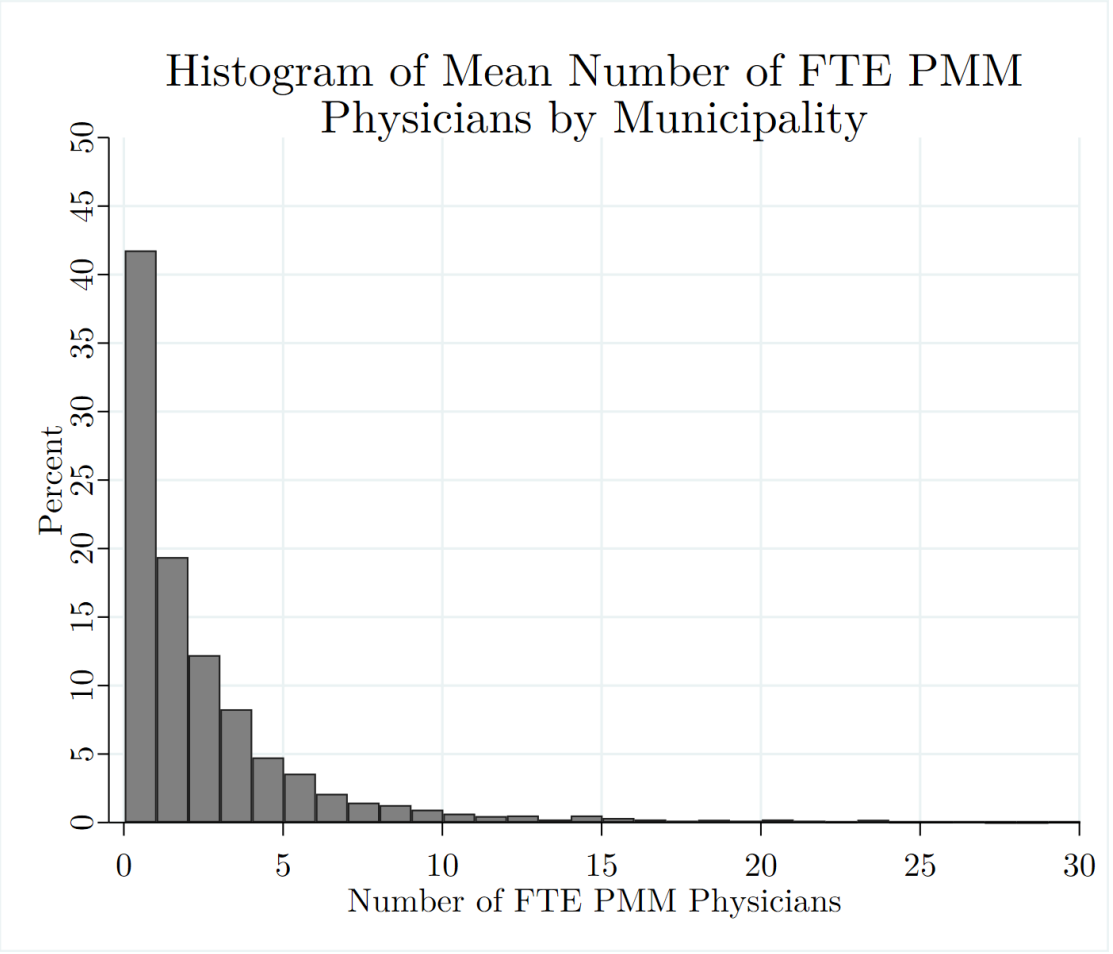


FIGURE 4.1: Histogram of Mean Number of FTE PMM Physicians by Municipality

Note: Histogram of the mean number of FTE PMM physicians by municipality for the period post-treatment. Municipalities that never received PMM doctors, or, municipalities that received more than 30 FTE PMM physicians are not included.

compared treated and non-treated municipalities. There is a concern that the estimates in previous studies are plagued by selection biases, because the policy was not intended to be random and municipalities were able to select into treatment. Instead, this paper seeks to exploit the variation in PMM numbers by municipality, and limit the estimating sample to those that received any PMM doctors, to estimate the marginal effect of receiving an PMM doctor on hospitalisations and mortality.

If we were seeking to estimate the marginal effect in this setting an obvious solution would be to use a difference-in-differences identification strategy with a continuous treatment, in place of the usual binary treatment. Indeed, previous studies used a difference-in-differences specification to estimate the effects. However, there are several issues with this empirical strategy in our setting.

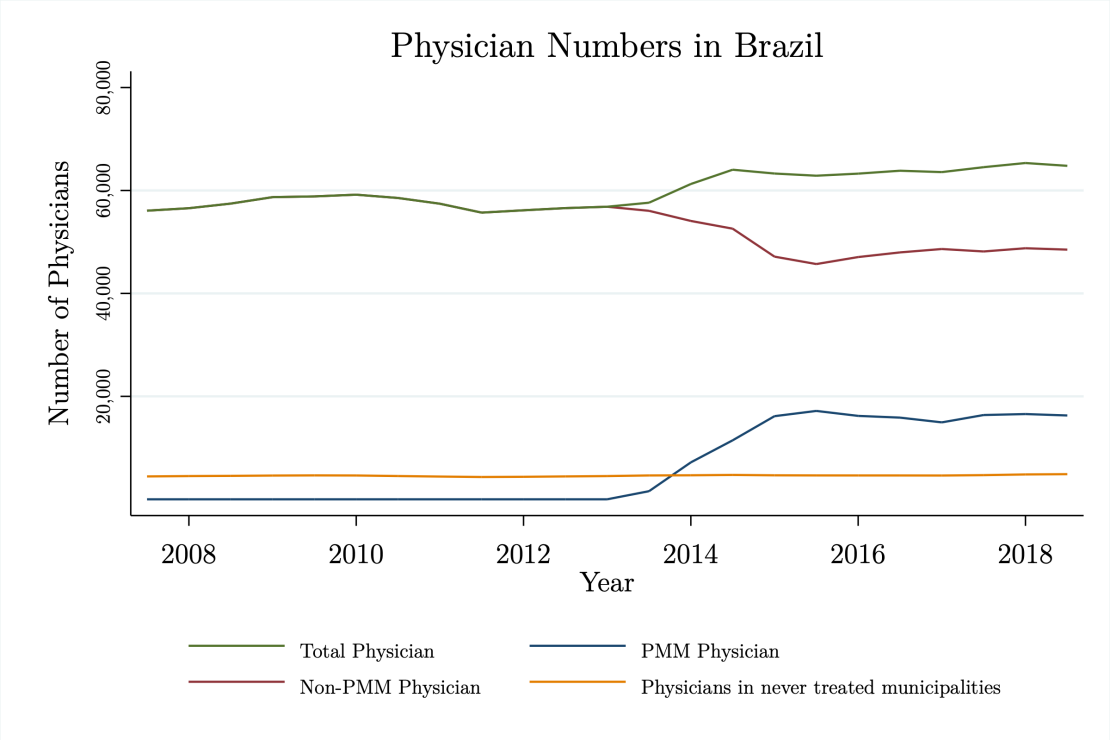


FIGURE 4.2: Number of Physicians in our study period, by physician type (PMM and non-PMM) and never-treated municipalities.

Note: Figure shows the number of physicians over time. Green line represents total physician numbers (both PMM and non-PMM) for municipalities that eventually received PMM-physicians. Blue line represents PMM physicians. Red line represents non-PMM physician numbers for municipalities that did receive PMM physicians. Yellow line represents the number of physicians in municipalities that never received any PMM doctors.

Firstly, the major concern when analysing the impact of the increase in PMM physicians is that the allocation to municipalities may not be random. The programme itself had a priority criteria, and although this criteria was not well adhered to (between 2013 and 2014, 22.3% of receiving municipalities did not meet the priority criteria (Oliveira et al., 2016)) the aim of the programme was to allocate PMM doctors to underserved areas, and therefore by design, the allocation was not intended to be random. There is also potential for a further form of self-selection into receiving PMM doctors, as municipalities were required indicate they wanted to receive PMM doctors (Mattos and Mazetto, 2019). The eventual allocation of the programme makes it somewhat more difficult to determine whether the assignment was indeed random or not. However, the difference-in-differences identification requires a less restrictive assumption than random assignment into treatment. Instead, for the difference difference-in-differences estimate to be an unbiased estimate, it is required that there are no time-varying municipality unobservables. This is usually called the "parallel trends" assumption and requires that municipalities would have followed the same

trajectory if they had not received the treatment. To assess whether this holds in a particular setting an event-study is usually estimated to check the pre-treatment time dynamics. To assess whether the the parallel trends assumption holds in our setting we estimate the following model:

$$Y_{i,t} = \beta_0 + \sum_{l=-2}^{-11} \beta^l_{pre} D^l_{i,t} \times Treatment_i + \beta_{post} Post_{i,t} \times Treatment_i + \lambda_t + \mu_i + e_{i,t} \quad (4.1)$$

Where $Y_{i,t}$ represents total hospitalisations or mortality per 1,000 population, $D^l_{i,t}$ is a dummy variable taking value 1 if municipality i is l periods pre-treatment in period t. $Treatment_i$ is an indicator variable which takes value 1 if municipality i ever received a PMM physician in our data. $Post_{i,t}$ is an indicator variable taking value 1 if municipality i is in a post-treatment period at time t. λ_t and μ_i are period and municipality fixed-effects respectively. We present results of this precise specification, as well as specifications including control variables, and linear time trends for the control variables in table 4.2 [3]

We find that there is indeed evidence that the parallel trends assumption does not hold in our setting. Looking at columns (1) to (3) of table 4.2, we find that in most pre-periods there is no evidence of differential trends in terms of hospitalisations. However, there is some evidence that two periods prior to receiving a PMM treatment and control groups do diverge, namely hospitalisation rates decrease in treated municipalities. This differing trend is more pronounced when looking at total mortality in columns (4) to (6). For mortality, there appears to be a significantly lower mortality rate in the five periods prior to receiving a PMM doctor. These results suggest that there are significant differences in the time trends of treatment and control groups, and therefore the parallel trends assumption is violated in our setting. This makes the difference-in-differences estimator problematic.

The second concern with regards to the use of a difference-in-differences in our setting is that the adoption of the policy was staggered, in other words, different municipalities received their first PMM doctor at different times, and indeed received a different number of PMM doctors. There is a substantial recent, and growing, literature on the issue of estimates being biased when using a two-way fixed-effects (TWFE) difference-in-difference in a staggered roll-out setting (Borusyak and Jaravel, 2017; Athey and Imbens, 2018; Callaway and Sant'Anna, 2020; de Chaisemartin and D'Haultfœuille, 2020; Sun and Abraham, 2020; de Chaisemartin and D'Haultfoeuille, 2021; Goodman-Bacon, 2021; Roth and Sant'Anna, 2021). Previous studies on the

[3] The time-varying control variables included are: density of non-PMM primary care doctors, density of nurses, beds and health facilities, and percentage with private healthcare insurance. Additionally, time-invariant control variables interacted with the linear-time trends are: percentage of population in extreme poverty, percentage of population in rural areas, per capita income, and total municipality population.

	Total hospitalisation rate			Total mortality rate		
	(1)	(2)	(3)	(4)	(5)	(6)
$Treatment \times Post$	0.056	0.095	0.204	-0.035**	-0.038**	-0.043***
	(0.189)	(0.191)	(0.189)	(0.013)	(0.013)	(0.013)
$Treatment \times T_{treat} - 2$	-0.262*	-0.247	-0.276*	-0.081***	-0.081***	-0.079***
	(0.133)	(0.132)	(0.132)	(0.015)	(0.015)	(0.015)
$Treatment \times T_{treat} - 3$	0.010	0.034	-0.017	-0.049**	-0.049**	-0.045**
	(0.196)	(0.196)	(0.194)	(0.015)	(0.015)	(0.015)
$Treatment \times T_{treat} - 4$	0.137	0.161	0.099	-0.057***	-0.058***	-0.051**
	(0.228)	(0.226)	(0.224)	(0.017)	(0.017)	(0.017)
$Treatment \times T_{treat} - 5$	0.136	0.153	0.084	-0.054**	-0.054**	-0.047**
	(0.269)	(0.267)	(0.265)	(0.018)	(0.018)	(0.018)
Municipality Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Period Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	Yes	Yes	No	Yes	Yes
Controls × Linear Time Trend	No	No	Yes	No	No	Yes
Number of Municipalities	3,542	3,542	3,542	3,542	3,542	3,542
Observations	70,844	70,844	70,844	70,844	70,844	70,844

$T_{treat} - k$ is a binary indicator denoting k periods prior to receiving municipalities first PMM doctor, which happened in period T_{treat}. Regressions are estimated using a balanced panel. Control variables included are: density of non-PMM primary care doctors, density of nurses, beds and health facilities, percentage of population in extreme poverty, percentage of population in rural areas, percentage with private healthcare insurance, per capita income, total municipality population. All specifications included eleven pre-treatment periods, however only five treatment-period interactions are presented here. Treated municipalities are "pruned" so that at most eleven pre-treatment periods are included in the regressions. Cluster-robust standard errors are presented in the parentheses, and are clustered at the municipality level. *** denotes P-value of 0.01 or less , ** denotes P-value of 0.05 or less, * denotes P-value of 0.10 or less

TABLE 4.2: Two-way Fixed Effects Estimates of the impact on total hospitalisations and mortality, including pre-trends

PMM by Fontes et al. (2018); Carrillo and Feres (2019); Mattos and Mazetto (2019); Hone et al. (2020) have overcome this issue by using the period of initial roll-out of the programme as the post treatment period for municipalities. Although this approach would likley provide unbiased estimates due to the variation in treatment timing being removed, it is also likely that the estimates would be an underestimate of the true effect. It is possible to estimate an unbiased treatment effect in a setting staggered roll-out, and much of this recent literature have proposed estimators to do so, however many of these estimators require a binary treatment, unlike the continuous treatment effect we seek to use in this paper. de Chaisemartin and D'Haultfoeuille (2021) is an example that does allow the use of a continuous treatment under staggered adoption. The issue with de Chaisemartin and D'Haultfoeuille's estimator is that it still requires the parallel trends assumption to hold, which, as shown, is not the case in our setting.

The final concern is that receiving a PMM doctor was not an "absorbing state", in the sense that some municipalities received PMM doctors and then at some point after receiving PMM doctors the municipality is observed to have no PMM doctors. The flow of PMM is doctors is somewhat more dynamic than what is assumed using a

difference-in-differences specification[4].

These issues make the difference-in-differences estimator problematic in our setting. Our solution to this is to use a generalised synthetic control estimator with a continuous treatment (Powell, 2021), to estimate the effect of a marginal increase in PMM doctors on hospitalisations and mortality. The GSC methodology is an attractive methodology for our setting, and has previously been found to outperform difference-in-differences and interactive fixed effects methodologies in simulations (O'Neill et al., 2020). The GSC estimator has a number of useful advantages over traditional synthetic control and difference-in-differences estimators, some of which are particularly salient in this context.

Firstly, Powell (2021)'s GSC estimator, and synthetic control estimators more generally, does not require the assumption of parallel trends (Abadie, 2021). Instead of comparing the trends of treated units to control units, and ensure they are reasonable comparisons by making sure their trends are parallel, synthetic control estimators optimally choose a weighted combination of other units, so that the synthetic control, by construction, is a reasonable control unit to the corresponding treated unit, and therefore follows a similar time trend.

Secondly, the GSC estimator does not suffer from biases that TWFE difference-in-differences estimators do when the treatment roll-out is staggered. The coefficient of a TWFE difference-in-differences estimator do not correspond to an average treatment effects (ATE) when treatment effects are heterogeneous across time or units (Roth et al., 2022), which is the case in a staggered roll-out setting. To correctly estimate an ATE, using a TWFE difference-in-differences, the treatment effect must be constant, which excludes staggered roll-out or event studies in many cases. The synthetic control estimator avoids these issues because it does not impose any restrictions on treatment effect heterogeneity. The synthetic control estimates a treatment effect for each treated unit by simply comparing the differences between the treated unit and its corresponding synthetic control. Therefore, technically the synthetic control estimator not only allows a different treatment effect to be estimated for each treated unit but also for each time unit. Given that the synthetic control estimator can estimate these heterogeneous treatment effects, the bias arising from staggered roll-out is not a concern.

Additionally, the GSC estimator allows us to limit our sample to only treated municipalities. Limiting our sample to only treated municipalities is a very useful advantage of this methodology. As municipalities themselves were able to apply to receive PMM doctors and assignment was designed not to be random, and indeed we show that the trends of treated and non-treated differ prior to receiving treatment.

[4] Again, de Chaisemartin and D'Haultfoeuille (2021) does propose a difference-in-differences estimator that does not require this restriction.

This is a concern when estimating treatment effects because the selection into the treatment may be inducing bias into our estimate of the treatment effect. If municipalities that received the treatment were systematically different to those that did not receive the treatment then our estimates would be biased. In our analysis, limiting the sample to only municipalities that received PMM doctors eliminates this concern, as we are comparing municipalities that received PMM doctors with each other, rather than estimating the treatment to be the difference between treated and untreated municipalities.

Finally, the GSC estimator is able to estimate a treatment effect for a continuous treatment variable, which is the primary aim of this paper.

4.4.1 Generalised Synthetic Control

The basic idea behind the traditional synthetic control estimator (Abadie et al., 2010, 2015, 2020; Abadie, 2021) is to generate a counterfactual unit from the units not exposed to the policy. The counterfactual unit, which is named the synthetic control, is a weighted average of untreated units, where the optimal weights are such that the synthetic control resembles the treated unit had it not been treated. The weights are chosen so that the synthetic control resembles the pre-treatment characteristics of the treated unit. The treatment effect is then simply the difference between the outcome of the treated unit and the synthetic control. [5]

Powell (2021)'s GSC extends this approach beyond a case study with single treated unit and instead allows for many treated units and a continuous or multiple discrete treatment variables, by estimating a synthetic control for each unit, whether treated or not. Unlike the traditional synthetic control estimator, because the GSC creates a synthetic control for each of many treated units, the variation in the treatment intensity can be exploited to estimate a continuous treatment effect.

Powell models outcomes in the following way:

$$Y_{i,t}^{N} = \delta_t + \boldsymbol{\lambda_t}\boldsymbol{\mu_i} + \epsilon_{it} \tag{4.2}$$

$$Y_{i,t} = Y_{i,t}^{N} + \boldsymbol{\alpha}\boldsymbol{D_{i,t}} \tag{4.3}$$

where $Y_{i,t}^{N}$ denotes the untreated outcome for municipality i in time t, δ_t is an unknown constant common factor, $\boldsymbol{\lambda_t}$ is a vector of unobservable common factors, $\boldsymbol{\mu_i}$ is a vector of unknown factor loadings, and the error term ϵ_{it} is the unobserved shocks

[5] A formal illustration of the traditional synthetic control is available in Abadie et al. (2010), Abadie et al. (2015) and Abadie (2021).

to municipality i in time t which has mean zero. $\boldsymbol{D_{i,t}}$ is a vector of treatment variables, which in our case is the continuous measure of PMM doctor numbers, and $\boldsymbol{\alpha}$ is a vector of treatment effects.

As with the standard synthetic control estimator (Abadie et al., 2010, 2015; Abadie, 2021), the estimation comes from the assumption that a counter-factual unit for each municipality can be estimated using a weighted sum of the donor municipalities. For municipality i the counter-factual, or the synthetic control, is the weighted average of all other municipalities:

$$Y_{i,t}^{N} = \sum_{j \neq i} \omega_i^j Y_{j,t}^{N} \tag{4.4}$$

Where ω_i^j denotes the weight assigned to municipality j to generate the synthetic control for municipality i, where the weights are usually constrained to be non-negative and sum to 1. Under this assumption, it is possible to estimate $\boldsymbol{\alpha}$ by minimising the following objective function:

$$\underset{\boldsymbol{b}, \phi_1 \in \omega_1, \ldots, \phi_N \in \Omega_N}{\operatorname{argmin}} \left\{ \frac{1}{2NT} \sum_{i=1}^{N} \sum_{t=1}^{T} \left[Y_{it} - \boldsymbol{D}_{it}^{'}\boldsymbol{b} - \sum_{j \neq i} (\phi_i^j (Y_{jt} - \boldsymbol{D}_{jt}^{'}\boldsymbol{b})) \right]^2 \right\} \tag{4.5}$$

Where $\boldsymbol{D}_{it}^{'}$ can be a a vector of indicator, continuous or interaction variables. $\sum_{j \neq i} (\phi_i^j (Y_{jt} - \boldsymbol{D}_{jt}^{'}\boldsymbol{b})$ is the untreated outcome of municipality i, or the counter-factual outcome $Y_{i,t}^{N}$ derived from equation 4.3.

4.4.1.1 Implementation

We now present how Powell's GSC was implemented in our setting. In our analysis we model outcomes in the same way as Powell:

$$Y_{i,t}^{N} = \delta_t + \boldsymbol{\lambda_t}\boldsymbol{\mu_i} + \epsilon_{it} \tag{4.6}$$

$$Y_{i,t} = Y_{i,t}^{N} + \alpha D_{i,t} \tag{4.7}$$

Where $Y_{i,t}$ denotes hospitalisation or mortality rates in municipality i in period t. $Y_{i,t}^{N}$ are the hospitalisations or mortality rates had the municipality not been treated. $D_{i,t}$ is PMM doctor density, and as above δ_t is an unknown constant common factor, $\boldsymbol{\lambda_t}$ is a vector of unobservable common factors, $\boldsymbol{\mu_i}$ is a vector of unknown factor loadings, and the error term ϵ_{it} is the unobserved shocks to municipality i in time t which has mean zero. Finally, α is the treatment effect of PMM doctors which we are seeking to

estimate. The counter-factual unit $Y_{i,t}^N$ exists by assumption, however we weight municipalities that have better fitting synthetic controls in the pre-period higher than those that fit less well when estimating α.

Powell (2021) states that it is possible to jointly estimate ω_i^j and $\boldsymbol{b}$ from equation 4.5 by minimising the objective function with respect to both parameters. Although this is possible in our setting, the computational burden of this approach as well as us observing a pre-treatment period for all municipalities leads us to diverge from this approach somewhat. Because we observe a pre-treatment period, we are able to estimate the synthetic control weights ω_i^j using only pre-treatment periods where we can guarantee that both $\boldsymbol{D}_{it}^{'}\boldsymbol{b} = 0$ and $\sum_{j\neq i}\phi_i^j \boldsymbol{D}_{jt}^{'}\boldsymbol{b} = 0$. Therefore, we estimate the synthetic control weights ω_i^j for all municipalities using pre-treatment periods. Candidates for the synthetic control of municipality i are municipalities which are in the same federal unit as municipality i [6]. We estimate synthetic controls for each municipality in the usual way proposed by Abadie et al. (2010, 2015, 2020) and we include a number of observable covariates to estimate these weights, named the predictor variables. These predictor variables include: density of nurses, beds and health facilities, percentage of population in extreme poverty, percentage of population in rural areas, percentage with private healthcare insurance, per capita income, total municipality population, as well as the outcome for each pre-treatment year and density of non-PMM primary care doctors for each pre-treatment year. Given these synthetic control weights ω_i^j, we are then able to estimate the treatment effect α from equation 4.7, and do so by regressing the difference between the observed and synthetic outcomes $Y_{i,t}^N - \sum_{j\neq i}\omega_i^j Y_{j,t}^N$ on the difference in treatment intensity (PMM physicians density) $D_{i,t}^N - \sum_{j\neq i}\omega_i^j D_{j,t}^N$.

In summary, our procedure is as follows:

1. Generate Synthetic control weights ω_i^j for each municipality i. These weights are estimated using the pre-treatment periods 2007-2012 and predictor variables. In periods 2007-2012 no municipality received PMM doctors and therefore: $\boldsymbol{D}_{it}^{'}\boldsymbol{b} = 0$ and $\sum_{j\neq i}\phi_i^j \boldsymbol{D}_{jt}^{'}\boldsymbol{b} = 0$. Municipality j is considered as a donor for municipality i, if i and j are in the same Federal unit, and that we have data on municipality j for at least as long as municipality i.

2. Use weights from step 1 to generate synthetic outcomes $\sum_{j\neq i}\phi_i^j Y_{jt}$ and synthetic treatment $\sum_{j\neq i}\phi_i^j \boldsymbol{D}_{jt}^{'}$ for each municipality and each period.

[6] We do this for two reasons. The first is computational. Estimating a synthetic control for each of 5570 municipality in our sample using all other municipalities in our sample is extremely computationally intensive. Secondly, municipalities within the same Federal Unit are likely better candidates for a synthetic control because many health related policies and administrative decisions are made at the federal level, therefore municipalities in the same Federal unit are more likely to be exposed to similar shocks. There is also a further concern that we may be over-fitting the synthetic control (Abadie, 2021). By minimising the pool of donors, this would be less of a concern.

3. Create "synthetic control fit" weights, so that better fitting synthetic controls are more heavily weighted in the estimation than synthetic controls that fit less well. We consider the pre-treatment fit of synthetic controls, and for each municipality i. The "synthetic control fit" weight is:
$\widehat{\Omega}_i = \frac{1}{\left(\sqrt{\frac{1}{T_0}\sum_{t=1}^{T_0}\left[Y_{it}-\sum_{j\neq i}\hat{w}_i^j Y_{jt}\right]^2}\right)}$. Where T_0 is period prior to the first municipality receiving a PMM doctor.

4. Finally, we estimate α by regressing $Y_{it} - \sum_{j\neq i}\hat{w}_i^j Y_{jt}$ on $D_{it}^{'} - \sum_{j\neq i}\hat{w}_i^j D_{jt}^{'}$, and using the "synthetic control fit" weights from step 3.

4.4.2 Covariate balance

The key identifying assumption of the GSC estimator is that a synthetic control can be estimated for each municipality, such that $Y_{i,t}^N = \sum_{j\neq i}\omega_i^j Y_{j,t}^N$ (Powell, 2021). As discussed in section 4.4.1, we do weight better fitting units higher than poorly fitting units, however this procedure does not *per se* ensure that this assumption holds when estimating the treatment effects. Therefore, we present the root mean squared error (RMSE) of the estimated synthetic control ($\sum_{j\neq i}\omega_i^j Y_{j,t}^N$) compared to the observed outcome ($Y_{i,t}^N$) for the years prior to initial PMM roll-out (T_0). In doing so, we provide evidence that the necessary assumption holds in our setting. We also present the weighted RMSE, which use the weights from step 3 of the procedure in section 4.4.1.1. In addition, we present the RMSE of predictor variables used for estimating the synthetic controls.

An additional concern is that our estimates are substantially impacted from leaving out important predictor variables when generating the synthetic controls. It is possible that if we leave out an important variable we may be generating synthetic controls that are poor comparisons for municipalities. To ensure that this is not the case, we estimate synthetic controls for each municipality using no additional predictor variables, only the outcomes and the number of non-PMM doctors for each year prior to T_0, and an alternative set of predictor variables [7].

Tables 4.3 and 4.4 presents the RMSE for total hospitalisation rate as the outcome, for the binary treatment and continuous treatment models respectively. Figures 4.3 and 4.4 show the RMSE of the hospitalisation rate for pre-PMM roll-out periods and each

[7] These variables were chosen by doing a prediction exercise, to predict whether a municipality would ever receive a PMM doctor. The alternative set of predictor variables are: whether ruling party was in opposition to PMM, municipality area, distance to municipality capital, health spending per person, average winter rainfall, average summer rainfall, ratio of population over 65 years of age, percentage of school attendance of the population aged 25 to 29 years old, percentage of the income of the poorest 20%, percentage employed in the trade sector, percentage of the population living in households with piped water, percentage of people in households with water supply and sanitation inadequate, and total rural population.

specification. For the outcome variables in all hospitalisation models the RMSE is approximately 11% of the mean value of hospitalisations. When we weight better fitting synthetic controls higher, we find that this difference reduces to approximately 2.5% of the mean outcome. The graphical representation of the RMSE of hospitalisation rate shows that the RMSE is stable for each pre-PMM roll-out period, aside from a substantial rise in RMSE in the period prior to PMM roll-out for non-weighted specifications. The main specification shows the biggest rise in both the binary and continuous models, but when we do the weighting procedure to improve fit, we not only substantially reduce RMSE overall but also remove this increase in RMSE in the period pre roll-out.

Total hospitalisations rate - Binary Treatment Model						
	Main Specification		No Predictor Variables		Alternative Predictor Variables	
	RMSE	Weighted RMSE	RMSE	Weighted RMSE	RMSE	Weighted RMSE
Total hospitalisations rate						
$T_0 - 11$	3.4037	0.8892	2.9068	0.8466	3.2337	0.9186
$T_0 - 10$	3.5396	0.8871	3.1651	0.8654	3.5868	0.9147
$T_0 - 9$	3.6947	0.8989	3.0556	0.8587	3.379	0.9092
$T_0 - 8$	3.735	0.8885	3.2339	0.8608	3.5348	0.915
$T_0 - 7$	3.8449	0.8865	3.3282	0.8662	3.4966	0.9348
$T_0 - 6$	3.6931	0.8811	3.1946	0.8681	3.1473	0.9071
$T_0 - 5$	3.7541	0.8793	3.1591	0.8558	3.1515	0.9082
$T_0 - 4$	3.6604	0.8577	3.2736	0.8566	3.3224	0.8904
$T_0 - 3$	3.6771	0.8722	3.0567	0.8629	2.9925	0.8908
$T_0 - 2$	3.7347	0.8757	3.3144	0.868	3.2881	0.9072
$T_0 - 1$	4.6275	0.9174	3.7366	0.8802	3.2306	0.9048
Predictor Variables						
non-PMM Physicians per 1,000	0.1842	0.1228	0.2606	0.166	0.1907	0.1297
Nurses per 1,000	0.2643	0.1593	0.278	0.1701	0.2409	0.1509
Hospital Beds per 1,000	1.8308	0.8395	2.1756	0.9018	1.8939	0.8049
Health Facilities per 1,000	3.3888	2.0184	3.4723	2.0629	3.1817	1.8971
% with Private Insurance	9.2928	2.2881	9.7077	2.2947	7.8919	1.7613
Income per capita	165.7035	19.6608	167.1864	44.6907	143.7351	32.377
% of extreme poor	7.2772	1.3443	6.8626	1.6619	7.5085	1.9038
Population	249,653.1	7,655.394	205,386.9	4,441.169	163,770.8	6,788.176
% Rural Population	18.4018	6.5344	20.4081	4.5862	7.8919	1.7613

RMSE between the observed municipality observation, and their corresponding synthetic control. T_0 denotes the period of initial PMM roll-out. Synthetic controls were estimated using the variables used for the main specification, with no predictor variables, and with an alternative set of controls. All synthetic controls were estimated using the outcome density of non-PMM primary care doctors for each year prior to roll-out. Main specification predictor variables additionally include: density of nurses, beds and health facilities, percentage of population in extreme poverty, percentage of population in rural areas, percentage with private healthcare insurance, per capita income, total municipality population. The alternative set of predictor variables are: whether ruling party was in opposition to PMM, municipality area, distance to municipality capital, health spending per person, average winter rainfall, average summer rainfall, ratio of population over 65 years of age, percentage of school attendance of the population aged 25 to 29 years old, percentage of the income of the poorest 20%, percentage employed in the trade sector, percentage of the population living in households with piped water, percentage of people in households with water supply and sanitation inadequate, and total rural population.

TABLE 4.3: Root Mean Squared Error (RMSE) of Total Hospitalisations and predictor variables using estimated binary treatment GSC weights for all municipalities

It is reassuring to note that, for hospitalisation outcome models that include no predictor variables, the RMSE are comparable in size. As may be expected the RMSE

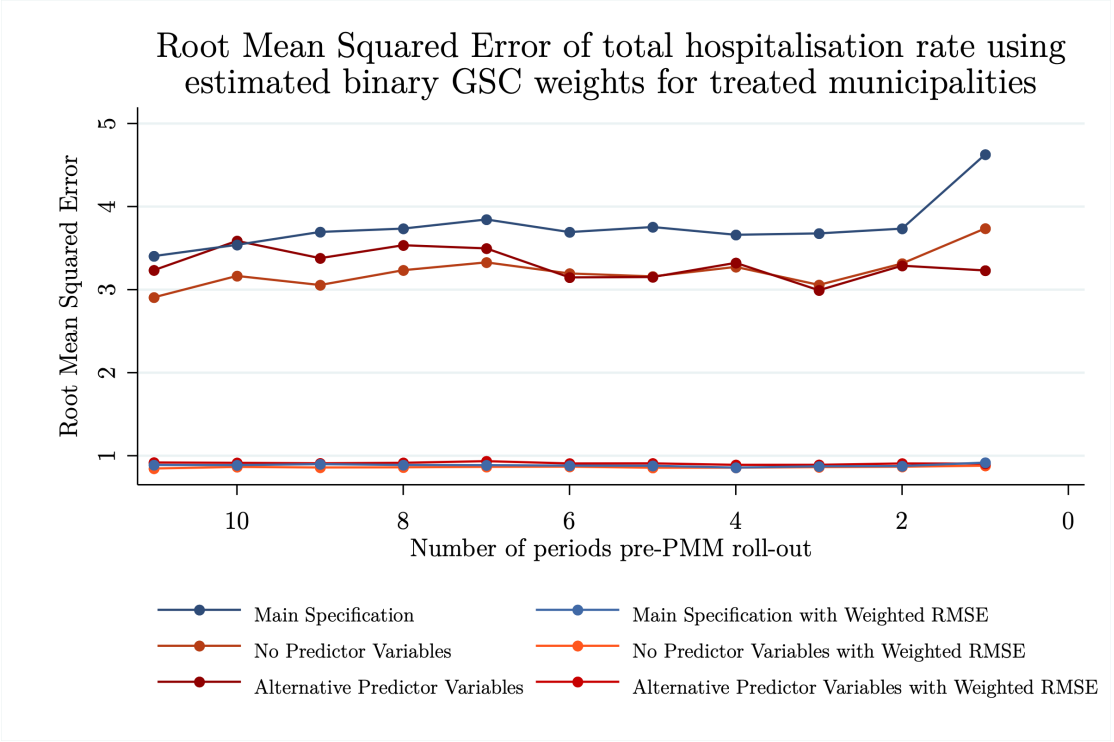


FIGURE 4.3: Root Mean Squared Error (RMSE) of Total Hospitalisation rate using estimated binary treatment GSC weights for all municipalities

Graphical representation of the RMSEs shown in Table 4.3. Each line shows the RMSE or weighted RMSE for either the main, no predictor variables, or alternative predictor variables specifications for each period pre-PMM roll-out.

of the outcomes are lower, however the predictor variables that are not included in those synthetic controls estimates also have RMSEs of similar size, albeit larger in magnitude. The synthetic controls using the alternative predictor variables also have RMSEs of a similar magnitude. Therefore we are confident that if we excluded a variable that was relevant for generating the synthetic control, or included one that was not relevant, the synthetic controls estimates would not differ substantially.

Tables 4.5 and 4.6, present the binary treatment and continuous treatment models respectively but for for total mortality rate. Figures 4.5 and 4.6 show the RMSE of the mortality rate for pre-PMM roll-out periods for each specification. For mortality models, the RMSE of the outcome variables follow a similar pattern as for hospitalisations. Where the RMSE is approximately 10% of the mean value. Using the weights, the RMSE reduces to approximately 3%. Once again, the RMSE when we use no predictor variables when generating the synthetic controls are comparable, however are larger in magnitude. The RMSE is also of a similar magnitude when we use the alternative predictor variables when generating the synthetic controls. The graphical analysis shows that the RMSE is quite stable. As with hospitalisations rate,

Total hospitalisations rate - Continuous Treatment Model						
	Main Specification		No Predictor Variables		Alternative Predictor Variables	
	RMSE	Weighted RMSE	RMSE	Weighted RMSE	RMSE	Weighted RMSE
Total hospitalisations rate						
$T_0 - 11$	3.5834	0.9215	2.9491	0.803	3.1199	0.88
$T_0 - 10$	3.7343	0.9113	3.1972	0.8252	3.4664	0.885
$T_0 - 9$	3.9457	0.9316	3.1224	0.8142	3.359	0.8862
$T_0 - 8$	3.8193	0.9014	3.14	0.7949	3.2884	0.8826
$T_0 - 7$	3.9913	0.9102	3.3428	0.8145	3.3019	0.8924
$T_0 - 6$	3.8814	0.9025	3.2342	0.8141	3.1218	0.887
$T_0 - 5$	3.872	0.8905	3.2462	0.7981	3.1941	0.887
$T_0 - 4$	3.7441	0.8702	3.1631	0.7917	3.092	0.8617
$T_0 - 3$	3.7704	0.8982	2.988	0.8014	2.9735	0.8765
$T_0 - 2$	3.8756	0.8997	3.2112	0.8117	2.9731	0.8785
$T_0 - 1$	4.7477	0.9421	3.7525	0.8231	3.0229	0.882
Predictor Variables						
non-PMM Physicians per 1,000	0.1821	0.1215	0.2429	0.1569	0.2142	0.1405
Nurses per 1,000	0.2574	0.1531	0.2612	0.161	0.2433	0.1527
Hospital Beds per 1,000	1.7186	0.814	1.996	0.8768	1.9176	0.8429
Health Facilities per 1,000	3.2659	1.945	3.3411	2.0001	3.2682	1.9686
% with Private Insurance	9.1561	2.2706	9.2646	2.2107	8.1003	1.9101
Income per capita	165.7127	19.9137	167.097	45.2937	156.6949	37.6506
% of extreme poor	7.4302	1.5257	7.2035	1.8015	7.31	2.0828
Population	266,738.8	9,990.104	227,588.3	6,008.666	166,107	6,554.554
% Rural Population	18.6184	6.6612	20.5535	7.5271	20.5	7.4808

RMSE between the observed municipality observation, and their corresponding synthetic control. T_0 denotes the period of initial PMM roll-out. Synthetic controls were estimated using the variables used for the main specification, with no predictor variables, and with an alternative set of controls. All synthetic controls were estimated using the outcome density of non-PMM primary care doctors for each year prior to roll-out. Main specification predictor variables additionally include: density of nurses, beds and health facilities, percentage of population in extreme poverty, percentage of population in rural areas, percentage with private healthcare insurance, per capita income, total municipality population. The alternative set of predictor variables are: whether ruling party was in opposition to PMM, municipality area, distance to municipality capital, health spending per person, average winter rainfall, average summer rainfall, ratio of population over 65 years of age, percentage of school attendance of the population aged 25 to 29 years old, percentage of the income of the poorest 20%, percentage employed in the trade sector, percentage of the population living in households with piped water, percentage of people in households with water supply and sanitation inadequate, and total rural population.

TABLE 4.4: Root Mean Squared Error (RMSE) of Total Hospitalisations and predictor variables using estimated continuous GSC weights for treated municipalities

in the period pre-PMM roll-out there is a slight rise in RMSE for the non-weighted specifications, whereas the weighted versions have RMSE which are much smaller in magnitude and do not show this rise just pre roll-out.

We believe, therefore, that these statistics provide evidence that the synthetic controls generated for each municipality are good matches due to the difference between the actual outcomes and synthetic control outcomes being relatively small on average. We are also confident that the synthetic controls are comparable in terms of characteristics to the municipalities characteristics themselves, due to the RMSE of these characteristics being relatively small. Reassuringly, excluding predictor variables, or including an alternative set of predictor variables does not appear to have a substantial impact of the fit of the synthetic controls for municipalities. Finally,

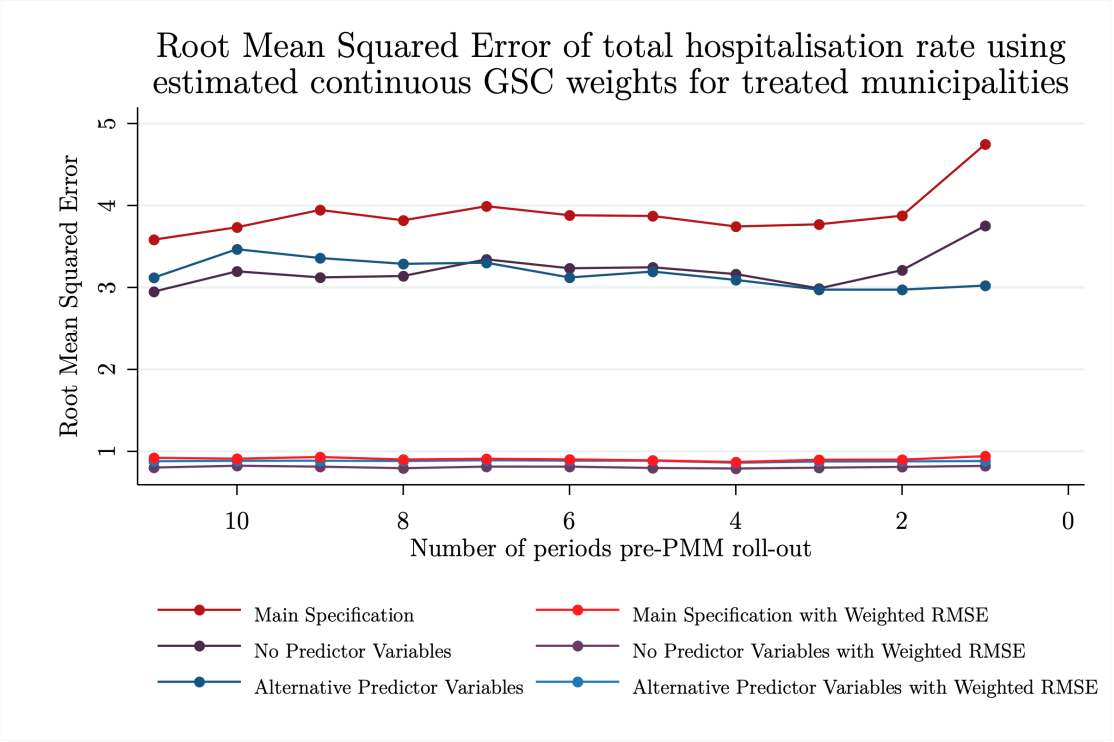


FIGURE 4.4: Root Mean Squared Error (RMSE) of Total Hospitalisations and predictor variables using estimated continuous GSC weights for treated municipalities

Graphical representation of the RMSEs shown in Table 4.4. Each line shows the RMSE or weighted RMSE for either the main, no predictor variables, or alternative predictor variables specifications for each period pre-PMM roll-out.

weighting better fitting does have a substantial impact on the RMSE, and is therefore a useful step in ensuring that a the required assumption holds when estimating the impact of PMM doctor density.

4.4.3 Inference

By design the synthetic control estimator creates correlations between units, because each municipality potentially contributes to a synthetic control of another municipality. Inference in this setting is potentially problematic, given that most inference methods assume clusters are asymptotically independent. We use the CAI inference procedure included in Powell (2021) [8]. The CAI inference procedure proposed by Powell allows for this within-unit dependence. Indeed, the procedure allows for dependence across all observations, not only within clusters. This approach is also preferred to the usual placebo-based inference usually used for SC, because

[8]Full details of the procedure are available in Powell (2021) and Powell (2020a).

Total mortality rate - Binary Treatment Model						
	Main Specification		No Predictor Variables		Alternative Predictor Variables	
	RMSE	Weighted RMSE	RMSE	Weighted RMSE	RMSE	Weighted RMSE
Total mortality rate						
$T_0 - 11$	0.2793	0.0891	0.3195	0.1121	0.2942	0.0954
$T_0 - 10$	0.285	0.09	0.3203	0.1125	0.2945	0.0944
$T_0 - 9$	0.2799	0.0905	0.3184	0.1119	0.2915	0.0937
$T_0 - 8$	0.2817	0.0897	0.326	0.1123	0.297	0.0936
$T_0 - 7$	0.2741	0.0888	0.3217	0.1135	0.2921	0.0955
$T_0 - 6$	0.2882	0.0916	0.3193	0.1126	0.293	0.0951
$T_0 - 5$	0.2848	0.0896	0.3241	0.1127	0.2985	0.0951
$T_0 - 4$	0.3028	0.0903	0.3312	0.1135	0.3076	0.096
$T_0 - 3$	0.2733	0.0891	0.3161	0.1123	0.285	0.094
$T_0 - 2$	0.2776	0.09	0.3197	0.1141	0.285	0.0945
$T_0 - 1$	0.2984	0.0923	0.331	0.1144	0.3061	0.0965
Predictor Variables						
non-PMM Physicians per 1,000	0.1838	0.1265	0.2606	0.166	0.1974	0.1393
Nurses per 1,000	0.2629	0.1568	0.278	0.1701	0.2341	0.1488
Hospital Beds per 1,000	2.0722	0.8976	2.1756	0.9018	1.9751	0.8083
Health Facilities per 1,000	3.3877	2.0128	3.4723	2.0629	3.178	1.9026
% with Private Insurance	9.4742	2.1716	9.7077	2.2947	7.8273	1.6172
Income per capita	166.4536	38.8098	167.1864	44.6907	143.5939	30.5537
Proportion of extreme poor	6.9481	1.2964	6.8626	1.6619	7.2008	1.7489
Population	253,665.2	6,721.92	205,386.9	4,441.169	171,320.6	4,429
% Rural Population	6.9481	1.2964	20.4081	4.5862	19.4969	6.5581

RMSE between the observed municipality observation, and their corresponding synthetic control. T_0 denotes the period of initial PMM roll-out. Synthetic controls were estimated using the variables used for the main specification, with no predictor variables, and with an alternative set of controls. All synthetic controls were estimated using the outcome density of non-PMM primary care doctors for each year prior to roll-out. Main specification predictor variables additionally include: density of nurses, beds and health facilities, percentage of population in extreme poverty, percentage of population in rural areas, percentage with private healthcare insurance, per capita income, total municipality population. The alternative set of predictor variables are: whether ruling party was in opposition to PMM, municipality area, distance to municipality capital, health spending per person, average winter rainfall, average summer rainfall, ratio of population over 65 years of age, percentage of school attendance of the population aged 25 to 29 years old, percentage of the income of the poorest 20%, percentage employed in the trade sector, percentage of the population living in households with piped water, percentage of people in households with water supply and sanitation inadequate, and total rural population.

TABLE 4.5: Root Mean Squared Error (RMSE) of Total Mortality and predictor variables using estimated binary GSC weights for all municipalities

those inference procedure rely on a binary treatment variable which we do not have in our case.

Briefly, the CAI works by imposing the null hypothesis ($\alpha = 0$) and generating a gradient for each observation. This gradient is the difference in the dependent variable (PMM doctor density) between municipality i and its synthetic control, multiplied by the residual of the model under the null hypothesis. Independent scores are generated for each cluster using principal component analysis, which are used to create the Wald statistic. The Wald statistic is simulated 999 times, using the Webb distribution weights (Webb, 2014). Finally, p-values are calculated by computing the proportion of simulated test statistics which are greater than the non-simulated test statistic.

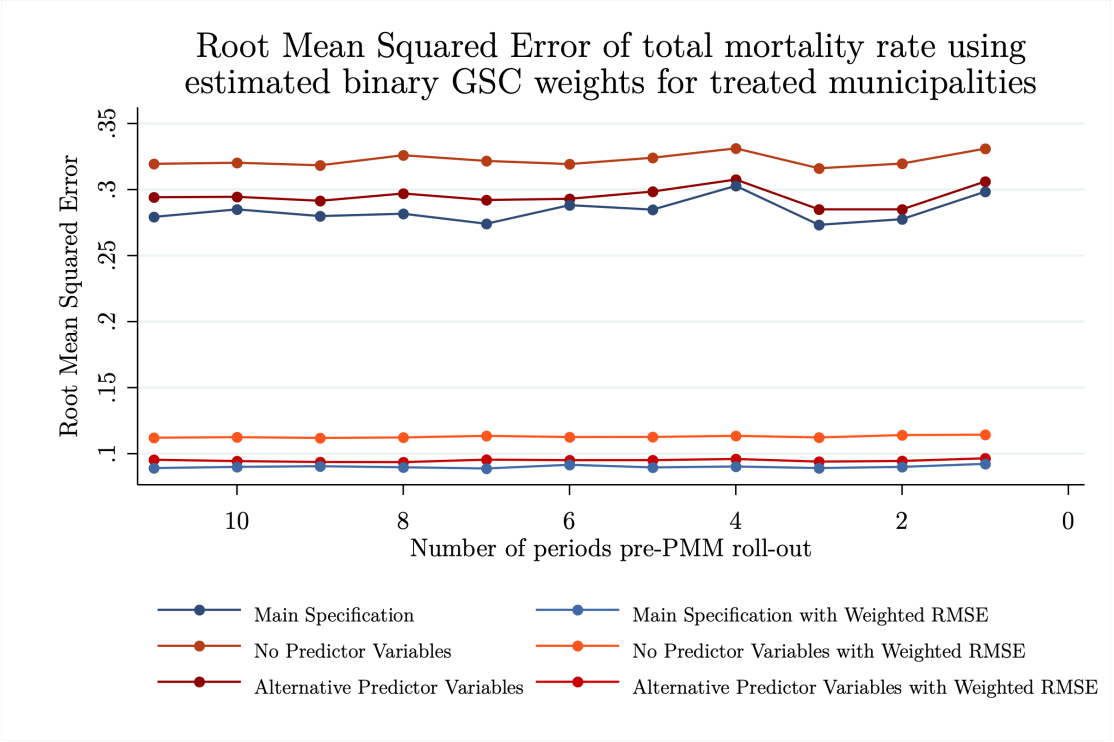


FIGURE 4.5: Root Mean Squared Error (RMSE) of Total Mortality and predictor variables using estimated binary GSC weights for all municipalities

Graphical representation of the RMSEs shown in Table 4.5. Each line shows the RMSE or weighted RMSE for either the main, no predictor variables, or alternative predictor variables specifications for each period pre-PMM roll-out.

In our setting we have a case where the number of municipalities N is large, however the number of time periods are fewer. Powell states that the CAI inference method is still valid in the circumstance, however time periods should be treated as the clusters, rather than each unit, or municipalities in this case. We report the p-values calculated using this procedure in our results table in parentheses.

4.5 Results

This section present results for both binary and continuous treatment for the impact of the PMM on hospitalisations and mortality rates. Alongside total hospitalisations and mortality rates, we also present hospitalisations and mortality rates by ambulatory care sensitive conditions (ACS) and non-ambulatory care sensitive conditions (non-ACS), both once again being per 1,000 population. As previously mentioned, *a priori* if we observe a decrease in total hospitalisations and mortality rates then we would expect these changes to be driven by ACS conditions, rather than non-ACS conditions. ACS conditions are ones in which mortality and hospitalisation risk can be

Total mortality rate - Continuous Treatment Model						
	Main Specification		No Predictor Variables		Alternative Predictor Variables	
	RMSE	Weighted RMSE	RMSE	Weighted RMSE	RMSE	Weighted RMSE
Total mortality rate						
$T_0 - 11$	0.279	0.0908	0.2959	0.0977	0.2895	0.0975
$T_0 - 10$	0.2751	0.09	0.2953	0.0986	0.2848	0.0962
$T_0 - 9$	0.2726	0.09	0.2911	0.0973	0.2875	0.0971
$T_0 - 8$	0.2726	0.0902	0.2979	0.0989	0.2866	0.0961
$T_0 - 7$	0.273	0.0894	0.2962	0.0982	0.29	0.0971
$T_0 - 6$	0.28	0.0914	0.2995	0.0984	0.2957	0.0976
$T_0 - 5$	0.2786	0.0907	0.2989	0.0981	0.2978	0.097
$T_0 - 4$	0.2891	0.0896	0.2987	0.0984	0.2819	0.0954
$T_0 - 3$	0.2781	0.0906	0.2914	0.0979	0.289	0.0968
$T_0 - 2$	0.2751	0.0897	0.2928	0.0988	0.287	0.097
$T_0 - 1$	0.2872	0.0921	0.3028	0.0989	0.2961	0.0966
Predictor Variables						
non-PMM Physicians per 1,000	0.1806	0.1238	0.2431	0.1644	0.2147	0.1478
Nurses per 1,000	0.2512	0.1489	0.256	0.1582	0.2414	0.1499
Hospital Beds per 1,000	1.9848	0.8252	2.3487	0.9088	2.0115	0.8199
Health Facilities per 1,000	3.2647	1.9462	3.34	1.9875	3.1818	1.8967
% with Private Insurance	9.2526	2.0626	9.277	1.9975	8.1085	1.6632
Income per capita	164.1567	39.5877	168.313	42.1729	150.212	33.5373
Proportion of extreme poor	7.1741	1.3108	6.9907	1.4201	7.1509	2.0219
Population	270,266.4	8,592.734	228,256.3	5,924.672	182,550.3	6,692.126
% Rural Population	18.4818	5.4354	20.6628	7.3585	20.5937	7.5368

RMSE between the observed municipality observation, and their corresponding synthetic control. T_0 denotes the period of initial PMM roll-out. Synthetic controls were estimated using the variables used for the main specification, with no predictor variables, and with an alternative set of controls. All synthetic controls were estimated using the outcome density of non-PMM primary care doctors for each year prior to roll-out. Main specification predictor variables additionally include: density of nurses, beds and health facilities, percentage of population in extreme poverty, percentage of population in rural areas, percentage with private healthcare insurance, per capita income, total municipality population. The alternative set of predictor variables are: whether ruling party was in opposition to PMM, municipality area, distance to municipality capital, health spending per person, average winter rainfall, average summer rainfall, ratio of population over 65 years of age, percentage of school attendance of the population aged 25 to 29 years old, percentage of the income of the poorest 20%, percentage employed in the trade sector, percentage of the population living in households with piped water, percentage of people in households with water supply and sanitation inadequate, and total rural population.

TABLE 4.6: Root Mean Squared Error (RMSE) of Total Mortality and predictor variables using estimated continuous GSC weights for treated municipalities

reduced through early detection, or preventative treatments, which are usually delivered through a primary care setting.

For the binary treatment estimates all municipalities in the same federal unit as municipality i are considered to be donors for the synthetic control. We also don't restrict the treatment to be an absorbing state for the binary treatment models, and municipalities are allowed to move in and out of treatment. For the continuous treatment effect, as donors for the synthetic control of municipality i we use only municipalities that are in the same federal unit and that received any PMM doctors.

Alongside the GSC results, we present two-way fixed effects difference-in-differences estimates (TWFE Diff-in-Diff). For the TWFE Diff-in-Diff we use the time in which a

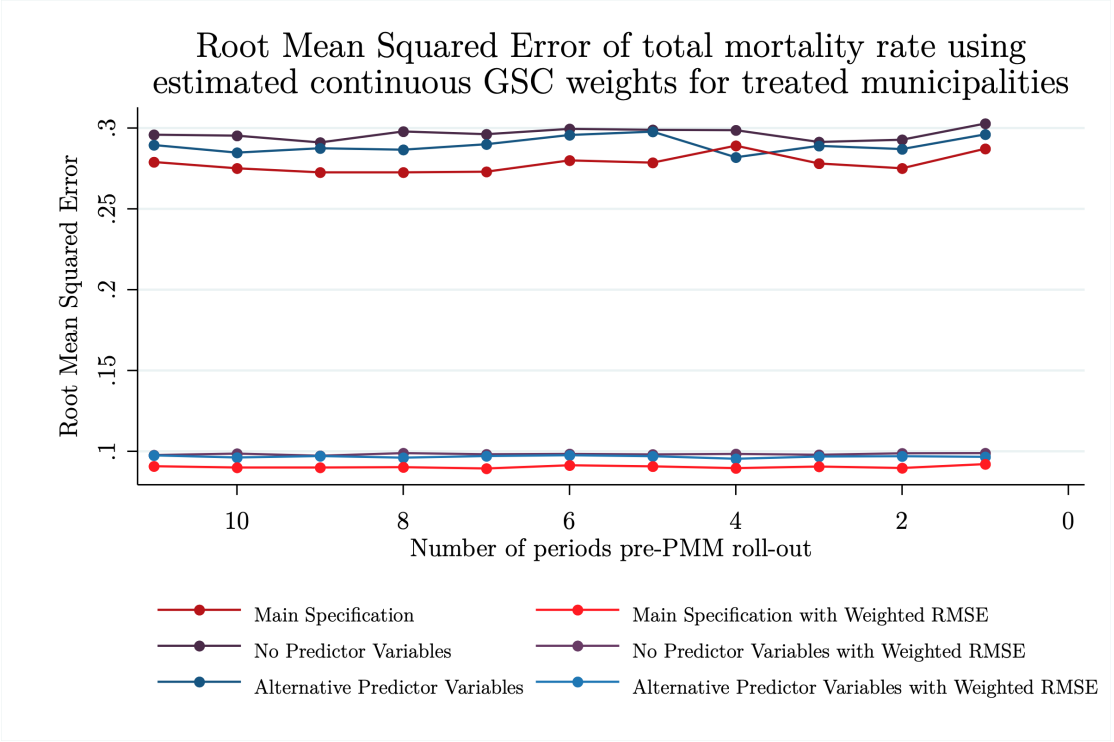


FIGURE 4.6: Root Mean Squared Error (RMSE) of Total Mortality and predictor variables using estimated continuous GSC weights for all municipalities

Graphical representation of the RMSEs shown in Table 4.6. Each line shows the RMSE or weighted RMSE for either the main, no predictor variables, or alternative predictor variables specifications for each period pre-PMM roll-out.

municipality received their first PMM doctor as the treatment period, rather than using the period of initial programme roll-out, therefore there is a staggered treatment. TWFE Diff-in-Diff estimates include the same set of predictor variables which are used in the synthetic control procedure. The time-varying control variables included are: density of non-PMM primary care doctors, density of nurses, beds and health facilities, and percentage with private healthcare insurance. Additionally, time-invariant control variables interacted with the linear-time trends are: percentage of population in extreme poverty, percentage of population in rural areas, per capita income, and total municipality population.

It is worth noting that the number of observations for the GSC estimates are approximately half the number of observations for the Diff-in-Diff estimates, this is due to only the period post PMM implementation being used for estimating the treatment effect. As discussed in section 4.4.1.1, the period before the PMM roll-out began is used to estimate the synthetic control weights for each municipality.

We report p-values for both the GSC and the TWFE Diff-in-Diff estimators in parentheses. Also, we report weighted RMSE of the pre PMM roll-out difference

between the synthetic control outcome and the observed outcome, using the weights calculated in step 3 of the procedure in section 4.4.1.1.

4.5.1 Hospitalisations

Table 4.7 presents estimates of the impact of the PMM on hospitalisations. Firstly, estimates of the binary treatment models are presented in panel (a). We find no evidence of an impact of the PMM itself on total number of hospitalisations in any of the specifications we present. We also find no evidence of an effect when we separately analyse ambulatory care sensitive conditions and non-ambulatory care sensitive conditions.

	Binary Treatment					
	TWFE Diff-in-Diff			GSC Estimates		
Panel (a)	Total	ACS	non-ACS	Total	ACS	non-ACS
PMM Treatment	0.0725	0.0433	0.0292	-0.7696	-0.1999	0.0170
P-values	*(0.788)*	*(0.759)*	*(0.864)*	*(0.286)*	*(0.298)*	*(0.232)*
Weighted RMSE				0.882	3.343	3.531
Number of Municipalities		3,542			3,522	
Number of Observations		77,924			38,742	
	Continuous Treatment					
	TWFE Diff-in-Diff			GSC Estimates		
Panel (b)	Total	ACS	non-ACS	Total	ACS	non-ACS
PMM Doctor Density	2.1834	-0.4126	2.5960	-3.8739	-0.2914	-2.9419
P-values	*(0.122)*	*(0.564)*	*(0.005)*	*(0.087)*	*(0.171)*	*(0.706)*
Weighted RMSE				0.903	3.259	3.359
Number of Municipalities		3,097			3,081	
Number of Observations		68,134			33,891	

Binary treatment estimates use a binary measure of PMM doctor (whether ever received a PMM doctor) and use all municipalities. Continuous treatment estimates use continuous measure of PMM doctors (per 1,000) and use only municipalities that received PMM doctors. GSC estimates use procedure outlined in 4.4.1.1, and p-values are calculated using the inference procedure outline in section 4.4.3. TWFE Diff-in-Diff estimates are estimated with both municipality and period fixed effects, and linear time trends of control variables. TWFE Diff-in-Diff estimates include the same set of control variables which are used in the synthetic control procedure: density of non-PMM primary care doctors, density of nurses, beds and health facilities, percentage of population in extreme poverty, percentage of population in rural areas, percentage with private healthcare insurance, per capita income, total municipality population. P-values for TWFE Diff-in-Diff specifications are calculated using cluster-robust standard errors, clustered at the municipality level.

TABLE 4.7: Binary and continuous treatment estimates of effect of PMM doctors on hospitalisations using difference-in-differences and generalised synthetic control

Panel (b) of table 4.7 presents estimates of our main results, the marginal impact of PMM doctors, where a continuous measure of physician density is used. For total hospitalisations, we find that the TWFE Diff-in-Diff estimates a statistically insignificant effect, whereas for the GSC there is suggestive evidence of a decrease of 3.9 total hospitalisations per 1,000 population for an increase in one PMM physician per 1,000 population. However, when we break down this effect by ACS and non-ACS conditions, our estimates are not statistically significant. This may be due to these models fitting less well to pre roll-out outcomes, as is shown by the weighted RMSE.

In general, we find some evidence of a decrease in hospitalisations from an increase in primary care physician supply, but we are unable to determine whether this is driven by ACS or non-ACS hospitalisations.

4.5.2 Mortality

Table 4.8 presents estimates of the impact of the PMM on mortality. As for hospitalisations, panel (a) presents the binary treatment models. Once again, we find no evidence that there are significant changes in total mortality as a result of municipalities receiving a PMM in any of our specifications. Indeed, we also find no evidence that mortality from ACS or non-ACS changes as a result of a municipality participating in the PMM.

When we analyse the impact of mortality using the continuous measure of PMM doctor density, we once again find no evidence of an impact on total or ACS mortalities. However, for our TWFE Diff-in-Diff there is evidence of a significant increase in non-ACS mortality. When using the GSC estimator, we find no evidence that there were significant changes in mortality as a result of an increase in PMM physicians.

4.6 Conclusion

This paper provides estimates of the impact of an increase in primary care physician supply on hospitalisations and mortality. We do so by exploiting an expansion of primary care physician in Brazil, which was mostly driven by importing Cuban doctors to work in areas considered to be underserved by primary care healthcare at the time. Using Powell's generalised synthetic control methodology, we estimate the policy effect of the PMM on hospitalisations and mortality, and also importantly, we estimate the marginal effect of an increase in primary care physicians.

Comparing our results to previous studies, in terms of mortality our results broadly concur with those of Carrillo and Feres (2019) who find no evidence of an impact of

	Binary Treatment					
	TWFE Diff-in-Diff			GSC Estimates		
Panel (a)	Total	ACS	non-ACS	Total	ACS	non-ACS
PMM Treatment	-0.0105	-0.0092	-0.0013	-0.0176	-0.0033	0.0299
P-values	*(0.360)*	*(0.150)*	*(0.886)*	*(0.305)*	*(0.304)*	*(0.262)*
Weighted RMSE				0.008	0.202	0.218
Number of Municipalities		3,544			3,524	
Number of Observations		77,968			38,764	
	Continuous Treatment					
	TWFE Diff-in-Diff			GSC Estimates		
Panel (b)	Total	ACS	non-ACS	Total	ACS	non-ACS
PMM Doctor Density	0.0614	-0.0498	0.1112	-0.0971	0.2051	-0.3436
P-values	*(0.326)*	*(0.151)*	*(0.040)*	*(0.985)*	*(0.420)*	*(0.518)*
Weighted RMSE				0.090	0.192	0.205
Number of Municipalities		3,099			3,083	
Number of Observations		68,178			33,913	

Binary treatment estimates use a binary measure of PMM doctor (whether ever received a PMM doctor) and use all municipalities. Continuous treatment estimates use continuous measure of PMM doctors (per 1,000) and use only municipalities that received PMM doctors. GSC estimates use procedure outlined in 4.4.1.1, and p-values are calculated using the inference procedure outline in section 4.4.3. TWFE Diff-in-Diff estimates are estimated with both municipality and period fixed effects, and linear time trends of control variables. TWFE Diff-in-Diff estimates include the same set of control variables which are used in the synthetic control procedure: density of non-PMM primary care doctors, density of nurses, beds and health facilities, percentage of population in extreme poverty, percentage of population in rural areas, percentage with private healthcare insurance, per capita income, total municipality population. P-values for TWFE Diff-in-Diff specifications are calculated using cluster-robust standard errors, clustered at the municipality level.

TABLE 4.8: Binary and continuous treatment estimates of effect of PMM doctors on mortalityusing difference-in-differences and generalised synthetic control

the PMM programme itself on health, and we find similar results in terms of mortality. Indeed, in terms of mortality our results are also similar to Mattos and Mazetto (2019) in that we find limited evidence of an impact on mortality. However, our results and Mattos and Mazetto are different to those of Hone et al. (2020), who found there to be a significant reduction in amendable mortality as a result of the PMM. In terms of hospitalisations, we find no evidence that participation in the programme itself had any impact on hospitalisations, and these results are in agreement with those of Mattos and Mazetto (2019), who find no evidence of an impact on total hospitalisations, although they do find participation in the programme decreased

hospitalisations from infectious and parasitic diseases. However, both ours and Mattos and Mazetto's results are in contrast to those of Fontes et al. (2018) who find a significant reduction in mortality from ACS conditions. In terms of the impact of PMM doctor density, we know of no other studies that have estimated this marginal effect and therefore we are unable to compare those results.

Although there is substantial discussion on the importance of primary care in improving national health and further claims that there are insufficient numbers of physicians to provide this care (World Health Organisation, 2006; Gladu, 2007; Gorman and Brooks, 2009; Petterson et al., 2012; Truglio et al., 2012; Frisch, 2013; Hoyler et al., 2014; Islam, 2014; Majeed, 2015; The Kaiser Family Foundation, 2020). Although we find there to be a decrease in hospitalisations from increasing primary care physician density, we find no evidence of primary care physicians impacting population mortality. Our results question the notion that expanding the number of primary care physicians would yield drastic improvements in health, and that doing so would be a cost-effective means of improving population health.

However, there are several explanations of our results or potential reasons why our results may not be generalisable beyond this setting. Firstly, it may be that the increase in physicians did not have an immediate impact but would materialise in the longer term, especially for mortality. In the long term, the expectation is that primary care physicians would decrease mortality through administering preventative treatment and managing long-term conditions that could eventually lead to premature death. This is especially relevant when we consider that the leading causes of mortality in our data were COPD (chronic obstructive pulmonary disease) and asthma, diabetes and hypertension, all of which would benefit from long-term management from a primary care physician. However because the the policy was short lived, due to Cuban doctors being removed from Brazil in 2018, it may be the that the programme was not given sufficient time to yield significant benefits.

Secondly, given that many of the PMM doctors were from Cuba it is possible that their effectiveness was limited. Although the PMM specifically required foreign doctors to undertake training, which included background on the Brazilian healthcare system and Portuguese language lessons (Mattos and Mazetto, 2019), it is possible that this training was not sufficient enough for the PMM doctors be a perfect substitute for Brazilian trained doctors. Indeed, although Silva et al. (2016) found that, in general, patients were satisfied with the service that the foreign doctors provided, they also note that many patients stated that communicating with their doctor was a barrier to receiving quality treatment. Although, this does not change the interpretation of our specific results, it may make the results less generalisable, in the sense that natively trained doctors may be more able to communicate with patients and engage with the healthcare system better. Therefore, we may expect that an increase in natively trained physicians would have lead to an effect of a larger magnitude.

Additionally, for an expansion of primary care doctors to indeed increase health outcomes it requires that there was a need for services that were otherwise unmet. Therefore, if in our setting the physician supply was in areas where there was no unmet need, then we would not expect changes in hospitalisations or mortality. This is unlikely to be the case in our setting however because the programme was specifically designed to increase primary care doctor numbers in areas that were underserved. Indeed, Hone et al. (2020) estimated heterogeneous impacts of the programme itself by physician density prior to the roll-out, and Carrillo and Feres (2019) do find that there was an increase in primary care doctor visits as a result of the PMM, and therefore it is unlikely that this explains our results. However, on a related point, we may not find a significant effect as a result of an increase in physician supply because health professionals were already servicing patients needs. In particular, Carrillo and Feres (2019) found that although prenatal visits by physicians increased, prenatal visits by nurses decreased as a result of the PMM. This finding suggests that PMM physicians, at least partly, substituted for the care given by nurses in receiving municipalities. Given that nurses are able to "produce as high quality care as primary care doctors and achieve as good health outcomes for patients" (Laurant et al., 2005), the increase in doctor numbers may not be cost-effective and comparable care was provided by nurses in these communities. Therefore, increasing physician supply in areas where nurses are sufficiently trained and are able to meet the healthcare needs of the population may not be an efficient use of public finance, if the policy objective is simply to improve population health.

Finally, Hone et al. (2020) found evidence that PMM doctors substituted for non-PMM doctors in municipalities that participated in the PMM, and indeed we show this in Figure 4.2. They found that that although primary care doctor density did rise in municipalities that received PMM doctors, the density of non-PMM doctors actually fell. In municipalities that received PMM doctors the total primary care doctor density rose by 5.74 per 100,000 but density of non-PMM doctors fell by 9.4 per 100,000. This suggest that there was a "pass through rate" of PMM doctors of less than one. A unit PMM doctor increase per 1,000 did not lead to a one unit increase in overall primary care physicians, but instead equated to a 0.38 unit increase in primary care physicians using Hone et al.'s estimates. However, although the pass through rate was less than one, there was clearly still an increase in the supply of primary care physicians overall, and again, we show this in Figure 4.2. Therefore we don't believe that this explains our null results. If there was indeed an impact of the increase in supply of physicians, then this substitution would have impacted the scaling, or magnitude of our results, but we would have still found statistically significant estimates.

The results in this paper question the notion that increasing primary care physician density has a significant impact on population health, and that the deficit in physician numbers would have a catastrophic impact on population health. We find no evidence

that mortality rates are impacted by increasing physician density, although we do find some evidence that hospitalisation rates are impacted beneficially. We do however acknowledge that our results are short-run estimates of increasing physician supply, and we are unable to estimate the long-term impacts on population health. From a public policy perspective, our findings suggest that increasing primary care physician numbers may not be a cost-effective means of immediately improving population health, however we are unable to make definitive claims regarding the cost-effectiveness over the long-terms.

From a methodological and econometric perspective, we show that choice of estimator plays a vital role in estimating the benefits of public policies. This paper shows that the carefully choosing the estimator, as well as considerations of the programme setting, and how the programme was administered, are necessary for correctly evaluating the benefits of programmes like the PMM. In future work it may be possible to use this paper as a starting point for comparing the different empirical methodologies used to analyse the PMM. This work may analyse how different choices in estimator, and the ways in which the nuances of the programme are handled in the analysis, impact the conclusions drawn about the effectiveness of the programme. Firstly this work would replicate previous work that uses a standard difference-in-differences estimator like those discussed in Section 4.1, before estimating a staggered roll-out version of the difference-in-differences, which is more applicable to this setting. One could then compare those results to a difference-in-differences estimator with a continuous treatment and the GSC we use in this paper, to assess whether considering the intensity of treatment makes a difference to the conclusions drawn. This would provide a useful reference for empirical researchers and would highlight the importance of careful empirical design.

Chapter 5

Conclusion

The main aim of this thesis was to explore important policy-relevant questions in health. This thesis pushes the envelop of applied health econometrics by using state-of-the-art econometric methods to assess two questions which are directly relevant to policymakers in health, those being: (i) whether a publicly financed healthcare expansion has redistributive impacts, and, (ii) whether an increase in primary physician supply has any impact on population health. This thesis also includes an essay on (iii) the impact of a diabetes diagnosis on health-related lifestyle behaviours of the diagnosed individual, and their partners, which although less directly relevant to policymakers, has significant implications for policy evaluation. I answered each of these questions using modern econometric techniques that have been developed recently and therefore were not able to be answered before now. In the absence of these econometric techniques, a causal analysis of the questions in this thesis would not have been possible. In particular, Lee, Dong, Card et al. and Powell have provided the methodological contributions necessary for the analysis of the questions in this thesis. However, more generally, this thesis relies heavily on decades of methodological contributions from many authors, who provided the groundwork for the econometric methods used in this thesis. Future methodological contributions will open the door for further analysis relating to the topics in this thesis. These contributions will also allow for further analysis on aspects of the questions in this thesis which I was not able to address.

In these concluding remarks, I will briefly discuss the findings of each essay, and discuss the relevance of their findings for future policy, as well as avenues of future research.

The first essay of this thesis analyses the redistributive effects of a publicly financed healthcare expansion. This essay, uses the Oregon Health Insurance Experiment to estimate the heterogeneous impacts of publicly financed health insurance. I estimate the impact of the publicly financed health insurance on disposable income, and then

separately for out-of-pocket medical expenditures and earned income, using a quantile regression to uncover the heterogeneous treatment effects. The results show that there is a clear heterogeneity in the impact, with those at the lower end of the income distribution benefiting substantially more than those on the right-hand side of the income distribution. The benefits of those on the left-hand side of the distribution are present for both out-of-pocket medical expenditure, as well as earned income. Although, as discussed in the essay, I have access to a limited sub-set of the entire income distribution, I still find strong evidence of a redistributive effect in this narrow income window. Given this, it is somewhat difficult to make generalisations to the entire distribution, however, it is likely that the redistributive effects will be even stronger when the entire distribution is considered.

Further analysis in chapter two shows that the health insurance expansion caused an increase in household employment, and this is likely driven by improvements in health. I also show that households that I consider to be "at high-risk" of needing intra-household care show a drastic increase in disposable income from receiving health insurance coverage. Grossman (1972)'s theory of health and human capital provided theoretical evidence of the impact of a publicly financed health insurance expansion, and our results support his theoretical prediction. Grossman theory claims that as health stock increases, the marginal benefit of health decreases, and those on the left-hand side of the distribution are expected to have the largest marginal benefit to healthcare as a result. The increase in working hours due to an increase in health stock is a key conclusion of the Grossman model, and indeed I find that an increase access to healthcare leads to the same result as the theoretical prediction for left-hand side of the distribution, whom have the lowest health stock *ex ante*. My results conclude that this increase in labour supply is a key driver in the redistributional effects of the Medicaid expansion.

As mentioned, general equilibrium effects are not analysed in this essay, and instead I present an alternative redistributive pathway specific for health. As I mention in the essay and in the introduction of this thesis, public goods are expected to cause a redistribution of wealth though taxation (Besley and Coate, 1991), and therefore, the results in this essay offer evidence of an additional source of redistribution resulting from the public financing of healthcare which has not been explore in substantial detail previously.

From a policy-making perspective these results are important for several reasons. Firstly, the findings in this essay do support the claim that Universal Healthcare Coverage, or publicly financed health insurance, is able to achieve the "goals of ending extreme poverty and increasing equity and shared prosperity" (World Bank, 2021a). The essay's main finding supports this claim, in that the results indicate that income inequality has the potential to be reduced, due to lower income household benefiting more than those on the right-hand side of the distribution, therefore redistributing

wealth in a way that reduces inequality. If policymakers seek to tackle income inequality, the findings in this essay provide evidence that publicly financed health insurance is able to do that. Further, the results in this essay also support the claim that poverty can be reduced, because the health insurance expansion increased income and reduced out-of-pocket expenditures of households at the highest risk of poverty.

There are several avenues for future research based on these results. The most obvious of these investigations would be to analyse the redistributive impacts for the entire income distribution. Given that the analysis in chapter two is conducted on only a sub-set of the entire income distribution, it is not possible to make claims with certainty regarding the impact on the entire distribution. Although the results do show a clear gradient, it is likely that the conclusions will be the same for the entire distribution, however it would be worthwhile estimating the redistributive impact for the entire distribution to be certain of this claim. Secondly, given that we find that low-income individuals increase labour supply as a result of the policy, it is possible that the publicly financed health insurance expansion would, to some extent, pay for itself. The individuals that increase their labour supply may, as a result, increase their income tax contributions and be less likely to take unemployment benefits. Future research should investigate whether the increase in tax contributions, and decrease in employment benefits paid to these individuals do indeed offset the cost of increasing health insurance coverage. If this is indeed the case, then this could profoundly impact policymakers decision to extend healthcare coverage to their populations. Finally, I estimate that households most at risk of needing intra-household care drive a substantial amount of the effect I estimate. Future research should further investigate the extent to which healthcare coverage reduces the burden of informal care on households, and the extent to which this changes labour market decisions of households.

The third chapter of this thesis analyses individual behaviours, and the importance of network in determining those behaviours. I analyse the impact a diabetes diagnosis has on the diagnosed individual, as well as their partners. I find that those diagnosed with diabetes have a reduced probability of smoking, and an increased probability of partaking in physical activity. There is no evidence that diet changes, based on the limited set of dietary information I have access to. I also find that partners' change their behaviours in a remarkably similar way, with the impact on physical activity being very similar in magnitude to the direct effect, and the impact on smoking being somewhat smaller in magnitude. These results are extremely robust to the choice of specification, and I present many alternative specifications in this chapter to ensure that the results are not specification dependent. The falsification tests also provide support of the identification strategy.

It is reassuring from a public health perspective that individuals' behaviours do change as a result of a diabetes diagnosis. In particular, increasing physical activity

and decreasing tobacco consumption suggests a compliance to diabetes treatment which has beneficial direct impacts, but also spillover benefits. However, from a public health and medical perspective, it is worrying that individuals do not appear to change their eating habits or their consumption of alcohol, and further work is required to understand how to induce changes in diet and alcohol consumption.

These main findings are important in terms of evaluating health-related public policy, as I find significant evidence that partners behaviours co-move with one another, and are causally related, rather than being driven by assortative matching. This finding is important for evaluating various public policies, as my results show that within-household spillovers should be both expected and considered in the policy evaluation process. Without considering that health-related policies can impact both the intended target and a wider network, policy makers may be substantially under or over-estimating the benefits of those policies.

I also analyse the impact by time since diagnosis, and attempt to decompose the spillovers into the health information causal channel and the joint participation causal channel. I find no evidence that there is heterogeneity by time since diagnosis, and the results suggest that the joint participation causal channel is the driver of these spillovers. These results have an additional policy-relevant contribution, in that I find very little evidence that information is driving our results. Instead, I find that the co-movement between partners' is the result of jointly participating in these behaviours. Additionally, the fact that I find no heterogeneity by time since diagnosis is reassuring from a medical perspective, and this result suggests that these behaviours are better modelled by habit formation, rather than salience. Given that salience (i.e. being closer to the health shock induces larger behavioural changes that eventually revert to pre-shock behaviours) doesn't explain these results, medical professionals can be more confident that individuals maintain positive changes in physical activity and tobacco consumption, which will have long-term health benefits for both themselves and their partners.

Although this essay presents clear evidence that some behaviours do change as a result of a health shock, it also presents many more questions for future research. Given that my results indicate that only some health-related behaviours change but not others, future research should continue to investigate which behaviours are malleable and which are inflexible. Further, I find that spillovers exist for all behaviours that I find evidence of a direct effect for, however this may not be the case for all behaviours. It would, therefore, be worthwhile to continue analysing spillovers in health to investigate whether this trend continues, and which behaviours induce spillovers and which do not. This work is necessary to correctly estimate the cost-effectiveness of various public policies.

Moving onto the secondary findings of this essay, my results differ to those of Fadlon and Nielsen in that I find no evidence of a salience effect. It is plausible that more "shocking" events, like that of heart attacks analysed by Fadlon and Nielsen are more likely to induce temporary lifestyle changes, than less shocking health changes like I investigate. However, it may not be the shock itself, but rather that some behaviours are more likely to be habit forming behaviours than others, and understanding this further would benefit future policy and medical care. In addition, given that I find no evidence to suggest that information is driving the behavioural changes in partners', future work is needed to investigate whether information, or what type of information is able to induce behavioural changes. This is policy relevant, as many public health campaigns aim specifically to increase health information in the targeted individual, and the underlying assumption is that the information itself will induce behavioural changes in the target. However, given that I find that information *per se* has no impact on behavioural changes, it somewhat undermines this assumption. The notion that public health campaigns, that rely on informing the target, have substantial impact on behavioural changes is not supported by our results, however a substantial amount more research is needed in this area to make concrete claims. Indeed, "only 1% of peer reviewed papers that looked at the impact of [public health] marketing campaigns assessed behavioural, rather than attitudinal, changes" (Mahony, 2015), and therefore much more research is needed in understanding the role of information in behavioural choices. Finally, I estimate that joint participation in behaviours is the driver of the spillover effects in my setting, and future work should analyse whether joint participation in health-related activities extends beyond the household onto the wider network. Joint participation within wider networks, and specifically neighbours or close friends may explain some of the empirical correlation within geographical regions, however one may hypotheses that the magnitude of this effect is likely to fade as connections become less strong.

The fourth chapter explores the relationship between primary care physicians supply and health outcomes. I use a unique policy setting in which there was a significant increase in primary care physicians which was predominantly driven by importing foreign doctors. In this regard the Brazilian *Programa Mais Médicos* is an ideal setting to analyse the impacts of an increase in primary care physicians because the increase in physician numbers within a municipality was mostly not the result of a redistribution of physicians within Brazil, but was instead due to Cuban doctors filling vacancies in these municipalities. Although there was some limited evidence to suggest that hospitalisations decreased significantly due to the change in physician density, I was unable to determine whether this result was due to changes in primary care preventable hospitalisations or other types of hospitalisations. Therefore, given the marginal significance of the main effect, and no evidence of an effect when analysing preventable hospitalisations and non-preventable hospitalisations separately, I conclude that the increase in physician density also had no impact on

hospitalisations. These results do question the notion that primary care physicians are able to immediately increase population health. Given that the increase in primary care physicians had no impact on mortality, if the policy objective is to improve population health, then increasing primary care physicians may not be a reliable means of doing so.

These results are vital evidence for policymakers seeking to improve public health by increasing the density of primary care physicians within regions, as such a policy does not appear to have substantial health benefits, based on my results. Some previous studies, however, have found that there to be a beneficial impact of physician density (Gulliford, 2002; Anand and Bärnighausen, 2004; Basu et al., 2019; Fadlon and Van Parys, 2020), whereas my results concur with those of Aakvik and Holmås (2006) in estimating insignificant effects. The essay does present several explanations of these results. Firstly, it is reasonably expected that the impact of primary care physicians on population health would take time to materialise. Given that the *Programa Mais Médicos* began in 2013 and had ended in 2018, there may not have been sufficient time for physicians to impact health outcomes of the population. Secondly, it is possible that given that many of the PMM doctors were from Cuba it is possible that their effectiveness was limited. Although language and background training was provided to these doctors, it is possible that the effectiveness of foreign doctors was not as high as the native doctors. This also makes these results less generalisable beyond this setting, as importing foreign doctors is a particularly unique aspect of this programme. Finally, for the doctors to have an impact, it is required that there is unmet need. For an expansion of primary care doctors to indeed increase health outcomes it requires that there was a need for services that were unmet prior to the programme roll-out. It is possible that there was no unmet need prior to roll-out, however this is not supported by Carrillo and Feres (2019), because they find that there was an increase in primary care doctor visits as a result of the *Programa Mais Médicos*. It is therefore unlikely that this explains the results. However, instead what may explain these results is that *Programa Mais Médicos* simply substituted for nurse visits, and this is a conclusion that is supported by Carrillo and Feres (2019). This point, in particular, should be explored by future research. The results in this essay should somewhat ease the concerns of policymakers regarding primary care physician deficits (World Health Organisation, 2006; Gladu, 2007; Gorman and Brooks, 2009; Hoyler et al., 2014; Petterson et al., 2012; Truglio et al., 2012; Frisch, 2013; Islam, 2014; Majeed, 2015; The Kaiser Family Foundation, 2020), while also recognising the limitations of this work and the need for further research analysing to understand the settings in which policymakers can expect increases in primary care physician supply to impact population health.

In terms of future research, there are several clear avenues of investigation. Firstly, as mentioned briefly above, an investigation of the impact of nurse density on

population health would be useful for policymakers. Given that the quality of care given by nurses and physicians are similar (Laurant et al., 2005), it is possible that increasing primary care nurse supply could be a more cost-effective means of improving population health than increasing physician supply. The impact of nurses also potentially explains the null results I estimate in this essay. If there was sufficient nurse density in these regions to meet patient need, then it would be unreasonable to expect health outcomes of the population to increase as a result of the increase in physician supply. On a related point, future research should empirically explore the *Programa Mais Médicos* further, to more generally understand the reasons for the null effects I estimate. This research would aid in understanding the settings in which policymakers can expect physicians to improve population health, and where such an expansion is unlikely to have any benefits. Finally, if increasing physician supply in primary care in is not an effective means of improving population health, further work is required in determining what tools policymakers have at their disposal to do so. Given that between 10% and 15% of mortality is avoidable by improving medical care (McGinnis et al., 2002) it may be unreasonable to expect investments in healthcare services to lead to significant changes in health. Instead, given that 50% of health is determined by behavioural factors (Hubbard, 2007) future work in improving health-related behaviours may be more fruitful in improving population health.

The aim of this thesis was to contribute knowledge for the purpose of informing policymakers in health settings. My results provide empirical evidence in favour of publicly financed healthcare for redistributive purposes, they show that spillovers between spouses exist and are economically meaningful, and also question the notion that primary care physicians are the key to improving population health. Further research on related topics is vital for ensuring that policy can contribute to the continued improvements in population health, and for dealing with the challenges facing modern healthcare systems.

Bibliography

Arild Aakvik and Tor Helge Holmås. Access to primary health care and health outcomes: The relationships between GP characteristics and mortality rates. *Journal of Health Economics*, 25(6):1139–1153, November 2006. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629606000439.

Alberto Abadie. Using Synthetic Controls: Feasibility, Data Requirements, and Methodological Aspects. *Journal of Economic Literature*, 59(2):391–425, June 2021. ISSN 0022-0515. . URL https://www.aeaweb.org/articles?id=10.1257/jel.20191450&&from=f.

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California's Tobacco Control Program. *Journal of the American Statistical Association*, 105(490):493–505, June 2010. ISSN 0162-1459. . URL https://doi.org/10.1198/jasa.2009.ap08746. Publisher: Taylor & Francis _eprint: https://doi.org/10.1198/jasa.2009.ap08746.

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Comparative Politics and the Synthetic Control Method. *American Journal of Political Science*, 59(2):495–510, 2015. ISSN 1540-5907. . URL https://onlinelibrary.wiley.com/doi/abs/10.1111/ajps.12116. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/ajps.12116.

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. SYNTH: Stata module to implement Synthetic Control Methods for Comparative Case Studies, May 2020. URL https://ideas.repec.org/c/boc/bocode/s457334.html. Language: en Publication Title: Statistical Software Components.

Heidi Allen, Katherine Baicker, Amy Finkelstein, Sarah Taubman, and Bill J. Wright. What The Oregon Health Study Can Tell Us About Expanding Medicaid. *Health Affairs*, 29(8):1498–1506, August 2010. ISSN 0278-2715. . URL https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2010.0191.

Andreas Ammermueller and Jörn-Steffen Pischke. Peer Effects in European Primary Schools: Evidence from the Progress in International Reading Literacy Study. *Journal of Labor Economics*, 27(3):315–348, July 2009. ISSN 0734-306X. . URL

https://www.journals.uchicago.edu/doi/full/10.1086/603650. Publisher: The University of Chicago Press.

Sudhir Anand and Till Bärnighausen. Human resources and health outcomes: cross-country econometric study. *The Lancet*, 364(9445):1603–1609, October 2004. ISSN 0140-6736. . URL https://www.sciencedirect.com/science/article/pii/S0140673604173133.

Joshua D. Angrist and Kevin Lang. Does School Integration Generate Peer Effects? Evidence from Boston's Metco Program. *American Economic Review*, 94(5):1613–1634, December 2004. ISSN 0002-8282. . URL https://www.aeaweb.org/articles?id=10.1257/0002828043052169.

Susan Athey and Guido W. Imbens. Design-based Analysis in Difference-In-Differences Settings with Staggered Adoption. Technical Report w24963, National Bureau of Economic Research, September 2018. URL https://www.nber.org/papers/w24963.

David I. Auerbach and Arthur L. Kellermann. A decade of health care cost growth has wiped out real income gains for an average US family. *Health Affairs (Project Hope)*, 30(9):1630–1636, September 2011. ISSN 1544-5208. .

Katherine Baicker and Amy Finkelstein. The Impact of Medicaid Expansion on Voter Participation: Evidence from the Oregon Health Insurance Experiment. *Quarterly Journal of Political Science*, 14(4):383–400, October 2019. ISSN 1554-0626, 1554-0634. . URL https://www.nowpublishers.com/article/Details/QJPS-19026. Publisher: Now Publishers, Inc.

Katherine Baicker, Sarah L. Taubman, Heidi L. Allen, Mira Bernstein, Jonathan H. Gruber, Joseph P. Newhouse, Eric C. Schneider, Bill J. Wright, Alan M. Zaslavsky, and Amy N. Finkelstein. The Oregon Experiment — Effects of Medicaid on Clinical Outcomes. *New England Journal of Medicine*, 368(18):1713–1722, May 2013. ISSN 0028-4793. . URL https://doi.org/10.1056/NEJMsa1212321.

Katherine Baicker, Amy Finkelstein, Jae Song, and Sarah Taubman. The Impact of Medicaid on Labor Market Activity and Program Participation: Evidence from the Oregon Health Insurance Experiment. *American Economic Review*, 104(5):322–328, May 2014. ISSN 0002-8282. . URL https://www.aeaweb.org/articles?id=10.1257/aer.104.5.322.

Katherine Baicker, Heidi L. Allen, Bill J. Wright, and Amy N. Finkelstein. The Effect Of Medicaid On Medication Use Among Poor Adults: Evidence From Oregon. *Health Affairs*, 36(12):2110–2114, December 2017. ISSN 0278-2715. . URL https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2017.0925. Publisher: Health Affairs.

Katherine Baicker, Heidi L. Allen, Bill J. Wright, Sarah L. Taubman, and Amy N. Finkelstein. The Effect of Medicaid on Dental Care of Poor Adults: Evidence from the Oregon Health Insurance Experiment. *Health Services Research*, 53(4):2147–2164, 2018a. ISSN 1475-6773. . URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-6773.12757. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-6773.12757.

Katherine Baicker, Heidi L. Allen, Bill J. Wright, Sarah L. Taubman, and Amy N. Finkelstein. The Effect of Medicaid on Management of Depression: Evidence From the Oregon Health Insurance Experiment. *The Milbank Quarterly*, 96(1):29–56, March 2018b. . URL https://www.milbank.org/quarterly/articles/effect-medicaid-management-depression-evidence-oregon-health-insurance-experiment/.

E. Barrett-Connor, L. Suarez, and M. H. Criqui. Spouse concordance of plasma cholesterol and triglyceride. *Journal of Chronic Diseases*, 35(5):333–340, 1982. ISSN 0021-9681. .

Ann Bartel and Paul Taubman. Health and Labor Market Success: The Role of Various Diseases. *The Review of Economics and Statistics*, 61(1):1–8, 1979. ISSN 0034-6535. . URL https://www.jstor.org/stable/1924825.

Anirban Basu and David Meltzer. Implications of spillover effects within the family for medical cost-effectiveness analysis. *Journal of Health Economics*, 24(4):751–773, July 2005. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629605000317.

Sanjay Basu, Seth A. Berkowitz, Robert L. Phillips, Asaf Bitton, Bruce E. Landon, and Russell S. Phillips. Association of Primary Care Physician Supply With Population Mortality in the United States, 2005-2015. *JAMA Internal Medicine*, 179(4):506–514, April 2019. ISSN 2168-6106. . URL https://doi.org/10.1001/jamainternmed.2018.7624.

Gary S. Becker. A Theory of Marriage: Part I. *Journal of Political Economy*, 81(4): 813–846, 1973. ISSN 0022-3808. URL https://www.jstor.org/stable/1831130.

Gary S. Becker. A Treatise on the Family. January 1981. URL https://www.nber.org/books/beck81-1.

Sascha O. Becker, Peter H. Egger, and Maximilian von Ehrlich. Absorptive Capacity and the Growth and Investment Effects of Regional Transfers: A Regression Discontinuity Design with Heterogeneous Treatment Effects. *American Economic Journal: Economic Policy*, 5(4):29–77, November 2013. ISSN 1945-7731. . URL https://www.aeaweb.org/articles?id=10.1257/pol.5.4.29.

Timothy Besley and Stephen Coate. Public Provision of Private Goods and the Redistribution of Income. *The American Economic Review*, 81(4):979–984, 1991. ISSN 0002-8282. URL http://www.jstor.org/stable/2006658.

Ayona Bhattacharjee, Jong Kook Shin, Chetan Subramanian, and Shailender Swaminathan. Healthcare investment and income inequality. *Journal of Health Economics*, 56:163–177, December 2017. ISSN 0167-6296. . URL http://www.sciencedirect.com/science/article/pii/S0167629617303600.

Asaf Bitton, Jocelyn Fifield, Hannah Ratcliffe, Ami Karlage, Hong Wang, Jeremy H. Veillard, Dan Schwarz, and Lisa R. Hirschhorn. Primary healthcare system performance in low-income and middle-income countries: a scoping review of the evidence from 2010 to 2017. *BMJ Global Health*, 4(Suppl 8):e001551, August 2019. ISSN 2059-7908. . URL https://gh.bmj.com/content/4/suppl_8/e001551. Publisher: BMJ Specialist Journals Section: Research.

Katia Vergetti Bloch, Carlos Henrique Klein, Nelson Albuquerque de Souza e Silva, Armando da Rocha Nogueira, and Lucia Helena Alvares Salis. Socioeconomic aspects of spousal concordance for hypertension, obesity, and smoking in a community of Rio de Janeiro, Brazil. *Arquivos Brasileiros De Cardiologia*, 80(2): 179–186, 171–178, February 2003. ISSN 0066-782X. .

Linda J. Blumberg, Timothy A. Waidmann, Fredric Blavin, and Jeremy Roth. Trends in Health Care Financial Burdens, 2001 to 2009. *The Milbank Quarterly*, 92(1):88–113, 2014. ISSN 1468-0009. . URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0009.12042. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1468-0009.12042.

K. Bolin, B. Lindgren, and P. Lundborg. Informal and formal care among single-living elderly in Europe. *Health Economics*, 17(3):393–409, 2008. ISSN 1099-1050. . URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.1275. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.1275.

Eric Bonsang. Does informal care from children to their elderly parents substitute for formal care in Europe? *Journal of Health Economics*, 28(1):143–154, January 2009. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629608001252.

Michael Boozer and Stephen E. Cacciola. Inside the 'Black Box' of Project Star: Estimation of Peer Effects Using Experimental Data. SSRN Scholarly Paper ID 277009, Social Science Research Network, Rochester, NY, June 2001. URL https://papers.ssrn.com/abstract=277009.

Bijan J. Borah and Anirban Basu. Highlighting Differences Between Conditional and Unconditional Quantile Regression Approaches Through an Application to Assess Medication Adherence. *Health Economics*, 22(9):1052–1070, 2013. ISSN 1099-1050. . URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.2927. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.2927.

Kirill Borusyak and Xavier Jaravel. Revisiting Event Study Designs. SSRN Scholarly Paper ID 2826228, Social Science Research Network, Rochester, NY, May 2017. URL https://papers.ssrn.com/abstract=2826228.

Caron F Bove, Jeffery Sobal, and Barbara S Rauschenbach. Food choices among newly married couples: convergence, conflict, individualism, and projects. *Appetite*, 40(1): 25–41, February 2003. ISSN 0195-6663. . URL http://www.sciencedirect.com/science/article/pii/S0195666302001472.

Nils Braakmann. The consequences of own and spousal disability on labor market outcomes and subjective well-being: evidence from Germany. *Review of Economics of the Household*, 12(4):717–736, December 2014. ISSN 1573-7152. . URL https://doi.org/10.1007/s11150-012-9164-7.

Mary A. Burke and Tim R. Sass. Classroom Peer Effects and Student Achievement. *Journal of Labor Economics*, 31(1):51–82, January 2013. ISSN 0734-306X. . URL https://www.journals.uchicago.edu/doi/full/10.1086/666653. Publisher: The University of Chicago Press.

Michael Caley. Remember Barbara Starfield: primary care is the health system's bedrock. *BMJ (Clinical research ed.)*, 347:f4627, July 2013. ISSN 1756-1833. .

Brantly Callaway and Pedro H. C. Sant'Anna. Difference-in-Differences with Multiple Time Periods. *arXiv:1803.09015 [econ, math, stat]*, December 2020. URL http://arxiv.org/abs/1803.09015. arXiv: 1803.09015.

David Card, Carlos Dobkin, and Nicole Maestas. The Impact of Nearly Universal Insurance Coverage on Health Care Utilization: Evidence from Medicare. *The American economic review*, 98(5):2242–2258, December 2008. ISSN 0002-8282. . URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2600774/.

David Card, David S. Lee, Zhuan Pei, and Andrea Weber. Inference on Causal Effects in a Generalized Regression Kink Design. *Econometrica*, 83(6):2453–2483, 2015. ISSN 1468-0262. . URL https://onlinelibrary.wiley.com/doi/abs/10.3982/ECTA11224.

Katherine Grace Carman and Lei Zhang. Classroom peer effects and academic achievement: Evidence from a Chinese middle school. *China Economic Review*, 23(2): 223–237, June 2012. ISSN 1043-951X. . URL https://www.sciencedirect.com/science/article/pii/S1043951X11001167.

F. Carmichael, S. Charles, and C. Hulme. Who will care? Employment participation and willingness to supply informal care. *Journal of Health Economics*, 29(1):182–190, January 2010. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629609001209.

Fiona Carmichael and Susan Charles. The opportunity costs of informal care: does gender matter? *Journal of Health Economics*, 22(5):781–803, September 2003. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629603000444.

Fiona Carmichael and Marco G. Ercolani. Unpaid caregiving and paid work over life-courses: Different pathways, diverging outcomes. *Social Science & Medicine*, 156: 1–11, May 2016. ISSN 0277-9536. . URL https://www.sciencedirect.com/science/article/pii/S0277953616301204.

Bladimir Carrillo and Jose Feres. Provider Supply, Utilization, and Infant Health: Evidence from a Physician Distribution Policy. *American Economic Journal: Economic Policy*, 11(3):156–196, August 2019. ISSN 1945-7731. . URL https://www.aeaweb.org/articles?id=10.1257/pol.20170619.

Matias D. Cattaneo, Rocio Titiunik, and Gonzalo Vazquez-Bare. The Regression Discontinuity Design. In *Handbook of Research Methods in Political Science and International Relations*, pages 835–857. Sage Publications, June 2020. URL http://dx.doi.org/10.4135/9781526486387.n47.

Pierre-André Chiappori, Sonia Oreffice, and Climent Quintana-Domeque. Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market. *Journal of Political Economy*, 120(4):659–695, August 2012. ISSN 0022-3808. . URL https://www.journals.uchicago.edu/doi/full/10.1086/667941. Publisher: The University of Chicago Press.

Thomas N. Chirikos and Gilbert Nestel. Further Evidence on the Economic Effects of Poor Health. *The Review of Economics and Statistics*, 67(1):61–69, 1985. ISSN 0034-6535. . URL https://www.jstor.org/stable/1928435.

Nicholas A. Christakis and James H. Fowler. The Collective Dynamics of Smoking in a Large Social Network. *New England Journal of Medicine*, 358(21):2249–2258, May 2008. ISSN 0028-4793. . URL https://doi.org/10.1056/NEJMsa0706154.

Andrea S. Christopher, David U. Himmelstein, Steffie Woolhandler, and Danny McCormick. The Effects of Household Medical Expenditures on Income Inequality in the United States. *American Journal of Public Health*, 108(3):351–354, January 2018. ISSN 0090-0036. . URL https://ajph.aphapublications.org/doi/abs/10.2105/AJPH.2017.304213. Publisher: American Public Health Association.

Andrew E. Clark and Fabrice Etilé. Don't give up on me baby: Spousal correlation in smoking behaviour. *Journal of Health Economics*, 25(5):958–978, September 2006. ISSN 0167-6296. . URL http://www.sciencedirect.com/science/article/pii/S0167629606000117.

Robin Clark. Substitution between Formal and Informal Care for Persons with Severe Mental Illness and Substance Use Disorders. *Center for Health Policy and Research (CHPR) Publications*, 4(3), April 2002. URL https://escholarship.umassmed.edu/healthpolicy_pp/27.

Laura K. Cobb, Mara A. McAdams-DeMarco, Rachel R. Huxley, Mark Woodward, Silvia Koton, Josef Coresh, and Cheryl A. M. Anderson. The Association of Spousal Smoking Status With the Ability to Quit Smoking: The Atherosclerosis Risk in Communities Study. *American Journal of Epidemiology*, 179(10):1182–1187, May 2014. ISSN 0002-9262. . URL https://academic.oup.com/aje/article/179/10/1182/2739155.

Norma B. Coe, Jing Guo, R. Tamara Konetzka, and Courtney Harold Van Houtven. What is the marginal benefit of payment-induced family care? Impact on Medicaid spending and health of care recipients. *Health Economics*, 28(5):678–692, 2019. ISSN 1099-1050. . URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.3873. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.3873.

Courtney C. Coile. Health Shocks and Couples' Labor Supply Decisions. Working Paper 10810, National Bureau of Economic Research, October 2004. URL https://www.nber.org/papers/w10810. Series: Working Paper Series.

Anna Condelius, Anna-Karin Edberg, Ingalill Rahm Hallberg, and Ulf Jakobsson. Utilization of medical healthcare among people receiving long-term care at home or in special accommodation. *Scandinavian Journal of Caring Sciences*, 24(2):404–413, 2010. ISSN 1471-6712. . URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1471-6712.2009.00725.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1471-6712.2009.00725.x.

Paul Contoyannis and Nigel Rice. The impact of health on wages: Evidence from the British Household Panel Survey. *Empirical Economics*, 26(4):599–622, November 2001. ISSN 1435-8921. . URL https://doi.org/10.1007/s001810000073.

David M. Cutler and Edward L. Glaeser. Social Interactions and Smoking. *Research Findings in the Economics of Aging*, pages 123–141, February 2010. URL https://www.nber.org/chapters/c8196. Publisher: University of Chicago Press.

Clément de Chaisemartin and Xavier D'Haultfoeuille. Difference-in-Differences Estimators of Intertemporal Treatment Effects. *arXiv:2007.04267 [econ]*, May 2021. URL http://arxiv.org/abs/2007.04267. arXiv: 2007.04267.

Clément de Chaisemartin and Xavier D'Haultfœuille. Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects. *American Economic Review*, 110 (9):2964–2996, September 2020. ISSN 0002-8282. . URL https://www.aeaweb.org/articles?id=10.1257/aer.20181169.

Angus Deaton. Health, Inequality, and Economic Development. *Journal of Economic Literature*, 41(1):113–158, March 2003. ISSN 0022-0515. . URL https://www.aeaweb.org/articles?id=10.1257/002205103321544710.

Department of Health and Human Services. 2009 Poverty Guidelines, Federal Register Notice, January 2009. URL https://aspe.hhs.gov/2009-poverty-guidelines-federal-register-notice.

Department of Health and Human Services. 2008 Poverty Guidelines, Federal Register Notice, December 2015. URL https://aspe.hhs.gov/2008-poverty-guidelines-federal-register-notice.

Augusto Di Castelnuovo, Gianni Quacquaruccio, Maria Benedetta Donati, Giovanni de Gaetano, and Licia Iacoviello. Spousal Concordance for Major Coronary Risk Factors: A Systematic Review and Meta-Analysis. *American Journal of Epidemiology*, 169(1):1–8, January 2009. ISSN 0002-9262. . URL https://academic.oup.com/aje/article/169/1/1/210447. Publisher: Oxford Academic.

Diabetes UK. Myth: I can't eat fruit if I have diabetes. URL https://www.diabetes.org.uk/guide-to-diabetes/enjoy-food/eating-with-diabetes/diabetes-food-myths/myth-fruit-diabetes.

Diabetes UK. Facts and Stats 2019, 2019. URL https://www.diabetes.org.uk/professionals/position-statements-reports/statistics.

Christian Dippel, Robert Gold, Stephan Heblich, and Rodrigo Pinto. Mediation Analysis in IV Settings With a Single Instrument. *Working Paper*, page 41, July 2020.

A Dixon, J Le Grand, J Henderson, R Murray, and E Poteliakhoff. Is the British National Health Service equitable? The evidence on socioeconomic differences in utilization, 2007. *Journal of Health Services Research & Policy*, 12(2):104–109, April 2007. .

Yingying Dong. Jumpy or Kinky? Regression Discontinuity without the Discontinuity. *University of California-Irvine, Department of Economics*, page 55, 2011. URL https://ideas.repec.org/p/irv/wpaper/111207.html.

Eddy van Doorslaer and Xander Koolman. Explaining the differences in income-related health inequalities across European countries. *Health Economics*, 13 (7):609–628, 2004. ISSN 1099-1050. . URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.918.

Eddy van Doorslaer, Cristina Masseria, and Xander Koolman. Inequalities in access to medical care by income in developed countries. *CMAJ*, 174(2):177–183, January 2006. ISSN 0820-3946, 1488-2329. . URL http://www.cmaj.ca/content/174/2/177.

Lena Dorin, Suzi C. Turner, Lea Beckmann, Jörg große Schlarmann, Andreas Faatz, Sabine Metzing, and Andreas Büscher. Which need characteristics influence healthcare service utilization in home care arrangements in Germany? *BMC Health Services Research*, 14(1):233, May 2014. ISSN 1472-6963. . URL https://doi.org/10.1186/1472-6963-14-233.

Arnaud Dupuy and Alfred Galichon. Personality Traits and the Marriage Market. *Journal of Political Economy*, 122(6):1271–1319, December 2014. ISSN 0022-3808. . URL https://www.journals.uchicago.edu/doi/abs/10.1086/677191.

Majid Ezzati and Elio Riboli. Can noncommunicable diseases be prevented? Lessons from studies of populations and individuals. *Science (New York, N.Y.)*, 337(6101): 1482–1487, September 2012. ISSN 1095-9203. .

Majid Ezzati and Elio Riboli. Behavioral and Dietary Risk Factors for Noncommunicable Diseases. *New England Journal of Medicine*, 369(10):954–964, September 2013. ISSN 0028-4793. . URL https://doi.org/10.1056/NEJMra1203528. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMra1203528.

Itzik Fadlon and Torben Heien Nielsen. Family Health Behaviors. *American Economic Review*, 109(9):3162–3191, September 2019. ISSN 0002-8282. . URL https://www.aeaweb.org/articles?id=10.1257/aer.20171993.

Itzik Fadlon and Torben Heien Nielsen. Family Labor Supply Responses to Severe Health Shocks: Evidence from Danish Administrative Records. *American Economic Journal: Applied Economics*, 13(3):1–30, July 2021. ISSN 1945-7782. . URL https://www.aeaweb.org/articles?id=10.1257/app.20170604&&from=f.

Itzik Fadlon and Jessica Van Parys. Primary care physician practice styles and patient care: Evidence from physician exits in Medicare. *Journal of Health Economics*, 71: 102304, May 2020. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629619308367.

Tracy A. Falba and Jody L. Sindelar. Spousal Concordance in Health Behavior Change. *Health Services Research*, 43(1p1):96–116, 2008. ISSN 1475-6773. . URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-6773.2007.00754.x.

Hai Fang, Mir M. Ali, and John A. Rizzo. Does smoking affect body weight and obesity in China? *Economics & Human Biology*, 7(3):334–350, December 2009. ISSN 1570-677X. . URL http://www.sciencedirect.com/science/article/pii/S1570677X09000537.

Lisa Farrell and Michael A. Shields. Investigating the economic and demographic determinants of sporting participation in England. *Journal of the Royal Statistical*

Society: Series A (Statistics in Society), 165(2):335–348, 2002. ISSN 1467-985X. . URL https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/1467-985X.00626.

Amy Finkelstein and Robin McKnight. What did Medicare do? The initial impact of Medicare on mortality and out of pocket medical spending. *Journal of Public Economics*, 92(7):1644–1668, July 2008. ISSN 0047-2727. . URL http://www.sciencedirect.com/science/article/pii/S0047272707001764.

Amy Finkelstein, Sarah Taubman, Bill Wright, Mira Bernstein, Jonathan Gruber, Joseph P. Newhouse, Heidi Allen, Katherine Baicker, and Oregon Health Study Group. The Oregon Health Insurance Experiment: Evidence from the First Year*. *The Quarterly Journal of Economics*, 127(3):1057–1106, August 2012. ISSN 0033-5533. . URL https://doi.org/10.1093/qje/qjs020.

Amy N. Finkelstein, Sarah L. Taubman, Heidi L. Allen, Bill J. Wright, and Katherine Baicker. Effect of Medicaid Coverage on ED Use — Further Evidence from Oregon's Experiment. *New England Journal of Medicine*, 375(16):1505–1507, October 2016. ISSN 0028-4793. . URL https://doi.org/10.1056/NEJMp1609533. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJMp1609533.

Jason Fletcher and Ryne Marksteiner. Causal Spousal Health Spillover Effects and Implications for Program Evaluation. *American Economic Journal: Economic Policy*, 9 (4):144–166, November 2017. ISSN 1945-7731. . URL https://www.aeaweb.org/articles?id=10.1257/pol.20150573.

Luiz Felipe Campos Fontes, Otavio Canozzi Conceição, and Paulo de Andrade Jacinto. Evaluating the impact of physicians' provision on primary healthcare: Evidence from Brazil's More Doctors Program. *Health Economics*, 27(8):1284–1299, August 2018. ISSN 10579230. . URL http://doi.wiley.com/10.1002/hec.3775.

Nita G Forouhi, Anoop Misra, Viswanathan Mohan, Roy Taylor, and William Yancy. Dietary and nutritional approaches for prevention and management of type 2 diabetes. *BMJ*, page k2234, June 2018. ISSN 0959-8138, 1756-1833. . URL https://www.bmj.com/lookup/doi/10.1136/bmj.k2234.

Brigham R. Frandsen. Party Bias in Union Representation Elections: Testing for Manipulation in the Regression Discontinuity Design when the Running Variable is Discrete. In *Advances in Econometrics*, volume 38, pages 281–315. Emerald Publishing Limited, May 2017. ISBN 978-1-78714-390-6 978-1-78714-389-0. . URL https://www.emerald.com/insight/content/doi/10.1108/S0731-905320170000038012/full/html.

Melissa M. Franks, Amy Mehraban Pienta, and Linda A. Wray. It takes two: marriage and smoking cessation in the middle years. *Journal of Aging and Health*, 14(3): 336–354, August 2002. ISSN 0898-2643. .

Esther M Friedman, Juleen Rodakowski, Richard Schulz, Scott R Beach, Grant R Martsolf, and A Everette James, III. Do Family Caregivers Offset Healthcare Costs for Older Adults? A Mapping Review on the Costs of Care for Older Adults With Versus Without Caregivers. *The Gerontologist*, 59(5):e535–e551, September 2019. ISSN 0016-9013. . URL https://doi.org/10.1093/geront/gny182.

Suzy Frisch. The primary care physician shortage. *BMJ*, 347:f6559, November 2013. ISSN 1756-1833. . URL https://www.bmj.com/content/347/bmj.f6559. Publisher: British Medical Journal Publishing Group Section: Feature.

Clemens Fuest. *Schools and the Equal Opportunity Problem*. MIT Press, 2007. ISBN 978-0-262-23257-9. Google-Books-ID: uI7wB7M1T5cC.

Alison A. Galbraith, Sabrina T. Wong, Sue E. Kim, and Paul W. Newacheck. Out-of-Pocket Financial Burden for Low-Income Families with Children: Socioeconomic Disparities and Effects of Insurance. *Health Services Research*, 40(6p1): 1722–1736, 2005. ISSN 1475-6773. . URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1475-6773.2005.00421.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1475-6773.2005.00421.x.

Alan M. Garber and Charles E. Phelps. Economic foundations of cost-effectiveness analysis. *Journal of Health Economics*, 16(1):1–31, February 1997. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629696005061.

Pilar García-Gómez, Hans van Kippersluis, Owen O'Donnell, and Eddy van Doorslaer. Long Term and Spillover Effects of Health Shocks on Employment and Income. *The Journal of human resources*, 48(4):873–909, 2013. ISSN 0022-166X. . URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4110210/.

Christina Gathmann, Kristiina Huttunen, Laura Jernström, Lauri Sääksvuori, and Robin Stitzing. In Sickness and in Health: Job Displacement and Health Spillovers in Couples. *IZA Working Paper Series*, (IZA DP No. 13329):65, June 2020.

Andrew Gelman and Guido Imbens. Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics*, 37(3):447–456, July 2019. ISSN 0735-0015. . URL https://amstat.tandfonline.com/doi/full/10.1080/07350015.2017.1366909.

François-Pierre Gladu. Perceived shortage of family doctors in Quebec: Can we do something about it? *Canadian Family Physician*, 53(11):1858–1860, November 2007. ISSN 0008-350X, 1715-5258. URL https://www.cfp.ca/content/53/11/1858. Publisher: The College of Family Physicians of Canada Section: Commentary.

Gerhard Glomm and B. Ravikumar. Public versus Private Investment in Human Capital: Endogenous Growth and Income Inequality. *Journal of Political Economy*,

100(4):818–834, 1992. ISSN 0022-3808. URL http://www.jstor.org/stable/2138689.

Ezra Golberstein, David C. Grabowski, Kenneth M. Langa, and Michael E. Chernew. The Effect of Medicare Home Health Care Payment on Informal Care. *Inquiry : a journal of medical care organization, provision and financing*, 46(1):58–71, 2009. ISSN 0046-9580. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3092523/.

Andrew Goodman-Bacon. Difference-in-differences with variation in treatment timing. *Journal of Econometrics*, June 2021. ISSN 0304-4076. . URL https://www.sciencedirect.com/science/article/pii/S0304407621001445.

Des F. Gorman and Peter M. Brooks. On solutions to the shortage of doctors in Australia and New Zealand. *The Medical Journal of Australia*, 190(3):152–156, February 2009. ISSN 0025-729X. .

K. Graham and K. Braun. Concordance of use of alcohol and other substances among older adult couples. *Addictive Behaviors*, 24(6):839–856, December 1999. ISSN 0306-4603. .

Michael Grossman. On the Concept of Health Capital and the Demand for Health. *Journal of Political Economy*, 80(2):223–255, March 1972. ISSN 0022-3808. . URL https://www.journals.uchicago.edu/doi/abs/10.1086/259880.

Frederico Guanais and James Macinko. Primary care and avoidable hospitalizations: evidence from Brazil. *The Journal of Ambulatory Care Management*, 32(2):115–122, June 2009. ISSN 0148-9917. .

Martin C. Gulliford. Availability of primary care doctors and population health in England: is there an association? *Journal of Public Health Medicine*, 24(4):252–254, December 2002. ISSN 0957-4832. .

Jens Gundgaard. Income-related inequality in utilization of health services in Denmark: Evidence from Funen County. *Scandinavian Journal of Public Health*, 34(5): 462–471, October 2006. ISSN 1403-4948. . URL https://doi.org/10.1080/14034940600554644.

Eric A. Hanushek, John F. Kain, Jacob M. Markman, and Steven G. Rivkin. Does peer ability affect student achievement? *Journal of Applied Econometrics*, 18(5):527–544, 2003. ISSN 1099-1255. . URL https://onlinelibrary.wiley.com/doi/abs/10.1002/jae.741. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/jae.741.

Matthew Harris and Matthew Harris. Mais Médicos (More Doctors) Program – a view from England. *Ciência & Saúde Coletiva*, 21(9):2919–2923, September 2016. ISSN 1413-8123. . URL http://www.scielo.br/scielo.php?script=sci_abstract&

pid=S1413-81232016000902919&lng=en&nrm=iso&tlng=en. Publisher: ABRASCO - Associação Brasileira de Saúde Coletiva.

Lynda Hayward, Sharon Davies, Roberta Robb, Margaret Denton, and Greg Auton. Publicly funded and family-friend care in the case of long-term illness: the role of the spouse. *Canadian Journal on Aging = La Revue Canadienne Du Vieillissement*, 23 Suppl 1:S39–48, 2004. ISSN 0714-9808. .

Susan P. Helmrich, David R. Ragland, Rita W. Leung, and Ralph S. Paffenbarger. Physical Activity and Reduced Occurrence of Non-Insulin-Dependent Diabetes Mellitus. *New England Journal of Medicine*, 325(3):147–152, July 1991. ISSN 0028-4793. . URL https://doi.org/10.1056/NEJM199107183250302. Publisher: Massachusetts Medical Society _eprint: https://doi.org/10.1056/NEJM199107183250302.

Karen Hodgson, Sarah R. Deeny, and Adam Steventon. Ambulatory care-sensitive conditions: their potential uses and limitations. *BMJ Quality & Safety*, 28(6):429–433, June 2019. ISSN 2044-5415, 2044-5423. . URL https://qualitysafety.bmj.com/content/28/6/429. Publisher: BMJ Publishing Group Ltd Section: Editorial.

John Holahan and Sheila Zedlewski. Who Pays for Health Care in the United States? Implications for Health System Reform. *Inquiry*, 29(2):231–248, 1992. ISSN 0046-9580. URL https://www.jstor.org/stable/29772294. Publisher: Sage Publications, Inc.

Bruce Hollingsworth, Asako Ohinata, Matteo Picchio, and Ian Walker. The Impacts of Free Universal Elderly Care on the Supply of Informal Care and Labour Supply*. *Oxford Bulletin of Economics and Statistics*, n/a(n/a), December 2021. ISSN 1468-0084. . URL https://onlinelibrary.wiley.com/doi/abs/10.1111/obes.12473. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/obes.12473.

Thomas Hone, Davide Rasella, Mauricio L. Barreto, Azeem Majeed, and Christopher Millett. Association between expansion of primary healthcare and racial inequalities in mortality amenable to primary care in Brazil: A national longitudinal analysis. *PLOS Medicine*, 14(5):e1002306, May 2017. ISSN 1549-1676. . URL https://journals.plos.org/plosmedicine/article?id=10.1371/journal.pmed.1002306. Publisher: Public Library of Science.

Thomas Hone, Timothy Powell-Jackson, Leonor Maria Pacheco Santos, Ricardo de Sousa Soares, Felipe Proenço de Oliveira, Mauro Niskier Sanchez, Matthew Harris, Felipe de Oliveira de Souza Santos, and Christopher Millett. Impact of the Programa Mais médicos (more doctors Programme) on primary care doctor supply and amenable mortality: quasi-experimental study of 5565 Brazilian municipalities.

BMC Health Services Research, 20(1):873, September 2020. ISSN 1472-6963. . URL https://doi.org/10.1186/s12913-020-05716-2.

Courtney Van Houtven, Fiona Carmichael, Josephine Jacobs, and Peter C. Coyte. The Economics of Informal Care, May 2019. URL https://oxfordre.com/economics/view/10.1093/acrefore/9780190625979.001.0001/acrefore-9780190625979-e-265. ISBN: 9780190625979.

Caroline Hoxby. Peer Effects in the Classroom: Learning from Gender and Race Variation. Working Paper 7867, National Bureau of Economic Research, August 2000. URL https://www.nber.org/papers/w7867. Series: Working Paper Series.

Caroline M. Hoxby and Gretchen Weingarth. *Taking race out of the equation: School reassignment and the structure of peer effects*. 2005.

Marguerite Hoyler, Samuel R. G. Finlayson, Craig D. McClain, John G. Meara, and Lars Hagander. Shortage of doctors, shortage of data: a review of the global surgery, obstetrics, and anesthesia workforce literature. *World Journal of Surgery*, 38(2): 269–280, February 2014. ISSN 1432-2323. .

Frank B. Hu, JoAnn E. Manson, Meir J. Stampfer, Graham Colditz, Simin Liu, Caren G. Solomon, and Walter C. Willett. Diet, Lifestyle, and the Risk of Type 2 Diabetes Mellitus in Women. *New England Journal of Medicine*, 345(11):790–797, September 2001. ISSN 0028-4793. . URL https://doi.org/10.1056/NEJMoa010492.

Thomas Hubbard. The Boston Paradox: Lots of Health Care, Not Enough Health. *Boston Foundation*, November 2007. URL https://folio.iupui.edu/handle/10244/730. Accepted: 2010-04-30T19:09:11Z Publisher: Boston Foundation.

Karin H Humphries and Eddy van Doorslaer. Income-related health inequality in Canada. *Social Science & Medicine*, 50(5):663–671, March 2000. ISSN 0277-9536. . URL http://www.sciencedirect.com/science/article/pii/S0277953699003196.

Stefan Hut and Emily Oster. Changes in Household Diet: Determinants and Predictability. Technical Report w24892, National Bureau of Economic Research, Cambridge, MA, August 2018. URL http://www.nber.org/papers/w24892.pdf.

Guido W. Imbens and Thomas Lemieux. Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, 142(2):615–635, February 2008. ISSN 03044076. . URL https://linkinghub.elsevier.com/retrieve/pii/S0304407607001091.

Nazrul Islam. The dilemma of physician shortage and international recruitment in Canada. *International Journal of Health Policy and Management*, 3(1):29–32, May 2014. ISSN 2322-5939. . URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4075100/.

Aljoscha Janssen and Elle Parslow. Pregnancy persistently reduces alcohol purchases: Causal evidence from scanner data. *Health Economics*, 30(2):231–247, 2021. ISSN 1099-1050. . URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.4188. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.4188.

Stephen P. Jenkins and Lars Osberg. Nobody to Play with? In Daniel S. Hamermesh and Gerard A. Pfann, editors, *The Economics of Time Use*, volume 271 of *Contributions to Economic Analysis*, pages 113–145. Emerald Group Publishing Limited, January 2004. ISBN 978-1-84950-838-4 978-0-444-51534-6. . URL https://doi.org/10.1016/S0573-8555(04)71005-6.

Sung-Hee Jeon and R. Vincent Pohl. Health and work in the family: Evidence from spouses' cancer diagnoses. *Journal of Health Economics*, 52:1–18, March 2017. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629616305720.

George A Kaplan, Elsie R Pamuk, John W Lynch, Richard D Cohen, and Jennifer L Balfour. Inequality in income and mortality in the United States: analysis of mortality and potential pathways. *BMJ*, 312(7037):999–1003, 1996. ISSN 0959-8138. . URL https://www.bmj.com/content/312/7037/999.

Patricia Ketsche, E. Kathleen Adams, Sally Wallace, Viji Diane Kannan, and Harini Kannan. Lower-Income Families Pay A Higher Share Of Income Toward National Health Care Spending Than Higher-Income Families Do. *Health Affairs*, 30(9): 1637–1646, September 2011. ISSN 0278-2715. . URL https://www.healthaffairs.org/doi/full/10.1377/hlthaff.2010.0712. Publisher: Health Affairs.

Ahmed Khwaja, Frank Sloan, and Sukyung Chung. The Effects of Spousal Health on the Decision to Smoke: Evidence on Consumption Externalities, Altruism and Learning Within the Household. *Journal of Risk and Uncertainty*, 32(1):17–35, January 2006. ISSN 1573-0476. . URL https://doi.org/10.1007/s10797-006-6664-5.

Hyuncheol Bryant Kim and Wilfredo Lim. Long-term care insurance, informal care, and medical expenditures. *Journal of Public Economics*, 125:128–142, May 2015. ISSN 0047-2727. . URL https://www.sciencedirect.com/science/article/pii/S0047272714002503.

Hyuncheol Bryant Kim, Suejin A. Lee, and Wilfredo Lim. Knowing is not half the battle: Impacts of information from the National Health Screening Program in Korea. *Journal of Health Economics*, 65:1–14, May 2019. ISSN 0167-6296. . URL http://www.sciencedirect.com/science/article/pii/S0167629617304629.

Henrik Kleven, Camille Landais, and Jakob Egholt Søgaard. Children and Gender Inequality: Evidence from Denmark. *American Economic Journal: Applied Economics*,

11(4):181–209, October 2019. ISSN 1945-7782. . URL https://www.aeaweb.org/articles?id=10.1257/app.20180010.

L. N. Kolonel and J. Lee. Husband-wife correspondence in smoking, drinking, and dietary habits. *The American Journal of Clinical Nutrition*, 34(1):99–104, January 1981. ISSN 0002-9165. . URL https://academic.oup.com/ajcn/article/34/1/99/4692800.

Kelvin J. Lancaster. A New Approach to Consumer Theory. *Journal of Political Economy*, 74(2):132–157, 1966. ISSN 0022-3808. URL https://www.jstor.org/stable/1828835.

Camille Landais. Assessing the Welfare Effects of Unemployment Benefits Using the Regression Kink Design. *American Economic Journal: Economic Policy*, 7(4):243–278, November 2015. ISSN 1945-7731. . URL https://www.aeaweb.org/articles?id=10.1257/pol.20130248.

Miranda Laurant, David Reeves, Rosella Hermens, Jose Braspenning, Richard Grol, and Bonnie Sibbald. Substitution of doctors by nurses in primary care. *Cochrane Database of Systematic Reviews*, (2), 2005. ISSN 1465-1858. . URL https://www.cochranelibrary.com/cdsr/doi/10.1002/14651858.CD001271.pub2/full. Publisher: John Wiley & Sons, Ltd.

Victor Lavy and Analia Schlosser. Mechanisms and Impacts of Gender Peer Effects at School. *American Economic Journal: Applied Economics*, 3(2):1–33, April 2011. ISSN 1945-7782. . URL https://www.aeaweb.org/articles?id=10.1257/app.3.2.1.

K. Leadley, C. L. Clark, and R. Caetano. Couples' drinking patterns, intimate partner violence, and alcohol-related partnership problems. *Journal of Substance Abuse*, 11(3): 253–263, 2000. ISSN 0899-3289. .

Sokbae Lee. Endogeneity in quantile regression models: A control function approach. *Journal of Econometrics*, 141(2):1131–1158, December 2007. ISSN 0304-4076. . URL http://www.sciencedirect.com/science/article/pii/S0304407607000279.

Kenneth E. Leonard and Pamela Mudar. Peer and partner drinking and the transition to marriage: A longitudinal examination of selection and influence processes. *Psychology of Addictive Behaviors*, 17(2):115–125, 2003. ISSN 1939-1501(Electronic),0893-164X(Print). .

Harold S. Luft. The Impact of Poor Health on Earnings. *The Review of Economics and Statistics*, 57(1):43–57, 1975. ISSN 0034-6535. . URL https://www.jstor.org/stable/1937862.

Li-ching Lyu, Su-hua Huang, Chi-yin Hsu, Meei-shyuan Lee, and Shieh-hwa Lin. Inter-relationships of nutrient intakes for urban Chinese spouses in Taiwan.

International Journal of Food Sciences and Nutrition, 55(3):227–236, May 2004. ISSN 0963-7486. . URL https://doi.org/10.1080/09637480410001733851.

E. Macario and G. Sorensen. Spousal similarities in fruit and vegetable consumption. *American journal of health promotion: AJHP*, 12(6):369–377, August 1998. ISSN 0890-1171. .

James Macinko, Inês Dourado, Rosana Aquino, Palmira de Fátima Bonolo, Maria Fernanda Lima-Costa, Maria Guadalupe Medina, Eduardo Mota, Veneza Berenice de Oliveira, and Maria Aparecida Turci. Major expansion of primary care in Brazil linked to decline in unnecessary hospitalization. *Health Affairs (Project Hope)*, 29(12):2149–2160, December 2010. ISSN 1544-5208. .

Chris Mahony. Public health marketing campaigns: who profits? *BMJ*, 350:h514, January 2015. ISSN 1756-1833. . URL https://www.bmj.com/content/350/bmj.h514. Publisher: British Medical Journal Publishing Group Section: Feature.

Azeem Majeed. Primary care: a fading jewel in the NHS crown. *London Journal of Primary Care*, 7(5):89–91, September 2015. ISSN 1757-1472. . URL https://doi.org/10.1080/17571472.2015.1082343. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/17571472.2015.1082343.

Willard G. Manning, Joseph P. Newhouse, Naihua Duan, Emmett B. Keeler, and Arleen Leibowitz. Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment. *The American Economic Review*, 77(3):251–277, 1987. ISSN 0002-8282. URL http://www.jstor.org/stable/1804094.

M. G. Marmot, S. Stansfeld, C. Patel, F. North, J. Head, I. White, E. Brunner, A. Feeney, M. G. Marmot, and G. Davey Smith. Health inequalities among British civil servants: the Whitehall II study. *The Lancet*, 337(8754):1387–1393, June 1991. ISSN 0140-6736. . URL http://www.sciencedirect.com/science/article/pii/014067369193068K.

Enlinson Mattos and Debora Mazetto. Assessing the impact of more doctors' program on healthcare indicators in Brazil. *World Development*, 123:104617, November 2019. ISSN 0305-750X. . URL http://www.sciencedirect.com/science/article/pii/S0305750X19302116.

Fernanda Mazzotta, Francesca Bettio, and Valentina Zigante. Eldercare hours, work hours and perceived filial obligations. *Applied Economics*, 52(21):2219–2238, May 2020. ISSN 0003-6846. . URL https://doi.org/10.1080/00036846.2019.1687839. Publisher: Routledge _eprint: https://doi.org/10.1080/00036846.2019.1687839.

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

0304-4076. . URL
http://www.sciencedirect.com/science/article/pii/S0304407607001133.

J. Michael McGinnis, Pamela Williams-Russo, and James R. Knickman. The Case For More Active Policy Attention To Health Promotion. *Health Affairs*, 21(2):78–93, March 2002. ISSN 0278-2715. . URL
https://www.healthaffairs.org/doi/full/10.1377/hlthaff.21.2.78.
Publisher: Health Affairs.

Emmanouil Mentzakis, Paul McNamee, and Mandy Ryan. Who cares and how much: exploring the determinants of co-residential informal care. *Review of Economics of the Household*, 7(3):283–303, September 2009. ISSN 1573-7152. . URL
https://doi.org/10.1007/s11150-008-9047-0.

Deanna Meyler, Jim P. Stimpson, and M. Kristen Peek. Health concordance within couples: A systematic review. *Social Science & Medicine*, 64(11):2297–2310, June 2007. ISSN 0277-9536. . URL
http://www.sciencedirect.com/science/article/pii/S0277953607000433.

Pierre-Carl Michaud, Axel Heitmueller, and Zafar Nazarov. A dynamic analysis of informal care and employment in England. *Labour Economics*, 17(3):455–465, June 2010. ISSN 0927-5371. . URL
https://www.sciencedirect.com/science/article/pii/S0927537110000023.

Corina Mommaerts. Are coresidence and nursing homes substitutes? Evidence from Medicaid spend-down provisions. *Journal of Health Economics*, 59:125–138, May 2018. ISSN 0167-6296. . URL
https://www.sciencedirect.com/science/article/pii/S0167629617300607.

Irina Mozhaeva. Informal caregiving and work: A high price to pay. The case of Baltic States. *The Journal of the Economics of Ageing*, 19:100318, June 2021. ISSN 2212-828X. . URL
https://www.sciencedirect.com/science/article/pii/S2212828X21000128.

Ha Trong Nguyen and Luke Brian Connelly. The effect of unpaid caregiving intensity on labour force participation: Results from a multinomial endogenous treatment model. *Social Science & Medicine*, 100:115–122, January 2014. ISSN 0277-9536. . URL
https://www.sciencedirect.com/science/article/pii/S0277953613005790.

NHS. Type 2 diabetes - Food and keeping active, March 2018. URL https://www.nhs.uk/conditions/type-2-diabetes/food-and-keeping-active/.
Section: conditions.

NICE. Type 2 diabetes prevention: population and community-level interventions. Technical report, NICE Guideline, May 2011. URL
https://www.nice.org.uk/guidance/ph35. Publisher: NICE.

NICE. Type 2 diabetes: prevention in people at high risk. Technical report, NICE Guideline, July 2012. URL https://www.nice.org.uk/guidance/ph38/chapter/Recommendations.

Owen O'Donnell and Carol Propper. Equity and the distribution of UK National Health Service resources. *Journal of Health Economics*, 10(1):1–19, May 1991. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/016762969190014E.

OECD. *Realising the Potential of Primary Health Care*. OECD Health Policy Studies. OECD Publishing, Paris, 2020. URL https://doi.org/10.1787/a92adee4-en.

João Paulo Alves Oliveira, Mauro Niskier Sanchez, Leonor Maria Pacheco Santos, João Paulo Alves Oliveira, Mauro Niskier Sanchez, and Leonor Maria Pacheco Santos. The Mais Médicos (More Doctors) Program: the placement of physicians in priority municipalities in Brazil from 2013 to 2014. *Ciência & Saúde Coletiva*, 21(9):2719–2727, September 2016. ISSN 1413-8123. . URL http://www.scielo.br/scielo.php?script=sci_abstract&pid=S1413-81232016000902719&lng=en&nrm=iso&tlng=en. Publisher: ABRASCO - Associação Brasileira de Saúde Coletiva.

Stephen O'Neill, Noemi Kreif, Matt Sutton, and Richard Grieve. A comparison of methods for health policy evaluation with controlled pre-post designs. *Health Services Research*, 55(2):328–338, 2020. ISSN 1475-6773. . URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1475-6773.13274. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1475-6773.13274.

Athanasios Orphanides and David Zervos. Rational Addiction with Learning and Regret. *Journal of Political Economy*, 103(4):739–758, 1995. ISSN 0022-3808. URL https://www.jstor.org/stable/2138580. Publisher: University of Chicago Press.

Emily Oster. Diabetes and Diet: Purchasing Behavior Change in Response to Health Information. *American Economic Journal: Applied Economics*, 10(4):308–348, October 2018. ISSN 1945-7782. . URL https://www.aeaweb.org/articles?id=10.1257/app.20160232.

Jairnilson Paim, Claudia Travassos, Celia Almeida, Ligia Bahia, and James Macinko. The Brazilian health system: history, advances, and challenges. *The Lancet*, 377 (9779):1778–1797, May 2011. ISSN 0140-6736, 1474-547X. . URL https://www.thelancet.com/journals/lancet/article/PIIS0140-6736(11)60054-8/abstract. Publisher: Elsevier.

Supasit Pannarunothai and Anne Mills. The poor pay more: Health-related inequality in Thailand. *Social Science & Medicine*, 44(12):1781–1790, June 1997. ISSN 0277-9536. . URL http://www.sciencedirect.com/science/article/pii/S0277953696002870.

Lucélia L. Pereira, Leonor M. P. Santos, Wallace Santos, Aimê Oliveira, and Daphne Rattner. Mais Médicos program: provision of medical doctors in rural, remote and socially vulnerable areas of Brazil, 2013-2014. *Rural and Remote Health*, 16(1):3616, March 2016. ISSN 1445-6354.

Petra Persson, Xinyao Qiu, and Maya Rossin-Slater. Family Spillover Effects of Marginal Diagnoses: The Case of ADHD. Working Paper 28334, National Bureau of Economic Research, January 2021. URL https://www.nber.org/papers/w28334. Series: Working Paper Series.

Stephen M. Petterson, Winston R. Liaw, Robert L. Phillips, David L. Rabin, David S. Meyers, and Andrew W. Bazemore. Projecting US primary care physician workforce needs: 2010-2025. *Annals of Family Medicine*, 10(6):503–509, December 2012. ISSN 1544-1717. .

Linda Pickard. Substitution between formal and informal care: a 'natural experiment' in social policy in Britain between 1985 and 2000. *Ageing & Society*, 32(7):1147–1175, October 2012. ISSN 1469-1779, 0144-686X. . URL https://www.cambridge.org/core/journals/ageing-and-society/article/substitution-between-formal-and-informal-care-a-natural-experiment-in-social-policy-in-br 4A02728B6309E5092DEBD77DA106C8BC. Publisher: Cambridge University Press.

David Powell. Panel Data Inference with Dependent Clusters. April 2020a. URL https://sites.google.com/site/davidmatthewpowell/inference-with-correlated-clusters.

David Powell. Quantile Treatment Effects in the Presence of Covariates. *The Review of Economics and Statistics*, 102(5):994–1005, December 2020b. ISSN 0034-6535. . URL https://doi.org/10.1162/rest_a_00858.

David Powell. Synthetic Control Estimation Beyond Comparative Case Studies: Does the Minimum Wage Reduce Employment?*. *Journal of Business & Economic Statistics*, 0(ja):1–39, May 2021. ISSN 0735-0015. . URL https://doi.org/10.1080/07350015.2021.1927743. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/07350015.2021.1927743.

Frank Puffer. Access to Primary Health Care: A Comparison of the US and the UK*. *Journal of Social Policy*, 15(3):293–313, July 1986. ISSN 1469-7823, 0047-2794. . URL https://www.cambridge.org/core/journals/journal-of-social-policy/article/abs/access-to-primary-health-care-a-comparison-of-the-us-and-the-uk/1536AFB861B79766D348FC9BCECB869E. Publisher: Cambridge University Press.

S. Purdy, T. Griffin, C. Salisbury, and D. Sharp. Ambulatory care sensitive conditions: terminology and disease coding need to be more specific to aid policy makers and clinicians. *Public Health*, 123(2):169–173, February 2009. ISSN 1476-5616. .

Mala Rao and Eva Pilot. The missing link – the role of primary care in global health. *Global Health Action*, 7, February 2014. ISSN 1654-9716. . URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3926992/.

Edith Rasell, Jared Bernstein, and Kainan Tang. The Impact of Health Care Financing on Family Budgets. *International Journal of Health Services*, 24(4):691–714, October 1994. ISSN 0020-7314. . URL https://doi.org/10.2190/MM38-P4HV-2W32-4KYR. Publisher: SAGE Publications Inc.

Davide Rasella, Michael O. Harhay, Marina L. Pamponet, Rosana Aquino, and Mauricio L. Barreto. Impact of primary health care on mortality from heart and cerebrovascular diseases in Brazil: a nationwide analysis of longitudinal data. *BMJ*, 349:g4014, July 2014. ISSN 1756-1833. . URL https://www.bmj.com/content/349/bmj.g4014. Publisher: British Medical Journal Publishing Group Section: Research.

Aart-Jan Riekhoff and Maria Vaalavuo. Health shocks and couples' labor market participation: A turning point or stuck in the trajectory? *Social Science & Medicine*, 276:113843, May 2021. ISSN 0277-9536. . URL https://www.sciencedirect.com/science/article/pii/S0277953621001751.

Jeanne S. Ringel, Susan D. Hosek, Ben A. Vollaard, and Sergej Mahnovski. The Elasticity of Demand for Health Care. A Review of the Literature and its Application to the Military Health System. Technical Report MR-1355-OSD, RAND NATIONAL DEFENSE RESEARCH INST SANTA MONICA CA, January 2002. URL http://www.dtic.mil/docs/citations/ADA403148.

David Roodman. Fitting fully observed recursive mixed-process models with cmp. *The Stata Journal*, 11(2):159–206, 2011.

Jonathan Roth and Pedro H. C. Sant'Anna. Efficient Estimation for Staggered Rollout Designs. *arXiv:2102.01291 [econ, math, stat]*, March 2021. URL http://arxiv.org/abs/2102.01291. arXiv: 2102.01291.

Jonathan Roth, Pedro H. C. Sant'Anna, Alyssa Bilinski, and John Poe. What's Trending in Difference-in-Differences? A Synthesis of the Recent Econometrics Literature. *arXiv:2201.01194 [econ, stat]*, January 2022. URL http://arxiv.org/abs/2201.01194. arXiv: 2201.01194.

L. B. Russell, M. R. Gold, J. E. Siegel, N. Daniels, and M. C. Weinstein. The role of cost-effectiveness analysis in health and medicine. Panel on Cost-Effectiveness in Health and Medicine. *JAMA*, 276(14):1172–1177, October 1996. ISSN 0098-7484.

Adam Sacarny, Katherine Baicker, and Amy Finkelstein. Out of the Woodwork: Enrollment Spillovers in the Oregon Health Insurance Experiment. Working Paper

26871, National Bureau of Economic Research, March 2020. URL https://www.nber.org/papers/w26871. Series: Working Paper Series.

Bruce Sacerdote. Chapter 4 - Peer Effects in Education: How Might They Work, How Big Are They and How Much Do We Know Thus Far? In Eric A. Hanushek, Stephen Machin, and Ludger Woessmann, editors, *Handbook of the Economics of Education*, volume 3, pages 249–277. Elsevier, January 2011. . URL https://www.sciencedirect.com/science/article/pii/B9780444534293000041.

Leonor Maria Pacheco Santos, Aimê Oliveira, Josélia Souza Trindade, Ivana CHC Barreto, Poliana Araújo Palmeira, Yamila Comes, Felipe OS Santos, Wallace Santos, João Paulo Alves Oliveira, Vanira Matos Pessoa, and Helena Eri Shimizu. Implementation research: towards universal health coverage with more doctors in Brazil. *Bulletin of the World Health Organization*, 95(2):103–112, February 2017. ISSN 0042-9686. . URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5327934/.

Leonor Maria Pacheco Santos, Christopher Millett, Davide Rasella, and Thomas Hone. The end of Brazil's More Doctors programme? *BMJ*, 363, December 2018. ISSN 0959-8138, 1756-1833. . URL https://www.bmj.com/content/363/bmj.k5247. Publisher: British Medical Journal Publishing Group Section: Editorial.

Hendrik Schmitz and Matthias Westphal. Informal care and long-term labor market outcomes. *Journal of Health Economics*, 56:1–18, December 2017. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629616304350.

Donald S. Shepard. Cost-effectiveness in Health and Medicine. By M.R. Gold, J.E Siegel, L.B. Russell, and M.C. Weinstein (eds). New York: Oxford University Press, 1996. *The Journal of Mental Health Policy and Economics*, 2(2):91–92, 1999. ISSN 1099-176X. . URL https://onlinelibrary.wiley.com/doi/abs/10.1002/%28SICI%291099-176X%28199906%292%3A2%3C91%3A%3AAID-MHP46%3E3.0.CO%3B2-I. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/%28SICI%291099-176X%28199906%292%3A2%3C91%3A%3AAID-MHP46%3E3.0.CO%3B2-I.

Leiyu Shi. The Impact of Primary Care: A Focused Review. *Scientifica*, 2012:e432892, December 2012. . URL https://www.hindawi.com/journals/scientifica/2012/432892/. Publisher: Hindawi.

Michele J. Siegel. Measuring the effect of husband's health on wife's labor supply. *Health Economics*, 15(6):579–601, 2006. ISSN 1099-1050. . URL https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.1084. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/hec.1084.

Tiago Rodrigues Bento da Silva, Jennifer do Vale e Silva, Andrezza Graziella Veríssimo Pontes, and Andrea Taborda Ribas da Cunha. Service users' perceptions of the Mais Médicos (More Doctors) Program in the Municipality of Mossoró, Brazil. *Ciência & Saúde Coletiva*, 21:2861–2869, September 2016. ISSN 1413-8123, 1413-8123, 1678-4561. . URL https://www.scielosp.org/article/csc/2016.v21n9/2861-2869/en/. Publisher: ABRASCO - Associação Brasileira de Saúde Coletiva.

J. Speight and C. Bradley. The ADKnowl: identifying knowledge deficits in diabetes care. *Diabetic Medicine*, 18(8):626–633, 2001. ISSN 1464-5491. . URL https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1464-5491.2001.00537.x. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1046/j.1464-5491.2001.00537.x.

Mark Stabile, Audrey Laporte, and Peter C. Coyte. Household responses to public home care programs. *Journal of Health Economics*, 25(4):674–701, July 2006. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629605001165.

Barbara Starfield. *Primary Care: Balancing Health Needs, Services, and Technology*. Oxford University Press, 1998. ISBN 978-0-19-512543-6. Google-Books-ID: QMm17oCEjrEC.

Barbara Starfield and Leiyu Shi. Policy relevant determinants of health: an international perspective. *Health Policy (Amsterdam, Netherlands)*, 60(3):201–218, June 2002. ISSN 0168-8510. .

Barbara Starfield, Leiyu Shi, and James Macinko. Contribution of Primary Care to Health Systems and Health. *The Milbank Quarterly*, 83(3):457–502, September 2005. ISSN 0887-378X. . URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2690145/.

Jim P. Stimpson, Meredith C. Masel, Laura Rudkin, and M. Kristen Peek. Shared health behaviors among older Mexican American spouses. *American Journal of Health Behavior*, 30(5):495–502, October 2006. ISSN 1087-3244. .

Liyang Sun and Sarah Abraham. Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, December 2020. ISSN 0304-4076. . URL https://www.sciencedirect.com/science/article/pii/S030440762030378X.

Leonie Sundmacher, Diana Fischbach, Wiebke Schuettig, Christoph Naumann, Uta Augustin, and Cristina Faisst. Which hospitalisations are ambulatory care-sensitive, to what degree, and how could the rates be reduced? Results of a group consensus study in Germany. *Health Policy*, 119(11):1415–1423, November 2015. ISSN 0168-8510. . URL https://www.sciencedirect.com/science/article/pii/S0168851015001943.

Kevin Sylwester. Can education expenditures reduce income inequality? *Economics of Education Review*, 21(1):43–52, February 2002. ISSN 0272-7757. . URL http://www.sciencedirect.com/science/article/pii/S0272775700000388.

Célia L. Szwarcwald, Paulo RB Souza-Júnior, and Giseli N. Damacena. Socioeconomic inequalities in the use of outpatient services in Brazil according to health care need: evidence from the World Health Survey. *BMC Health Services Research*, 10(1):217, July 2010. ISSN 1472-6963. . URL https://doi.org/10.1186/1472-6963-10-217.

Sarah L. Taubman, Heidi L. Allen, Bill J. Wright, Katherine Baicker, and Amy N. Finkelstein. Medicaid Increases Emergency-Department Use: Evidence from Oregon's Health Insurance Experiment. *Science*, January 2014. . URL https://www.science.org/doi/abs/10.1126/science.1246183. Publisher: American Association for the Advancement of Science.

Joseph Terza, Anirban Basu, and Paul J. Rathouz. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, 27(3):531–543, 2008a.

Joseph V. Terza, Anirban Basu, and Paul J. Rathouz. Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling. *Journal of Health Economics*, 27(3):531–543, May 2008b. ISSN 0167-6296. . URL http://www.sciencedirect.com/science/article/pii/S0167629607001063.

The Kaiser Family Foundation. Primary Care Health Professional Shortage Areas (HPSAs), November 2020. URL https://www.kff.org/other/state-indicator/primary-care-health-professional-shortage-areas-hpsas/.

Joseph Truglio, Michelle Graziano, Rajesh Vedanthan, Sigrid Hahn, Carlos Rios, Brett Hendel-Paterson, and Jonathan Ripp. Global Health and Primary Care: Increasing Burden of Chronic Diseases and Need for Integrated Training. *The Mount Sinai journal of medicine, New York*, 79(4):464–474, 2012. ISSN 0027-2507. . URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4362673/.

U.N. General Assembly. Resolution adopted by the General Assembly on 12 December 2012, March 2013.

U.N. General Assembly. Transforming our World: The 2030 Agenda for Sustainable Development, September 2015. URL https://sustainabledevelopment.un.org/post2015/transformingourworld/publication.

United States Census Bureau. American Community Survey 5-Year Estimates 2005-2009, 2009. URL https://www.census.gov/data/developers/data-sets/acs-5year/2009.html.

United States Census Bureau. American Community Survey 1-Year Estimates 2010, 2010. URL https://data.census.gov/cedsci/.

Martin Valdivia. Public health infrastructure and equity in the utilization of outpatient health care services in Peru. *Health Policy and Planning*, 17(S1):12–19, December 2002. ISSN 0268-1080. .

Eddy van Doorslaer and Adam Wagstaff. Equity in the delivery of health care: some international comparisons. *Journal of Health Economics*, 11(4):389–411, December 1992. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/016762969290013Q.

Eddy van Doorslaer, Adam Wagstaff, Han Bleichrodt, Samuel Calonge, Ulf-G. Gerdtham, Michael Gerfin, José Geurts, Lorna Gross, Unto Häkkinen, Robert E. Leu, Owen O'Donell, Carol Propper, Frank Puffer, Marisol Rodríguez, Gun Sundberg, and Olaf Winkelhake. Income-related inequalities in health: some international comparisons. *Journal of Health Economics*, 16(1):93–112, February 1997. ISSN 0167-6296. . URL http://www.sciencedirect.com/science/article/pii/S0167629696005322.

Eddy van Doorslaer, Adam Wagstaff, Hattem van der Burg, Terkel Christiansen, Guido Citoni, Rita Di Biase, Ulf-G Gerdtham, Mike Gerfin, Lorna Gross, Unto Häkinnen, Jürgen John, Paul Johnson, Jan Klavus, Claire Lachaud, Jørgen Lauritsen, Robert Leu, Brian Nolan, João Pereira, Carol Propper, Frank Puffer, Lise Rochaix, Martin Schellhorn, Gun Sundberg, and Olaf Winkelhake. The redistributive effect of health care finance in twelve OECD countries. *Journal of Health Economics*, 18(3): 291–313, June 1999. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629698000435.

Eddy van Doorslaer, Adam Wagstaff, Hattem van der Burg, Terkel Christiansen, Diana De Graeve, Inge Duchesne, Ulf-G Gerdtham, Michael Gerfin, José Geurts, Lorna Gross, Unto Häkkinen, Jürgen John, Jan Klavus, Robert E Leu, Brian Nolan, Owen O'Donnell, Carol Propper, Frank Puffer, Martin Schellhorn, Gun Sundberg, and Olaf Winkelhake. Equity in the delivery of health care in Europe and the US. *Journal of Health Economics*, 19(5):553–583, September 2000. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629600000503.

Courtney Harold Van Houtven and Edward C. Norton. Informal care and health care use of older adults. *Journal of Health Economics*, 23(6):1159–1180, November 2004. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629604000852.

Courtney Harold Van Houtven, Norma B. Coe, and Meghan M. Skira. The effect of informal care on work and wages. *Journal of Health Economics*, 32(1):240–252, January 2013. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629612001671.

M. H. Venters, D. R. Jacobs, R. V. Luepker, L. A. Maiman, and R. F. Gillum. Spouse concordance of smoking patterns: the Minnesota Heart Survey. *American Journal of Epidemiology*, 120(4):608–616, October 1984. ISSN 0002-9262. .

Adam Wagstaff and Eddy van Doorslaer. Equity in the finance of health care: Some international comparisons. *Journal of Health Economics*, 11(4):361–387, December 1992. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/016762969290012P.

Adam Wagstaff, Eddy Van Doorslaer, and Pierella Paci. Equity in the finance and delivery of health care: some tentative cross-country comparisons. *Oxford Review of Economic Policy*, 5(1):89–112, March 1989. ISSN 0266-903X. . URL https://doi.org/10.1093/oxrep/5.1.89.

Adam Wagstaff, Eddy van Doorslaer, Hattem van der Burg, Samuel Calonge, Terkel Christiansen, Guido Citoni, Ulf-G Gerdtham, Mike Gerfin, Lorna Gross, Unto Häkinnen, Paul Johnson, Jürgen John, Jan Klavus, Claire Lachaud, Jørgen Lauritsen, Robert Leu, Brian Nolan, Encarna Perán, João Pereira, Carol Propper, Frank Puffer, Lise Rochaix, Marisol Rodrıguez, Martin Schellhorn, Gun Sundberg, and Olaf Winkelhake. Equity in the finance of health care: some further international comparisons1This paper derives from the project 'Equity in the finance and delivery of health care in Europe' (the so-called ECuity Project), funded by the European Union's Biomed I Programme (contract BMH1-CT92). Eddy van Doorslaer and Adam Wagstaff jointly coordinated the project. We are grateful to the EU for financial support and to two referees for their helpful comments on an earlier version of the paper. We alone are responsible for any remaining errors.1. *Journal of Health Economics*, 18(3):263–290, June 1999. ISSN 0167-6296. . URL https://www.sciencedirect.com/science/article/pii/S0167629698000447.

Adam Wagstaff, Magnus Lindelow, Gao Jun, Xu Ling, and Qian Juncheng. Extending health insurance to the rural population: An impact evaluation of China's new cooperative medical scheme. *Journal of Health Economics*, 28(1):1–19, January 2009. ISSN 0167-6296. . URL http://www.sciencedirect.com/science/article/pii/S0167629608001392.

Hugh R. Waters. Measuring equity in access to health care. *Social Science & Medicine*, 51(4):599–612, August 2000. ISSN 0277-9536. . URL https://www.sciencedirect.com/science/article/pii/S0277953600000034.

France M. Weaver and Bryce A. Weaver. Does availability of informal care within the household impact hospitalisation? *Health Economics, Policy and Law*, 9(1):71–93, January 2014. ISSN 1744-1331, 1744-134X. . URL https://www.cambridge.org/core/journals/health-economics-policy-and-law/article/does-availability-of-informal-care-within-the-household-impact-hospitalisation/E0C0473A230786C1A732AB0579637931. Publisher: Cambridge University Press.

Matthew D. Webb. Reworking Wild Bootstrap Based Inference For Clustered Errors. Technical Report 1315, Economics Department, Queen's University, November 2014. URL https://ideas.repec.org/p/qed/wpaper/1315.html. Publication Title: Working Paper.

Diane Whitmore. Resource and Peer Impacts on Girls' Academic Achievement: Evidence from a Randomized Experiment. *American Economic Review*, 95(2):199–203, May 2005. ISSN 0002-8282. . URL https://www.aeaweb.org/articles?id=10.1257/000282805774670158.

WHO. Diabetes. URL https://www.who.int/westernpacific/health-topics/diabetes.

WHO. Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus. Technical report, 2011. URL https://www.who.int/diabetes/publications/diagnosis_diabetes2011/en/.

WHO. Global report on diabetes. Technical report, 2016. URL https://www.who.int/publications-detail/global-report-on-diabetes.

WHO. Universal health coverage (UHC), 2019a. URL https://www.who.int/news-room/fact-sheets/detail/universal-health-coverage-(uhc).

WHO. Who - What is universal coverage?, 2019b. URL http://www.who.int/health_financing/universal_coverage_definition/en/.

WHO. Primary health care, January 2021. URL https://www.who.int/news-room/fact-sheets/detail/primary-health-care.

Richard G. Wilkinson. Socioeconomic determinants of health: Health inequalities: relative or absolute material standards? *BMJ*, 314(7080):591, February 1997. ISSN 0959-8138, 1468-5833. . URL https://www.bmj.com/content/314/7080/591. Publisher: British Medical Journal Publishing Group Section: Education and debate.

Carole Willi, Patrick Bodenmann, William A. Ghali, Peter D. Faris, and Jacques Cornuz. Active smoking and the risk of type 2 diabetes: a systematic review and meta-analysis. *JAMA*, 298(22):2654–2664, December 2007. ISSN 1538-3598. .

World Bank. Universal Health Coverage, 2021a. URL https://www.worldbank.org/en/topic/universalhealthcoverage.

World Bank. Universal Health Coverage, 2021b. URL https://www.worldbank.org/en/topic/universalhealthcoverage.

World Health Organisation, editor. *Working together for health*. Number 2006 in The world health report. World Health Organisation, Geneva, 2006. ISBN 978-92-4-156317-8. OCLC: 254951518.

World Health Organization, editor. *The World health report: health systems financing: the path to universal coverage*. World Health Organization, Geneva, 2010. ISBN 978-92-4-156402-1. OCLC: ocn676729081.

World Health Organization. SDG 10: Health and reduced inequalities. Policy Brief, Regional Office for Europe, 2019. URL https://www.euro.who.int/en/health-topics/health-policy/sustainable-development-goals/publications/2019/policy-briefs-on-health-and-the-sustainable-development-goals/sdg-10-health-and-reduced-inequalities.

World Health Organization. 10 global health issues to track in 2021, 2021a. URL https://www.who.int/news-room/spotlight/10-global-health-issues-to-track-in-2021.

World Health Organization. Global Health Estimates, 2021b. URL https://www.who.int/data/gho/data/themes/mortality-and-global-health-estimates.

World Health Organization. What is PHC?, 2021c. URL https://www.who.int/activities/what-is-PHC.

Byung-Kwang Yoo, Jay Bhattacharya, Kathryn M McDonald, and Alan M Garber. Impacts of Informal Caregiver Availability on Long-term Care Expenditures in OECD Countries. *Health Services Research*, 39(6 Pt 2):1971–1992, December 2004. ISSN 0017-9124. . URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1361108/.

John S. Yudkin and Victor M. Montori. The epidemic of pre-diabetes: the medicine and the politics. *BMJ*, 349:g4485, July 2014. ISSN 1756-1833. . URL https://www.bmj.com/content/349/bmj.g4485.