UNIVERSITY OF SOUTHAMPTON

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

Three essays in health economics: Using prescription data to explore questions on variation in medical care and population health

by

Matías Francisco Ortiz de Zárate Rodríguez
BSc, MSc

A thesis for the degree of
Doctor of Philosophy

July 2022

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. Parts of this work have been published as:
   Matias Ortiz de Zarate, Emmanouil Mentzakis, Simon DS Fraser, Paul Roderick, Paul Rutter, and Carmine Ornaghi. Price versus clinical guidelines in primary care statin prescribing: a retrospective cohort study and cost simulation model. *Journal of the Royal Society of Medicine*, 2021

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

Jointly Authored Work Statement

Chapter titled "Mental health and the COVID-19 pandemic: evidence from antidepressant prescriptions in general practices in England" was co-written with Professor Carmine Ornaghi and Dr. Emmanouil Mentzakis (University of Southampton). The contribution of each author is outlined below:

- I, EM and CO contributed to the design of the study
- I extracted the data and wrote the statistical programmes
- I wrote the first draft
- All authors contributed to further drafts, conceptualisation and approved the final manuscript

Chapter titled "Price versus clinical guidelines in primary care *statin* prescribing: a retrospective cohort study and cost simulation model" was co-written with Professor Carmine Ornaghi, Dr. Emmanouil Mentzakis, Dr. Simon DS Fraser, Professor Paul Roderick (University of Southampton), and Professor Paul Rutter (University of Portsmouth). The contribution of each author is outlined below:

- I, EM and CO contributed to the design of the study
- I extracted the data and wrote the statistical programmes
- I, EM and CO wrote the first draft
- All authors contributed to further drafts, conceptualisation and approved the final manuscript

Acknowledgements

First and foremost, to my two supervisors Carmine Ornaghi and Emmanouil Mentzakis. For their mentorship, advice, time, patience, interest and friendship.

To Corrado Giulietti, Michael Vlassopoulos and Brendon Mcconnell for taking part in my progression reviews and commenting on my work.

To Simon Fraser, Paul Roderick, and Paul Rutter (University of Portsmouth) for the joint work we did resulting in my first publication.

To the South Coast Doctoral Training Partnership (ESRC) for funding my studies and training.

For Lee

Chapter 1

Introduction

Health spending is growing globally across countries of all income levels, and it continues to expand faster than the economy, relying less on out-of-pocket spending and more on public funding (World Health Organization, 2019). Greater life expectancy with the resulting increasing ageing population, the higher prevalence of long-term chronic conditions, such as cardiovascular diseases, cancer, diabetes and chronic respiratory diseases, growing technology costs and rising demand are among the contributing factors behind its rise. It has been suggested that strengthening the supply of primary care practitioners, population coverage and engagement, and its preventive character may be one of the most cost-effective ways to reduce health care costs and unnecessary medical interventions in specialized (secondary) care (Scott, 2000; Starfield et al., 2005). In tax-funded health services, such as the English National Health Service (NHS), the reduction of unwarranted variation in care, and the reduction of pharmaceutical expenditure (growing at a considerably higher rate than the overall NHS budget), are identified as key aspects to improve efficiency and to tackle increasing health spending.[1] However, to address these latter challenges, evidence is still needed to understand the scope of the variation in the use of treatments and interventions across all levels of health care, for example, the management of chronic illnesses and medicine prescriptions in primary care, and to empirically show if the observed variation is justified in terms of improving patients outcomes.

While facing these challenges, countries are experiencing the aftermath of the COVID-19 pandemic starting in early 2020. Notwithstanding its claiming of millions of lives and rising infections, economic and social implications, the pandemic has tested the capacity of health care systems, affected the way people seek health, and changed the underlying health of the population, in particular, that of mental health. Regarding the latter, the evidence points to an exacerbation of the risk factors associated with poor mental health (e.g. economic insecurity, unemployment), a deterioration of the

[1] See https://www.england.nhs.uk/five-year-forward-view/next-steps-on-the-nhs-five-year-forward-view/funding-and-efficiency/.

protective factors (e.g. social interaction, routine, access to health care), resulting in its worsening (OECD, 2021). However, most of the evidence relies on survey data using self-reported measures of anxiety and depression as a proxy for mental health,[2] which in turn may be subject to a number of issues such as lack of representativeness, measurement error, reporting heterogeneity, and respondent's attrition. Little evidence comes from outcomes reflecting, not only the individuals' subjective experience about her or his mental health status but also how health care systems internalize the underlying changes in their treated population health, e.g. the prescription of psychotropic medication, referrals to specialized psychological therapies, etc.

Against this background, this dissertation presents three research articles in health economics that explore empirical questions on the topics of variation in medical care and patients health following a large-scale shock, all approached by the analysis of prescription data from general practice.

Specifically, in two related chapters, I analyse the prescribing behaviour of general practitioners in the context of patent expirations and the introduction of public clinical guidelines, and its consequences on costs to the health service and patients outcomes using the market of cholesterol-lowering medication in England as an exemplar. The development of this thesis has been crossed by the contingency of the COVID-19 pandemic. As it unfolded, I decided to use the knowledge and techniques acquired from the above-mentioned investigation to produce original evidence on the consequences of such a large-scale shock on mental health, departing from traditional measures that capture this aspect of health. Accordingly, in one self-contained chapter, I use antidepressant prescriptions data from general practitioner practices to investigate the impact of the COVID-19 pandemic on population mental health in England during the first year since its onset.

Hereafter, I outline the general content of each essay.

In the essay **"Price versus clinical guidelines in primary care *statin* prescribing: a retrospective cohort study and cost simulation model"**, we analyse the impact of generic entry following patent expiration of drugs, and the publication of national clinical guidance from the National Institute for Health and Care Excellence (NICE), on prescribing behaviour of general practitioners (GP) practices in primary care in England. The setting is the prescription of statins, cholesterol-lowering medication for the treatment of cardiovascular disease (CVD) risk, prescribed for new patients starting treatment between 2004 and 2018. We find that the trends in drug uptake are largely driven by a decrease in acquisition costs triggered by patent expiration, preceding the publication of NICE guidelines, which themselves did not seem to affect prescription trends. Significant heterogeneity is observed in the prescription of the most cost-effective statin across GP practices over time. Additionally, we perform a cost simulation exercise to

[2] For example, Patient Health Questionnaire 9 (PHQ-9) and General Anxiety Disorder (GAD-7).

estimate the cost-saving that could have been achieved by assuming perfect therapeutic substitution among alternative treatments. It shows that, between 2004 and 2018, the NHS could have saved £2.8bn, i.e. around 40% of the £6.3bn spent on statins during this time, if all GP practices had prescribed only the most cost-effective treatment. We conclude that there is potential for large savings for the NHS if new and, whenever possible, ongoing patients are promptly switched to the first medicine that becomes available as generic within a therapeutic class as long as it has similar efficacy to still-patented medicines.

This paper has been published in the Journal of the Royal Society of Medicine under: *Ortiz De Zarate M, Mentzakis E, Fraser SD, Roderick P, Rutter P, Ornaghi C. Price versus clinical guidelines in primary care statin prescribing: a retrospective cohort study and cost simulation model. Journal of the Royal Society of Medicine. November 2021,*[3] and the content of the essay is an exact reproduction of the published version.

In the essay **"Prescribing behaviour and patient health outcomes: evidence from *statin* prescriptions in general practices in England"** I extend previous the analysis and explore the implications of the observed variation in treatments on patient outcomes. Specifically, I use the market of statins to study the relationship between heterogeneous prescribing behaviour, and costs to the health service as well as patients' health outcomes. Based on the drug treatment choices made by general practitioners for new patients treated between 2004 and 2018, I characterize prescribing behaviour along two dimensions: the propensity to follow cost-effective and guidance-based recommendations, what I call *compliance*, and the extent to which health providers react to their patients' health characteristics when allocating treatments, which I call *responsiveness*. I find that least compliant providers have significantly higher costs than most compliant and that such difference comes at no additional benefit, as there is no statistically significant difference in terms of the probability of patients experiencing an adverse health outcome throughout the treatment course. Regarding responsiveness, I find that this dimension does not result in relevant differences in costs or health outcomes. The results ultimately show that the most cost-effective prescribing behaviour is displayed by the most compliant and least responsive providers, as no prescribing style appears to improve outcomes while this specific group is associated with the lowest costs. Overall, the prescription of anti-cholesterol drugs in England has similar features to those found in other health markets intensive in the use of surgical procedures, i.e. large variation across providers, significant differences in costs but no improvements in health. The findings imply that further efforts could have been made to convince practitioners to follow existing recommendations.

In the essay **"Mental health and the COVID-19 pandemic: evidence from antidepressant prescriptions in general practices in England"**, we use data of antidepressants

[3]Link to the online version: https://journals.sagepub.com/doi/full/10.1177/01410768211051713.

prescribed by general practices to investigate the impact of COVID-19 on population mental health in England during the first year of the pandemic. The challenge in estimating such an effect is to isolate it from the simultaneous reduction in access to general practice health services – another consequence of the pandemic. To do this, we construct a counterfactual group inspired by the synthetic control method. To construct the counterfactual group we rely on the prescription series of drug classes other than antidepressants provided that they satisfy two requirements: first, that the incidence of the conditions for which these other classes are prescribed for was not directly affected by the pandemic but that nevertheless experienced a reduction in prescription levels due to the contraction in general practice services, and second, that these classes can be shown to satisfactorily predict the series of antidepressants in the pre-pandemic period. By combining a set of drugs that satisfy these requirements we provide an estimate of the effect of the COVID-19 pandemic on depression and related diagnoses, net of the access effect. Two falsification tests support the validity of the approach. The results show an average increase of 152 daily doses of antidepressant medication per month per general practice – the equivalent of treating five additional adult patients per practice during the first year of the pandemic. Compared to the national prevalence of depression (on average 828 patients per practice or 11.6 per cent of the population in 2019) this figure suggests a moderate impact of the pandemic on mental health and a deviation from the public perception and findings of early studies of a sizeable negative impact. Notwithstanding, our analysis uncovers significant geographical heterogeneity across local authority districts in the estimated effect on antidepressants. We find that antidepressants increased more in districts associated with reduced mobility, and higher infection incidence of coronavirus, and that they increased less in district greater number of unemployment benefit claimants, potentially highlighting the benefits of the social safety net on mental health.

This essay was submitted for publication on December 13, 2021.

Before continuing with the three core chapters of this dissertation, two sections now follow the remaining of this introductory chapter that provide a preliminary conceptual and contextual background against which the ensuing papers can be approached. The first one, Section 1.1, summarizes the main ideas in the handbook chapter *Who Ordered That? The Economics of Treatment Choices in Medical Care* by Chandra et al. (2012), and the second one, Section 1.2, describes main aspects of the institutional setting in which general practices operate.

1.1 Conceptual Framework on the Economics of Treatment Choice

As it was suggested earlier, an important area of work in health economics is the study of variations in medical care, the so-called area variations. That is, how medical treatment choices are made and why they are so diverse? A first step in answering these questions comes from having a common conceptual/theoretical framework that aims at identifying the driving factors behind the observed behaviour and outcomes. In this sense, I next summarize the main ideas outlined in the handbook chapter by Chandra et al. (2012) on the economics of treatment choice.

There are three main categories of factors influencing treatment choices. The first is demand-side factors: price, income and preferences. In most health market purchasing of care by one of these three elements. Yet, universal health care systems, like the NHS, there is no significant role for prices and income as almost all health care is free at the point of delivery. Preferences can and do vary. Some treatments reduce greatly the probability of an adverse outcomes, or they increase the length (quantity) of life, while they can have detrimental effects to quality of life, side-effects, etc. Patients care and weigh these aspects to a different degree. The authors conjecture that differences in preferences do not explain a great part of treatment variation, not because preferences do not differ, but because physicians do not fully take them into account. In summary, patients' utility function depends on the benefit they can asses for the different alternative treatments, conditional on their severity, income net of the cost of treatment (when it applies), and the patient-specific idiosyncratic error term capturing heterogeneity in the benefits of each treatment, and preferences over side-effects, for example.

The second category corresponds to supply-side factors. In here, elements in the physician's utility function that can influence the choice of treatment include: the perfect/imperfect way in which they can assess the benefit a patient can derive from the different alternative treatments; the fees associated to performing them; the prior training, experience, procedural skill, and specialization; the different technologies available in the health care organization in which they operate, which makes them to adopt different faith in different treatments; and the internalization of medical literature and clinical guidelines; among others. Then, heterogeneity in treatment choice can be explained by how the benefits to the patients may differ, the different weight doctors place on the patient's benefit relative to the profits (professionalism), the different fees and how they may induce physician to provide more of one treatment, and other factors including specialization and training, adoption of guidelines, defensive medicine (the idea that providers perform unnecessary procedures in order to protect themselves from lawsuits).

A third category relates to the influences specific to the clinical situation in which the treatment determination is made. These are situational factors. The following examples illustrate the diversity and randomness of medical situations: *"A particular man*

may have preferred radiation to surgery, but the urologist on call that day happened to talk the man into surgery. The tragic or salient death of a man during surgery the previous week may lead a primary care physician to refer her next patient in similar condition to a radiation oncologist instead. The lengthy wait for a radiation appointment relative to surgery may lead a patient to choose surgery." In this sense, physicians and patients would be following psychological heuristics, or rules of thumb, which affect clinical decision making. It has been theorized that physicians deploy heuristics-based (or ready-to-wear) treatments because they minimize a number of behavioural costs relative to patient-specific (or tailor-made) treatments, related to the availability heuristic to make decisions,[4] communication regarding treatment choices and their risk (framing), channel factors that affects the way in which benefits and costs of an specific treatment choice are perceived, and the tendency of not to change their behaviour thus leading to status quo and confirmation bias. Other situational factors do not relate exclusive to either the physician or the patients, or a combination of the two, and may depend on the health care organization, or the insurer.

Finally, there is the gray area of medicine. This refers to clinical situations in which there is no authoritative guidelines or consensus on how to manage or treat certain conditions, where there is no black or white. Even in some areas of medicine where there is evidence regarding the effectiveness of alternative treatments (e.g. from clinical trials), where we learn from average effects, there is still little or no knowledge that tells physicians about a patients' potential marginal benefits, or about the physician's own impact due procedural skill in utilizing such treatment on the potential benefit. In this kind of situations, economic incentives are likely to have the largest impacts. The scenario can result in that patients with a similar set of observable characteristics (severity or baseline risk) end up receiving widely different treatments.

1.2 General Practice Institutional Background

The UK's national healthcare system is the National Health Service (NHS), it is funded primarily by general taxation, and is free at the point of use. However, patients pay a charge for prescription-only-medicines that is currently set at £9.35 per item, while some items are always free, including contraceptives and medicines prescribed for hospital inpatients. Children, the elderly, and certain medically exempt individuals do not pay a prescription charge.

To receive primary care health services from the NHS patients must first register, for free, with a general practice. General Practice is usually the first point of contact with

[4] In diagnosis, the availability heuristic suggests that a physician who has just seen a patient with influenza may be more likely to make the diagnosis of influenza for the next patient who walks through the door with a cough, even if this latter patient has a rare lung disease.

the health service and can refer patients to other levels of care such as elective (non-emergency) hospital care, acting as gate-keepers for higher-level health care settings. Patients can register with any surgery that is closer to where they live or work (or where their children go to school). However, practices may refuse to accept patients for two main reasons. First, it can refuse to register patients who live outside the practice boundary (or catchments area) agreed with their Clinical Commissioning Group (CCGs) – formerly Primary Care Trust.[5] Second, practices can notify their CCGs that their patient list is closed, in which case no new patients can be accepted for a period between 3-12 months. Despite this, a practice cannot refuse to accept patients on the grounds of gender, age, race, social class, religion, sexual orientation, appearance, disability or pre-existing medical conditions.

In 2019 in England, there were approximately 60 million patients registered at around 6,800 GP practices, resulting in around 8,800 patients per practice, on average. In that same period the GP workforce is constituted by a total of around 34,300 full-time equivalent (FTE) practitioners, resulting in 57 FTE GPs per 100,000 patients, or 5 FTE GPs per practice.[6]

GP practice contracts and funding. General practices are small to medium-size firms most of which are limited liability partnerships owned by a partnership of GPs or an individual GP. The NHS commissioners (either NHS England or CCGs) contracts practices (not GPs) to provide generalist medical services in a specific geographical or population area. Each GP practice must enter into one of three types of NHS GP contracts to run NHS-commissioned general practice, which set the mandatory requirements and other provisions for the services to provide.

The first type is the General Medical Services (GMS), the national standard GP contract, whose terms are negotiated nationally every year between NHS England and the British Medical Association (the doctors' trade union). In 2018/19 around 70 per cent of GP practices operated under it. The second type corresponds to the Personal Medical Services (PMS), which unlike the GMS contract, is negotiated and agreed locally by CCGs or NHS England with a general practice or practices. This contract offers commissioners an alternative route with more flexibility to tailor requirements to local needs while also keeping within national guidelines and legislation. The PMS contract is being phased out, but in 2018/19, 26 per cent of practices held one. The third type is the Alternative Provider Medical Services (APMS) contract, which offers greater flexibility than the other two contract types. The APMS framework allows contracts with organisations (such as private companies or third sector providers) other than general

[5]The local NHS organisations responsible for the administration of primary care in their area

[6]See https://digital.nhs.uk/data-and-information/publications/statistical/general-practice-workforce-archive/final-31-march-2020.

practitioners/partnerships of GPs to provide primary care services. In 2018/19, 2 per cent of practices held this type of contract.

The core elements of the contracts agree the geographical area that a practice will cover, require the practice to maintain a list of patients for the area, establish the essential medical services the practice must provide, set standards for premises and workforce, and requirements for inspection and oversight, among others.

The funding of a general practice depends on lump-sum, quality incentive payments, and items of service. The lump-sum payments, also know as Global sum payments, represent about half the money a practice receives for delivering the core part of its contract. These payments are based on an estimate of a practice's patient workload and certain unavoidable costs, and are determined by a formula (the Carr-Hill formula) that takes as input the list size of the practice, patients' demographic mix, local morbidity measures, local market factors, rurality, among others. Quality incentives from the Quality and Outcomes Framework (QOF) scheme accounts for around 10 per cent of a practice's income. The QOF is a voluntary programme that practices can opt in to in order to receive payments based on good performance against a number of indicators. Although is voluntary, participation rates are high, and in 2018/19 more than 95 per cent of practices took part. The framework covers a range of clinical areas, for example, management of hypertension, cardiovascular disease, depression, and asthma; prescribing safety; vaccinating; and screening target population. Each area has a range of indicators that equate to a QOF score, which at the end of the financial year, is equated to an amount of money. It is important to emphasize that QOF is based on the use of resources, and not on the patients health outcomes. Additionally, if a practice is leasing its premises, rent is generally reimbursed in full in arrears, but have to fund all other expenses such as paying its workforce - including salaried GPs, nurses, health care assistants and administrative staff – from their revenue.[7]

Pricing of medicines and GP prescribing. In the UK, drug manufacturers are in principle able to set their own list prices. However, the total income of manufacturers is regulated by the voluntary Pharmaceutical Price Regulation Scheme (PPRS) – although voluntary, about 90% of manufacturers of patent protected medicines consent to its application. This scheme constrains what they can charge for their products, as it requires manufacturers to make quarterly rebate payments at pre-agreed levels, and makes them unable to exceed profitability limits set on both Return on Capital and Return on Sales measures. This has the effect of controlling prices, and tightens the circumstances under which prices can be increased.

[7]For further information on the contracts and funding of general practices see Santos et al. (2017), https://www.england.nhs.uk/contact-us/privacy-notice/how-we-use-your-information/our-services/primary-care-commissioning/ and https://www.kingsfund.org.uk/publications/gp-funding-and-contracts-explained.

The stability of the list price of patented medicines is a relevant aspect of the cost-effectiveness evaluation made by the National Institute for Health and Care Excellence (NICE). Indeed, a favourable recommendation in the form of Technology Appraisal, for example, is critical for the success of a drug, as the NHS is under legal obligation to reimburse NICEs' recommended medicines.

Prescription drugs, issued by GP practitioners, are dispensed by community pharmacies and subsequently reimbursed by the NHS, based on the reimbursement prices set in the Drug Tariff,[8], which for patented medicines is the manufacturers' set list price, and for generic products is based on a weighted average of the list prices from wholesalers and generic manufacturers producing an off-patent drug. Pharmacies often purchase stock directly from pharmaceutical companies and wholesalers at lower prices and benefit from the marginal difference between the purchase price and the reimbursement price set in the Drug Tariff.[9]

CCGs are responsible for setting for each GP practice within their organization a prescribing budget. However, this requirement, as well as other aspects that organise prescribing in general practice, are not legally nor economically binding, and have the effect of nudging more than restricting prescribing behaviour, while always allowing the use of their clinical judgement. The cost of what practices prescribed, and then pharmacies dispense, is charged against the prescribing budget. The GP practice contract prevents practices charging their patients, although dispensing practices (with an in-house dispensary), which operate mostly in rural areas with fewer pharmacies and correspond to roughly 1 in 8 practices in the NHS, may do. Practices may be instructed by their CCG to switch drug treatments of their patients to save money from the budget, if they feel is clinically appropriate. Also practices or practitioners considered to be engaging in excessive prescribing may be challenged by their CCG, and additionally a CCG may define list of drugs they wish to stop providing within their area, as the cost-effective status of medicines changes over time. Finally, CCGs may implement Prescribing incentive schemes to incentivise and reward GPs to change practice and improve quality and cost effectiveness in prescribing.

Although the order in which the essays have been presented in this introductory chapter corresponds to the chronological order in which they were conceived, the remaining of the dissertation is structured as follows: Chapter 2 presents "Mental health and the COVID-19 pandemic: evidence from antidepressant prescriptions in general practices in England", Chapter 3 presents "Price versus clinical guidelines in primary care *statin* prescribing: a retrospective cohort study and cost simulation model", and Chapter 4 presents "Prescribing behaviour and patient health outcomes: evidence from *statin* prescriptions in general practices in England". Finally, Chapter 5 concludes.

[8]The Drug Tariff outlines what will be paid to pharmacy contractors for NHS services provided either for reimbursement or for remuneration

[9]See https://uk.practicallaw.thomsonreuters.com/3-500-9763?transitionType=Default&contextData=(sc.Default) for further information on medicinal product regulation in the UK.

Chapter 2

Mental health and the COVID-19 pandemic: evidence from antidepressant prescriptions in general practices in England

Abstract

We use data of antidepressants prescribed by general practices to investigate the impact of COVID-19 on population mental health in England during the first year of the pandemic. The challenge in estimating such an effect is to isolate it from the simultaneous reduction in access to general practice health services – another consequence of the pandemic. To do this, we construct a counterfactual group inspired by the synthetic control method. By combining prescription data for other drug classes whose incidence was not directly affected by the pandemic but that nevertheless experienced a reduction in prescription levels due to the contraction in general practice services, we provide an estimate of the effect of the COVID-19 pandemic on depression and related diagnoses, net of the problems of access. Two placebo tests support the validity of the proposed method. Our results show an average increase of 152 daily doses of antidepressant medication per month per practice – the equivalent of treating five additional adult patients per practice during the first year of the pandemic. Compared to the national prevalence of depression (on average 828 patients per practice or 11.6 per cent of the population in 2019) this figure suggests a moderate impact of the pandemic on mental health, and a deviation from the public perception and findings of early studies of a sizeable negative impact. Furthermore, our analysis uncovers significant geographical heterogeneity across local authority districts in the estimated effect on antidepressants, which in turn appears associated with reduced mobility, greater number of unemployment benefit claimants and higher infection incidence of coronavirus.

JEL Codes: I10, I18, C5

Keywords: Mental health, Antidepressant, Prescription, Depression, General Practice, Artificial Counterfactual, COVID-19

2.1 Introduction

In this paper, we use antidepressant (AD) prescriptions data from general practitioner (GP) practices in England to investigate the impact of the COVID-19 pandemic on population mental health during the first year since its onset.

Mental health is the leading cause of disability, with one in four adults in England experiencing at least one diagnosable mental health problem in any given year.[1] Mental health conditions impose a significant societal loss,[2] with the total costs of mental ill-health for England reaching 119 billion per annum.[3] Depression is the most common mental health condition, which can lead to alcohol and drug abuse, as well as an increased risk of suicide, and consequently has significant implications for social and economic inequalities.

The severe impact of the COVID-19 pandemic on health and the economy is likely to have exacerbated the already mounting burden of mental health. Factors such as fear of infection, job insecurity, isolation and lack of social interactions may have triggered significant psychological issues and, in turn, increased the number of people turning to medical treatment. Recent studies using self-reported measures of mental illnesses (Etheridge and Spantig, 2020; Proto and Quintana-Domeque, 2021) confirm the widely-held belief as depicted in the media of a significant deterioration in population mental health.[4] However, studies using hard outcomes have found that prescriptions of AD and mental health episodes recorded in primary care decreased during the first months of the pandemic with respect to expected trends (Carr et al., 2021).

One explanation for these contradictory findings is that individuals with new mental health episodes may have opted not to seek primary care owing to fear of infection if they ventured outdoors. At the same time, patients may have been prevented from accessing primary care services because of restrictions on movements introduced during the first months of the pandemic, coupled with redeployment of resources towards the vaccination campaign later on (Williams et al., 2020). In short, while the pandemic is likely to have had a negative impact on mental health and well-being, it is conceivable that prescriptions of AD actually decreased because of a fall in the demand for or restrictions in the supply of primary care mental health services.[5] In the rest of the paper,

[1] See https://www.england.nhs.uk/wp-content/uploads/2016/02/Mental-Health-Taskforce-FYFV-final.pdf.

[2] As a result of increased health care utilization, social support cost and economic losses due to work incapacity.

[3] See https://www.centreformentalhealth.org.uk/publications/spending-review-wellbeing. For comparison, the total burden of cardio-vascular disease and cancer in the UK are estimated at 23.2bn and 7.6bn, respectively.

[4] See, e.g. https://on.ft.com/3hyJYvI.

[5] For example, from Carr et al. (2021): "[H]ealth services were required to balance infection control with access to care for patients and GPs were advised to minimise the number of face-to-face contacts" and that"public health messages encouraged patients to avoid attending general practices and hospitals to help control the virus".

we will refer to the net impact of demand and supply shocks to GP practice services as *problems of access*.

Starting from the realisation that the observed levels of actual prescriptions of the AD class during the pandemic period are the likely result of two effects working in opposite directions, i.e. a *mental health effect* which increases them and the problems of access which decreases them, this paper proposes a methodology inspired by the main features of synthetic control method (SCM),[6,7] to causally identify the effect of the COVID-19 pandemic on depression and related diagnoses at the population level. Namely, we construct an artificial counterfactual group for each GP practice in our sample that provides a reasonable measure of the contraction in AD prescriptions attributable to problems of access only. To achieve this we use prescription data for drugs in *therapeutic classes* other than AD medications,[8] provided they satisfy two requirements. The first is that the incidence of the conditions for which these drugs are prescribed for is only affected by problems of access and not by a direct effect of the pandemic on these conditions.[9] The second is that a combination of these other classes proves to satisfactorily reproduce, i.e. predict, the evolution of the AD class in a period immediately before the start of the pandemic. Subsequently, these drug classes and the parameters governing such combination, i.e. the fitted values, are then used to define a counterfactual group that is meant to reproduce how prescriptions of AD would have evolved during the pandemic period as if these were affected only by problems of access. Finally, we estimate the average treatment effect of the pandemic on mental health by computing the mean difference across GP practices between the actual AD prescription time series and the control group.

To check the validity of our approach we performed two placebo tests. First, a temporal-placebo test where we applied the same methodology to a time frame prior to the start of the pandemic. In principle, where no major social event that would have affected the prescription of drugs in England, the methodology should find no significant differences between the actual series of AD and the constructed counterfactual group. Second, a therapeutic class-placebo test was conducted where we kept the original study

[6]Essentially, the SCM is based on the idea that a combination of untreated units (the donor pool) can provide a better comparison for a unit exposed to the treatment. This combination is generated by finding weights using pre-treatment data only such that, if constructed successfully, in the pre-treatment period should reproduce almost exactly the outcome variable of the treated unit, and in the treatment period serves to outline what would have happened to the treated unit had the treatment never occurred. For seminal works, see Abadie and Gardeazabal (2003); Abadie et al. (2010, 2015).

[7]In Appendix 2.7 we explain why the SCM, in its original formulation, is not feasible to implement.

[8]A therapeutic or drug class is defined as a set of medications and other active substances that have a similar chemical structures, the same mechanism of action (i.e. binding to the same biological target), a related mode of action, and/or are used to treat the same disease. In this paper, we rely on the latter definition. Explained in more detail in Section 2.4.

[9]For example, as noted in the report 'Direct and Indirect health impacts of COVID-19 in England', "[s]elf-isolation, home working and social distancing reduce the spread of infectious diseases, as well as [] the spread of COVID-19". Then, there was a direct effect or change in the underlying health need for antibacterial and/or antiviral treatment during the initial period of the pandemic. Accordingly, antibacterial and antiviral drugs would not satisfy this first requirement. Explained in more detail in Section 3.2.

time frame and applied the same methodology to each of the therapeutic areas used to predict AD. If the effect of the pandemic on psychological illnesses was indeed more intense than for physiological conditions treated by these other drug classes, then the estimated effect for AD should be larger and more extreme relative to the distribution of the placebo effects. We show that both placebo tests support the credibility of our results.

In a further analysis, we investigated geographical variation, at the local authority district level, in the effect of interest and its correlation with three relevant social, economic and public health circumstances in the context of the pandemic, namely, reduced mobility, increased share of unemployment benefit claimants and of infection incidence, i.e. number of new COVID-19 cases.[10] The choice of these three variables was motivated by the following considerations. First, during the initial twelve months of the pandemic, England entered into three national lock-downs (on 23rd March 2020, 5th November 2020, and 6th January 2021) with differing duration and levels of compliance. By reducing mobility and social interactions significantly these policy measures may have further negatively affected mental well-being. Second, the implementation of lock-downs resulted in closure of non-essential businesses and disruptions to economic activities, which caused an increase in unemployment, however moderated by furlough schemes putted in place, and thus in the number of people who experienced economic insecurity and hardship. Since unemployment data at the district-month level are not available, we proxied it by the number of unemployment benefit claimants. However, this variable can act both as an indicator of hardship and as a measure of the benefit of a social security net in the face of hardship.[11] Hence, the sign of the effect becomes an empirical matter. Finally, the higher the number of cases in an area, the greater the fear of being infected which arguably can adversely affect mental health.

Using administrative prescription data for main therapeutic classes prescribed by 95 per cent of all GP practices in England (around 6.6 thousand), we estimate an average

[10] Relatedly, see Brooks et al. (2020) for a literature review on the psychological impacts of quarantine, based on studies across ten countries including people with SARS, Ebola, the 2009 and 2010 H1N1 influenza pandemic, Middle East respiratory syndrome, and equine influenza. The study highlights that stressors of poorer mental health include duration of the quarantine, frustration and boredom, fears of infection, financial loss, among others.

[11] During the first 18 months of the pandemic, the UK government implemented the Coronavirus Job Retention Scheme (CJRS) designed to support employers to retain and continue to pay staff while businesses were closed. The scheme appears to have had a certain degree of success given that in April 2020, the Office for Budget Responsibility predicted that unemployment would peak at 10% in 2020 when it actually peaked at 5.2% (see https://commonslibrary.parliament.uk/examining-the-end-of-the-furlough-scheme/). Despite this, some risk to the labour market have been identified, such as the potential job losses when the scheme ends, lower re-employment rates, and increased number of redundancies, which in 2020 were grater than in the previous year (see https://ifs.org.uk/publications/15644). Most likely this public policy had the effect of increasing less the level of unemployment that would have occurred otherwise, yet the increased risk in economic insecurity and hardship still persists, and accordingly, its relation with mental health is of interest. In this research we attempt to capture this consequential aspect of the pandemic by using data strongly correlated with unemployment, such as of those individuals claiming unemployment benefits.

increase of 152 *daily doses*[12] of AD per month across practices during the 12-month period since the start of the COVID-19 pandemic in England in February 2020.[13] Such increase is equivalent to an average of five more adult patients per practice treated every month for depression during the first year of the pandemic. This suggests a relatively moderate impact of the pandemic on mental health when compared to the national prevalence of depression in 2019, estimated to be an average of 828 patients per practice (around 11.6 per cent of the registered adult population). Furthermore, our analysis uncovers significant heterogeneity across the country, and we find that prescriptions of AD in GP practices increased more in districts where the number of new infection cases and time spent at residential places were greater, thus exhibiting the expected positive correlation between poorer mental health outcomes and fears of infection and isolation. However, we find a lesser increase in AD prescriptions in districts with a higher share of unemployment benefit claimants, potentially highlighting the benefits of the social safety net on mental health.

The rest of the paper is structured as follows. Section 2.2 discusses relevant literature and provides a general overview of depression and AD prescriptions in general practice in England. Section 3.2 presents the proposed econometric methodology and falsification tests. Section 2.4 introduces the data and the variable definitions used in the analysis. Section 2.5 presents the results, and Section 2.6 discusses the results and concludes.

2.2 Background

2.2.1 Related literature

A report from the Office of National Statistics, using the self-reported Patient Health Questionnaire (PHQ-8) variable from the Opinions and Lifestyle Survey, examines depression in adults during January to March 2021 and compares it to pre-pandemic estimates.[14] A two percentage points (from 19% to 21%) increase in the prevalence of depression is found between early 2021 and November 2020 and more than 10 percentage points compared to pre-pandemic levels. Younger adults, women, disabled, clinically extremely vulnerable, those renting and those living in the most deprived areas of England are more likely to experience some form of depression.

[12]The (defined) daily dose is a standard statistical measure of drug consumption, defined by the World Health Organization (WHO), and it enables comparison of drug usage between different drugs in the same class or between different health care environments, or to look at trends in drug utilisation over time. See https://www.whocc.no/ddd/definition_and_general_considera/.

[13]This given that the first case of COVID-19 reported in England occurred in late January 2020. See Moss et al. (2020).

[14]See https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/articles/coronavirusanddepressioninadultsgreatbritain/januarytomarch2021.

Similar findings are reported by a number of studies using the self-reported Generalized Health Questionnaire (GHQ-12) data from the UK Household Longitudinal Study. Etheridge and Spantig (2020) document an overall decline in mental well-being, as well as gender disparities explained by family, time use and caring. Women, the young and those with high childcare duties report larger drops. Proto and Quintana-Domeque (2021) further decomposing the effect find that the gender gap in mental health increases only among White British individuals but not other ethnicities. Banks and Xu (2020) confirm substantial decline in mental health with significantly bigger effects for young adults and women, as do Daly et al. (2020). Cheng et al. (2021) find that declines in mental well-being are worse for working parents who manage competing time demands between work and home (i.e. childcare and home schooling) and are mainly driven by increased financial insecurity, while the burden is not shared equally between men and women, nor between richer and poorer households.

However, studies relying on survey or longitudinal data and employing self-reported health outcomes face a number of limitations. Analyses largely rely on fixed effects estimations lacking causal identification in the absence of exogenous variation and in the presence of confounding factors and omitted variables. Similarly, self-reported health questions are prone to measurement errors (Bound et al., 1991) and reporting heterogeneity (Butler et al., 1987) also biasing estimates. Finally, given the survey nature of the data, deteriorating mental health may be a contributing factor to respondents' attrition or selection leading to further biases of any estimated effects (Fitzgerald et al., 1998).

Moving away from self-reported outcomes and using primary care electronic health records from GP practices, Carr et al. (2021) study trends in primary care-recorded mental illness, self-harm episodes and psychotropic medication in the UK, between April and September 2020. They find initial (up to May) sharp reductions in the incidence of primary care-recorded depression, anxiety disorders and first antidepressant prescribing in English GPs, with subsequent increases in the period up to June and by September figures returned back to their expected levels. However, failing to account for the well-documented problem of access to health care services since the start of the pandemic biases any before-after comparison of trends.

This paper addresses the methodological challenges of the past literature and contributes causal estimates of the consequences of the pandemic on mental health. Contrary to most existing studies who zoom in on the initial phase of the pandemic, we offer a longer term perspective on the issue of mental illness. We focus on measures not previously considered in this strain of the economic literature and away from self-reported scales of subjective well-being,[15] which mostly measure symptoms rather than

[15]Various studies have utilized self-reported scales like the Patient Health Questionnaire (PHQ-8), 12-item Generalized Health Questionnaire (GHQ-12), Center for Epidemiologic Studies Depression Scale

actual diagnoses.[16] Our health outcomes measures reflect not only an individual perspective but also how a health care system internalises the underlying health changes of the population during the pandemic context. At the same time we recognize and control for problems of access, either due to demand- or supply-side factors, that would affect all variables that capture the points of contact between patients and health care services.

2.2.2 Context

Depression in general practice. The clinical identification and recognition of depression typically starts in primary care (general practice) and depending on the severity, the management and treatment includes psychological or psychosocial interventions, drug therapy, and referrals to specialized mental health (secondary) care services. Regarding drug treatment, the current national guidance suggests that AD should be considered for those people with a history of moderate to severe depression; for those with fewer depressive symptoms present for a period of typically 2 years; with mild to moderate depression that persist after other interventions have not been beneficial; and those with moderate and severe depression. However, they should not be routinely used to treat mild depression. We note also that AD are not only the effective drug treatment for moderate to severe depression, but are also used for other mental health illnesses such as chronic anxiety, generalised anxiety disorder (a form of chronic anxiety), panic disorder, obsessive-compulsive disorder, post-traumatic stress disorder, and phobic states such as social anxiety disorder. So, in the context of COVID-19, overall prescriptions of AD can provide a broad picture of the impact of the pandemic on mental health, as internalized by primary care.

The Quality Outcome Framework report for the 12-month period between April 2019 and March 2020 (the year before the start of the COVID-19 pandemic) indicates that depression is the second most prevalent condition treated in general practice, as on average 11.6% of patients have currently this diagnosis.[17] Table 2.1 provides characteristics of depression as recorded in GP. A growing trend over the past few years is present for prevalence (i.e. number of patients with a current diagnosis of depression),

(CES-D), Mental Health Inventory 5 (MHI-5), the WHO-5 index, and other ad-hoc questionnaires in surveys (see, e.g. Etheridge and Spantig (2020), Banks and Xu (2020), Cheng et al. (2021), Davillas and Jones (2020), Proto and Quintana-Domeque (2021), Daly et al. (2020), Giuntella et al. (2021), Lu et al. (2021), Siflinger et al. (2021), Adams-Prassl et al. (2020)). Others have used measures of revealed mental health problems, e.g. helpline calls related to domestic and sexual violence, economic insecurity, preganancy and abortion, and mental health (see, e.g. Armbruster and Klotzbücher (2020), Leslie and Wilson (2020), Silverio-Murillo et al. (2021)), and internet search on terms related to death, suicide, anxiety, depression, loneliness, etc. (see, e.g. Brodeur et al. (2021); Tubadji et al. (2020)).

[16] Although both symptoms and diagnoses are important, the difference between them is relevant in the pandemic context. See Taquet et al. (2021).

[17] The first most prevalent condition is hypertension (14.1%) and the third is obesity (10.5%). See https://digital.nhs.uk/data-and-information/publications/statistical/quality-and-outcomes-framework-achievement-prevalence-and-exceptions-data/2019-20.

incidence (i.e. number of patients with a new diagnosis of depression) and the number of daily doses of AD prescribed (columns 3, 4 and 5, respectively).

TABLE 2.1: Mean characteristics of depression in general practice in England

Year	(1) GP practices (#)	(2) List Size (#)	(3) Prevalence (#)	(4) Incidence (#)	(5) AD (daily dose)	(6) Prevalence *wrt* List Size (%)	(7) Incidence *wrt* Prevalence (%)	(8) AD per patient per month (daily dose)
2016-17	7,392	6,235.8	566.5	73.0		8.9	12.7	
2017-18	7,100	6,540.7	646.4	79.4	291,773	9.8	12.1	41.2
2018-19	6,873	6,876.8	738.7	89.4	320,488	10.7	11.9	39.6
2019-20	6,720	7,164.7	828.2	86.4	357,144	11.6	10.3	43.1

NOTES. The table presents mean characteristics of depression across GP practices. 'Year' column corresponds to the QOF reporting period (12-month period from 1 April to 31 March). Column (1) shows the number of practices. Column (2) the adult list size (number of patients +18 registered at the practice). Column (3) the average prevalence of depression. Column (4) the average incidence of depression. Column (5) the average number of daily doses of antidepressants prescribed in the period. Column (6) is obtained by dividing column (3) by (1) times 100%. Column (7) is obtained by dividing column (4) by (3) times 100%. Column (8) represents the number of number of daily doses of AD per depressive patient per month; it is obtained by diving number of AD (column (5)) by prevalence (column (3)), and then divided by 12.
SOURCES. Quality Outcome Framework (QOF) series and English Prescribing Dataset (EPD).

Antidepressants before and during the pandemic. Figure 2.1 shows the average monthly trend of AD before and after February 2020 across GP practices. Pre-pandemic these prescriptions were growing at a rate of 235 daily doses per month. During the pandemic we observe a significant decrease with respect to previous levels, of 70 daily doses per month fewer than before. Taken at face value this fact could imply that depression and poor mental health outcomes have decreased since the start of the pandemic. However, for reasons delineated previously, the observed decrease in the slope of AD may not represent a lack of effect of the pandemic on mental health prescriptions but may reflect instead that the effect the pandemic had on access to GP services dominates over its effect on mental health.[18]

Preliminary analysis of GP practice prescription data indicates that the absolute decrease in the AD trend during the pandemic is lesser than for other large therapeutic areas such as lipid-regulating, hypertension, antisecretory, and antianginal drugs. In relative terms (that is, reduction in the post-trend as a proportion of the pre-trend) AD have experienced one of the lowest decreases, only superseded by drugs used for psychosis, rheumatic disease, and nausea & vertigo.[19] This may suggest that, net of the

[18]The report 'Direct and Indirect health impacts of COVID-19 in England' published in September 2021 acknowledges the indirect impact of the pandemic on accessing health care services in all settings (primary care, secodary specialised care, and hospital care – elective, urgent, etc.). For example, they show that in primary care, consulations fell significantly after the start of the pandemic and only fully recovered by May 2021. This drop is explained by both demand-side and supply-side factors. On the former, these correspond to *changes in underlying health need* as a result of voluntary or mandatory behaviour changes (e.g. fewer road traffic accidents during the pandemic, may have led to a decline in the demand for trauma care) and *changes in health-seeking behaviour* (e.g. people avoiding making appointments to reduce the burden on the NHS or avoiding health care facilities for fear of catching the virus). And on the latter, these factors correspond to *changes in the provision of healthcare* to respond to COVID-19. These health system adaptations were put in place to minimise the spread of COVID-19 and for the reallocation of resources to manage urgent care of COVID-19 patients. In many healthcare settings, interventions were postponed, conducted online, or cancelled to enable providers to focus on urgent care needs at peaks during the pandemic. See DHSC and ONS (2021).

[19]See Appendix Table 2.6 for these results, and Appendix Figure 2.7 for the trends before and during the pandemic for selected therapeutic areas prescribed in general practice.

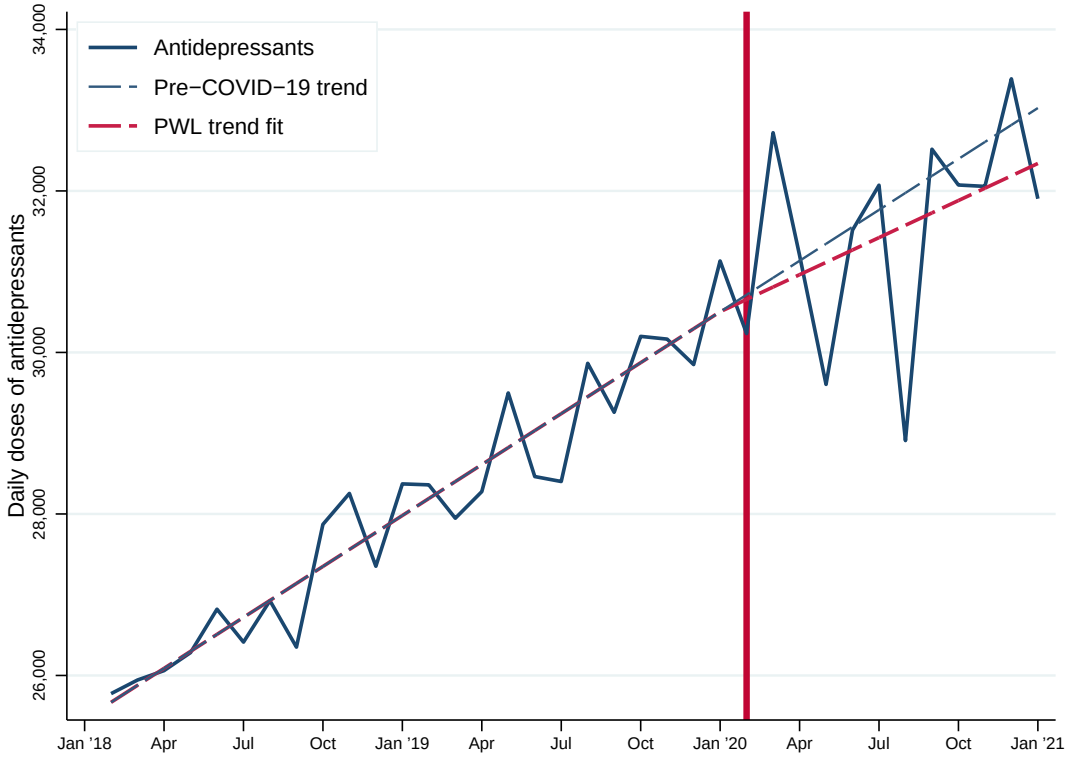


FIGURE 2.1: Antidepressants in general practice before and after COVID-19

NOTES. The figure shows monthly antidepressants prescriptions levels averaged across GP practices in England between February 2018 and January 2021. y-axis represents the number of daily doses of AD and x-axis months. Thick vertical red line indicates February 2020 which marks the start of the pandemic. The figure also shows two fitted trends using the following piece-wise linear (PWL) regression model, that allows to model shifts in trends for a given breakpoint: $Y_{gt}^{AD} = \alpha_0 + \beta_1 t_1 + \beta_2 t_2 + \epsilon_{gt}$, where Y_{gt}^{AD} denotes daily doses of AD prescribed by GP practice g in year-month t. The time regressor t_1 runs from 1 to 36 for the 36 months between February 2018 and January 2021 (inclusive), and t_2 takes the value of 0 for the months prior to February 2020, and from then onwards it runs from 1 to 12 up to January 2021. The coefficient β_1 represents the monthly change in AD before the break, and β_2 represents the change in β_1 after the break. The blue dashed line corresponds to the pre-trend, then projected after the break, i.e. $\hat{Y}_{gt}^{AD} = \hat{\alpha}_0 + \hat{\beta}_1 t_1$. The red dashed line correspond to the full fitting of the model, i.e. $\hat{Y}_{gt}^{AD} = \hat{\alpha}_0 + \hat{\beta}_1 t_1 + \hat{\beta}_2 t_2$, that allows to model a shift after February 2020.
SOURCE. English Prescribing Dataset (EPD).

decrease in prescription levels as consequence of problems of access, AD may have been subject to an additional positive shock that would have increased these prescriptions. As a corollary other drug classes, not subject to a direct effect of the pandemic on the underlying health condition for which these drugs are prescribed for but still subject to access shocks, can be used to define a valid counterfactual for AD.

The dynamics of the prescription series of AD after February 2020 shown in Figure 2.1 deserve some further discussion. Immediately after the pandemic, in March, there is a significant upward spike. This can hardly be explained by the sudden and supposedly relevant impact of the pandemic on mental health. Instead, anticipatory behaviour from general practitioners and practices treating patients under repeated prescription may be in place. That is, in anticipation of the disruptions that the recently declared

pandemic might have on public health and on accessing the health service, general practitioners considered that repeated prescriptions programmed to be issued in the future, given the contingency, could be better placed earlier and as soon as possible, to avoid patients going either to the pharmacy, to get their next round of prescriptions, or to follow-up appointments, for an assessment of their current drug treatment. Another notable data point corresponds to the downward spike in August, which is explained by the fact that this month is a time in the middle of the summer, in which appointments and therefore prescriptions go down. These two abnormal episodes are not exclusive to the AD class, as when looking at other main therapeutic groups prescribed in general practice these spikes are present as well (see Appendix Figure 2.7).[20]

2.3 Methodology

2.3.1 Construction of counterfactual

Theoretical approach. Let Y_{gt}^{AD} denote the number of daily doses of the AD class of drugs prescribed by GP practice $g = 1, ..., G$ in month-year $t = 1, ..., T$. Let $T_0 + 1$ to indicate February 2020 marking the start of the novel coronavirus pandemic in England. Then, the pre-COVID-19 (or pre-treatment) period is given by $t \in \{1, 2, ..., T_0\}$, and the COVID-19 (or treatment) period given by $t \in \{T_0 + 1, T_0 + 2, ..., T\}$. Next, we assume that the monthly daily doses of AD can be described by the following structure:

$$Y_{gt}^{AD} = \begin{cases} \tau_{gt}^{AD} + \varepsilon_{gt}^{AD} & \text{if } t \leq T_0 \text{ (\textbf{Pre-treatment period})} \\ \tau_{gt}^{AD} + [\mu_{gt}^{AD} + \chi_{gt}^{AD} + \varepsilon_{gt}^{AD}] & \text{if } t > T_0 \text{ (\textbf{Treatment period}),} \end{cases} \tag{2.1}$$

where τ_{gt}^{AD} represents a linear deterministic trend component and ε_{gt}^{AD} a white-noise error term. Consequently, in the pre-treatment period, Y_{gt}^{AD} is defined as a trend-stationary process. As a consequence of the pandemic, in the treatment period, AD prescriptions are affected by two additional different shocks, indistinguishable in themselves. The first, denoted by μ_{gt}^{AD}, represents the direct effect on mental health and related diagnoses, which would increase the prescriptions of AD (i.e. $E[\mu_{gt}^{AD}] \geq 0$). And the second, denoted by χ_{gt}^{AD}, represents the effect on access to GP services, which would decrease them (i.e. $E[\chi_{gt}^{AD}] \leq 0$). Hence, the diminished upward trend of AD (i.e. a decrease in incidence with respect to the expected levels) revealed by before-after comparison does not necessarily mean that the pandemic did not have any significant impact on mental health. Instead, it could be that the effect of the pandemic on deterring demand for or restricting access to GP services dominates over the effect on mental health, i.e. $|E[\chi_{gt}^{AD}]| > |E[\mu_{gt}^{AD}]|$.

[20] A discussion on how these dynamics might be driving our results is presented in the corresponding Section 2.5.

In order to isolate the mental health effect we would ideally need a counterfactual group showing the same evolution as Y_{gt}^{AD} in the pre-treatment period and that can emulate prescriptions levels in the treatment period as if these were affected only by problems of access, absent of a direct effect on diagnoses treated with AD. That is:

$$C_{gt}^{AD} = \begin{cases} \tau_{gt}^{AD} + \varepsilon_{gt}^{AD} & \text{if } t \leq T_0 \\ \tau_{gt}^{AD} + [\chi_{gt}^{AD} + \varepsilon_{gt}^{AD}] & \text{if } t > T_0. \end{cases} \tag{2.2}$$

The difference between the actual observed level of AD prescriptions and the counterfactual, $Y_{gt}^{AD} - C_{gt}^{AD}$, is equal to μ_{gt}^{AD}, the effect of COVID-19 on AD' prescriptions and the main interest of our research.

Given this theoretical framework we now turn to the implementation of the empirical strategy that involves the construction of an artificial counterfactual group for AD that can approximate the behaviour in eq. (2.2).

Empirical approach. Our empirical strategy consists of finding and combining a set of therapeutic classes other than AD that satisfy two requirements. First, that the incidence of the diseases for which these other classes are prescribed for, were only affected by problems of access and not by a direct effect of the pandemic resulting in changes to the underlying health needs; second, that a linear combination of these other classes can proves to satisfactorily reproduce (predict) the prescriptions of AD in the pre-treatment period. Having met these two requirements, the counterfactual group would provide an estimate of AD prescriptions as if these were affected only by problems of access.

Similarly to eq. (2.1), let Y_{gt}^{j}, denoting the number of daily doses of drugs in therapeutic class $j \in J$ prescribed by practice g in month t, to be described by the following structure:

$$Y_{gt}^{j} = \begin{cases} \tau_{gt}^{j} + \varepsilon_{gt}^{j} & \text{if } t \leq T_0 \\ \tau_{gt}^{j} + [\chi_{gt}^{j} + \varepsilon_{gt}^{j}] & \text{if } t > T_0, \end{cases} \tag{2.3}$$

where the terms τ_{gt}^{j}, ε_{gt}^{j} and χ_{gt}^{j} represent the linear deterministic trend, white noise error term and the pandemic's effect on access to prescriptions services, respectively (specific to class j and practice g).[21] Consistent with the first requirement of no direct effect we assume that $\mu_{gt}^{j} = 0 \; \forall j$.

To define the set J, we start from an initial subset of the 24 main therapeutic classes (including AD) in general practice representing altogether 98% of the total volume of drugs prescribed in 2019.[22] The selection of the J therapeutic classes used to generate

[21] Note that, just as with Y_{gt}^{AD}, Y_{gt}^{j} for each j in the pre-treatment period is defined as a trend-stationary processes.

[22] A list of these 24 main therapeutic areas prescribed in general practice is presented in Table 2.7 in the Appendix.

our control group begins by excluding those that do not meet the first requirement set above. Ten are excluded for the following reasons. First, we exclude antibacterials and antihistamines. In the context of the pandemic the population as a whole were urged to increase personal hygiene (e.g. constantly washing of hands, use face-covering masks, etc.) and to stay in-doors. This might have prevented a greater number of infections and allergy cases, compared to normal expected levels, as personal hygiene improved and exposure to the outdoors decreased. Next, we exclude analgesics. This class includes drugs used for palliative and end-of-life care. The pandemic brought not just COVID-19-related deaths but also excess of deaths from other causes. As a results, it is very likely this drug class has been prescribed more than expected. Also we exclude bronchodilators and corticosteroids (for respiratory diseases). Given that COVID-19 affects the respiratory system these drugs might have been used to treat suspected cases. Indeed, prescription data for these classes shows above-normal levels in March and April 2020. There is another reason why this classes should be excluded. Reductions to traffic lead to fewer road accidents as well as reduced air-pollution, which has been found to decrease the incidence of asthma and, therefore, the use of bronchodilators. Finally, we exclude drugs used for more acute conditions, such as laxatives and drugs used for nausea and vertigo, as well as other drugs treatments for mental health problems, such as hypnotics and anxiolytics, and for psychoses and related disorders.

The final set of 14 therapeutic areas is presented in Table 2.9 and their series shown in Figure 2.7, both in the Appendix. It includes the largest therapeutic areas used for long-term or chronic conditions such as cardiovascular diseases, hypertension, gastrointestinal tract disorders and osteoporosis. Evidence from the report 'Direct and Indirect health impacts of COVID-19 in England' indicates that the incidence of chronic conditions in the short-run was mostly affected by changes in health-seeking behaviour and health care system adaptations, and less so due to changes in the underlying health needs (as is the case of mental health and depression, for example), while the management of already diagnosed long-term conditions appears less negatively affected. In line with expectations following our first requirement, this set of classes represent highly prevalent conditions, imply long-term pharmacological treatment (similar to depression and AD treatment) and most importantly that the reduced incidence during the pandemic is mainly attributed to problems of access and not related to a direct effect of the pandemic on the conditions for which they are prescribed for.

Having defined the classes available in generating the counterfactual group we now discuss how to operationalise the second condition (i.e. to predict AD prescriptions in the pre-treatment period), which is done in three steps.

The **first step** in the construction of the counterfactual for AD is to de-trend the actual prescription series of all therapeutic classes, including AD, for each GP practice. As the construction of our control group involves time-series regressions, to avoid spurious

results we require stationary data, which will be rendered by removing the trend (de-trending) from the trend-stationary processes assumed throughout.[23] We do this first by estimating a linear trend term, using pre-treatment period data only, and subtracting it from the actual prescription series in every period. This rests on the assumption that the same trend would have been observed in 2020, in the absence of the pandemic. Specifically, we estimate the following model for each therapeutic class prescribed by GP $g = 1, ..., G$:

$$Y^j_{gt} = \tau^j_{gt} + \varepsilon^j_{gt} = (\alpha_{jg} + \beta_{jg} \cdot t) + \varepsilon^j_{gt}, \tag{2.4}$$

for the pre-treatment period, i.e. for $t \leq T_0$. Subtracting the estimated linear trend, $\hat{\tau}^j_{gt} = \hat{\alpha}_{jg} + \hat{\beta}_{jg} \cdot t$, from each actual prescription series in every period results in the following estimated de-trended series, denoted by $\tilde{y}^{AD}_{gt}$ and $\tilde{y}^j_{gt}$:

$$\tilde{y}^{AD}_{gt} = Y^{AD}_{gt} - \hat{\tau}^{AD}_t = \begin{cases} \hat{\varepsilon}^{AD}_{gt} & \text{if } t \leq T_0 \\ \widehat{[\mu^{AD}_{gt} + \chi^{AD}_{gt} + \varepsilon^{AD}_{gt}]} & \text{if } t > T_0, \end{cases} \tag{2.5}$$

$$\tilde{y}^j_{gt} = Y^j_{gt} - \hat{\tau}^j_{gt} = \begin{cases} \hat{\varepsilon}^j_{gt} & \text{if } t \leq T_0 \\ \widehat{[\chi^j_{gt} + \varepsilon^j_{gt}]} & \text{if } t > T_0, \end{cases} \tag{2.6}$$

for AD, in eq. (2.5), and for drug class $j \in J$, in eq. (2.6), respectively. Through the trend-stationary assumptions made for Y^{AD}_{gt} and $\{Y^j_{gt}\}_{j \in J}$, the resulting residuals $\hat{\varepsilon}^{AD}_{gt}$ and $\{\hat{\varepsilon}^j_{gt}\}_{j \in J}$ (or estimated shocks) for $t \leq T_0$ (pre-treatment period) are now a stationary processes. In Section 2.5, results from unit-root tests are presented to support this assumption.

The **second step** of the procedure involves the estimation of GP-specific linear regressions between the residuals of AD $\hat{\varepsilon}^{AD}_{gt}$ *on* the J residuals of other therapeutic areas $\{\hat{\varepsilon}^j_{gt}\}_{j \in J}$ in the pre-treatment period to predict the shocks of AD. It is important to emphasize that this step does not require a causal interpretation of the coefficients associated with the J other therapeutic classes, and instead is performed to capture a statistical relation between different prescribing tasks that take place in any given practice in the pre-treatment period. Such statistical relation can be rationalised as the existence of common shocks that affect all prescriptions. For example, a practice may have decide not to open for a certain number of days, or practitioners may be on leave, delaying appointments, diagnosis and treatments; months with greater number of working days are more likely to result in larger appointments and prescriptions volumes; patients with comorbidities (that is, more than one disease or condition is present in the same person at the same time) can be given drug treatment (and therefore be prescribed) for

[23] Note that this transformation is consistent with the theoretical framework above, in which, by removing the trend term τ^{AD}_{gt} to both eqs. (2.1) and (2.2), their difference still results in μ^{AD}_{gt}, the effect of interest.

several conditions on the same appointment. Specifically, we estimate the following linear model for each GP $g = 1, ..., G$ in the pre-treatment period:

$$\tilde{y}_{gt}^{AD} = \sum_{j \in J} \omega_{jg} \cdot \tilde{y}_{gt}^{j} + v_{gt}, \text{ for } t \leq T_0$$

$$\Rightarrow \hat{\varepsilon}_{gt}^{AD} = \sum_{j \in J} \omega_{jg} \cdot \hat{\varepsilon}_{gt}^{j} + v_{gt}. \tag{2.7}$$

The estimated coefficients $\{\hat{\omega}_{jg}\}_{j \in J}$, specific to each practice g, represent the above mentioned statistical relation and the weights placed upon the shocks of the J therapeutic classes combined to reproduce the AD series in the pre-treatment period.

In the **third step**, we define the counterfactual group at the GP g-level, denoted by κ_{gt}^{AD}, as the fitted values of the model in eq. (2.7) predicted in- and out-of-sample:

$$\kappa_{gt}^{AD} = \begin{cases} \sum_j \hat{\omega}_{jg} \cdot \hat{\varepsilon}_{gt}^{j} & \text{if } t \leq T_0 \\ \sum_j \hat{\omega}_{jg} \cdot \widehat{[\chi_{gt}^{j} + \varepsilon_{gt}^{j}]} & \text{if } t > T_0. \end{cases} \tag{2.8}$$

Eq. (2.8) makes clear that in the treatment period the effects of access to GP services, specific to each therapeutic class j, are combined using the estimated coefficient that reflect the statistical relationship between AD and each of the J therapeutic classes in the pre-treatment period. If the cyclical changes to AD in the pre-treatment period can be satisfactorily predicted by a combination of the shocks of the J clases, with weights given by coefficients $\hat{\omega}_{jg}$, then we can apply these same coefficients to the de-trended series of these therapeutic areas in the treatment period to construct conterfactual group for AD, containing only the problems of access to primary care services – just as with the drugs used to construct it, i.e. $\chi_{gt}^{AD} \sim \sum_j \omega_{jg} \cdot \chi_{gt}^{j}$.

Subsequently, we take the difference between the actual residuals of AD and the constructed counterfactual, i.e. the difference between eqs. (2.5) and (2.8):

$$\tilde{y}_{gt}^{AD} - \kappa_{gt}^{AD} = \begin{cases} \hat{\varepsilon}_{gt}^{AD} - \sum_j \hat{\omega}_{jg} \cdot \hat{\varepsilon}_{gt}^{j} & \text{if } t \leq T_0 \\ \widehat{[\mu_{gt}^{AD} + \chi_{gt}^{AD} + \varepsilon_{gt}^{AD}]} - \sum_j \hat{\omega}_{jg} \cdot \widehat{[\chi_{gt}^{j} + \varepsilon_{gt}^{j}]} & \text{if } t > T_0. \end{cases}$$

For the pre-treatment period ($t \leq T_0$) provided that the fit of eq. 2.7 is sufficiently good we have that $\hat{\varepsilon}_{gt}^{AD} - \sum_j \hat{\omega}_{jg} \cdot \hat{\varepsilon}_{gt}^{j} \approx 0$. For the treatment period ($t > T_0$), we obtain:

$$\tilde{y}_{gt}^{AD} - \kappa_{gt}^{AD} = \widehat{[\mu_{gt}^{AD} + \chi_{gt}^{AD} + \varepsilon_{gt}^{AD}]} - \sum_j \hat{\omega}_{jg} \cdot \widehat{[\chi_{gt}^{j} + \varepsilon_{gt}^{j}]},$$

and again using eq. (2.7)

$$= \hat{\mu}_{gt}^{AD} + \hat{\chi}_{gt}^{AD} - \sum_j \hat{\omega}_{jg} \cdot \hat{\chi}_{gt}^{j}.$$

Then $\tilde{y}_{gt}^{AD} - \kappa_{gt}^{AD} = \hat{\mu}_{gt}^{AD}$ under the fundamental assumption that $\hat{\chi}_{gt}^{AD} = \sum_j \hat{\omega}_{jg} \cdot \hat{\chi}_{gt}^{j}$ for each practice g. That is, we assume that the access shock to AD – which cannot be observed in itself – is equal to the weighted sum of the J class-specific shocks, with weights $\hat{\omega}_{jg}$ estimated using pre-treatment information only.

To measure the average treatment effect of the pandemic on AD prescribing in general practice, we estimate the following difference-in-difference model:

$$\Delta Y_{gt}^{AD} = \gamma_0 + \mu \text{COVID-19}_t + \psi_{gt}, \tag{2.9}$$

where ΔY_{gt}^{AD} is defined as the difference between the de-trended residuals of AD and its counterfactual, i.e. $\Delta Y_{gt}^{AD} \equiv \tilde{y}_{gt}^{AD} - \kappa_{gt}^{AD}$, COVID-19$_t$ is a dummy variable that takes a value of 1 for $t > T_0$, and of 0 otherwise, and ψ_{gt} is the error term. Coefficient γ_0 represents the difference between the shock to AD and the counterfactual in the pre-treatment period. μ is the main coefficient of interest representing the average monthly change in daily doses of the AD class prescribed across GP practices relative to the counterfactual – net of the problems of access – for the 12-month period since the start of the pandemic (February 2020), that is:

$$\mu = E[\Delta Y_{gt}^{AD} \mid \text{COVID-19}_t = 1] - E[\Delta Y_{gt}^{AD} \mid \text{COVID-19}_t = 0].$$

We also estimate the model in eq. (2.9) using as dependent variable the number of depressive adult patients to which the number of daily doses is equivalent to, denoted by P_{gt}^{AD}. Furthermore, we use a normalisation of both daily doses and number of patients to account for differences in GP practices' adult population size (i.e. the registered adult population), denoted by nY_{gt}^{AD} and nP_{gt}^{AD}, respectively. These latter two variables represent our preferred measures of the use of AD and associated number of users. See Appendix 2.7 for the details on the construction of these variables.

2.3.2 Placebo tests

The soundness of our methodology is examined through two placebo (or falsification) tests, which provide an alternative mode of inference. This alternative mode of inference is based on the premise that our confidence that a particular artificial control estimate reflects the impact of the intervention under scrutiny would be severely undermined if we obtained estimated effects of similar or even greater magnitudes in cases where the intervention did not take place.[24]

First, we conduct a temporal–placebo test where the treatment period is reassigned to a time frame before the pandemic actually took place. If the proposed methodology is applied to past periods were no major event has occurred and therefore prescriptions

[24]See Abadie et al. (2015), p. 499.

of all therapeutic areas have evolved according to their expected trends, then the actual series of AD should in principle be not significantly different from the constructed counterfactual group. To conduct this placebo test we assume a pre-treatment period of January 2017–December 2018 and a treatment period of January 2019–December 2020.

Second, we conduct a therapeutic class–placebo test in which we proceed as if we were estimating the effect of the pandemic for each of the J therapeutic areas prescribed in general practice used to construct the counterfactual group. This allows us to evaluate the effect estimated for the prescriptions of AD against the distribution of the placebo effects. Then, the effect of the COVID-19 pandemic on AD prescription will be deemed statistically significant if the estimated effect for the AD class is unusually large relative to the distribution of the placebo effects, estimated for the other 14 therapeutic areas that where not directly affected by the pandemic. Following the SCM approach, we carry out this second mode of inference by calculating the root mean squared prediction errors (RMSPE),[25] for the pre-treatment and treatment periods, for each therapeutic class j in the donor pool, and for every practice g. Remember that the RMSPE measures the magnitude of the gap between the path of the actual series of prescriptions for any particular therapeutic class and its artificial counterfactual counterpart. The idea here is that a large post-treatment RMSPE is not indicative of a large effect of the intervention if the artificial control does not closely reproduce the outcome of interest prior to the treatment. That is, a large treatment effect may be artificially created by a lack of pre-treatment fit, rather than by the effect of the pandemic on AD. Then, one way to evaluate the AD gap relative to the gaps obtained from the placebo tests is to look at the distribution of the ratios of treatment/pre-treatment RMSPE. Specifically, we will observe where does the RMSPE ratio for AD, averaged across practices, i.e. $Ave_g[RMSPEratio_{gAD}]$, lies in the distribution of the corresponding ratios associated to the other 14 therapeutic classes in the predictors pool, i.e. $Ave_g[RMSPEratio_{gj}]$ for each j.

2.3.3 Heterogeneity

Finally, we examine the way that the estimated effect of the pandemic on the prescriptions of AD vary with new social, economic and public health-related circumstances caused by the pandemic, namely changes in (1) mobility, (2) number of unemployment benefits claimants, and (3) incidence of coronavirus (i.e. new infection cases). To do

[25] The pre-treatment RMSPE for any class j prescribed by practice g is defined as:

$$RMSPE_{gj}^{pre-treatment} = \left(\frac{1}{T_0}\sum_{t=1}^{T_0}\left(\Delta Y_{gt}^{j}\right)^2\right)^{\frac{1}{2}},$$

with $\Delta Y_{gt}^{j} \equiv \tilde{y}_{gt}^{j} - \kappa_{gt}^{j}$. The RMSPE can be analogously defined for the treatment period.

this we redefine coefficients γ_0 and μ in eq. (2.9) in the following way:

$$\gamma_0 \equiv \bar{\gamma}_0 + \theta_M M_{l,t-p} + \theta_C C_{l,t-p} + \theta_F F_{l,t-p} \tag{2.10}$$

$$\mu \equiv \bar{\mu} + \delta_M M_{l,t-p} + \delta_C C_{l,t-p} + \delta_F F_{l,t-p} \tag{2.11}$$

where $M_{l,t-p}$, $C_{l,t-p}$ and $F_{l,t-p}$ denote mobility, claimant count for unemployment benefits, and number of new infections, respectively, specific to local authority district l to which GP $g(l)$ belongs to, in month $t-p$, where p is the order of lags. Further details of these variables are given in Section 2.4. For mobility and new infections rates, only the coefficient δ_M and δ_F are identified as mobility and infection rates data only exists since February 2020, i.e. the start of the treatment period. Claimant Count data is available prior to and during the pandemic, in which case both δ_C and θ_C are identified. In order to conduct this examination the following model is estimated:

$$\begin{aligned}\Delta Y^{AD}_{g(l)t} = \bar{\gamma}_0 + \bar{\mu}\text{COVID-19}_t + \delta_M(M_{l,t-p} \times \text{COVID-19}_t) \\ + \theta_C C_{l,t-p} + \delta_C(C_{l,t-p} \times \text{COVID-19}_t) + \delta_F(F_{l,t-p} \times \text{COVID-19}_t) + \psi_{g(l)t}.\end{aligned} \tag{2.12}$$

For this exercise our interest lies in the coefficients:

$$\delta_X = \frac{\partial\{E[\Delta Y^{AD}_{g(l)t} \mid \text{COVID-19}_t = 1] - E[\Delta Y^{AD}_{g(l)t} \mid \text{COVID-19}_t = 0]\}}{\partial X_{l,t-p}},$$

with $X = \{M, C, F\}$, interpreted as the variation in AD prescriptions, as consequence of the pandemic, associated with changes in these three factors at the practice's local district.

As is clear from the definitions in (2.10) and (2.11), we assume that variation in the three observables would manifest in mental health outcomes not contemporaneously but with lag. Then, in Section 2.5, we present different specifications of model in eq. (2.12) combining the first and second lag, i.e. $p = \{1, 2\}$, of the variables of interest.

2.4 Data and sample

GP practices and prescription data. Administrative prescription data at the GP practice-level comes from the English Prescribing Dataset (EPD), a publicly available source containing detailed information on all medicines prescribed and dispensed each month by approximately 7 thousand GP practices currently operating in England.[26] GP practices are identified by their national code (by which we can link prescription data to other sources such as the Quality Outcomes Framework series containing data on

[26] See https://opendata.nhsbsa.net/dataset/english-prescribing-data-epd.

prevalence of depression), and medicines are identified by their British National Formulary (BNF) code (by which we can aggregate medicines that belong to a given therapeutic class). To capture prescriptions levels, we use the Average Daily Quantity (ADQ) field in the dataset,[27] representing the number of daily doses prescribed every month for each medicine for each practice.[28] This variable allows to aggregate and compare the prescribed quantities of different medicines with different strengths that are used to treat a common condition or disease.[29] Our analysis, however, is not conducted at the medicine level, but at the therapeutic class level. That is, we are not interested in how the prescriptions of a particular medicine used to treat depression has changed because of the pandemic, but rather on how the whole class of AD has varied. To do this we use the BNF drug classification system and aggregate the daily doses of all medicines that belong to each Section in the hierarchical structure of the BNF.[30] For example, AD are contained within Chapter 4 'Central Nervous System' Section 3 'Antidepressant Drugs' of the BNF. The sections of the BNF define the therapeutic classes and they correspond to the j dimension in Y^j_{gt}.

After imposing data requirements based on (1) the number of observations per practice and (2) on having additional information on the size and prevalence of depression for each one of them, our final sample consists of 6,627 GP practices, corresponding roughly to 90% percent of the total number of practices in the EPD for the February 2018 to January 2021 period (7,334), and to 98.6% of the practices participating in the Quality Outcome Framework scheme in 2019 (6,720), main secondary source from which we obtain the additional measures. These GP practices issue all the relevant therapeutic classes considered in our methodology for most of the 36-month period between Feb 2018 and Jan 2021,[31] and there is complete information on size and prevalence. (See Appendix 2.7 for an explanation regarding the derivation of the final sample based on the proposed data requirements.) Main summary statistics and graphical representation

[27] Average Daily Quantity (ADQ) represents the typical daily dose of a medication, prescribed to adult patients by GP practices. It is the English equivalent of the World Health Organizations defined daily dose (DDD).

[28] As prescription data in their raw form refer to prescription levels in months of differing lengths, we normalise the variable Y^j_{gt}, for each g and j, by diving it by the total number of days in month t and then multiply it by 365/12. In this way the values of Y^j_{gt} reflect now prescription levels in months having the same number of days.

[29] For example, assume that a GP prescribes to a patient 28 tablets of sertraline 50 mg. for a 28-day treatment period (one tablet per day) in a given month (sertraline is the most commonly prescribed AD in general practice). Given that the typical daily dose of sertraline is 50 mg. (a day), this corresponds to 28 daily doses of sertraline in such month. Now, assume that the same practice prescribes to another patient 28 tablets of citalopram 40 mg. for the same period (citalopram is the second most commonly prescribed AD), for which the typical daily dose is 20 mg. Then this practice prescribes 28 daily doses of sertraline for the first patient, 56 daily doses of citalopram for the second, which sums up to 84 daily doses of drugs used to treat depression in such month. Accordingly, now we do not have to deal with quantities (milligrams) of different medicines to treated depression, but with standard daily doses of AD.

[30] See https://openprescribing.net/bnf/ for the BNF classification system and https://openprescribing.net/chemical/ for a list of the medicines in the BNF.

[31] 6,598 practices are observed for 36 months and 29 for 35 months, resulting in 238,543 practice-months pairs.

of the prescription series Y_{gt}^{j}, and the de-trended series $\tilde{y}_{gt}^{j}$ (as constructed according to the first step in Section 3.2) for AD and for each therapeutic class j are presented in Appendix 2.7.

Using address information of each GP practice, we can identify 312 lower tier local authority (LTLA) districts in England to which they belong to, geographies by which we can link prescription data with mobility, unemployment benefit claimants and coronavirus new infection cases data. The LTLAs correspond to local government structures such as London Borough councils, county and district councils, metropolitan district councils, and unitary authority, and are responsible for the provision of a range of services for their local communities including NHS services, social services and education. Table 2.2 presents some characteristics on their population, number of GP practices within each, and on the number of patients registered at these. On average, a LTLA has 179 thousand population, the smallest has roughly 40 thousand and the largest 1.14 million. With regard to number of practices, the mean corresponds to 21, the minimum to 4 and the maximum to 165. And with respect to the number of patients registered, the figures are in line with those of the population: on average 152 thousand, with a minimum of 31 thousand, and a maximum of 989 thousand.

TABLE 2.2: Characteristics of LTLA districts

	(1) Mean	(2) SD	(3) Min.	(4) Median	(5) Max.
Population	179,108.2	120,145	39,930	140,769.7	1,138,478
No. GP practices	21.2	17.5	4	14.5	165
Sum List Size of GP practices	151,871.8	107,596.4	31,217	114,663.5	988,990

NOTES. Summary statistics for LTLA. $N = 312$. The figures in this Table are obtained by averaging yearly values of these characteristics over the years 2018-2020.

We note that LTLA areas might not be the best medium to link prescription data with the contextual factors considered for our heterogeneity analysis (i.e. mobility, unemployment benefit claimants and coronavirus new infection cases). Ideally, we would want to link prescription information of each practice to the values of these variables realizing themselves exclusively in the catchment area of the practice, that is, where their registered patient list resides, so to obtain a more accurate picture about what were they exposed to. However and unfortunately, catchment area data is not readily available nor easily manageable by traditional statistical software, and even if it were, it would not be straightforward to link it with other smaller statistical geographies, such as Middle Layer Super Output Area (MSOA),[32] for which we could obtain data on the factors of interest. Given the differences in the compositional size of the geographical areas we dispose and choose to conduct our heterogeneity analysis (i.e. LTLA), all

[32]MSOAs are statistical geographies that are constructed to produce spatial units with comparable population and household numbers, making them particularly suitable for econometric exercises and reducing the need for population normalisations or weighting of data.

variables used for it are normalized so as to reflect differences in size. (A further discussion of the limitations of relying on LTLA geographies is presented in the concluding Section 2.6.)

Next, we describe mobility, unemployment benefit claimants and coronavirus new infection cases data.

Mobility. To capture changes in mobility we use the Google COVID-19 Community Mobility Reports (Google, 2021) providing data for each LTLA district in England.[33] In its original form these data shows how time spent in certain categorized places changes daily compared to baseline days. A baseline day represents a normal value for that day of the week, and is defined as the median value, for the corresponding day of the week, during the 5-week period between Jan 3 and Feb 6, 2020 (a period before widespread disruption as communities responded to COVID-19). For the analysis we focus on mobility data on residential- and work-places. The values of the mobility variables represent the percentage variation in each categorized place, compared to the baseline, for each district and day. Higher values implies greater time spent at the corresponding place, and vice versa. Given the difference in temporal frequency between prescription and mobility data (which are monthly and daily, respectively) we collapse the daily changes in mobility at the month-level and obtain for each district the median value of the daily mobility distribution in each month.[34] Then, the mobility variable denoted by $M_{l,t}$ represents the median mobility value (for either residency or workplace), with respect to the baseline, for local authority district l in month t.

Claimant Count. Job insecurity and economic disruption are captured by data on unemployment benefits claimants. Specifically we use the experimental statistics Claimant Count from the UKs Office for National Statistics.[35] These figures consists of Jobseekers Allowance and some Universal Credit claimants by local authority district and month.[36] The values of the claimant count are represented in levels (i.e. number of people claiming unemployment benefits) and as percentages of the LTLA's population aged from 16 to 64 (based on mid-year 2018 population estimates). For our exercise, we

[33] See https://www.google.com/covid19/mobility/index.html?hl=en for data downloading and further information. Additionally, the mapping between the local authority district coding used in the reports, and the official national LTLA codes (by which mobility data can be linked with prescription data) is done using the resources generated by the Office for National Statistics (ONS) and Data Science Campus. See https://datasciencecampus.ons.gov.uk/supporting-the-response-to-coronavirus-covid-19/ and https://github.com/datasciencecampus/google-mobility-reports-data.

[34] Using the median or the mean results in very similar distribution. We use the median as it results in a distribution that exhibits a slightly higher variation across districts and time than the mean.

[35] See https://www.ons.gov.uk/employmentandlabourmarket/peoplenotinwork/unemployment/datasets/claimantcountbyunitaryandlocalauthorityexperimental.

[36] The UC claimants that are included are (1) those that were recorded as not in employment, and (2) those claimants of Universal Credit who are required to search for work.

use the latter. Then, the claimant count variable denoted by $C_{l,t}$ represents the percentage of people claiming unemployment benefits in local authority district l in month t.

COVID-19 new infections. Data on the number of COVID-19 new infection cases comes from the 'Coronavirus (COVID-19) in the UK' Dashboard from the UK government.[37] This publicly available dataset contains the number of daily infection cases by LTLA.[38] Just as with the claimant count these data are presented in levels and in rates expressed as per 100,000 population.[39] The latter measure is the one we use. In order to link cases to prescription data we sum all new daily cases (i.e. incidence) in a given month for each LTLA. Then, infection rate variable denoted by $F_{l,t}$ represents the number of new infections per 100,000 population in local authority district l in month t.

Main summary statistics and graphical representation of mobility, claimant count and new infections variables are presented in Appendix 2.7.

2.5 Results

2.5.1 Counterfactual group and the COVID-19 pandemic effect on antidepressants

Figure 2.2 shows the de-trended series for AD and other main therapeutic classes averaged across all GPs in the sample, after implementing the first step. A visual inspection of the relation between the shocks of AD and of other therapeutic areas in the pre-treatment period (to the left of the red line) suggests a positive correlation. If this is the case then the counterfactual group, being a weighted linear combination of the other classes, should exhibit a greater decrease than the actual series of AD in the treatment period provided that the classes given a larger weight experienced a greater decrease in their trends after the pandemic. Such preliminary analysis would suggest that in the treatment period there is an excess of AD prescriptions relative to the counterfactual group which, by definition, accounts for the decrease in prescriptions caused only by problems accessing primary care services.

[37] See https://coronavirus.data.gov.uk/ for information and see https://api.coronavirus.data.gov.uk/v2/data?areaType=ltla&metric=cumCasesBySpecimenDate&metric=newCasesBySpecimenDate&metric=cumCasesBySpecimenDateRate&format=csv for data downloading.

[38] Specifically it contains the number of people with a positive COVID-19 virus test when the sample was taken from the person being tested. In England, cases data includes all positive lab confirmed virus test results plus positive rapid lateral flow tests that do not have negative confirmatory lab-based polymerase chain reaction (PCR) tests taken within 72 hours.

[39] The count is divided by the LTLA population and then multiplied by 100,000 without any adjustment for other factors. Populations used are Office for National Statistics 2019 mid-year estimates.

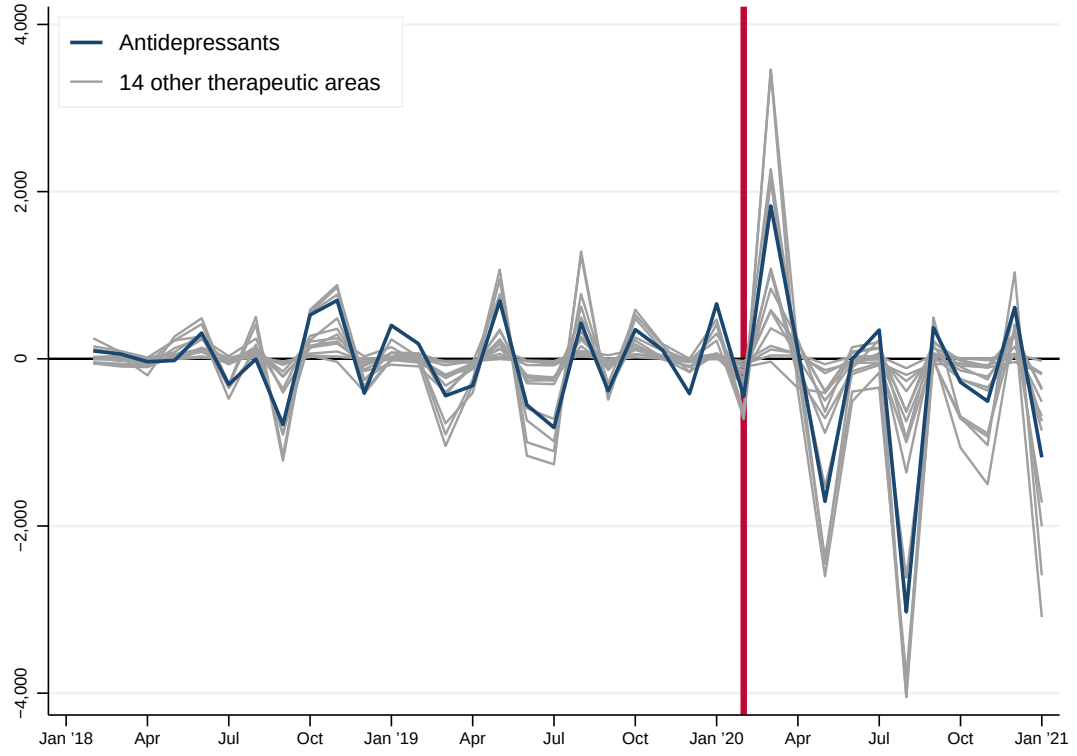


FIGURE 2.2: De-trended series (shocks) for AD and other therapeutic classes

NOTES. The figure shows the de-trended series for the class of AD (in thick blue) and for 14 therapeutic classes in the predictor pool (in light gray), obtained according to eqs. (2.5) and (2.6) in Section 3.2, averaged across practices.

Before implementing the second step (i.e. performing regressions with de-trended time-series) we conduct unit-root tests to show that these de-trended data are in fact stationary. Applying Dickey-Fuller tests on the series $Y^j_{gt} - \hat{\tau}^j_{gt}$ for AD and the 14 therapeutic classes j for each practice g on the pre-treatment period (2 year period between Feb 2018 to Jan 2020), we reject the null hypothesis of unit root for 97% of the AD series across practices, and for 95% or more for the 14 therapeutic classes, suggesting that the vast majority of practice level de-trended prescriptions series can be regarded as stationary.

Next, we discuss the results of the second step in which we perform GP-specific regressions of the de-trended residuals of AD on those of 14 therapeutic classes, according to model in eq. (2.7), in the pre-treatment period (February 2018 to January 2020). Table 2.3 presents the distribution of the *standardized* $\hat{\omega}_{jg}$ coefficients associated with the 14 therapeutic classes used to predict the shocks to AD, averaged across the 6,627 practices in the sample.[40] Remember that standardized (or beta) coefficients represent how many standard deviations the dependent variable (AD doses) will change, per standard

[40]The corresponding Table with non-standardized coefficients in prsented in Appendix Table 2.11.

deviation increase in the independent variable, instead of unit-level variations. Main statistics for these j distributions across practices are presented in columns (1)-(5), and the Table is sorted according to the values of column (1). Mean and median coefficients are all positive across all 14 therapeutic classes, thus confirming the positive correlation between the different shocks shown in Figure 2.2. These models have an average R^2 value of 0.84, median of 0.86, and 90 percent of these regression have an R^2 greater that 0.67 (10^{th}percentile). The high R^2 of these models gives strong support to using prescription drugs in other therapeutic classes to predict and reproduce the behaviour of AD in the pre-treatment period.

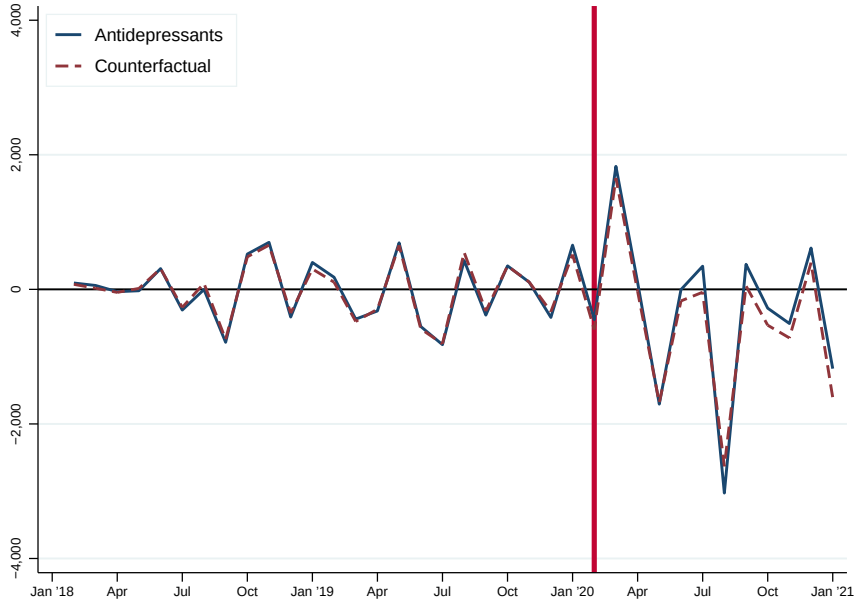
From column (1), and according to the interpretation of beta coefficients, we observe that the therapeutic classes most correlated with AD correspond to antisecretory, antiepileptics, thyroid & antithyroid, lipid-regulating, and anti-anginal drugs. On one hand, we note that lipid-regulating, antisecretory, and anti-anginal drugs correspond also to some of the largest groups prescribed in general practice (see Appendix Table 2.7), which indicates that the predictors of therapeutic classes are related by their prescribed volume. And on the other hand, anti-epileptics and thyroid & antithyroid are also strongly associated with AD, which suggests that predictors of therapeutic classes are also related by their comorbidity, as it has been identified that depression is the most frequent psychiatric comorbidity in patients with epilepsy (Kanner, 2005), and similarly, patients with thyroid disorders are more prone to develop depressive symptoms and conversely depression may be accompanied by various subtle thyroid abnormalities (Hage and Azar, 2012). (The relevance of the remaining information of Table 2.3, i.e. columns (6) to (12), will be made clear when we present main results and robustness checks, below.)

TABLE 2.3: Distribution of **standardized** coefficients $\hat{\omega}_{jg}$

	Panel A						Panel B					
	Models with all J therapeutic classes						Models with therapeutic classes pre-selected by linear LASSO					
	GP-specific OLS regressions					Pooled OLS	GP-specific OLS regressions					Pooled OLS
	(1) Mean	(2) SD	(3) p10	(4) Median	(5) p90	(6)	(7) Mean	(8) SD	(9) p10	(10) Median	(11) p90	(12)
Antisecretory	0.239	0.459	-0.319	0.243	0.783	0.318***	0.241	0.383	-0.235	0.245	0.705	0.313***
Antiepileptics	0.130	0.315	-0.248	0.131	0.512	0.198***	0.135	0.267	-0.191	0.135	0.467	0.198***
Thyroid & Antithyroid	0.093	0.346	-0.338	0.089	0.518	0.117***	0.093	0.295	-0.274	0.097	0.454	0.116***
Lipid-Regulating	0.090	0.572	-0.580	0.090	0.783	0.076***	0.103	0.445	-0.44	0.111	0.637	0.068***
Nit, Calc-Ch Block & Antianginal	0.078	0.460	-0.464	0.075	0.639	0.048***	0.081	0.387	-0.388	0.078	0.547	0.047***
Beta-Adrenoceptor Blocking	0.069	0.403	-0.426	0.066	0.548	0.079***	0.072	0.337	-0.349	0.073	0.489	0.076***
Rheumatic Diseases & Gout	0.055	0.279	-0.285	0.055	0.389	0.080***	0.058	0.237	-0.234	0.057	0.353	0.081***
Hypertension	0.047	0.541	-0.593	0.051	0.691	0.063***	0.060	0.445	-0.485	0.068	0.599	0.062***
Diabetes	0.042	0.375	-0.415	0.042	0.486	0.009						
Chronic Bowel	0.031	0.272	-0.296	0.031	0.362	0.031***	0.029	0.233	-0.253	0.031	0.315	0.031***
Drugs Affecting Bone Metabolism	0.025	0.291	-0.331	0.022	0.384	0.015***	0.029	0.244	-0.275	0.028	0.333	0.014***
Parkinsonism & Rel. Disorders	0.024	0.274	-0.312	0.024	0.354	0.015***	0.024	0.235	-0.264	0.022	0.315	0.015***
Diuretics	0.013	0.373	-0.438	0.015	0.462	0.000						
Antiplatelet	0.004	0.434	-0.523	0.004	0.541	-0.027***						
N						159,046						159,046
R^2	0.835	0.12	0.665	0.861	0.966	0.782	0.766	0.155	0.545	0.791	0.948	0.782

NOTES. The table presents the distribution of the OLS standardized coefficients from estimating GP-specific regressions of model in eq. (2.7). That is, they represent how many standard deviations the dependent variable (AD doses) will change, per standard deviation increase in the independent variable. Table is sorted according to the values of column (1). Mean and SD of the variables used in these regressions are presented in Table 2.9 columns (2)-(3). Panel A columns (1)-(5) presents main statistics of $\hat{\omega}_{jg}$, computed for each class j across 6,627 GP. Column (6) presents coefficients $\hat{\omega}_g$ estimated from the following pooled OLS regression: $\hat{\varepsilon}^{AD}_{gt} = \sum_{j \in J} \omega_j \cdot \hat{\varepsilon}^{j}_{gt} + v_{gt}$. N corresponds to $(G \times T_0 = 6,627 \times 24)$. Panel B show the same information as Panel A in which the GP-specific regressions of model in eq. (2.7) are estimated on a subset of therapeutic classes pre-selected by fitting a linear LASSO (with lasso penalty parameter selected by *adaptive* method), allowing to select from the initial set of 14 those that best predict the shocks to AD in the pre-treatment period. As can be seen from the table, the LASSO excludes the following classes: Antiplatelet, Diabetes and Diuretics. Statistical significance stars associated to coefficients in columns (6) and (12) come from robust standard errors clustered at the GP practice-level. *** $p < 0.001$.

The constructed counterfactual group is plotted in Figure 2.3. Panel (A) shows the counterfactual group as it was estimated according to the three steps and the de-trended series of actual AD prescriptions before and after the start of the COVID-19 pandemic, both averaged across practices. Panel (B) shows the same series adding the pre-treatment linear trend term, $\hat{\tau}^{AD}_{gt}$. Visual inspection suggests that our procedure can closely reproduce the average behaviour before the pandemic. During the pandemic, to the right of the red vertical line, the figure reveals a trend for the counterfactual group that is less steep than the actual series. This suggests that once controlling for the problem of access – what the counterfactual is meant to account for – there is an increase in the AD class. The precise quantification of this effect is given next.



(A) De-trended series

(B) De-trended series plus trend

FIGURE 2.3: AD prescriptions and counterfactual group

NOTES. Panel (A) shows the counterfactual, i.e. κ_{gt}^{AD}, as defined in eq. (2.8), and the de-trended residuals for AD, i.e. $\tilde{y}_{gt}^{AD}$, averaged across practices. Panel (B) shows the same series plus the estimated pre-treatment trend $\hat{\tau}_{gt}^{AD}$, i.e. $\kappa_{gt}^{AD} + \hat{\tau}_{gt}^{AD}$ and Y_{gt}^{AD}.

The results from model in eq. (2.9), and for the three related transformations of the doses variable, are presented in Table 2.4. According to the results in column (1), for the 12 month period since the start of the pandemic in England, between February 2020 and January 2021, the number of AD prescriptions have increased on average by 169

daily doses per month across GP practices. In terms of equivalent patients, column (2) shows that such increase corresponds to having around 5.7 more adult patients being treated for depression across practices monthly. Furthermore, by using our preferred measures of AD prescription levels and users, that is, normalised variables that account for differences in GP practices' adult population size, columns (3) and (4), suggests virtually the same results: number of daily doses and of equivalent depressive patients is 152 and 5.2, respectively.

TABLE 2.4: Average effect of the COVID-19 pandemic on AD across GP practices

	(1) Daily Doses	(2) Patients	(3) Normalised Daily Doses	(4) Normalised Patients
COVID-19	168.714***	5.689***	151.680***	5.194***
	(24.629)	(0.749)	(31.240)	(0.904)
Constant	0.038	0.001	0.065	0.001
	(8.208)	(0.250)	(10.411)	(0.301)
N	238543	238543	238543	238543
R^2	0.135	0.137	0.152	0.133

NOTES. The table presents the OLS estimates from model in eq. (2.9). Dependent variable in each model is given by the column-name. Normalised in models (3) and (4) refers to a normalisation by GP practices' adult population size. All models control for GP fixed effects. Clustered standard errors at the GP level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. N represents number of GP-months pairs ($G \times T$). See Appendix 2.7 for the details on the construction of the dependent variables in models (2)-(4).

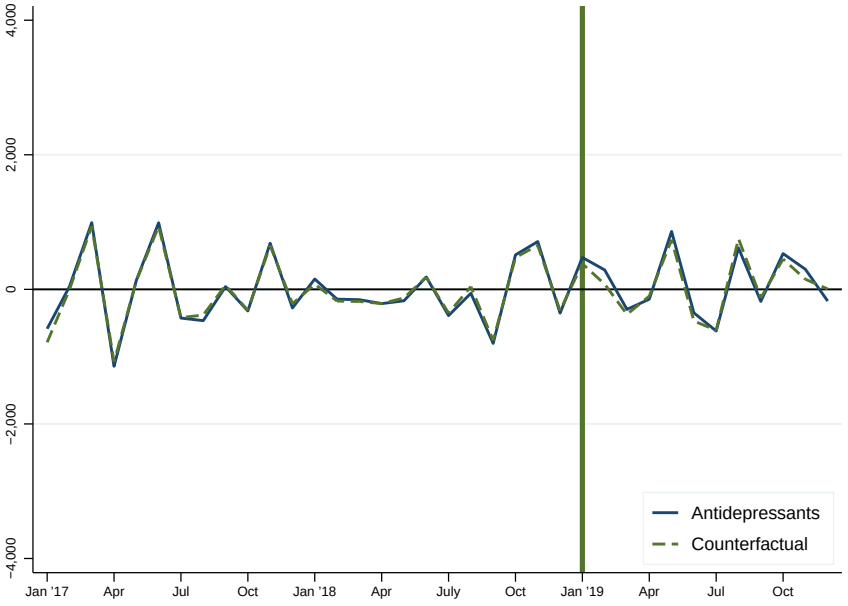
In the Background Section 2.2, we comment on certain dynamics of the prescription series and left pending the discussion on how these might be affecting our results. Therapeutic classes used to treat chronic (long-term) conditions, requiring indefinite treatment most of the times, like the ones considered in our methodology, rely on repeated prescriptions. That is, for example, a patient does not have to pick up his prescription every month from his local community pharmacy, but instead is given prescription for the next six-month twice a year, with the corresponding six-month follow-up appointment for assessment of the condition and the drug treatment. And as we commented earlier, the spikes observed in March 2020 very likely are the result of an anticipatory effect. And this anticipatory effect by itself, that is, without any additional effect that would change the expected trends of prescription, should have a null effect when taking average monthly prescription levels over a long period of time, as what is given as excess of prescription today taken from the future, would result in the same amount not been given tomorrow. (In other words, the specific values of the numerator over which an average is being calculated change, but not the total sum of those values). Additionally, if this anticipatory effect is present in all other therapeutic areas in the same intensity as for AD, and provided that the constructed counterfactual group is correctly specified and estimated, then it should also capture this anticipatory effect.

By this reasoning, we consider the concern that this anticipatory effect might be significantly affecting the results to be negligible.

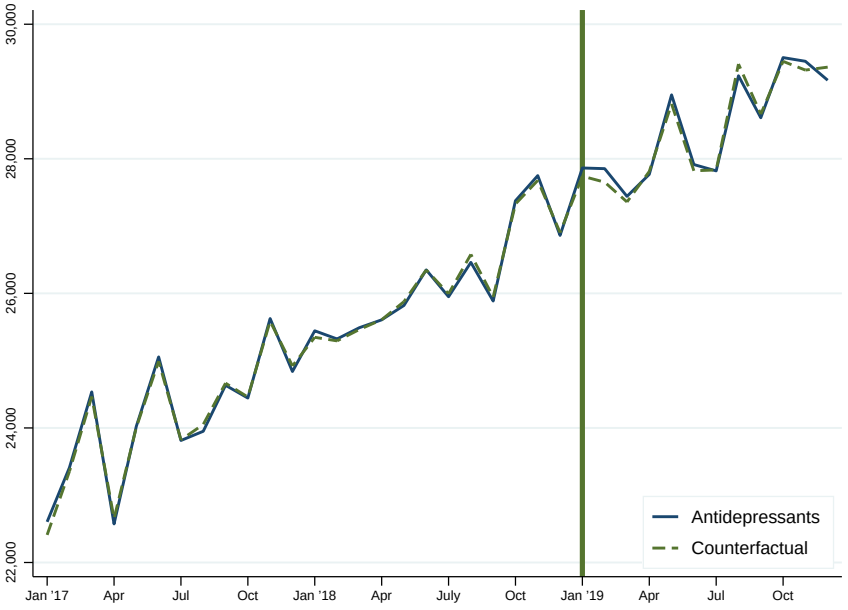
Robustness check. As a robustness check we compare the previous results with alternative ways of performing the second step in the construction of the counterfactual group. As described in the methodology, the second step involves the estimation of GP practice-specific linear regressions (by OLS) between the estimated shocks to AD and those of 14 other therapeutic areas, resulting in practice-specific coefficients which then are applied to the actual values of these in the treatment period to define the control group. Three different ways are considered. First, to perform a pooled OLS regression (instead of practice-specific regressions), to obtain a unique set of coefficients associated to each of the 14 therapeutic classes, which are then applied to the actual series of each practice. The second and third way are are just as the original second step and the latter one just considered, but instead of using the initial set of 14 therapeutic classes, we let a linear LASSO regression (pooling all GP practices) to select those therapeutic areas that best predict the shocks to AD in the pre-treatment period (selection method for the penalty parameter correspond to *adaptive* method). This procedure selects eleven therapeutic classes and excludes three: Antiplatelet, Diabetes and Diuretics. The distribution of coefficients $\hat{\omega}_{jg}$ associated with these alternatives way of performing the second step are presented in Table 2.3 (standardized) and Appendix Table 2.11 (non-standardized), columns (6)-(12), and the corresponding results are presented in Appendix Table 2.12. Under the four different ways of constructing the control group for AD the main findings hold, and are virtually the same.

2.5.2 Placebo Tests

Figure 2.4 shows the results from our temporal placebo study. Here, we re-assigned the pre-treatment and treatment periods 1 year before the pandemic event. Visually, the constructed counterfactual reproduces almost exactly the behaviour of AD in the pre-treatment period as well as its evolution during the 2019 period, as no major difference can be detected between both series. Applying the model in eq. (2.9) for this placebo treatment period results in a not statistically significant effects. This result supports our methodology as the difference between AD and the constructed counterfactual group are not statistically different when assigning the treatment to a period where no relevant public health shocks has taken place.



(A) De-trended series



(B) De-trended series plus trend

FIGURE 2.4: Temporal-Placebo Test

NOTES. The figure shows the results of the temporal-placebo test. Panel (A) shows the counterfactual, i.e. κ_{gt}^{AD}, as defined in eq. 2.8, and the de-trended residuals for AD, i.e. $\tilde{y}_{gt}^{AD}$, averaged across practices. Pane (B) shows the same averaged series plus the estimated pre-treatment trend $\hat{\tau}_{gt}^{AD}$, i.e. $\kappa_{gt}^{AD} + \hat{\tau}_{gt}^{AD}$ and Y_{gt}^{AD}.

Figure 2.5 shows the gap between the actual prescriptions and the counterfactual group for AD and 14 placebo gaps. The figure suggests that the treatment period effect for AD is the largest, when compared those of the 14 classes. In fact, results presented in

Appendix Table 2.13 show that, relative to the other 14 therapeutic classes, the gap for AD is the largest in magnitude.

Given that the effect found for AD is not necessarily indicative of a large if the constructed counterfactual does not closely reproduce the actual prescription series in the pre-treatment period, we also compute the RMSPE ratio and check whether this ratio for AD is indeed larger than the resulting corresponding ratios associated to the other therapeutic classes. Appendix Table 2.14 presents the RMSPE ratios for each therapeutic class averaged across practices. The results indicate that the ratio for AD is the largest. Consequently, the effect estimated for AD overall appears extreme in the distribution of these placebo effects; and if one were to chose a therapeutic class at random, the chances of obtaining an average RMSPE ratio as high as for AD, it would be $1/15 \simeq 0.066$, the lowest possible.

FIGURE 2.5: Gap between actual prescriptions and counterfactual group for AD and placebo gaps for 14 therapeutic classes

NOTES. The figure shows the gap between the actual prescriptions and the counterfactual group for AD and 14 placebo gaps after applying our methodology to each of the 14 therapeutic classes, averaged across practices.

2.5.3 Heterogeneity

The results presented so far represent the average effect of the COVID-19 pandemic on AD prescriptions across the whole GP sector. However, the features of our data allow us to explore geographical heterogeneity in this effect and more interestingly to study its correlates. We first observe graphically the extent of heterogeneity in the change in AD prescriptions across the 312 lower tier local authority (LTLA) districts to then show the estimation results from the model in eq. (2.12). Figure 2.6 plots the geographical distribution of the change in AD prescriptions in the 312 districts. The figure reveals significant heterogeneity across districts, where some exhibit increases and other decreases in the prescription of AD as a consequence of the pandemic once accounting for the problem of access. Possible explanations for this observed heterogeneity are the new social, economic and health related circumstances brought by the pandemic, such as reduced mobility, increased job insecurity and risk of infection.

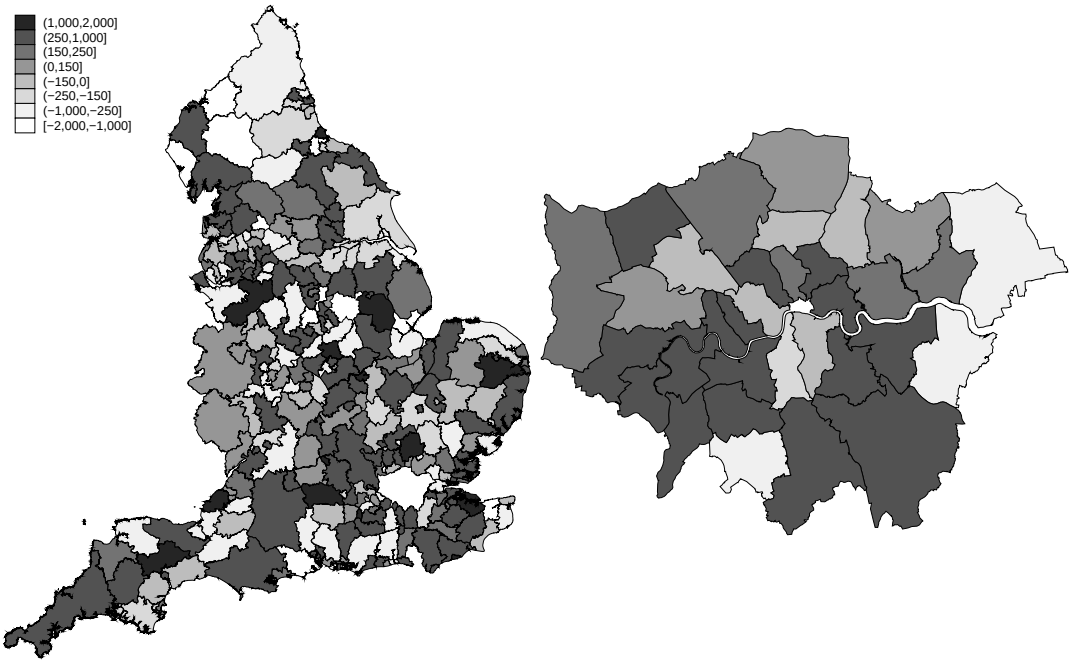


FIGURE 2.6: District-level variations in AD prescriptions after COVID-19

NOTES. The figure plots estimates of μ from model in eq. (2.9) performed for each LTLA district l, using as a dependent variable the number of daily doses of AD normalized by the practice's adult population size.

Results in Table 2.5 presents six models combining different lags of the variables of interest. With regard to mobility, we find that AD prescriptions increase more in practices located in districts where time spend in residential places was greater. This result

points towards the positive link between increased isolation and poorer mental health outcomes. When substituting the mobility-in-residency variable with workplace mobility we find, consistently, the opposite effect, namely AD prescriptions increase less in district where time spend in workplaces is greater. These relations comes with lag, as only the two-month lag coefficients associated to these variables are statistically significant. With respect to the number of new COVID-19 infection cases we find that AD prescriptions increase more in districts with a higher incidence rate and that these effects are statistically significant for both the first- and second-month lag coefficients. This may assert the positive relations between higher risk of infection and poorer mental health. In a different direction, we find that AD increased less in districts where the number of unemployment benefits claimants is greater, correlation that is statistically significant for both the first- and second-month lag coefficients. This could be considered tentative evidence of the potential benefits of the social security net in contributing to better mental health.

Finally, we comment on the actual interpretation of the coefficients of our preferred model specification in column (3). With regard to mobility, 1 percentage point increase in time spent inside residential places (with respect to the baseline) is associated to an average increase of 8.2 daily doses of AD per month, and a 1 SD increase (8.89 pp) in mobility in residential places is associated with a 73 daily doses increase (in terms of equivalent patients this increase corresponds to approximately 1.6 more patients treated for depression). As for the claimant count of unemployment benefits a 1 percent increase in the rate of people claiming benefits is associated to a reduction of 47.7 daily doses of AD per month, and a 1 SD increase (2.32 pp) is associated with a 110.7 daily doses decrease (this decrease in terms of equivalent patients corresponds to approximately 2.1 less patients treated for depression). Lastly, for new infections, a 1 new case per 100,000 population increase is associated to an increase of 0.086 daily doses of AD, and a 1 SD increase (600.4 per 100,000 population) is associated to a 52 daily doses increase (in terms of equivalent patients such increase corresponds to approximately 1.6 more adult patients treated for depression).

TABLE 2.5: Correlates of change in AD prescriptions

	(1)	(2)	(3)	(4)	(5)	(6)
Mobility: Residency (t-1)	-2.0488					
	(1.5952)					
Mobility: Residency (t-2)			8.2142***		3.8895*	
			(1.9750)		(1.6087)	
Mobility: Workplace (t-1)		0.9691				
		(0.7733)				
Mobility: Workplace (t-2)				-4.5305***		-2.4126**
				(0.9681)		(0.8124)
Claimant Count rate (t-1)	-50.9366**	-50.7577**	-47.7466*	-47.1439*		
	(19.6430)	(19.6025)	(19.6475)	(19.5834)		
Claimant Count rate (t-2)					-50.9584*	-50.8916*
					(20.0593)	(20.0353)
Infection Cases rate (t-1)	0.0673***	0.0637**	0.0866***	0.0965***		
	(0.0200)	(0.0206)	(0.0210)	(0.0218)		
Infection Cases rate (t-2)					0.0856*	0.0970**
					(0.0357)	(0.0364)
N	231912	231912	225283	225283	225287	225287
R^2	0.157	0.157	0.162	0.162	0.162	0.162

NOTES. The table presents OLS estimated from the model in eq. (2.12). Only δ coefficients are shown. Dependent variable corresponds to the number of daily doses of AD normalized by the practice's adult population size, that is $\Delta nY^{AD}_{g(l)t}$. Mean and SD of the independent variables used in these models are presented in Table 2.10. All models control for GP fixed effects. Clustered standard errors at the GP level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. N represents number of GP-months pairs ($G \times T$).

2.6 Discussion and conclusions

In this paper we investigate the mental health consequences of the COVID-19 pandemic in England using AD prescriptions data from GP practices. The main challenge we face comes from the recognition that the observed prescription levels are the result of two effects working in opposite directions, namely the direct mental health effect, which should increase AD prescriptions, and the problems of access, which should decrease them. To address this, we construct a counterfactual group based on other therapeutic classes of drugs that, on one hand, were arguably only (or mostly) affected by problems of access and not by a direct effect of the pandemic on the specific conditions treated by these drugs, and on the other hand, are able to predict satisfactorily the prescriptions of AD prior to COVID-19. Furthermore, we analyse geographical variation in AD prescriptions and explore its correlation with fear of infection (measured by the incidence of the virus), isolation (measured by mobility in residential and work-places) and economic hardship (measured by unemployment benefits claimants).

Our findings suggest that during the first year, the pandemic implied an average increase in AD prescriptions equivalent to 5 more adult patients treated for depression

per practice. We consider this to be a moderate effect relative to the national prevalence (i.e. existing patients) and incidence (i.e. new patients) of adult depression, which in 2019 was 828 and 86 patients per practice, respectively. This result is in sharp contrast with the general sentiment of severe impact on mental health, as reported in the popular press, grey literature, as well as from academic studies that have used self-reported measures of mental health.[41] Furthermore, we find that AD prescriptions increased more in districts where number of infections and time spent at home are greater, but they decrease in districts where number of unemployment benefits claimants is greater. The latter suggests that the social safety net might be effective in reducing economic hardship and, in turn, mental health problems.[42]

Most of the existing studies on COVID-19 and mental health have used self-reported measures of subjective well-being. One exception is Armitage (2021) where the author, after estimating that the number of items of AD prescribed in primary care between April and September 2020 was 3.94% higher than the corresponding period a year before, concludes that there was an increase of AD use as consequence of the COVID-19 pandemic.[43] We note that this analysis is misleading as it does not consider that in previous years the change in AD has been much larger than the one found for the 2019-2020 period: 4.8% in 2017-2018 and 6.05% in 2019-2019. Our data and findings are consistent with the study by Carr et al. (2021) which looks at trends in primary care-recorded mental illness, self-harm episodes and psychotropic medication in the UK, between April and September 2020. Up to May, they find a sharp reduction in first antidepressant prescribing in English GP practices, an increase up to June and then rates similar to expected levels by September. Yet, as the comparison between actual and expected trends, based on past information only, does not considers the indirect effect of the pandemic on access to health care services, their study is mainly a descriptive analysis of variables capturing the point of contact between patients and mental health care services, rather than an empirical analysis of the mental health effect caused by the pandemic, as we do in the present work.

Our study is not without limitations. First, the main variables used correspond to aggregate prescription levels in a given month, and as such we cannot distinguish the patient-type composition of the estimated increase. That is, whether the increase in AD corresponds to prescriptions issued for new patients that started treatment during

[41] See https://www.ons.gov.uk/peoplepopulationandcommunity/wellbeing/articles/coronavirusanddepressioninadultsgreatbritain/june2020.

[42] This finding is related to the literature on the impact of unemployment insurance benefits on health-care social needs, access to health care and mental health. Works by Kuka (2020), Berkowitz and Basu (2021), Kaufman et al. (2020) for the US and by Reichert and Tauchmann (2011) for Germany, find that being recipient of these benefits increases coverage, access and improves self-reported health. See also works by Janke et al. (2020) and Currie et al. (2015) on the effects of the economic cycle and financial crises on chronic illnesses, including mental health conditions, for the UK and the US, respectively.

[43] He also documents reductions in the number of people in contact with NHS-funded mental health services (specialized secondary care and psychological therapy services) and in the number of referrals from primary care to mental health services, pointing to a decrease in the demand (or access) to health services.

the COVID-19 period (i.e. an extensive margin), or to prescriptions for existing patients whose treatment length has been extended, e.g. a treatment that was expected to end during 2020, now is extended to last more, or to prescriptions for existing patients whose treatment regime has increased in intensity, e.g. a drug was given at 10mg, now is given at 20mg, (i.e. two forms of intensive margins); or to a combination of these.

Second, the heterogeneity analysis assumes that the district-level variation in the three covariates we correlate to AD, namely number of infection cases, mobility in residence places and number of unemployment benefits claimants, is related only to individuals and communities living within the borders of each district, and that they are registered with the local GP practice in those districts. As much of a reasonable assumption this might be, there will be cases where individuals may be receiving primary care services not necessarily in the district where they reside but in which they labour; or that a practice may have a catchment area that crosses borders between two adjacent districts.

Third, we acknowledge that antidepressants prescriptions and depression represents one of the many aspects in which poor mental health can manifested itself and that in turn it might not capture the full scope of the deterioration of psychological states during the pandemic. The variables used in this paper capture patients whose subjective evaluation of symptoms is such that they seek primary health care from their GP, are assessed by a professional practitioner, with whom the decision to start pharmacological treatment is made, or not, depending on the assessed severity of the condition. Other forms of affected mental health could also be observed and captured by, for example, prescription data for drugs used to treat substance dependence, anxiety, and psychosis; phone calls to domestic violence help lines; referrals to psychological talking therapies or related interventions; psychiatric interventions; suicide statistics; and even by self reported measures from surveys that may capture true symptoms of poorer mental health that simply do not require clinical attention. A more comprehensive study of the effects of COVID-19 on mental health would have to follow a more inclusive approach acknowledging all such possible outcomes and how they are internalized and registered by the health care system (NHS).

Finally, we note that our results are specific to the English health care system, socioeconomic conditions and governmental public health response to COVID-19. Further studies are needed to understand if the moderate impact we find for England, relative to the already high prevalence of depression, can be generalized to other countries or settings.

2.7 Appendix

Note on the relation between the proposed methodology and the SCM. Our approach is in the spirit of the Synthetic Control Method (SCM) where a counterfactual group for a treated unit (city, region, country, etc.) is constructed to estimate the effect of an intervention (treatment, shock, event, etc.) on an outcome of interest.[44] Yet, the nature of our problem and the available data do not lend themselves to the full use of this approach.

First, the SCM requires to have available data on the same outcome variable for treated and untreated units. In our case, this would mean having the antidepressants prescriptions series (our outcome of interest) for a set of GP exposed to the COVID-19 pandemic and for another set not exposed. This is impossible as the English society as a whole experienced the pandemic. Even if this were possible, we would still not be able to separate the problem of access from the mental health effect as the treated units would still contain the double effect and the untreated ones would reflect the evolution of AD in absence of the pandemic – a useless counterfactual in this case. As we explain also in the paper, the type of counterfactual for AD that is needed is not one that was not exposed to the pandemic, but one that contains only the effect on access to GP services, to then be able to isolate the mental health effect. This is why we look at other therapeutic areas that, one could argue, were untreated by a direct effect of the pandemic on the conditions for which they are prescribed (unlike the case of depression and antidepressants), but only treated by the problem of access, as all of them were. And our method relies on using these to construct such an ideal counterfactual.

Next, we would encounter two problems related to the two types of variables that the SCM requires to be implemented. The first type of variables correspond to outcome variables and the second type to pre-intervention characteristics that can predict the outcome of interest. For example, in Abadie et al. (2010), to estimate the effect of a tobacco control program in California, the outcome variable is per-capita cigarette sales, and the predictors are GDP per capita, the share of 15-24 aged population, retail price of cigarette, beer consumption per capita, and lagged cigarettes sales, for US states. In Abadie et al. (2015,) to estimate the effect of the reunification of Germany on West

[44] In summary, the idea behind the SCM is that a combination of few units that were not exposed to an intervention (i.e. the donor pool) can provide a better comparison for the unit exposed to the intervention, than a single unit alone (e.g. traditional difference-in-difference approach). Then, the SCM models for the aggregate outcome of the treated unit in a treatment period a counterfactual group, called the "synthetic control". The synthetic control is generated by finding optimal weights (all non-negative and adding up to 1) for the untreated units by matching as closely as possible a set pre-intervention characteristics (or predictors) of the treated unit with the values of these same characteristics for the units in the donor pool. These optimal weights, estimated using pre-intervention information only, are then applied to the same outcome variables of the untreated units to define the synthetic control as a weighted average of these which, if constructed successfully, is such that, in the pre-intervention period, reproduces almost exactly the outcome variable of the treated unit, and in the post-intervention period serves to outline what would have happened to the treated unit had the treatment never occurred. See Abadie and Gardeazabal (2003); Abadie et al. (2010, 2015).

Germany's growth, the outcome variable is GDP per capita, and the predictors are lagged GDP per capita, trade openness, inflation rate, schooling, etc., for 16 OECD member countries.

Given this, the first problem would be that the outcome variables we are dealing with are not expressed in the same unit of measurement. That is, in our case, we have doses of antidepressants as the main outcome for the treated unit, and doses of cholesterol-lowering medication, anti-hypertensives, anti-diabetics, etc., for the untreated ones. Leaving aside this minor issue, we would face a second problem regarding the availability and quality of predictors of prescriptions. Ideal predictor variables would be prevalence rates (i.e. existing patients with conditions for which the drugs in the specific class are prescribed for), incident rates (i.e. new patients), recovery rates, and lagged values of prescription rates, for example, for every GP practice and in every month. Unfortunately, such data are not available in this ideal format for every therapeutic class in our donor pool (although prevalence and incidence are available for the most common conditions treated in primary care). Additionally, there's the issue that one therapeutic class is not strictly associated to one only diagnosis but in most cases to several diagnoses and conditions, for which, again, the information is not readily and publicly available.

Despite these problems, we did attempted to implement the SCM (`synth` in STATA) aggregating GP prescription data at the month-level using some of the predictor variables just mentioned. This resulted in a pre-treatment period fit (between the AD class and its synthetic counterpart) that we judged to be unsatisfactory.

Finally, as explained in Section 3.2, by using other therapeutic areas to predict AD as the core of our methods to generate an appropriate counterfactual, we are treating these other drug classes, in themselves, as both untreated units and as pre-intervention characteristics used to predict the outcome of interest, by obtaining the optimal weights (linear regression coefficients) that match as closely as possible those of the treated unit, in the pre-treatment period. Despite the concerns pointed out here, the idea and intuitions behind the synthetic control method remain and are what we use for our setting, data and strategy.

TABLE 2.6: Prescription trends for main therapeutic areas in general practice before and after COVID-19

(1)	(2)	(3)	(4)	(5)	(6)
	Trends		Absolute		Relative
Therapeutic Area	Pre	Post	Change	Therapeutic Area	Change
	(daily doses/month)				(%)
Bronchodilators	202.2	-256.2	-458.4	Drugs Affecting Bone Metabolism	-374.09
Corticosteroids (Resp)	189.6	-100.1	-289.7	Diuretics	-265.96
Hypertension	145.3	-38.5	-183.8	Bronchodilators	-226.67
Lipid-Regulating	295.0	123.8	-171.2	Corticosteroids (Resp)	-152.80
Antisecretory	194.5	72.8	-121.7	Diabetes	-143.23
Nit, Calc-Ch Block & Antianginal	145.7	24.4	-121.2	Analgesics	-129.86
Antidepressants	234.8	165.0	-69.9	Hypertension	-126.48
Thyroid & Antithyroid	43.3	1.0	-42.4	Parkinsonism & Rel. Disorders	-124.21
Diabetes	25.0	-10.8	-35.8	Beta-Adrenoceptor Blocking	-103.11
Diuretics	13.4	-22.2	-35.6	Thyroid & Antithyroid	-97.78
Analgesics	24.9	-7.4	-32.3	Laxatives	-93.71
Beta-Adrenoceptor Blocking	25.1	-0.8	-25.8	Antiplatelet	-91.37
Antiplatelet	24.1	2.1	-22.0	Nit, Calc-Ch Block & Antianginal	-83.22
Antiepileptics	38.0	23.2	-14.8	Antisecretory	-62.57
Parkinsonism & Rel. Disorders	8.8	-2.1	-11.0	Lipid-Regulating	-58.04
Drugs Affecting Bone Metabolism	2.7	-7.4	-10.1	Hypnotics & Anxiolytics	-54.67
Laxatives	-10.5	-20.4	-9.9	Chronic Bowel	-40.69
Chronic Bowel	8.3	4.9	-3.4	Antiepileptics	-38.97
Hypnotics & Anxiolytics	-4.3	-6.6	-2.3	**Antidepressants**	-29.75
Psychoses & Rel. Disorders	6.6	4.8	-1.8	Psychoses & Rel. Disorders	-26.89
Nausea & Vertigo	0.7	3.2	2.5	Rheumatic Diseases & Gout	95.79
Rheumatic Diseases & Gout	-18.0	-0.8	17.3	Nausea & Vertigo	362.20

NOTES. The table presents trends (measured in daily doses per month) pre– and post–February 2020 in columns (2) and (3), respectively, averaged across practices, for 24 main therapeutic classes. Additionally, absolute and relative change in trends are presented in columns (4) and (6), respectively. Therapeutic classes in column (1) sorted according to the values of absolute change; those in column (5), according to values of relative change, i.e. the absolute change over the pre-trend value (times 100%). For example, with regard to the antidepressants class: pre-COVID it grows at a rate of 234.8 daily doses per month on average across practices; post-COVID it does at 165, an absolute decrease of 69.9. In relative terms it corresponds to a 29.75 percent decrease with respect to the pre-pandemic levels.

TABLE 2.7: Main therapeutic classes prescribed in general practice

British National Formulary (BNF)				
Code	Chapter Name	Section Name	Daily Doses	Share
0212	Cardiovascular system	Lipid-Regulating Drugs	503,435	13.0%
0205	Cardiovascular system	Hypertension and Heart Failure	484,690	12.5%
0103	Gastro-intestinal system	Antisecretory Drugs + Mucosal Protectants	404,301	10.4%
0403	Central nervous system	**Antidepressant Drugs**	362,291	9.3%
0206	Cardiovascular system	Nit, Calc Block & Other Antianginal Drugs	300,496	7.7%
0301	Respiratory system	Bronchodilators	248,474	6.4%
0407	Central nervous system	Analgesics	199,510	5.1%
0302	Respiratory system	Corticosteroids (Respiratory)	194,682	5.0%
0601	Endocrine system	Drugs Used In Diabetes	175,177	4.5%
0202	Cardiovascular system	Diuretics	152,144	3.9%
0602	Endocrine system	Thyroid And Antithyroid Drugs	114,807	3.0%
0209	Cardiovascular system	Antiplatelet Drugs	97,444	2.5%
0204	Cardiovascular system	Beta-Adrenoceptor Blocking Drugs	94,555	2.4%
0408	Central nervous system	Antiepileptics	90,725	2.3%
0304	Respiratory system	Antihist, Hyposensit & Allergic Emergen	69,268	1.8%
0106	Gastro-intestinal system	Laxatives	66,920	1.7%
0501	Infections	Antibacterial Drugs	52,222	1.3%
1001	Musculoskeletal & joint diseases	Drugs Used In Rheumatic Diseases & Gout	48,897	1.3%
0606	Endocrine system	Drugs Affecting Bone Metabolism	33,368	0.9%
0401	Central nervous system	Hypnotics And Anxiolytics	27,285	0.7%
0406	Central nervous system	Drugs Used In Nausea And Vertigo	22,339	0.6%
0402	Central nervous system	Drugs Used In Psychoses & Rel. Disorders	21,715	0.6%
0409	Central nervous system	Drugs Used In Park'ism/Related Disorders	18,833	0.5%
0105	Gastro-intestinal system	Chronic Bowel Disorders	16,905	0.4%
Total 24 selected BNF sections			3,800,482	97.8%
Total all BNF sections			3,885,979	

NOTES. The table presents the top 24 therapeutic classes (out of 102), or BNF sections, prescribed in general practice during the year 2019. Under 'Code' is the BNF Chapter-Section pair code associated to each therapeutic class (given by Section name). For example, Antidepressant Drugs belongs to Chapter 4 Section 3 of the BNF. Under 'Daily Doses' is shown the number of daily doses prescribed of each therapeutic class averaged across practices in 2019. Under 'Share' is the percentage with respect to 'Total all BNF sections'.

SOURCE. English Prescribing Dataset (EPD).

Daily doses into number of patients. To transform doses into patients we proceed as follows. First, from the yearly Quality Outcome Framework series we obtain practice-level data on the prevalence of depression (i.e. number of adult patients with a current diagnosis of depression) in 2019 (denoted by $P_{g,2019}$). Second, we aggregate monthly prescription data (daily doses) of AD over the same period for each practice (denoted it by $D_{g,2019} \equiv \sum_{t \in 2019} Y_{gt}^{AD}$). By computing the doses/patients ratio over the year 2019 and diving it by 12 (months), that is $DtP_{g,2019} \equiv \frac{D_{g,2019}}{P_{g,2019}} \cdot \frac{1}{12}$, we obtain a measure of the average number of daily doses of AD prescribed per patient with depression per month for each GP practice g in 2019. Finally, by dividing Y_{gt}^{AD} by $DtP_{g,2019}$ we transform the number of monthly daily doses of AD into the number of equivalent monthly patients treated for depression in each practice, denoted by $P_{gt}^{AD} \equiv \frac{Y_{gt}^{AD}}{DtP_{g,2019}}$. The implicit assumption made here is that patients registered as having depression are also on antidepressants treatment.

Normalization by practice's adult population list size. To account for differences in list sizes of each practice (i.e. the total registered adult population) we normalize both the number daily doses and equivalent patients (i.e. Y_{gt}^{AD} and P_{gt}^{AD}) by the adult list size of each practice in the year 2019, denoted by $LS_{g,2019}$. Specifically, we divide doses and equivalent patients by the list size and then multiply it by the average list size in 2019, denoted by $\overline{LS}_{2019} \equiv \sum_{g \in G} \frac{LS_{g,2019}}{G}$, where G is the total number of GP practices in 2019. Then we obtain normalised variables of both daily doses and equivalent patients, denoted by $nY_{gt}^{AD} \equiv \frac{Y_{gt}^{AD}}{LS_{g,2019}} \times \overline{LS}_{2019}$ and $nP_{gt}^{AD} \equiv \frac{P_{gt}^{AD}}{LS_{g,2019}} \times \overline{LS}_{2019}$, respectively. These have the interpretation of representing the number of daily doses and patients per average practice adult population size – which in 2019 corresponds to 7,165 patients, on average.

Derivation of final sample and data requirements. For the 36-month period between February 2018 and January 2021, the English Prescribing Dataset (EPD) has information from 7,334 GP practices. However, not all these practices can be used to constitute our final sample because of issues of missing and incomplete data. In order to have a meaningful and representative sample, we require the data to satisfy the following data requirements: (1) that every practice is observed issuing prescriptions for most of the 36-month period; and (2) that we can obtain, from a secondary source (e.g. the Quality and Outcomes Framework series), complete information on the prevalence of depression and the list size, i.e. registered adult population, information that is needed for the purpose of generating the number of depressive adult patient variable and implementing the normalisations discussed above at the end of Subsection 2.3.1.

Table 2.8 presents summary information for the full EPD dataset (Panel A), the final sample satisfying the requirements (Panel B), and the excluded sample not satisfying them (Panel C), computed over the three-year period from Feb 2018 to Jan 2021. As it can be observed, the final sample closely resembles the full EPD dataset in terms of the average number of AD doses per month across practices (Y_{gt}^{AD}), the number of months GP practices are observed issuing strictly positive quantities of prescriptions (T_g), depression prevalence and the list size of the practice. On the other hand, the excluded sample does not appear to be a random sample, rather it prescribes a third of what the full EPD data and the final data does, its practices are observed issuing prescription for less than two years, and they are significantly smaller in size.

TABLE 2.8: Summary of final sample

	(1) Mean	(2) SD	(3) Median	(4) GP × Months	(5) GP
A. Full EPD dataset					
AD Doses (Y_{gt}^{AD})	28,361.1	22,487.5	23,435.7	250,512	7,334
No. Monthly Observations (T_g)	34.2	6.5	36	-	7,334
Depression Prevalence$_{gt}$	801.2	634.4	652.0	248,599	7,173
List Size$_{gt}$	7,024.3	4,724.8	6,122	248,599	7,173
B. Final Sample – Satisfying criteria					
AD Doses (Y_{gt}^{AD})	29,243.5	22,510.4	24,326	238,543	6,627
No. Monthly Observations (T_g)	35.996	.066	36	-	6,627
Depression Prevalence$_{gt}$	816.5	639.2	667	238,543	6,627
List size$_{gt}$	7,150.6	4,751.4	6,243	238,543	6,627
C. Excluded Sample – Not satisfying criteria					
AD Doses (Y_{gt}^{AD})	10,774.2	12,664.4	6,454.9	11,969	707
No. Monthly Observations (T_g)	16.9	10.4	16	-	707
Depression Prevalence$_{gt}$	440.5	347.4	354	10,056	707
List size$_{gt}$	4,033.2	2,648.9	3,360.5	10,056	707

NOTES. The table presents summary information for the full EPD dataset (Panel A), the final sample satisfying the requirements (Panel B), and the excluded sample not satisfying them (Panel C), computed over the three-year period from Feb 2018 to Jan 2021. The variables correspond to the monthly number of AD Doses (Y_{gt}^{AD}), the number of months GP practices are observed issuing strictly positive quantities of prescriptions of any drug class (T_g), the depression prevalence levels and the list size. The Mean, SD and Median are presented in columns (1)-(3), respectively, and the number of GP practices×months and number of GP practices are presented in columns (4) and (5), respectively. The different values for these last two statistics within Panels A and C need further clarification. The full EPD dataset contains monthly information on AD (Y_{gt}^{AD}), and other therapeutic classes, for 7,334 GP practices observed for 34.2 months on average (T_g), resulting in 250,512 GP practices×months observations. Once this dataset is merged with the Quality and Outcomes Framework series, we obtain data on Depression Prevalence and List size. As it can be seen, the merge is not completely successful and we can only obtain additional information for 7,173 GP practices, resulting in 248,599 GP practices×months observations. The same explanation applies for Panel C. Given that Panel B shows characteristics of the final sample satisfying both requirements, there is no difference in the values of columns (4) and (5), between AD doses and no. of monthly observations, on one hand, and depression prevalence and list size, on the other.

SOURCES. English Prescribing Dataset (EPD) merged with Quality Outcome Framework (QOF) series.

Prescriptions data. To summarise these prescription data, Figure 2.7 shows the evolution of all therapeutic classes considered in the methodology, and Table 2.9 presents main summary statistics of daily doses variables (Y^j_{gt}) and de-trended series ($\tilde{y}^j_{gt}$). Column (1) of the table shows the average monthly prescriptions, during the pre-treatment period, of each therapeutic class. The largest ones correspond to lipid-lowering drugs, hypertensives, antisecretory, antidepressants and nitrates, calcium-channel blockers and other antianginal drugs. And the smallest to drugs used for parkinson and related disorders, chronic bowel disorders, drugs affecting bone metabolism, rheumatic disease and gout, and anti-epileptics. Despite the fact that, as shown in column (4), most therapeutic area exhibit higher levels in the treatment period, this does not take into account the change in trends. To address this, we also show the distribution of the de-trended series of each therapeutic class, constructed according the the first step in the methodology. As shown in column (2), and as expected, in the pre-treatment period the de-trended series of each therapeutic class has mean of zero (as de-trended trend-stationary processes are centred around zero). During the pandemic most therapeutic areas exhibit an average decrease in levels (see column (5)), with the greatest decrease observed for drug groups representing high volume of prescriptions. Additionally, columns (3) and (6) shows the volatility of the residuals (or estimated shock) which, once again, are associated with higher prescription volumes classes.

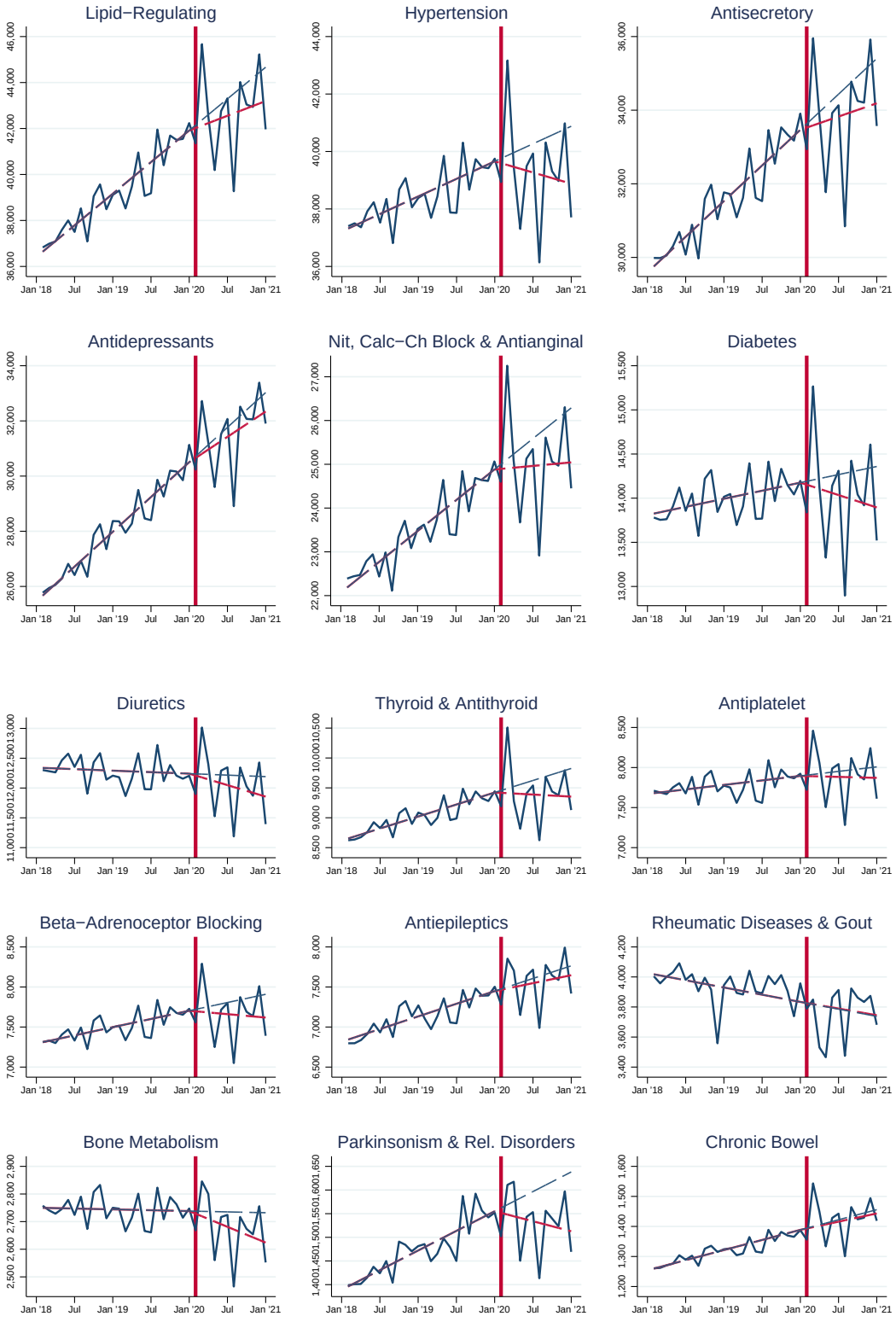


FIGURE 2.7: Evolution of selected therapeutic classes before and after COVID-19

NOTES. The figure shows the evolution of 14 selected therapeutic areas averaged across GPs for the 3-year period between February 208 and January 2021. That is $Ave_g[Y^j_{gt}]$ for each $j = 1, ..., J$ and $t = 1, ..., T$. They are sorted from left to right and from top to bottom in terms of volume prescribed in 2019.
SOURCE. English Prescribing Dataset (EPD).

TABLE 2.9: Summary statistics prescription data

	(1)	(2)	(3)	(4)	(5)	(6)
	Pre-treatment period ($t \leq T_0$)			Treatment period ($t > T_0$)		
	Y_{gt}^j	$\tilde{y}_{gt}^j = \hat{\varepsilon}_{gt}^j$		Y_{gt}^j	$\tilde{y}_{gt}^j = \widehat{[\chi_{gt}^j + \varepsilon_{gt}^j]}$	
	Mean	Mean	SD	Mean	Mean	SD
Antidepressants (AD)	28,065.10	0.00	2,660.22	31,578.16	-263.56	5,547.40
Antiepileptics	7,140.16	0.00	766.24	7,577.34	-29.99	1,497.98
Antiplatelet	7,775.66	0.00	779.75	7,915.46	0.36	1,515.19
Antisecretory	31,603.65	0.00	2,895.62	33,910.58	-660.70	6,145.21
Beta-Adrenoceptor Blocking	7,498.70	0.00	766.57	7,682.32	-110.95	1,526.09
Chronic Bowel	1,320.89	0.00	253.84	1,426.49	15.02	385.18
Diabetes	13,989.21	0.00	1,519.56	14,062.61	-191.67	3,028.07
Diuretics	12,270.30	0.00	1,278.59	12,085.39	-61.38	2,497.48
Drugs Affecting Bone Metabolism	2,742.19	0.00	366.89	2,683.15	-50.06	678.63
Hypertension	38,437.27	0.00	3,676.19	39,385.49	-771.22	7,496.46
Lipid-Regulating	39,220.16	0.00	3,696.16	42,774.10	-437.02	7,638.35
Nit, Calc-Ch Block & Antianginal	23,490.38	0.00	2,287.80	25,084.14	-401.84	4,782.68
Parkinsonism & Rel. Disorders	1,475.37	0.00	276.76	1,534.45	-68.92	457.15
Rheumatic Diseases & Gout	3,939.67	0.00	519.7	3,762.49	-103.93	952.03
Thyroid & Antithyroid	9,027.33	0.00	970.8	9,418.35	-190.52	1,923.07
N		159,046			79,497	

NOTES. The table presents summary statistics of the prescription variables Y_{gt}^j and $\tilde{y}_{gt}^j$, the latter constructed according to the first step of the methodology. Pre-treatment period runs from February 2018 to January 2020 (24 months), and treatment period from February 2020 to January 2021 (12). Column (1) shows mean monthly prescription levels (in daily doses) across practice, for each therapeutic class, in the pre-treatment period, i.e. $Ave_{gt}[Y_{gt}^j]$ for each j and for the pre-treatment period. Columns (2) and (3) show the mean and SD of the de-trended monthly prescription levels for each class in the pre-treatment period, i.e. $Ave_{gt}[\tilde{y}_{gt}^j]$ for each j and for $t \leq T_0$. Columns (4)-(6) show the same information as columns (1)-(3), respectively, but now computed for the treatment period. Number of observations N correspond to practice-months pairs, i.e. $G \times T_0$ for columns (1)-(3), and $G \times (T - T_0)$ for columns (4)-(6).

Mobility, Claimant Count and Infection cases data. Table 2.10 presents main summary statistics for mobility-in-residence, mobility-in-workplaces, claimant count and coronavirus infections variables. The values presented in this table correspond to the contemporary values (that is, in t) as well as lagged values (t-1 and t-2) as in the way they enter the model in eq (2.12). Columns (1) shows only the mean monthly value of the claimant count rate in the treatment period, which is 2.5 per cent of the districts population. The values of mobility and new infection cases in the pre-treatment period are set equal to zero as their data is only available from February 2020 onwards (thus the zero mean for these two variables). In the treatment period, the mean monthly mobility-in-residential places is 12 per cent higher with respect to baseline levels; mobility-in-workplaces decrease by 35 per cent, with respect to baseline levels. The claimant count during the pandemic increase on average to 5.6 per cent. And the new infection cases rate was 357 per 100,000 population on average, in any month.

TABLE 2.10: Summary statistics on mobility, claimant count and new coronavirus infections

	(1)	(2)	(3)	(4)
	Pre-treatment period ($t \leq T_0$)		Treatment period ($t > T_0$)	
	Mean	SD	Mean	SD
Mobility: Residency (t)	N/A	N/A	13.81	7.77
Mobility: Residency (t-1)	N/A	N/A	12.14	8.30
Mobility: Residency (t-2)	N/A	N/A	11.04	8.89
Mobility: Workplaces (t)	N/A	N/A	-38.90	19.46
Mobility: Workplaces (t-1)	N/A	N/A	-34.80	21.72
Mobility: Workplaces (t-2)	N/A	N/A	-31.95	23.63
Claimant Count rate (t)	2.54	1.21	5.87	2.24
Claimant Count rate (t-1)	2.52	1.20	5.58	2.32
Claimant Count rate (t-2)	2.50	1.19	5.29	2.36
Infection Cases rate (t)	N/A	N/A	357.82	597.69
Infection Cases rate (t-1)	N/A	N/A	348.11	600.41
Infection Cases rate (t-2)	N/A	N/A	221.33	413.15
N for vars in t	159,046		79,497	
N for vars in t-1	152,418		79,494	
N for vars in t-2	145,791		79,496	

NOTES. The table presents summary statistics of the mobility in residential place, mobility in workplaces, unemployment benefits claimant count, and COVID-19 new infection cases. These statistics are computed at the LTLA-month level, that is, they are averaged across the 312 local authority districts $\times$ the number of months comprising both pre- and treatment periods.

Figure 2.8 shows the evolution of mobility-in-residence, claimant count and coronavirus infections cases across the 312 (LTLA) districts in every month. Thick solid lines represent the mean value for each variables in every month, long-dashed lines represent the mean values $\pm$ 1 standard deviation and short-dashed lines represent the mean values $\pm$ 2 standard deviations. Up to April 2020 mobility inside residential places experience a 27 percent increase with respect to the baseline levels as the implementation of the first national lock-down began. From then on mobility is reduced toward the summer, to a 9 percentage increase (wrt baseline) in September, and then increases to a 20 percent increase (wrt baseline) in January 2021, coinciding with the second national lock-down. Variation in residence mobility is relatively stable and the monthly standard deviation (sd) from April onwards is around 3 percentage points. We do not show the evolution of mobility in workplaces as both have a similar opposite distribution (in fact they have a correlation of -0.9377). For the two-year period before February 2020 the claimant count went from 1.7 to 2.6 percent of the districts population (with sd of around 1.1), and during the pandemic the claimant count more than doubles reaching a level of 5.8 from May 2020 onwards (with sd of 1.8). Finally, between February and September 2020 monthly infection cases per 100,000 population averages 82 across districts (sd of around 63). During the following months the figures increase to an average of 734 in October, 869 in November, and to 1,475 in December (with sd of 623); then it decreases significantly to 104 in January 2021 (sd of 49).

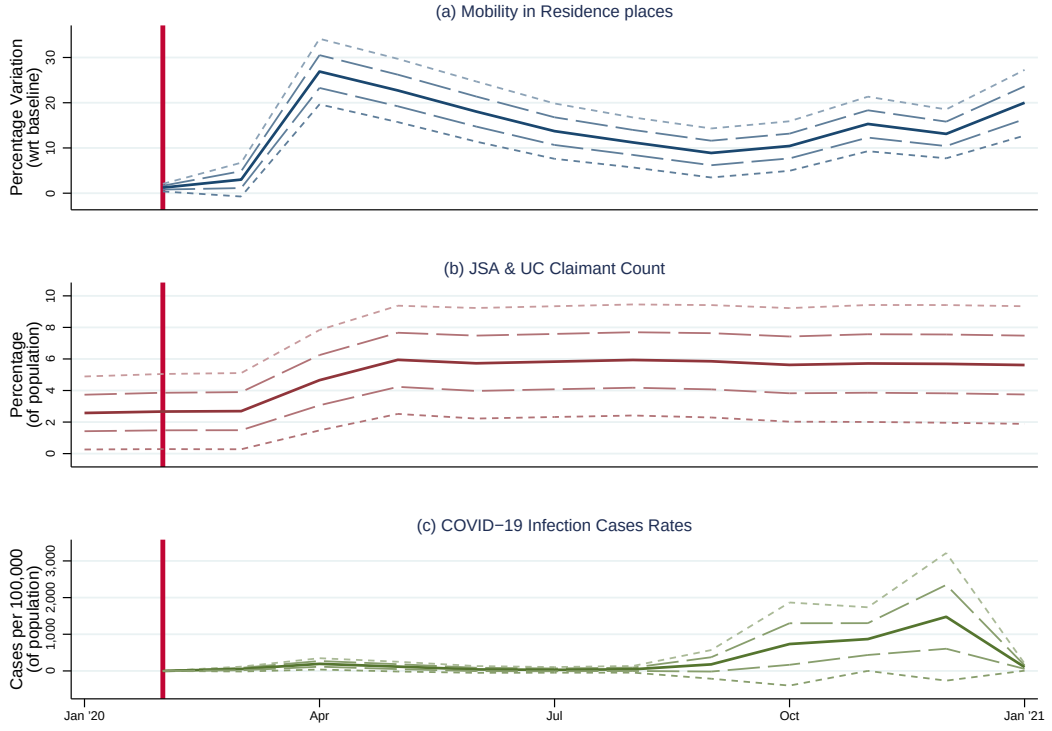


FIGURE 2.8: Mobility, Claimant Count and COVID-19 Infection Cases at the district level

NOTES. The figure shows mean values of the variables: mobility in residential places, JSA and UC claimant count and COVID-19 new infection cases rate, averaged across 312 LTLA districts. Thick solid lines represent the mean value for each variables in every month, long-dashed lines represent the mean values $\pm$ 1 standard deviation and short-dashed lines represent the mean values $\pm$ 2 standard deviations.

TABLE 2.11: Distribution of coefficients $\hat{\omega}_{jg}$

	Panel A						Panel B					
	Models with all J therapeutic classes						Models with therapeutic classes pre-selected by linear LASSO					
	GP-specific OLS regressions					Pooled OLS	GP-specific OLS regressions					Pooled OLS
	(1) Mean	(2) SD	(3) p10	(4) Median	(5) p90	(6)	(7) Mean	(8) SD	(9) p10	(10) Median	(11) p90	(12)
Antiepileptics	0.393	0.968	-0.714	0.376	1.52	0.687***	0.407	0.84	-0.554	0.379	1.384	0.687***
Antiplatelet	0.005	1.504	-1.668	0.011	1.651	-0.092***						
Antisecretory	0.225	0.443	-0.271	0.21	0.748	0.292***	0.227	0.375	-0.199	0.21	0.668	0.288***
Beta-Adrenoceptor Blocking	0.253	1.394	-1.296	0.204	1.864	0.275***	0.264	1.182	-1.083	0.219	1.615	0.264***
Chronic Bowel	0.215	2.159	-2.115	0.215	2.521	0.323***	0.210	1.831	-1.775	0.196	2.227	0.323***
Diabetes	0.070	0.638	-0.624	0.055	0.781	0.015						
Diuretics	0.020	0.733	-0.817	0.028	0.852	0.000						
Drugs Affecting Bone Metabolism	0.095	2.224	-1.891	0.116	2.174	0.107***	0.118	1.857	-1.578	0.142	1.901	0.098***
Hypertension	0.033	0.405	-0.425	0.034	0.488	0.046***	0.042	0.337	-0.337	0.046	0.43	0.045***
Lipid-Regulating	0.060	0.461	-0.417	0.061	0.541	0.055***	0.072	0.378	-0.309	0.074	0.447	0.049***
Nit, Calc-Ch Block & Antianginal	0.093	0.545	-0.522	0.081	0.72	0.055***	0.091	0.458	-0.424	0.079	0.626	0.055***
Parkinsonism & Rel. Disorders	0.188	2.676	-2.239	0.166	2.614	0.147***	0.201	2.567	-1.902	0.158	2.256	0.139***
Rheumatic Diseases & Gout	0.250	1.220	-1.093	0.212	1.656	0.410***	0.256	1.04	-0.923	0.219	1.478	0.413***
Thyroid & Antithyroid	0.234	0.906	-0.788	0.204	1.284	0.320***	0.239	0.768	-0.643	0.215	1.128	0.319***
N						159,046						159,046
R^2	0.835	0.12	0.665	0.861	0.966	0.782	0.766	0.155	0.545	0.791	0.948	0.782

NOTES. The table presents the distribution of the OLS coefficients from estimating GP-specific regressions of model in eq. (2.7). Mean and SD of the variables used in these regressions are presented in Table 2.9 columns (2)-(3). Panel A columns (1)-(5) presents main statistics of $\hat{\omega}_{jg}$, computed for each class j across 6,627 GP. Column (6) presents coefficients $\hat{\omega}_g$ estimated from the following pooled OLS regression: $\hat{\varepsilon}^{AD}_{gt} = \sum_{j \in J} \omega_j \cdot \hat{\varepsilon}^{j}_{gt} + v_{gt}$. N corresponds to ($G \times T_0 = 6,627 \times 24$). Panel B show the same information as Panel A in which the GP-specific regressions of model in eq. (2.7) are estimated on a subset of therapeutic classes pre-selected by fitting a linear LASSO (with lasso penalty parameter selected by *adaptive* method), allowing to select from the initial set of 14 those that best predict the shocks to AD in the pre-treatment period. As can be seen from the table, the LASSO excludes the following classes: Antiplatelet, Diabetes and Diuretics. Statistical significance stars associated to coefficients in columns (6) and (12) come from robust standard errors clustered at the GP practice-level. *** $p < 0.001$.

TABLE 2.12: Robustness check: alternative ways of implementing second step

	(1) Daily Doses	(2) Patients	(3) Normalised Daily Doses	(4) Normalised Patients
Panel A: From GP practice-specific OLS regressions with full set of 14 therapeutic classes				
COVID-19	168.714***	5.689***	151.680***	5.194***
	(24.629)	(0.749)	(31.240)	(0.904)
Constant	0.038	0.001	0.065	0.001
	(8.208)	(0.250)	(10.411)	(0.301)
N	238543	238543	238543	238543
R^2	0.135	0.137	0.152	0.133
Panel B: From Pooled OLS regression with full set of 14 therapeutic classes				
COVID-19	171.537***	5.307***	161.380***	5.122***
	(16.850)	(0.590)	(19.191)	(0.696)
Constant	0.049	0.007	0.074	0.012
	(5.615)	(0.197)	(6.396)	(0.232)
N	238543	238543	238543	238543
R^2	0.092	0.104	0.095	0.095
Panel C: From GP practice-specific OLS regressions with therapeutic classes pre-selected by linear LASSO				
COVID-19	160.442***	5.513***	150.542***	5.319***
	(21.585)	(0.655)	(28.177)	(0.780)
Constant	0.042	0.001	0.075	0.001
	(7.194)	(0.218)	(9.390)	(0.260)
N	238543	238543	238543	238543
R^2	0.128	0.130	0.147	0.124
Panel D: From Pooled OLS regression with therapeutic classes pre-selected by linear LASSO				
COVID-19	160.517***	5.022***	150.850***	4.869***
	(16.827)	(0.585)	(19.215)	(0.690)
Constant	0.048	0.007	0.074	0.012
	(5.608)	(0.195)	(6.404)	(0.230)
N	238543	238543	238543	238543
R^2	0.092	0.103	0.095	0.094

NOTES. The table presents the results from four different ways of implementing the second step for the construction of the counterfactual group as described above. Panel A presents the exact results shown in Table 2.4. Panel B shows the results where the second step is implemented from estimating a pooled OLS regression and applying the same set of coefficients to the de-trended prescription series of each practice. These coefficients are presented in Table 2.11 column (6). Panel C shows the results from the original second step as described in the methodology using, not the set of 14 therapeutic classes, but 11 pre-selected by estimation of a linear LASSO. The resulting distribution of performing practice-specific regressions using these 11 classes is presented in Table 2.11 columns (7)-(11). Finally, Panel D shows the results where the second step is implemented from estimating a pooled OLS regression, with the 11 therapeutic classes pre-selected by LASSO, and applying the same set of coefficients to the de-trended prescription series of each practice. These coefficients are presented in Table 2.11 column (12). The table presents the OLS estimates from model in eq. (2.9). Dependent variable in each model is given by the column-name. All models control for GP fixed effects. Clustered standard errors at the GP level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. N represents number of GP-months pairs ($G \times T$). See Appendix 2.7 for the details on the construction of the dependent variables in models (2)-(4).

TABLE 2.13: Therapeutic Class-Placebo Test

	(1) Antiepileptics	(2) Antiplatelet	(3) Antisecretory	(4) Beta-Adrenoceptor Blocking	(5) Chronic Bowel	(6) Diabetes	(7) Diuretics
COVID-19	70.833***	78.691***	-345.232***	-5.224	26.438***	80.321***	146.706***
	(11.762)	(8.912)	(25.858)	(7.903)	(5.317)	(20.574)	(17.626)
Constant	-0.001	-0.003	0.012	0.004	-0.003	-0.018	-0.011
	(3.920)	(2.970)	(8.618)	(2.634)	(1.772)	(6.857)	(5.874)
N	238543	238543	238543	238543	238543	238543	238543
R^2	0.135	0.150	0.141	0.126	0.129	0.131	0.152

	(8) Drugs Affecting Bone Metabolism	(9) Hypertension	(10) Lipid Regulating	(11) Nit, Calc-Ch Block & Antianginal	(12) Parkinsonism & Rel. Disorders	(13) Rheumatic Diseases & Gout	(14) Thyroid & Antithyroid
COVID-19	-37.076***	-304.378***	-21.043	-14.042	-48.089***	-29.128**	-58.257***
	(10.009)	(32.418)	(31.496)	(22.856)	(8.721)	(10.370)	(12.356)
Constant	-0.003	-0.004	0.057	0.007	0.007	-0.001	-0.003
	(3.336)	(10.804)	(10.496)	(7.617)	(2.906)	(3.456)	(4.118)
N	238543	238543	238543	238543	238543	238543	238543
R^2	0.184	0.146	0.142	0.141	0.181	0.147	0.127

NOTES. The table presents the results of therapeutic class-placebo test, after applying the methodology to each of the J (14) therapeutic classes used to construct the counterfactual for antidepressants. It shows OLS coefficients from the model in (2.9), using as dependent variable daily doses normalised by the practice's adult population size, i.e. ΔnY^j_{gt} for each therapeutic class j given by the column-name. All models control for GP fixed effects. Clustered standard errors at the GP level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. N represents number of GP-months pairs ($G \times T$).

TABLE 2.14: RMSPE ratio

	Average RMSPE ratio
Antidepressants	**4.72**
Antiepileptics	4.53
Antiplatelet	4.59
Antisecretory	4.65
Beta-Adrenoceptor Blocking	4.47
Chronic Bowel	4.53
Diabetes	4.58
Diuretics	4.61
Drugs Affecting Bone Metabolism	4.62
Hypertension	4.57
Lipid-Regulating	4.69
Nit, Calc-Ch Block & Antianginal	4.58
Parkinsonism & Rel. Disorders	4.63
Rheumatic Diseases & Gout	4.46
Thyroid & Antithyroid	4.45

NOTES. The table presents the RMSPE ratio for each therapeutic class averaged across GP practices.

Chapter 3

Price versus clinical guidelines in primary care *statin* prescribing: a retrospective cohort study and cost simulation model

Abstract

In this paper we provide a descriptive analysis whose objective is to investigate the relative impact of generic entry following patent expiration, on one hand, and the publication of clinical guidelines, on the other, on General Practice's prescribing behaviour using statins as an exemplar. For this exercise we retrieve data from the Royal College of General Practitioners Research & Surveillance Centre (RCGP RSC) database, containing patient-level information from a representative sample of general practices in England between 2004-2018, issuing first-time statin prescriptions for new patients at risk of cardiovascular diseases. Our findings indicate that the general trends of statin' prescriptions were largely driven by a decrease in acquisition costs triggered by patent expiration, preceding the guidelines issued by the National Institute for Health and Care Excellence (NICE), which themselves did not seem to affect prescription trends. Additionally, significant heterogeneity is observed in the prescription of the most cost-effective statin across general practices. A cost simulation exercise shows that, between 2004 and 2018, the National Health Service (NHS) could have saved £2.8bn (around 40% of the £6.3bn spent on statins during this time) if all GP practices had prescribed only the most cost-effective treatment. Our results suggest that there is potential for large savings for the NHS if new and, whenever possible, ongoing patients are promptly switched to the first medicine that becomes available as generic within a therapeutic class as long as it has similar efficacy to still-patented medicines.

Keywords: Statins, Prescribing Behaviour, Clinical Guidelines, Patent Expiration

3.1 Introduction

In a context where national health systems of all high- and medium-income countries are confronted by ballooning costs of caring for an ageing population and an increase in prevalence of long-term conditions, promoting cost-effective prescribing represents an important part of controlling healthcare expenditure (World Health Organization, 2019; Dall et al., 2013). In the English National Health Service (NHS), the National Institute for Health and Care Excellence (NICE) publishes national guidance aimed at promoting clinical and cost-effective evidence-based recommendations for the clinical management of different conditions. In therapeutic markets where treatments have similar safety and effectiveness, NICE recommendations may vary over time following changes in acquisition costs, e.g. due to patent expirations and the ensuing entry of generics. However, persistence of prescribing habits and prescribers' lack of awareness of medicines' actual cost may mean that the uptake of NICE recommended medicines can vary substantially across general practitioners and practices, despite efforts at local level, including Clinical Commissioning Groups (CCG), to encourage more cost-effective prescribing (Mason, 2008; Wathen and Dean, 2004; Harrison et al., 2003; Bedson et al., 2013; Walker et al., 2019). Since low responsiveness to adopt NICE recommendations can substantially undermine NHS efforts to contain drug expenditure, it is important to understand the pervasiveness of such behaviour.

Statins represent an ideal market to investigate the relative importance of prices and clinical guidelines in explaining prescribing behaviour for at least two reasons. First, statins are the most widely used treatment for primary and secondary prevention of cardiovascular disease, conditions with an estimated cost to the NHS of roughly £7.4 billion a year (UK Health Security Agency, 2019; Luengo-Fernandez et al., 2006). Second, there are five main events that have shaped the statins market over the last two decades. In May 2003, simvastatin (brand name *Zocor*) lost patent protection and became available as a generic drug. In January 2006, NICE published Technological Appraisal 94 (TA-94), stating that all statins were equally effective from a therapeutic point of view, advising general practitioners to consider costs of statins when choosing the initial treatment and advising that simvastatin was the cheapest (NICE, 2006). Clinical Guideline 67 (CG-67), released in May 2008, stated that treatment initiation should start with simvastatin (NICE, 2008). In May 2012, atorvastatin (brand name *Lipitor*) lost patent protection and became available generically. Finally, two years later, in May 2014, NICE published Clinical Guidelines 181 (CG-181) recommending atorvastatin as initial treatment (NICE, 2014). The reduced cost after patent expiration coupled with its relatively greater potency made atorvastatin the most cost-effective statin in the market.

Using statins as an exemplar, this study investigated the prescription dynamics in a representative sample of the English primary care sector between 2004 and 2018. First,

we explored the relationship between aggregate prescription trends and two sets of events that shaped the statin market: patent expirations and generic entry on the one hand, and publication of national clinical guidelines on the other. Second, we investigated variation in prescribing activity across general practices. Third, we quantified the forsaken savings for the NHS by assuming perfect therapeutic substitution, that is by comparing actual treatment choices to a hypothetical scenario where only the most cost-effective treatments are prescribed.

3.2 Methods

Data. Our data are retrieved from Royal College of General Practitioners Research and Surveillance Centre (RCGP R&SC) database, a nationally representative sample of 243 general practices in England. The population representativeness of this database has been addressed in previous studies, including its representativeness of the distribution of cardiovascular disease in England (Correa et al., 2016; Hinton et al., 2018). From this database, we retrieved all first prescription episodes for more than 400,000 patients treated with statins between Q3-2003 and Q4-2018. This database contains complete information of each prescription issued and an anonymised identity code of the general practice issuing the prescription.

We also retrieved from the Prescription Cost Analysis (PCA) database yearly statistics on the total quantities of each drug prescribed in primary care in England, with the corresponding total spending, obtained from net ingredient cost (NIC).[1] This is the basic cost of a drug as used in primary care. NIC is used in Prescription Services reports and other analyses, as it standardises cost throughout prescribing nationally, and allows comparisons of data from different sources. By aggregating the total spending of each strength of statin prescribed and dividing it by the corresponding total quantity, we obtained a measure of the average acquisition cost per strength of each statin in each year. All cost figures are expressed in constant 2018 GBP using the GDP deflators at market prices, and money GDP.[2]

In Appendix A, we compare the prescription data in the PCA dataset to those in the RCGP R&SC dataset to demonstrate that the latter constitutes a representative sample of national prescription of statins.

Trends and heterogeneity in statins prescription. Since 2003, the statins market has experienced five exogenous changes to prices and clinical standards as explained above

[1] See https://digital.nhs.uk/data-and-information/publications/statistical/prescription-cost-analysis.

[2] See https://www.gov.uk/government/statistics/gdp-deflators-at-market-prices-and-money-gdp-march-201

that may have triggered changes in general practitioners' prescribing choice. To document how prescription trends change in proximity of those events, we plotted the average proportion (across the 243 practices in the RCGP R&SC dataset) of new patients starting with one of the five statins for the period 2004 to 2018, as well as the average acquisition cost per defined daily dose for each statin. To explore heterogeneity in prescription patterns, we split general practices in the RCGP R&SC dataset into quintiles for every month in the data according to their share of new patients treated with simvastatin and plotted the average shares of new patients treated with simvastatin in each of the resulting five groups. Although this offers an insight into the evolution of overall heterogeneity in the data, it does not allow us to characterise persistency in general practitioners' prescribing choices. Hence, we additionally plotted the average share of new patients treated with simvastatin, keeping the composition of groups fixed at the quintile computed at Q3-2003.

Cost savings simulation. According to NICE, all statins are equally therapeutically effective. As stated in TA-94 (2006), *'from the evidence available [. . .] [and] for the purpose of initiating therapy, there were no data on clinical events to suggest the superiority of any one statin over all the others in reducing cardiovascular events'* (NICE, 2006; Law et al., 2003; Unit, 2005; Moon and Bogle, 2006; Armitage, 2007; Amarenco and Labreuche, 2009). Under the assumption that general practitioners cannot consistently anticipate whether a new patient would benefit from starting treatment with any given statin different from the one recommended by NICE, we evaluated prescription decisions according to a cost-minimisation criterion. Specifically, we quantified the potential savings for the NHS by comparing the actual cost of the observed prescription decisions with a hypothetical cost constructed by substituting the actual original treatments with a therapeutically similar treatment containing either simvastatin (for drug treatments initiating before May 2012) or atorvastatin (for those initiating from May 2012 onwards). By computing the difference between actual and hypothetic costs, we obtained a measure of the potential savings, both in absolute and relative terms.

The cost simulation was performed under two different scenarios. In the *first scenario*, the analysis was limited to the first 28 days of treatment for patients newly treated with statins. By focusing only on the first prescription episode, we compared the evolution of the spending on statin treatments using the same unit on analysis in different time periods, leaving aside the problem of following patients throughout their drug-treatment history. Clearly, the absolute value of savings obtained by considering only the first prescription episode for new patients is a partial account of the overall potential savings, as patients treated for cardiovascular disease risk will usually be on treatment indefinitely.

For this reason, we considered a *second scenario* where we computed hypothetical savings if practitioners had changed all (new and ongoing) patients to simvastatin (up

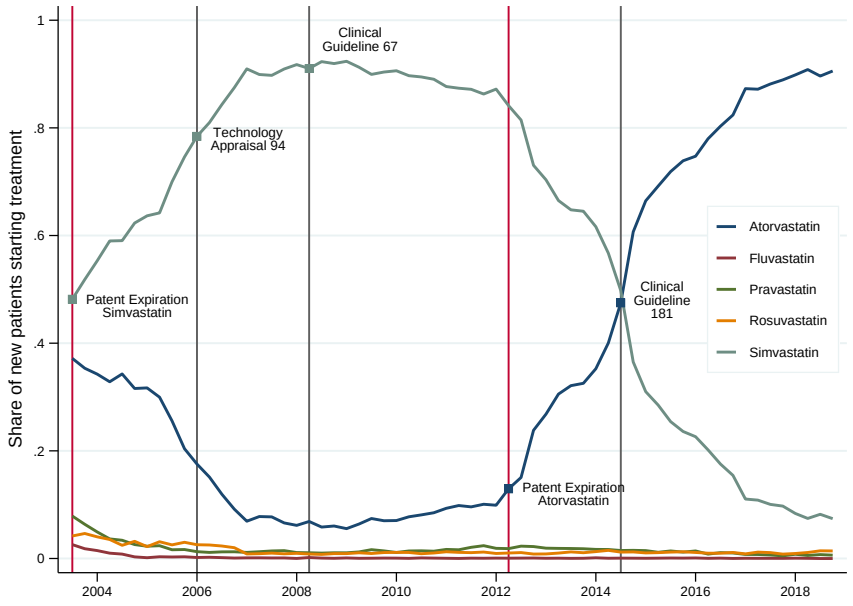
to May 2012) and atorvastatin (after May 2012). This second simulation can be considered an upper bound to the absolute savings under the strong assumption that existing patients could be immediately switched to the NICE-recommended treatments, regardless of any patient's preference or professional decision that led to the observed prescription choices. A detailed explanation of the methodology used for our cost simulation is presented in Appendix B.

3.3 Results

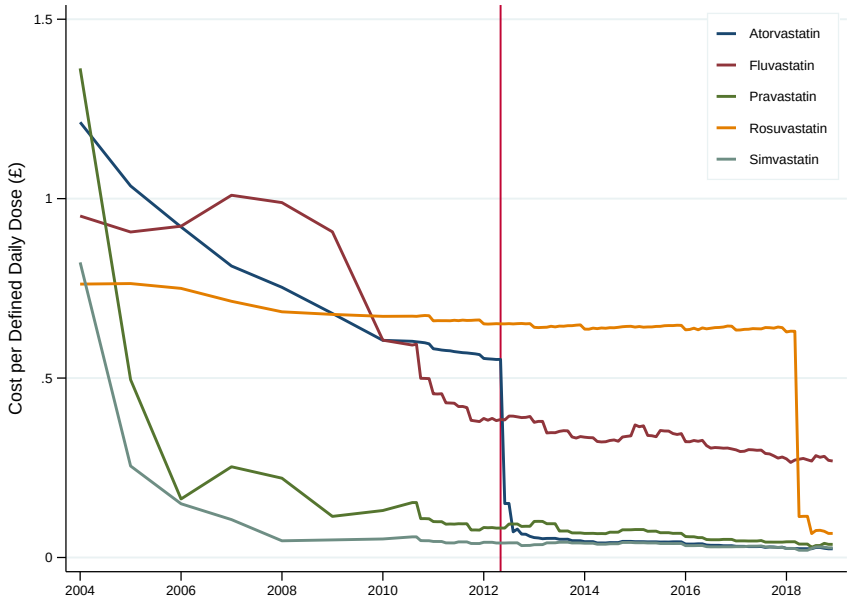
Trends in prescription and price. Figure 3.1 (A) plots the evolution of the market shares for new patients starting treatment with statins between 2004 and 2018, using the RCGP R&SC database. In the time window considered, simvastatin and atorvastatin were the most frequently prescribed among the five statins, representing approximately 96% of all initial prescriptions. The dominance of these two drugs in treating cardiovascular diseases resulted in the evolution of their shares following mirror image patterns.

The share of simvastatin increased rapidly after its patent expiration in 2003, from around 50% to more than 90% in May 2008, when NICE published CG-67. While this guideline explicitly recommended simvastatin for treatment initiation, the percentage of new patients prescribed simvastatin stayed constant over the next four years up to May 2012, and, if anything, slightly decreased. We also note that the introduction of TA-94 in 2006 failed to accelerate the uptake of simvastatin. Upon atorvastatin's patent protection expiration in May 2012, simvastatin's share started decreasing steadily from around 85% in 2012 to around 10% in 2018. Once again, publication of CG-181 in 2014, updating the recommendation for treatment initiation to atorvastatin, had minimal effect in speeding up the declining trend of simvastatin.

Figure 3.1 (B) shows the average acquisition costs per daily defined dose of each statin over time. The figure makes apparent the large drop in the acquisition cost of simvastatin soon after patent expiration of Zocor. Similarly, a sharp drop in acquisition cost for atorvastatin (virtually similar to simvastatin) was observed shortly after Lipitor patent expiration.



(A) Share of new patients on each drug.

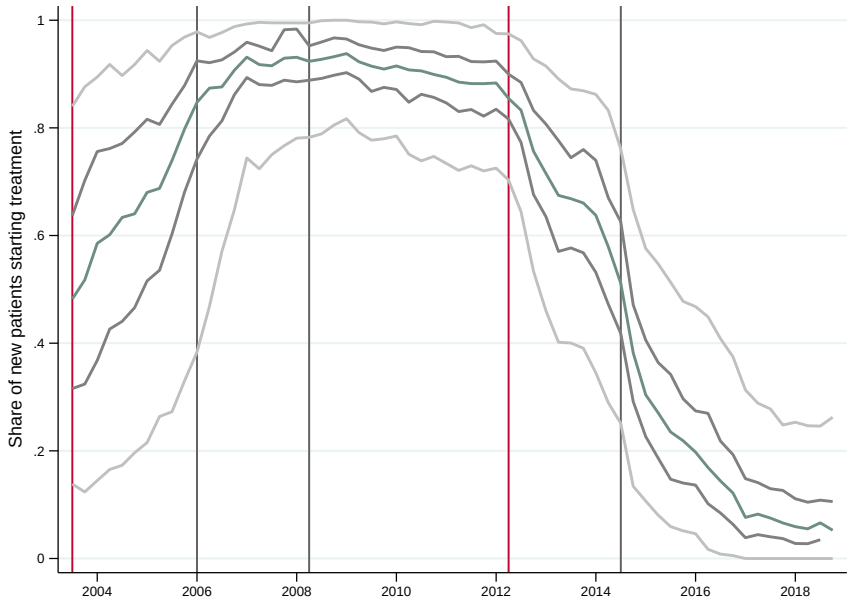


(B) Average acquisition cost per defined daily dose (DDD) of each drug.

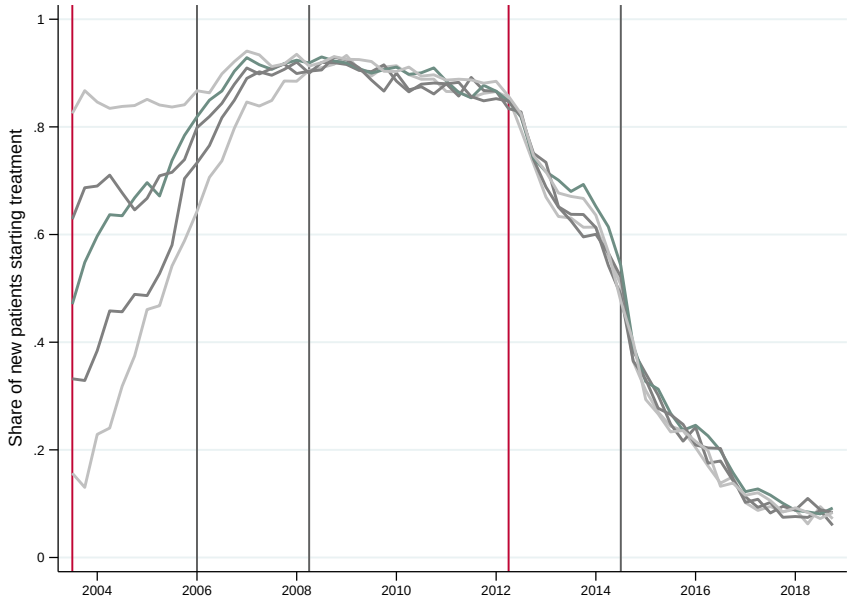
FIGURE 3.1: Trends in the statins prescribed for drug treatment initiation.

NOTES. Panel (A) shows the proportion of new patients starting drug treatment with each statin over time. The main five events are marked with vertical lines and small squares. The two vertical red lines marks the patent expiration of Zocor (simvastatin) and Lipitor (atorvastatin) in 2003 and May 2012, respectively; and the grey vertical lines indicate the publishing date of NICE's statin-related national guidance. Panel (B) shows the average acquisition cost per DDD for each statin over time. Statins' DDDs (or daily strength per day of treatment) established by the WHO are the following: for atorvastatin, 20mg; fluvastatin 60mg; pravastatin 30mg; rosuvastatin 10mg; and simvastatin 30mg.
SOURCE. Panel (A) from RCGP R&SC database, and panel (B) from Prescription Cost Analysis data series.

Heterogeneity in prescriptions across general practices. Figure 3.2 (A) presents average shares of new patients treated with simvastatin for each of the five quintiles of the general practices' prescription distribution. The figure reveals significant heterogeneity in prescribing choices across general practices during the time window of our study. When the simvastatin patent expired in May 2003, the proportion of patients treated with simvastatin ranged from less than 20% for general practices in the bottom quintile to more than 80% for the top quintile. The period up to 2006 saw an increase in the proportion of new patients treated with simvastatin across all general practices. At the time of the TA-94 introduction (January 2006), the difference between the second and fifth quintiles was around 20 percentage points, while the difference between first and fifth quintiles was still more than 50 percentage points. Following the introduction of CG-67 in May 2006, differences across practices fluctuated around 25 percentage points with most compliant with the NICE guideline practices treating almost all of their new patients with simvastatin, while least compliant practices prescribed simvastatin to less than 80% of their patients. Heterogeneity in prescription increased again following atorvastatin patent expiration in May 2012, when the difference in the share of new patients being prescribed simvastatin between the top and bottom quintiles reached about 50 percentage points. The subsequent CG-181 further reduced the overall levels of simvastatin prescriptions across the distribution but did little to reduce heterogeneity in the share of patients treated with simvastatin in the following years, with the difference between the top and bottom quintiles remaining at about 25 percentage points.



(A) Share of new patients on simvastatin with quantiles calculated quarterly.



(B) Share of new patients on simvastatin with fixed quantile composition calculated on Q3-2003.

FIGURE 3.2: Heterogeneity in initial prescriptions at the general practice level.

NOTES. Panel (A) shows the average proportion of new patients treated with simvastatin within five quintiles of general practices ranked by proportion of simvastatin prescriptions (e.g. the top line represents the average proportion of patients initially treated with simvastatin, by the top 20th percent of general practices, etc.), where the quintiles of practices are obtained separately for each month (i.e. practices in each quantile may be different). Panel (B) shows the average proportion of new patients treated with simvastatin but for quintiles of practices obtained at Q3-2003, (i.e. the practices in each quintile are the same).
SOURCE. RCGP R&SC database.

Figure 3.2 (B) tracks the evolution of prescription for five quintiles of general practitioners as constructed in Q3-2003. The dynamics up to 2006 suggest the uptake of cost-effective prescribing for general practitioners in the lower quintiles is rather slow. However, the disappearance of major differences in prescribing among the five groups from 2007 onwards indicates that any given general practice does not systematically deviate from prescribing the cost-effective statin. These dynamics suggest that the overall heterogeneity observed in panel (A) is due to slow learning and fluctuation between cost-effective and non-cost-effective prescribing.

Spending simulations. Spending simulation results for the *first scenario*, which considered only the first prescription episode, are presented on the left-hand side panel of Table 3.1. Column (1) shows the estimated number of new patients taking statins in every year from Q3-2003 to Q4-2018. We observed a decrease in the number of new patients from 1.15 million in 2004 to 782,000 in 2018. The total cost of the first prescription episode for these new patients decreased from £26.9 million in 2004 to £862,000 in 2018 (a 96.8 percentage decrease), due to the reduction in the number of new patients as well as the acquisition costs of statins.

TABLE 3.1: Spending simulation exercise.

	First Scenario First Prescription Episode (first 28-days of treatment)					Second Scenario All prescriptions				
Year	(1) New patients (#)	(2) Actual spending (£)	(3) Hypothetical spending (£)	(4) Savings (£)	(5) Savings *wrt* actual spending (%)	(6) All patients (#)	(7) Actual spending (£m)	(8) Hypothetical spending (£m)	(9) Savings (£m)	(10) Savings *wrt* actual spending (%)
2004	1,150,730	26,896,100	26,611,542	284,559	1.06	3,834,190	992.88	927.00	65.88	6.64
2005	1,024,085	13,164,249	10,332,805	2,831,444	21.51	4,457,033	758.85	586.94	171.91	22.65
2006	1,099,334	8,623,615	6,651,031	1,972,584	22.87	5,139,742	707.44	508.42	199.02	28.13
2007	905,206	4,771,157	3,128,028	1,643,129	34.44	5,548,000	627.31	288.83	338.48	53.96
2008	894,058	3,136,696	1,401,840	1,734,856	55.31	5,968,499	540.70	132.41	408.30	75.51
2009	873,406	3,254,101	1,415,472	1,838,629	56.50	6,351,952	527.68	126.14	401.54	76.10
2010	797,556	3,237,395	1,377,895	1,859,500	57.44	6,580,198	513.58	131.90	381.69	74.32
2011	733,701	3,145,095	1,029,339	2,115,756	67.27	6,670,907	494.57	104.61	389.95	78.85
2012	812,435	2,667,155	1,093,112	1,574,043	59.02	6,873,148	318.38	102.40	215.97	67.84
2013	804,797	1,416,539	1,402,155	14,384	1.02	7,103,662	164.28	132.90	31.38	19.10
2014	795,653	1,397,926	1,218,241	179,685	12.85	7,198,232	156.88	116.77	40.11	25.57
2015	792,381	1,486,687	1,310,250	176,437	11.87	7,316,021	163.59	124.42	39.16	23.94
2016	772,918	1,212,744	1,054,905	157,839	13.02	7,426,410	143.89	104.48	39.41	27.39
2017	801,680	1,150,362	940,154	210,208	18.27	7,526,470	138.76	92.57	46.19	33.29
2018	782,504	862,076	783,514	78,561	9.11	7,600,486	95.12	79.60	15.52	16.31
		76,421,897	59,750,283	16,671,614	21.82		6,343.91	3,559.39	2,784.51	43.89

Overall, a saving of £16.67 million, or 21.8% of the total actual cost could have been realised for the first 28 days of treatment alone if general practitioners had prescribed simvastatin as initial drug treatment before May 2012 and atorvastatin after May 2012. Most of the hypothetical savings accrued over the period 2008-2012 when cheap generic versions of simvastatin became available and atorvastatin was still under patent protection. After May 2012, once generics of atorvastatin also entered the market, hypothetical savings on first prescription episodes were mainly accredited to switching from rosuvastatin to atorvastatin. However, the implied savings were not large as rosuvastatin only held a small share of the market.

Results for the *second scenario* regarding total prescriptions for all existing patients are presented in the right-hand side panel of Table 3.1. The total number of patients on treatment every year increased from 3.8 million in 2004 to 7.6 million in 2018. However, the significant drop in price due to generic entries led to a drastic drop in spending over the same period: from just under £1 billion in 2004 to £95 million in 2018, a 90-percentage decrease. Cumulate spending on statins over the period 2004-2018 totalled £6.3 billion and estimated potential savings were £2.8 billion, or 43.9% of the actual spending on statins. As previously, large savings could have been obtained in the period 2004-2012 by switching patients from atorvastatin, which was available only as a branded drug, to simvastatin, that was generic.

3.4 Discussion

In resource-constrained healthcare systems, promoting cost-effective prescribing behaviour is an important component of their cost-containment strategy. Using data on statins, we investigated how general practitioners' prescription choice in England changed in the face of (i) a large reduction in the cost of available treatments and (ii) the introduction of specific clinical guidelines. We demonstrated substantial increases in market shares for simvastatin and atorvastatin as their patents expired and generics entered the market, but well before the introduction of NICE guidelines recommending their use.

Those trends suggest that practitioners in primary care are sensitive to the price of alternative treatments, and that their choices even anticipate the recommendation of future clinical guidelines. Indeed, it took four years from the generic availability of simvastatin for NICE to explicitly recommend it for treatment initiation, by which point the share of new patients being prescribed the drug was already at 90%. Similarly, migration from simvastatin to atorvastatin started soon after a generic became available in May 2012, despite the fact that atorvastatin was only recommended as the preferred treatment in the updated NICE guideline two years afterwards.

Previous studies have shown that medicine management teams from CCGs do play a role in informing and influencing practices' and practitioners' prescription choices (Mason, 2008; Wathen and Dean, 2004; Harrison et al., 2003). Whether prescriptions are autonomously chosen by general practitioners or are influenced by the different actors within the primary care sector, our results suggest that ultimately prescribing decisions are more responsive to the acquisition cost of alternative treatments than clinical guidelines.

Although our analysis shows that, on average, practitioners treating patients at risk of cardiovascular events prescribed cost-effectively, we also identified substantial heterogeneity in prescription across practices, which remained even after the publication of

official guidelines. Our descriptive analysis indicates that some general practitioners took longer to adopt cost-effective prescribing and some switched in and out of cost-effective prescribing throughout the study period, generating large overall heterogeneity. A number of explanations can be offered for such behaviours. For example, medicolegally, general practitioners may have felt inhibited to change prescribing habits simply on the basis of cost, without having had guidelines to justify the decision (Avery et al., 2012). Equally, general practitioners' and patients' characteristics, practices' characteristics, geo-social conditions, as well as local CCG prescribing guidelines and monitoring activities may influence prescribing decisions (Goldacre et al., 2019). Moreover, statins have been widely perceived as causing side effects such as muscle pains (with intermittent media coverage influencing prescribing behaviour. See Matthews et al. (2016)). General practitioners and patients may have been reluctant to switch statins for fear of inducing adverse events (Ju et al., 2018).

Under the plausible assumption that general practitioners cannot consistently anticipate whether a new patient would benefit from taking a drug other than the one recommended by NICE, we evaluated prescription choices in this market according to a cost-minimisation criterion where choosing statins other than simvastatin (before May 2012) or atorvastatin (after May 2012) can be considered suboptimal. Our cost-savings simulation analysis suggested that low responsiveness comes at a high price for the NHS. Namely, if all new patients had received the most cost-effective treatment (as later recommended in the guidelines), the NHS could have saved around 22% of the actual spending on initial prescriptions. Looking at all prescriptions for new and on going patients, we compute savings of £2.7bn, mainly between 2004 and 2012, representing roughly 44% of total spending on statins during this period. We acknowledge that this figure is an upper bound of potential savings, based on the strong assumption that all existing patients on drug treatment could be seamlessly switched to other statins, without considering side effects (e.g. myopathy) or other practicalities (e.g. planned-patient reviews). Looking at the data, we found that around 7% and 12% of patients were switched to simvastatin and then switched away from it in the next 4 and 12 months, respectively. Although these numbers suggest that simvastatin cannot be used by a non-insignificant proportion of patients, there is no doubt that there were still large savings to be made by prescribing the most cost-effective statins.

We anticipate that the experience of statins would be similarly observed in other therapeutic areas where treatments have similar modes of action and comparable levels of efficacy, for instance angiotensin-converting enzyme (ACE) inhibitors and proton pump inhibitors. Looking ahead, our analysis suggests that cost-conscious centralised public health systems could save substantial sums if new and, whenever possible, on going patients are promptly switched to cost-effective alternatives, in particular when the first medicine in a therapeutic class loses patent protection. The observed heterogeneity in prescribing behaviour suggests that an important step forward towards achieving

this goal would be a timely dissemination of best practices, with the aim of promoting cost-conscious prescribing behaviour. In the UK, where general practices are grouping into Primary Care Networks and there is growing co-working and co-location with pharmacists, such collaborative efforts are likely to drive future prescribing. Given general practitioners' limited time available to acquire information on market developments (e.g. new medicines coming into the market or brand-name medicines losing patent protection) across all drug classes they prescribe, there is an important role for academic detailing as well as online/computerised systems and prompts such as ScriptSwitch, rather than paper-based (e.g. Prescribing Outlook), to educate and offer updated advice on cost-effective medicines while preserving physicians' freedoms to prescribe and patients' ability to discuss their preferred choice of treatment (Hire and Rushforth, 2013). It is of note that the recently announced NICE strategy for 2021 to 2026 aims to provide dynamic, living guideline recommendations that are useful, useable and rapidly updated' (NICE (2021), p. 19).

Strengths and limitations. We study prescription dynamics of statins, a class of drugs widely prescribed in primary care, using a representative dataset of English practices for the period 2004-2018, a time window that includes patent expiration of brand-name statins and publications of new NICE guidelines. There is no reason to believe that the large forsaken savings we have identified would not generalise to other important therapeutic areas of the English NHS or to other healthcare systems. Admittedly, the extent of the savings is an empirical matter and crucially depends on the structure of healthcare systems, the penetration of generics within them and the incentives of different players in prescribing, dispensing, and reimbursing pharmaceutical treatments.

We acknowledge a few limitations for this study. We only observe a first prescription issued to patients treated in primary care, without being able to account for prescribing influences coming from other settings. For example, patients experiencing a first cardiovascular event may have received their first statin prescription in secondary care, and such decision might have influenced ongoing prescribing in primary care. However, it is unlikely that this would explain all of the heterogeneity in prescribing choices and the large forsaken savings shown in Table 3.1. Further, we had access only to limited data on general practitioner characteristics to examine factors associated with the observed heterogeneity, while analysis of free text from clinical records to explore documented decisions related to statin prescription choice was beyond the scope of this project.

3.5 Conclusions

The fact that general practitioners react to prices illustrates the strengths of a healthcare system that pays attention to cost-effectiveness. There is potential for large savings for the NHS if new and, whenever possible, ongoing patients are promptly switched to the first medicine that becomes available as generic within a therapeutic class where all other medicines have similar efficacy. On going efforts to create a system infrastructure to support and monitor general practitioner prescribing locally could prove effective in aligning incentives to select cost-effective treatments while preserving physicians' freedoms to prescribe and patients' ability to discuss their preferred choice of treatment.

3.6 Appendices

Appendix A

In this Appendix, we compare the data in the Prescription Cost Analysis (PCA) dataset to those in the RCGP R&SC dataset to demonstrate that the latter constitutes a representative sample of national prescription. One advantage of the PCA database is that prescription data go back to the year 1998. However, the PCA database cannot be used to investigate heterogeneity in prescription choice because data are available only at national level, not at general practice level.

Figure 3.3 compares the data over time in our two data sources: Panel (A) on the left shows the figures from the PCA dataset between 1998 and 2018. Panel (B) on the right shows the figures from RCGP R&SC database from 2004 to 2018. Top panels display the total quantity in terms of daily defined doses (DDDs) while the bottom panels display the shares of each of the five statins in the market. The similarity in the trends reported in panel (A) and (B) confirms that the RCGP R&SC database is a representative sample of national data of statins prescription.

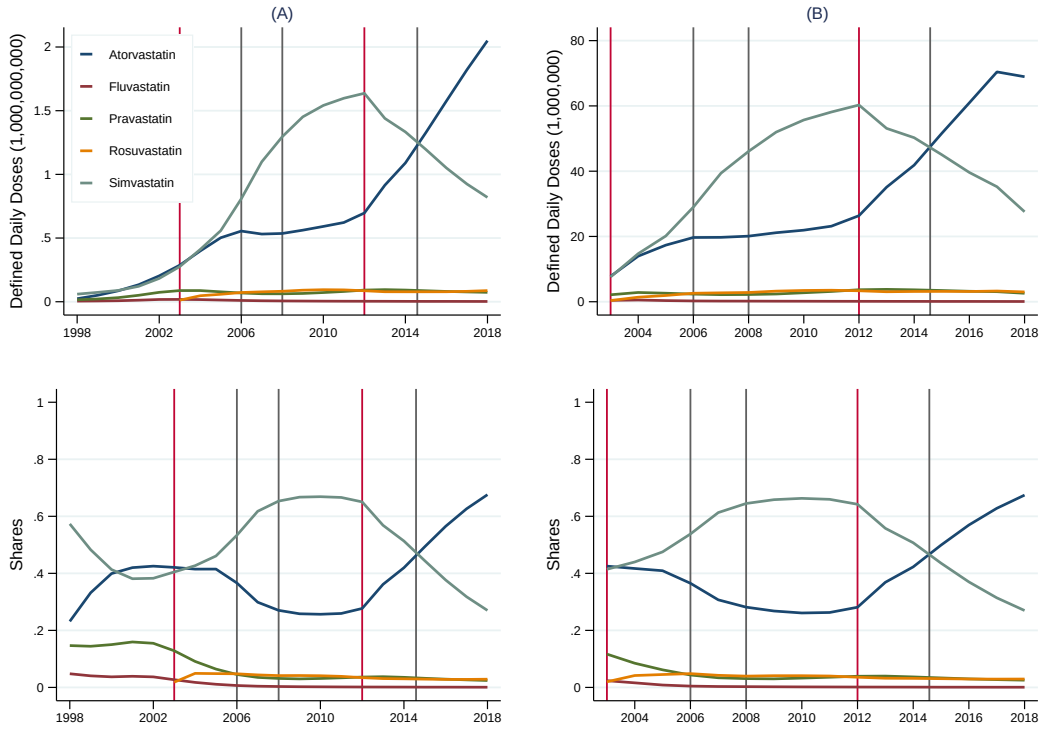


FIGURE 3.3: Volume of statins prescribed in main data sources.

NOTES. Panel (A) corresponds to Prescription Cost Analysis data. Panel (B) to RCGP R&SC data.

Appendix B

Spending savings simulation exercise methodology. This appendix describes the methodology used for the spending savings simulation exercise, by which we estimate the potential savings for the NHS that could have been achieved if general practitioners had prescribed simvastatin or atorvastatin as active ingredients, whenever these two medicines were the prescribing standard in this market according to the observed preferences of general practitioners and the recommendations in national guidance. We start by describing the computations of actual and hypothetical cost for the first scenario, in which the analysis refers to the first prescription episode, i.e. the first 28 days of drug treatment, for new patients only; and then the second scenario, in which we apply the same methodology to all prescriptions issued to all existing patients being treated in every period.

First scenario. To compute actual and hypothetical cost, we use information on the number of new patients treated, their initial drug treatment (i.e. a specific statin and strength), and a measure of each treatment's acquisition cost to the NHS per day of treatment. From the RCGP R&SC database, we count the number of new patients being prescribed statin treatment s in period t for the first time, denoted by n_{st}.

From the Prescription Cost Analysis (PCA) series, containing cost data on all medicines prescribed and dispensed and their corresponding cost to the NHS at the national level, we retrieve a measure of the actual acquisition cost of each statin treatment. To compute the average acquisition cost of treatment s in period t, denoted by C_{st}, we take the ratio between the Net Ingredient Cost (NIC_{st}) and the corresponding Total Quantity (Q_{st}) prescribed of each different strength of statin, that is $C_{st} = \frac{NIC_{st}}{Q_{st}}$.

Since data on the total number of new patients starting treatment on each statin nationally are not publicly available, we estimate such figure by combining information from the RCGP R&SC database (which is a nationally representative sample of general practices in England) with national aggregated data from the Prescription Cost Analysis series. Concretely, we compute the total number of new patients nationally, denoted by $\hat{N}$, as follows: $\hat{N}_{st} = \frac{n_{st}}{q_{st}} \times Q_{st}$, where q_{st} denotes the total quantity of each statin treatment prescribed in every period in the RCGP R&SC database. Indeed, since the RCGP R&SC sample of GP practices is representative of the English general practice sector, then the ratio of new patients to total quantities prescribed in both data sources should be equivalent. Aggregating $\hat{N}_{st}$ over all treatments at the year-level, $\hat{N}_t = \sum_s \hat{N}_{st}$ for each t, results in the figure reported in column (1) of Table 3.1, i.e. the estimated number of new patients treated with statins in each year.

Finally, we compute the actual cost of first prescription episodes for each statin treatment s in every period t by multiplying the total number of new patients on each treatment $\hat{N}_{st}$ with the cost per day of treatment C_{st} times 28, that is $AC_{st} = \hat{N}_{st} \times C_{st} \times 28$.

Then we aggregate AC_{st} over all treatments at the year-level, $AC_t = \sum_s AC_{st}$ for each$t$, which is the figure reported in column (2) of Table 3.1.

As explained above, practitioners' preferences when treating patients for cardiovascular disease risk moved towards simvastatin from its patent expiration (May 2003) until atorvastatin's patent expiration (May 2012); and from then onwards, they tended towards atorvastatin. Our cost simulation exercise *extremes* this observed behaviour by asking what would have been the cost savings if either simvastatin or atorvastatin had been the active ingredients originally prescribed to new patients, whenever these two medicines were the prescribing standard in specific periods. Accordingly, the hypothetical cost is constructed by substituting the originally prescribed treatment s, with a *therapeutically similar* one, denoted by s^*, containing simvastatin for those first-time prescriptions issued between 2004 and May 2012, or atorvastatin for those issued after May 2012.

The therapeutic similarity criteria we use is based on the ability of each strength of each drug (e.g. 1 tablet of atorvastatin 20 mg a day, 1 tablet of simvastatin 40 mg a day, etc.) in reducing low-density lipoprotein cholesterol levels per day of treatment. The percentage reduction in low density lipoprotein cholesterol is used in NICE's CG-181 to group the five statins (and each of their corresponding strengths) according to their intensity. The relationship between the strengths of the statins and reduction in low-density lipoprotein cholesterol is stated in NICE's CG-181, which in turn is based on the paper by (Law et al., 2003). A reproduction of this information is presented in Table 3.2.

TABLE 3.2: Percentage Reduction in low-density lipoprotein cholesterol.

Statin	Dose [mg.] per day				
	5	10	20	40	80
Fluvastatin			21%	27%	33%
Pravastatin		20%	24%	29%	
Simvastatin		27%	32%	37%	42%
Atorvastatin		37%	43%	49%	55%
Rosuvastatin	38%	43%	48%	53%	

NOTES. 20%-30%: low intensity. 31%-40%: medium intensity. Above 40%: high intensity.

To make this operative, for each level of percentage reduction in low-density lipoprotein cholesterol achieved by the originally prescribed treatment, i.e. a drug-strength pair, we look for the closest strength of both simvastatin and atorvastatin that achieves a similar level in low-density lipoprotein reduction to the originally prescribed one. The correspondence between original treatments and the substitutes is presented in Table 3.3. Columns (3) and (4) show the strength of simvastatin and atorvastatin, respectively, that achieves the closest percentage reduction in low-density lipoprotein cholesterol than the original treatments listed in columns (1) and (2). For example, if

a patient was prescribed atorvastatin 10mg a day for treatment initiation before 2012, the hypothetical prescription for this patient is a treatment of simvastatin 40mg a day, as both achieve a reduction of 37% in low-density lipoprotein cholesterol. Second example, if a patient was prescribed rosuvastatin 10mg for treatment initiation after 2012, then the hypothetical prescription for this patient is a treatment of atorvastatin 20mg, as both achieve a reduction of 43%.

TABLE 3.3: Correspondence between all statins treatments based on low-density lipoprotein cholesterol reduction.

(1)	(2)	(3)	(4)
Original Treatments		Similar Treatments	
Statin	Strength [mg.]	Simvastatin	Atorvastatin
Atorvastatin	10	40	
	20	80	
	30	80	
	40	80	
	60	80	
	80	80	
Fluvastatin	20	10	10
	40	10	10
	80	20	10
Pravastatin	10	10	10
	20	10	10
	40	10	10
Rosuvastatin	5	40	10
	10	80	20
	20	80	40
	40	80	80
Simvastatin	10		10
	20		10
	40		10
	80		20

NOTES. This table is based on the Grouping from Table 3.2.

Finally, the hypothetical cost is computed by multiplying the total number of new patients times the cost of the therapeutically similar treatments C_{s^*t} times 28, that is $HC_{st} = \hat{N}_{st} \times C_{s^*t} \times 28$. Then we aggregate HC_{st} over all treatments at the year-level, $HC_t = \sum_s HC_{st}$ for each t, which is the figure reported in column (3) of Table 3.1.

Second scenario. The second scenario considers not only first-time prescriptions for new patients, but all prescriptions for all existing patients treated with statins. For this, we use the information on total quantity and spending from the PCA database. The actual cost is obtained by aggregating spending on all statins prescribed in each year. The hypothetical cost is computed by replacing the per unit cost of the original treatment (statins and strength) with the corresponding cost of the therapeutically similar treatment (either simvastatin or atorvastatin), as described above. Additionally, to provide an estimate of the total number of all existing patients treated with statins in every period, $\hat{P}_{st}$ (the figure reported in column (6) of Table 3.1), we compute $\hat{P}_{st} = \sum_s \frac{p_{st}}{q_{st}} \times Q_{st}$,

where p_{st} denotes the total number of all existing patients using treatment s at time t as reported in the RCGP R&SC database.

Chapter 4

Prescribing behaviour and patient health outcomes: evidence from *statin* prescriptions in general practices in England

Abstract

We use the market of statins – cholesterol-lowering medications to lower cardiovascular risk – to study the relationship between prescribing behaviour, and cost to the health service as well as patients' health outcomes. Using treatment choices made by general practitioners for new patients treated between 2003 and 2018, we characterize prescribing behaviour along two dimensions: the propensity to follow cost-effective and guidance-based recommendations (*compliance*) and the extent to which health providers responds to their patients' health characteristics when allocating treatments (*responsiveness*). We find that least compliant providers have significantly higher costs than most compliant and that such difference comes at no additional benefit, in terms of patients' probability of experiencing adverse health outcomes. Concerning responsiveness, we find that this aspect of prescribing decision making does not result in sizeable differences in costs or health outcomes, despite findings suggesting that patient characteristics explain approximately one-third of the variation in treatment choices. Our results show that the most cost-effective prescribing behaviour is achieved by the most compliant and least responsive providers, as no prescribing style appears to improve outcomes while this group exhibits lower costs. Overall, the prescription of anti-cholesterol drugs in England has similar features to those found in other health care markets intensive in surgical procedures: large variation across providers, significant differences in costs but no improvements in health. The findings imply that further efforts could have been made to convince practitioners to follow existing recommendations.

4.1 Introduction

Variation in medical care at the geographical level, such as hospital referral regions (or any other type of health market area), has been the focus of a large body of literature, at least since the 1930s.[1] This vast literature has not only described the extent and scope of heterogeneity across health areas, but also tried to find the root causes and to evaluate its consequences on both payers and, most importantly, patient outcomes. The results show the importance of supply-side factors (e.g. physicians skill, experience, adoption of information, adherence to guidelines, or inertia in the use of procedures, medical devices or drugs) over demand-side factors (e.g. price & income effects, patients preferences and underlying health status) in explaining observed variation and that, such variation, is not associated with "improved satisfaction, outcomes or survival but [...] with significantly higher costs".[2]

Most of the existing evidence comes from medical interventions intensive in the use of surgical procedures (for example, coronary angioplasty and stent insertion for heart attacks), performed in hospital settings, and few studies come from primary care, which mostly deals with family (general) practice consultations, drug prescriptions, and the long-term management of chronic conditions (e.g. cardiovascular disease, hypertension, depression). This is probably due to the fact that larger spending is associated to secondary (i.e acute and urgent) and tertiary (specialized) health care,[3] and that, methodologically, better causal inference can be made in those settings – as there is less room for a patient selection of provider, and vice versa. Furthermore, while there is a growing literature on prescribing behaviour taking place in primary care, most of available studies are focused on describing and explaining it (e.g. Coscelli (2000), Rosenthal et al. (2002), Cleanthous (2002), Kyle and Williams (2017), Carey et al. (2020), Parker-Lue (2020)), with less attention to its implication on patients health outcomes (some exceptions include Currie and MacLeod (2020)).

In this paper, we study how two dimensions of prescribing behaviour, namely the propensity to follow cost-effective and guidance-based prescribing, what we call *compliance*, and the extent to which health providers responds to their patients' health characteristics when allocating treatments, which we call *responsiveness*, are associated to costs and health outcomes. Our setting is the prescription of cholesterol-lowering medication in general practice in England. This drug market has a number of features that

[1] For example, Glover (1938) first documented a more than four-fold variation in the incidence of tonsillectomy on children among British school districts.

[2] Chandra and Staiger (2007). Additionally, see Chandra et al. (2012) for an overview on the economics of treatment choice, and Phelps (2000) on the economic literature on the role of information in explaining geographic variation in medical practice style.

[3] For example, in England it has been estimated that general practice (primary care) provides over 300 million patient consultations each year, compared to 23 million A&E (accidents & emergency) visits, and yet 1-year worth of GP care per patient costs less than two A&E visits. See https://www.england.nhs.uk/five-year-forward-view/next-steps-on-the-nhs-five-year-forward-view/primary-care/.

make it an appropriate case study. First, cholesterol-lowering drugs, the most prescribed class in England, are used to prevent the risk of cardiovascular disease (CVD), one of the leading causes of mortality in the UK (after cancer) and which implies one of the largest health spending items to the national health service (NHS). Second, although there are 23 drugs indicated for the prevention of CVD according to the British National Formulary, one group, namely the *statins*, attracts more than 95 per cent of all prescriptions within the class since the early 2000s. Within this group, two drugs, simvastatin and atorvastatin, represent virtually the binary choice made by practitioners to treat their patients, which implies that we can more simply characterize clinical decision-making around these two alternatives. Last but not least, during the sample period this drug market was subject to five exogenous shocks that significantly affected their use: the patent expiration of the two above-mentioned treatments and the introduction of three clinical guidelines explicitly recommending their use for drug treatment initiation. These shocks generated significant variation across health providers and over time, which we exploit to study the consequences of the observed heterogeneity.

This work extends the analysis in Chapter 3 (i.e. Ortiz de Zarate et al. (2021)) by investigating how patient characteristics explain prescribing heterogeneity, on one hand, and evaluating the consequences of this heterogeneity (captured in the dimensions of compliance and responsiveness) on health outcomes, on the other. In that paper, we document the main features of this drug market, discuss the influences behind the observed trends and their variation, and provide an estimate of the magnitude of the inefficiency associated with the heterogeneity in treatment choices across practices, in terms of cost savings to the health service. This latter analysis is based on the assumption of perfect therapeutic substitution between available treatments, consistent with evidence from clinical trials, allowing us to abstract from any consideration regarding patient demand-side factors – if treatments are equivalent then it does not matter who receives them and only cost consideration should drive treatment assignment. However, results from clinical trials inform about the average benefits of medical intervention and are not the same as actual clinical practice, where physicians must learn for themselves about a particular patient's potential marginal benefit. Accordingly, it is relevant to investigate empirically the role that patients themselves play in explaining the observed heterogeneity, how to capture practitioners responses to patients health information, and what are the consequences of this aspect of prescribing behaviour on outcomes.

Our analysis is divided into two stages. In the first stage, we construct a characterization of prescribing behaviour that reflects compliance with the recommendations stated in the guidelines and with cost-effective prescribing more generally. Specifically, based on their relative propensity to prescribe the recommended treatments to new patients, we classify each practice in every period as being most or least compliant. Then, we

use this characterization to examine whether types of compliance are associated with costs to the health service and patient health outcomes. If patients from least compliant providers (with respect to most compliant) significantly and consistently exhibit better outcomes, then heterogeneity, and therefore higher costs, is justified. If otherwise, then it is pure inefficiency and the discretion of practitioners could be further rectified without negatively affecting outcomes.

In the second stage, we explore how both compliance and responsiveness to patients characteristics affect costs and outcomes. Following a framework developed in Currie et al. (2016), we add to compliance a second characterization that reflects sensitivity to patients health information. Such information is contained in an index that captures (i.e. predicts) patients' likelihood to be prescribed a given treatment, conditional on their observable characteristics, and it represents an empirical clinical practice norm according to how the general practice sector as a whole value observable aspects of the underlying health of patients. Based on their relative degree of responsiveness, we classify each practice in every period as being most or least responsive. Then we use the double characterization, i.e. of compliance and responsiveness, to examine whether these types of prescribing behaviour are associated with costs to the health system and patient health outcomes. If most responsive providers (with respect to least responsive) exhibit better outcomes, then this implies that such norm was beneficial, and physician discretion should be allowed and encouraged. If not, further efforts could be made to regulate such behaviour. Finally, we ask what is the role of patients underlying health in explaining heterogeneity in prescriptions choices made by practices.

We use a rich dataset containing prescription records and health status characteristics from patients taking statins treated by a representative sample of general practices in England between 2003 and 2018. These data allow us to compute the cost of treatment for each patient to the health service and to observe whether patients developed an adverse health condition at any given time since treatment initiation. The results of the first stage show that differences in compliance imply significant and sizeable differences in cost to the health service, but that providers prescribing the more expensive treatments do not exhibit significantly better outcomes for their patients, measured as the 3-year probability of experiencing cardiovascular events (what statin treatment seeks to prevent). We also find that least compliant providers are more likely to switch the initial drug of their patients later on in the course of treatment: this suggests a complying behaviour that does not manifest itself initially but with a lag, which can attenuate differences in cost that could have been larger otherwise. Lastly, we explore all-cause mortality as it is a likely consequence of patients with cardiovascular disease. Although we do not find a consistent link between this outcome and differences in compliance, we do find a trend that suggests that patients whose providers are intensive in the use of one particular treatment (i.e. atorvastatin) are associated with lower

mortality rates. We interpret this result with caution as, it is clear, the cause of death cannot be attributed directly to cardiovascular disease risk.

The results from the second stage are in-line with the first stage. That is, most compliant providers have significantly lower costs of treatment, differences that are greater within those practices that are also least responsive to patients characteristics. We also find that differences in responsiveness do not affect costs greatly, as compliance does. With regard to health outcomes, no significant differences are found between either type of compliance or responsiveness. This suggests that being most compliant and least responsive (to the norm), i.e. strict compliance, results in the most cost-effective prescribing behaviour, as this type is associated with the lowest cost levels, while there are no significant differences in health outcomes among the other types of prescribing style. We also estimate that patients' underlying health, characterized by a rich set of observables (including existing diagnoses and risk factors) explains roughly one-third of the variation in treatments that we can account for. Thus, although relevant in explaining variation in the allocation of initial treatments, acting differently upon different patients do not associate with greater differences in costs nor in outcomes.

Overall, our results suggest that the prescription of anti-cholesterol drugs in England has similar features as those found in other health care markets intensive in surgical procedures: large variation across providers, significant differences in costs but no improvements in health. The implications of the findings are that further efforts could have been made to contain heterogeneity as it is not justified. Costs can be reduced without sacrificing outcomes.

The structure of the paper is laid out as follows. Section 4.2 describes the data. Section 4.3 covers the methods. Results are presented in Section 4.4, and Section 4.5 concludes.

4.2 Data

For our empirical analysis, we extract data from the Royal College of General Practitioners Research and Surveillance Centre (RCGP R&SC) database.[4] This is one of the longest established primary care networks containing comprehensive and longitudinal computerised medical records on routine patient care data in the context of GP consultations in England, including risk factors (e.g. gender, age, smoking status), medical diagnosis (e.g. diabetes, hypertension), physiological measurements (e.g. body mass index) and prescription history (e.g. for each prescription episode it identifies the drug name, strength, number of tablets, and issue date). From this source, we obtain every prescription record for approximately 500 thousand patients taking statins between

[4]See https://www.rcgp.org.uk/clinical-and-research/our-programmes/research-and-surveillance-centre.aspx.

2003 and 2018 registered at one of the 243 general practices members of the network.[5] Patients are identified by a unique (anonymized) code and GP practices by their national official codes. It is unfortunately impossible to identify the actual practitioner responsible for each prescription decision, and instead, our unit of analysis becomes the general practice from which prescriptions are issued. This implies that what we are capturing and attributing is actually the mean behaviour of all practitioners within a GP surgery. As the main focus of this work is on characterizing prescribing behaviour at the decision-maker level and to link it with patients outcomes actually treated by them, we include only patients treated by the same GP practice throughout their observed drug treatment history.

The three clinical guidelines published during our study time frame were explicit about a drug recommendation for drug treatment initiation. Accordingly, the final sample is constituted of only new patients starting treatment with a statin for whom we can identify: (1) their treating practice, (2) the drug and date of the first prescription episode, (3) a set of characteristics representing baseline health risk, and (4) a set of outcomes related to the treatment process and to their health prognosis that realize themselves in the future (e.g. whether a patient experience an adverse health outcome within 3 years since starting treatment). These baseline risk factors and outcomes are chosen given their importance with the initiation and objectives of drug therapy for the prevention of cardiovascular events, according to the published clinical guidelines and the literature.[6]

Our final sample consists of 478,041 adult patients, between 18 and 90 years of age, being prescribed statins and treated by one of the 243 GP practices during the 61 year-quarters between 2003q3 and 2018q3. Out of the total number of practices, 96.35% are observed taking new patients every quarter in the sample period, while the rest are between 40-60 quarters. Table 4.1 presents the mean number of patients per practice per quarter and during the sample period. On average, a practice will receive 32 new patients every quarter adding up to 1,967 during the study time frame.

TABLE 4.1: Number of new patients across practices

	Mean	SD	Min.	Median	Max.
Quarterly no. new patients	32.41	22.77	1	28	397
Total No. new patients	1,967.25	1,061.44	291	1,745	6,771

[5]The representativeness of this database with respect to the English epidemiology has been assessed in Correa et al. (2016). Additionally, and relevant to our study, Hinton et al. (2018) establishes the representativeness of the database regarding the incidence and prevalence of cardiovascular disease.

[6]See NICE (2006), NICE (2008), NICE (2014), Wilson et al. (1998), Hippisley-Cox et al. (2008).

The available patient characteristics include gender, age, ethnicity, socio-economic status, smoking and alcohol consumption, and the date in which comorbidities and cardiovascular events developed throughout their observed series of prescriptions. Sample means of patient characteristics are presented in column (1) of Table 4.2. The average new patient starting statin treatment is a 64-year-old white male and belongs to the most deprived socio-economic group (according to the index of multiple deprivations). Out of the total, 22% are active smokers and 41% ex-smokers, around 30% have a risky alcohol consumption behaviour (either alcoholism or hazardous intake), 77% is overweight or obese, 44% have moderate to severe chronic kidney disease (stage 2 or higher), 24% has diabetes, half hypertension, and 28% starts treatment having experienced a cardiovascular disease (CVD), the most common being ischaemic stroke (10%) and acute myocardial infarction (7%). Additionally, the table shows mean characteristics for the two main periods of the study time frame, namely before and after 2012q2, in columns (2) and (3), respectively. Despite some minor differences, the average patient is similar over time.

TABLE 4.2: Mean patient characteristics – baseline risk

	(1) All new patients	(2) New patients starting treatment in $t < 2012q2$	(3) $t \geq 2012q2$
Male	0.55	0.55	0.55
Age	63.66 (12.74)	63.65 (12.75)	63.69 (12.72)
White	0.90	0.92	0.88
Asian, Black, Mixed & Other	0.10	0.08	0.12
Index of multiple deprivation			
Quintile 1 (most deprived)	0.16	0.16	0.17
Quintile 2	0.16	0.16	0.17
Quintile 3	0.19	0.19	0.19
Quintile 4	0.23	0.23	0.23
Quintile 5 (least deprived)	0.25	0.25	0.24
Smoking status			
Active smoker	0.22	0.23	0.20
Ex-smoker	0.41	0.39	0.43
Alcohol consumption			
Alcoholism	0.04	0.03	0.04
Hazardous	0.25	0.23	0.28
Non-drinker	0.20	0.21	0.18
Safe	0.51	0.53	0.49
Obesity (BMI$\geq$25)	0.77	0.77	0.76
Chronic Kidney Disease			
Stage 1	0.56	0.55	0.58
Stage 2	0.34	0.34	0.35
Stages 3-5	0.10	0.12	0.06
Diabetes (T1 & T2)	0.24	0.24	0.24
Hypertension	0.50	0.53	0.47
Cardiovascular disease diagnoses	0.28	0.29	0.25
Angina	0.06	0.08	0.04
Acute myocardial infarction	0.07	0.08	0.07
Atrial fibrillation	0.06	0.05	0.06
Congestive cardiac failure	0.03	0.03	0.03
Coronary artery disease	0.04	0.04	0.05
Ischaemic stroke & transient Ischaemic attack	0.10	0.10	0.09
Peripheral arterial disease	0.03	0.04	0.03
$N_{patients}$	478,041	285,450	192,591

NOTES. The table present the sample mean of patient baseline characteristics. Each variable is an indicator for the characteristics, except for age which is continuous, for which the standard deviation is shown in parenthesis. Column (1) shows the mean for the cross-section of new patients, and columns (2) and (3) for the two main period of the time frame, i.e. for the periods before and after 2012q2, respectively.

We measure outcomes using prescription and health information during the patient's drug treatment course. We consider four main outcomes for each patient: (1) cost of treatment, (2) probability of switching the initial treatment, (3) probability of developing an adverse health outcome, and (4) probability of dying from any cause (i.e. all-cause mortality). We describe these measures and their construction next.

Costs. In England, prescription-only medicines are financed by the national health service (NHS) and different health bodies (e.g. NICE) make substantial efforts in promoting cost-effective prescribing. Given this, we construct a variable that represents the cost that each patient implies to the NHS, as a result of the prescribing decisions made by their treating GP practice. By computing the total volume associated with every prescription episode and multiplying it with the corresponding unitary cost per drug we can obtain the total cumulative cost up to any point in the course of treatment.[7] We compute the patient-specific cumulative cost of treatment for three cut-off points since the first prescription, namely for one, two and three years.

Switching. By tracing forward into the drug treatment history we can identify when there was a first switch from the initial drug-assigned to a patient. Despite not being the focus of this research, the switching decision has been shown to be associated with poorer patients outcomes (Phillips et al., 2007). However, the switching decision may also reflect a potentially cost-effective choice every time that the change from the originally prescribed treatment goes to the recommendation stated in clinical guidance; it is this aspect that we are interested in.

Adverse health condition. The main objective of lipid-lowering drug therapy is to decrease dangerous cholesterol levels to prevent the onset of, or a repeated, cardiovascular event (primary and secondary prevention, respectively). Given this, we explore whether patients developed any adverse health conditions since treatment initiation. We generate a composite measure representing whether a patient experienced any of the CVDs listed in the lower part of Table 4.2 including hypertension, as a first occurrence after the start of drug treatment.[8]

All-cause mortality. As our data contains the date of death of patients (when it applies) we also explore all-cause mortality as an outcome, constructed in the same way as the adverse health condition variable is generated.

The sample means of the four main outcomes across time are shown graphically in Figure 4.1. Each panel shows the average of the corresponding outcome, for each year, computed over all new patients starting drug treatment in such year, and for three cut-offs. We briefly comment on their trends. The dramatic decrease in cost over time (panel A) is explained by the drop in acquisition cost of statins due to patent expiration and the growing cost-conscious prescribing revealed by the trends in Figure 3.1. The decision to switch initial treatments (panel B) is consistent also with patent expiration

[7] For example, a prescription consisting of 28 tablets of simvastatin 40mg in 2005 is equal to $28 \times (£0.31) = £8.6$, where £0.31 corresponds to the acquisition cost to the NHS of simvastatin 40mg in 2005. Cost data for each strength of each drug in each year is obtained from the Prescription Costs Analysis data; see https://digital.nhs.uk/data-and-information/publications/statistical/prescription-cost-analysis. All cost figures are expressed in constant 2018 GBP using the GDP deflators at market prices, and money GDP. See https://www.gov.uk/government/statistics/gdp-deflators-at-market-prices-and-money-gdp-march-2019-spring-statement.

[8] These conditions were also considered in the article by Hinton et al. (2018) when assessing the incidence and prevalence of cardiovascular disease in England using the RCGP&RSC database. In their paper's supplementary appendix there is a complete list of disease codes (Read Code list) comprising each of the conditions used here.

dates: is higher after the first patent expiration in mid-2003 (i.e. simvastatin) and it peaks again in the year of the second patent expiration (i.e. atorvastatin). The growth in the switching share within the first 3 years of treatment (blue line) since 2009 indicates that from then onwards it will become more likely that practitioners switch the initial treatment of their patient in response to the change in the cost-effective status of the drug that becomes generic in 2012 (atorvastatin). After such event, the switching decreases. With regard to the development of an adverse CVD event including hypertension (panel C), for most of the time frame, there is a stable trend: around 8 per cent of patients experience a first adverse outcome within year one, 12 per cent within year two, and 15 per cent within year three. The specific individual conditions underlying this composite measure are presented in Appendix Figure 4.8. Finally, all-cause mortality (panel D) is relatively stable up to 2012, after which year the patterns exhibit a nearly 3-fold increase towards the end of the data period.[9]

[9]The drastic increase in all-cause mortality after 2012 deserves an explanation. Yet, despite efforts to obtain one with the data holders, the RCGP R&SC, it has not been possible. The data requested to them was specifically tailored for this project, and once the research relation with them was over, they proceed to delete the data, as is the custom procedure. Unfortunately we did not raised the issue earlier, and generating the data again just for the purpose of answering this query was more costly to them than beneficial to this research purpose. Nevertheless, we looked for data and published articles in the academic and gray literature that showed the long-term trends for all-cause mortality associated to cardiovascular diseases (see, for example Bhatnagar et al. (2016)). In all cases reviewed, all-cause mortality from CVD diseases appears to be decreasing. Hence a discrepancy with these data, most likely explained by some form of informatic recording errors. However, the data used in this paper corresponds not to the national population nor a represenative sample of it, but to the subpopulation (or a sample thereof) of patients that are on statin medication. In other words, we only observe patients, their health outcomes, diagnosis, co-morbidities, and all-cause death, as long as they were being prescribed statin medication by their GP practice. And to the extent that unobservable aspects of patients' behaviour related to adherence and continuation of treatment explains the increasing all-cause mortality figures in our data, is not possible to know with certainty. We also note that this increase occurs after the patent expiration of atorvastatin and the subsequent drastic increase in its uptake and switching-to, as well as after a period of intense media coverage on the risks and benefits of statins, epsiode that has been shown to had an effect in patients behaviour, specifically in older adults stopping their statisn treatment (see Matthews et al. (2016)). This, we claim, could potentially explain an increase in mortality rates, unlikely for the national population but likely for the sub-population of on-medication patients.

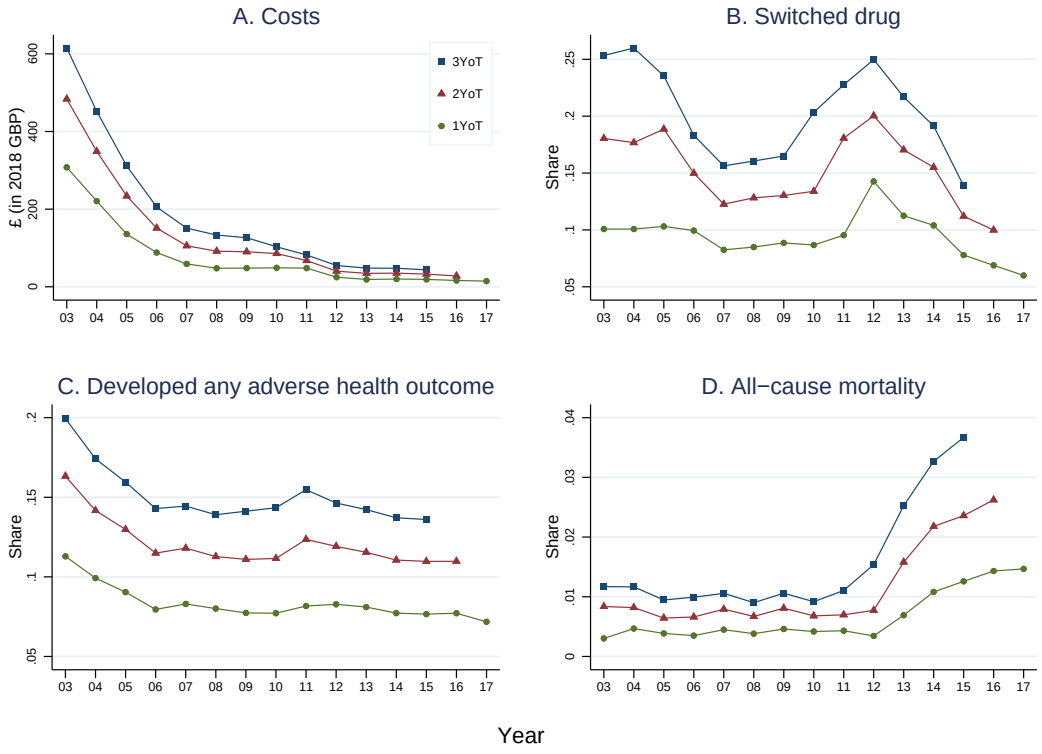


FIGURE 4.1: Sample mean of main outcomes

NOTES. The figure shows the mean for the four main outcomes for each year and for three lengths of time since the start of drug treatment, first (1YoT), second (2YoT) and third (3YoT) year. x-axis represents years and y-axis the outcomes measure. Each point in the graphs represents the mean of the specific outcome for a given length, computed over all new patients starting treatment in each year. For example, in panel A, the point (03, 600) on the blue line indicates that the average cumulative cost during the first 3 years of treatment for those patients starting therapy in 2003 is £600. In panel B. the point (03, .25) on the blue line indicates that within the first three years of treatment 25 per cent of new patients starting therapy in 2003 experienced a switch of their first drug. Etc. The individual adverse conditions underlying the composite outcome measure in panel C, are shown in Appendix Figure 4.8. The missing values for the 3-year (in 2016, 2017, and 2018), 2-year (in 2017 and 2018), and 1-year outcomes (2018), is explained by the fact that as we have data up to 2018q3, full 3-year outcomes can only be observed for those new patients starting treatment up to 2015q3, and therefore we lose those that started treatment after. Similarly, full 2-year outcomes can only be observed for those new patients starting treatment up to 2016q3, and then we lose those starting after such date. And so on. Then, the number of observations becomes: $N_{3YoT} = 388,749$, $N_{2YoT} = 417,762$ and $N_{1YoT} = 448,174$.

4.3 Methods

4.3.1 Compliance with recommended treatments

We start our investigation of the differences in prescribing behaviour and their associations to patient outcomes by constructing a measure that reflects the propensity to prescribe the recommended drug to new patients. To do this, we estimate the following linear model for the probability of prescribing the recommended treatment for new

patient i starting drug treatment in year-quarter t:

$$RT_{igt} = \pi_{gt} + \varepsilon_{igt}, \tag{4.1}$$

where RT_{igt} is a binary variable that takes the value of 1 if the initial drug prescribed to patient i by its GP g in quarter t corresponds to the recommended treatment in that period, defined according to the rule:[10]

$$RT_{igt} = \begin{cases} \text{simvastatin}_{igt} & \text{if } t < 2012q2 \\ \text{atorvastatin}_{igt} & \text{if } t \geq 2012q2. \end{cases} \tag{4.2}$$

The coefficient π_{gt} is a constant capturing the unconditional probability of GP g to prescribe the recommended treatment to new patients starting drug therapy in every quarter, or in other words, the share of compliant decision being made (i.e. $\pi_{gt} = Pr(RT_{igt} = 1) = E[RT_{igt}]$. This model is estimated for each practice g and each moving 3-quarter period up to t, i.e. for $\{t\text{-}2, t\text{-}1, t\}$.[11] In practice, the estimation of eq. (4.1) is performed twice, for each drug in the definition in (4.2). This results in two sets of parameters π^s_{gt} for drugs $s = \{\text{simvastatin, atorvastatin}\}$. However, for simplification, in what follows we retain the notation π_{gt} to denote the propensity to prescribe recommended treatments by each practice in each period while keeping in mind the definition of RT. Note that this distinction is practically immaterial as both sets of parameters have a correlation above -0.97, reflecting that, despite there being 5 statins available, the decision to initiate drug treatment was virtually a binary choice made between simvastatin and atorvastatin.

Having estimated $\hat{\pi}_{gt}$, we now characterise the general practices in our sample based on the relative propensity to comply with the prescribing standards set in this market.[12] Specifically, we generate two mutually exclusive groups of GP practices, by taking the median of the distribution of $\hat{\pi}_{gt}$ in every period t to define those most and least compliant with the rule. We define the binary variable C_{gt} to take the value of 1 if the propensity to prescribe the recommended treatment of practice g in quarter t is greater than the median in t (i.e. if $\hat{\pi}_{gt} \geq \text{median}_t[\hat{\pi}_{gt}]$, for each t): we call these most compliant, and the opposite group least compliant. Note that this definition allows for the

[10] In Chapter 3 Section 3.3, we show that between 2003q3 and 2012q2 the preferences of practitioners tended towards simvastatin for drug treatment initiation, influenced by its patent expiration, and the publication of two NICE guidelines explicitly recommending its use. From 2012q2 onwards, this preference immediately switched towards atorvastatin, after its patent expiration and going through the following clinical guideline changing the previous recommendation to the newly generic drug. Hence, the rule in eq. (4.2).

[11] An alternative to this approach would be to estimate this model for each quarter t. However, we prefer not to use this as it results in a distribution of general practice's share of new patients starting on the recommended treatments that is more volatile than with the moving quarter alternative. Such volatility is explained mainly by the fact that some practices have a lower frequency of new patients starting treatment in every quarter than others, creating greater variation than the moving quarter alternative which generates smoother long term trends for each practice.

[12] See Appendix Figure 4.9 for the distribution of $\hat{\pi}_{gt}$ centred around the median. Also see Figure 3.2 in Chapter 3 for the distribution of the median, and Section 3.3 for a discussion of this trend.

complying behaviour to change over time.[13] Table 4.3 shows that this procedure yields approximately equally-sized groups of patients between GP practices complying types.

TABLE 4.3: Distribution of patients into complying types

most compliant	237,272	49.63%
least compliant	240,769	50.37%
Total	478,041	100%

Next, we use this characterization to examine the associations between differences in complying prescribing behaviour and patients outcomes, by estimating the following model:

$$Y^j_{igt} = \alpha_0 + \sum_{y=2003}^{T^j} \beta_y \times \mathbb{1}_{\{t \in y(t)\}} \times C_{gt} + \Phi X_{igt} + \eta_g + \theta_{y(t)} + \varphi_{igt}, \tag{4.3}$$

where Y^j_{igt} denotes outcome j for patient i, starting drug treatment with GP g in year-quarter t, C_{gt} represents the complying type of practice g in quarter t, and $\mathbb{1}_{\{t \in y(t)\}}$ is an indicator variable that takes the value of 1 if quarter t in which patient i starts treatment belongs to year $y(t)$. (The notation $y(t)$ is used to make clear that for each quarter t in which a patient begins drug treatment, there is, of course, an associated year y.) T^j is the last year for which the each outcome j is observed (see the NOTES in Figure 4.1.) The model also includes the vector of patient characteristics capturing baseline risk, X_{igt}, and GP practice and year fixed effects, η_g and $\theta_{y(t)}$, respectively. (Note that an equivalent way to write the year fixed effects $\theta_{y(t)}$ is $\sum_{y=2003}^{T^j} \lambda_y \times \mathbb{1}_{\{t \in y(t)\}}$.) φ_{igt} is the error term.

The main coefficients of interest are the estimated $\hat{\beta}_y$, representing the differences in patients' outcomes between most and least compliant practices, i.e. $E[Y^j_{igt} \mid C_{gt} = 1] - E[Y^j_{igt} \mid C_{gt} = 0]$, for each year y, conditional on patients' baseline health, GP practice and year fixed effects. Thus, we will be comparing the outcomes of patients at risk of CVD taking statins with the same set of observable baseline health characteristics starting treatment in the same year but treated by GP practices that differ in their degree of compliance with prescription standards.

The specification in (4.3) allows us to observe the consequences of different complying prescribing behaviour in every period (net of the year-effects, that is, net of the mean evolution of each health outcomes), instead of a fixed constant value across time. As the statin drug market was subject to several events or shocks that shaped its evolution and because the treatment choices that underlay our characterization of practices

[13] In Appendix Tables 4.6 and 4.7 we show the extent of the variation in compliance status, i.e. of C_{gt}, over the full sample period and for each year, across the GP practices in our sample. During the full sample period, 61 quarters between 2003q3 and 2018q3, a GP practice will change 16% of the time between most and least complaint status, with a SD of 6%. When measuring this variation on year-by-year basis, it shows an increase from a 7% in 2003 to a 30% in 2012, the year in which atorvastatin becomes generic to then decrease down to 11% in 2018.

changed over time, the specification permits us to assess whether differences in outcomes contrast with the undergoing patent expirations and/or introduction of clinical guidelines.

4.3.2 Responsiveness to patient's characteristics

The way in which we investigate the link between heterogeneous prescribing decisions and patients outcomes above abstracts from the role that patient health status itself plays in influencing the treatment that he or she will end up receiving. To incorporate this role we study the extent to which observable patient characteristics affect the decision made for them by their treating practice.[14]

To do this, first, we estimate the following model for the probability that a new patient receives the recommended treatment:

$$RT_{igt} = \Gamma X_{igt} + \lambda_{gt} + \xi_{igt}, \tag{4.4}$$

where the variables RT_{igt} and X_{igt} are defined as above, and λ_{gt} represents practice$\times$quarter fixed effects, and ξ_{igt} is an error term. Omitting the practice fixed effects may bias the coefficients of patients characteristics X_{igt} every time providers' propensity to prescribe the recommended treatment is related to the composition and characteristics of their newly treated patients. This model is estimated by polling all new patients treated by all practices in each moving 3-quarter period. We estimate this model period by period as the clinical management of the conditions treated by statins may evolve over time given, for example, by experience, the publication of further clinical trial results about safety and side-effects, etc. This results in weights placed on certain aspects of the patient initial health risk (gender, age, smoking and alcohol consumption, comorbidities, and presence of existing cardiovascular conditions) that change over time.

Using eq. (4.4) we can identify those patients who appear, given their baseline risk, to be more likely to be prescribed the recommended treatments. Specifically, with the fitted values excluding the terms of the fixed effects, we obtain an index that ranks patients according to their likelihood of being prescribed the recommended treatment

[14]The methods that follow are an adaptation of the framework developed in Currie et al. (2016). In their application, they study the practice style of cardiologists and whether differences in aggressiveness, defined as the propensity to use invasive procedures, or responsiveness, defined as tailoring treatment decisions to the characteristics of individual patients, matter for outcomes of patients with heart attacks arriving at the emergency room. To do this, in a first step, they define a patient-specific index that is meant to capture their degree of appropriateness for an invasive procedure, made operative by estimating a model for the probability that a patient receives an invasive procedure on the full set of observable characteristics, using only data from hospitals with accredited teaching programs (assuming that these institutions define the standard of care, and what is appropriate). In a second step, they estimate for each cardiologist, models for the probability that they perform invasive procedures on their treated patients on an intercept (aggressiveness) and a slope associated to the variable representing patients' appropriateness index (responsiveness).

based on their initial set of characteristics only.[15] We denote this index by Z_{igt} and is defined as:

$$Z_{igt}(X_{igt}) \equiv \hat{\Gamma} X_{igt}. \tag{4.5}$$

This measure represents an empirical clinical practice norm across the general practice sector, as it captures the mean way in which practitioners value each patient observable attribute in order to prescribe treatments, in every period. This index summarises all of the patient's observable information into an a unidimensional measure that captures (predicts) the treatment that is more likely to be prescribed.[16]

It is important to emphasize that the relationship implied by the index does not necessarily reflect an optimal or appropriate drug assignment, i.e. one that minimises the probability of an adverse health outcome because of treatment, but an average clinical practice norm of the primary care sector, that changes as the market forces, institutional drivers and clinical information in this therapeutic class changes. That is, in it are contained all the influences that determine how patients characteristics are valued by practitioners when deciding how to assign drugs. And, in what follows, it is used to capture how sensitive are prescribers to such information.

With this new variable, we now estimate the following model to capture both degrees of compliance with prescription standards and responsiveness to patients characteristics:

$$RT_{igt} = \pi_{gt} + \rho_{gt} z_{igt} + \psi_{igt}, \tag{4.6}$$

where z_{igt} denotes the patient index Z_{igt}, expressed as deviation from the mean for each g and t.[17] This model is estimated for each practice g and each moving 3-quarter period centred around t. Here π_{gt} captures GP g's propensity to prescribe the recommended

[15] A technical note on the estimation of eq. (4.4) and the obtainment of the fitted values. Although not explicit, this model is estimated with a constant, e.g. $RT_{igt} = \gamma_0 + \Gamma X_{igt} + \lambda_{gt} + \xi_{igt}$. And in practice, the model is estimated by absorbing the fixed effects using the STATA command `areg`. With this, the fitted values excluding the fixed effects result in $\hat{\gamma}_0 + \hat{\Gamma} X_{igt}$, with $\hat{\gamma}_0 = E[RT_{igt}] - \hat{\Gamma} E[X_{igt}]$. Having used `regress`, it produces the same parameter estimates $\hat{\Gamma}$, but the constant will be different depending on the reference category used for the fixed effect term. Additionally, note that the correlation between our preferred way of constructing the index and one that excludes the fixed effects is 0.997. Then, in practice, our preferred choice is empirically irrelevant.

[16] The use of a measure relating suitability for medicial procedures and patients health characteristics has also been used in other works. In Chandra and Staiger (2007), they perform regression for the probability of receiving cardiac catheterization for patients experiencing heart attack on patients comorbidities and fixed effects for the hospital referral regions in the US, which delimits the area-level at which cardiac surgeries are performed. In their application, they use the fitted values of such regression, excluding the fixed effects, as a measure of clinical appropriateness; they include the fixed effects in their estimation because omitting them "will bias the coefficients on comorbidities as regions with more appropriate patients will tend to do more catheterizations". This procedure is the one we follow. In Barrenho et al. (2019) they estimate the probability that a patient receives laparoscopic surgery for colon cancer on a vector of patient characteristics, in a period of time that reflects accepted and best practices, to then characterize the mean type of patients that physicians treat, in a study of the effect of peers and networks on the uptake of innovation for such procedure.

[17] That is, $z_{igt} = Z_{igt} - E_{gt}[Z_{igt}]$.

treatment on the average patient and is equivalent to the constant in eq. (4.1),[18] and ρ_{gt} captures GP g's sensitivity to the patient's baseline risk summarized in their estimated index. Eq. (4.6) makes clear that the probability to prescribe the recommended treatment is modelled as a mean propensity and a deviation component that captures the willingness to deviate from the mean and prescribe according to what the patient index suggests.

In principle, the coefficient ρ_{gt} can take any value, but some key cases exemplify the intuition of the model. If $\rho_{gt} = 0$, practitioners in g place no weight on the characteristics of their patients (summarized by their index scores). Then this practice prescribes the recommended treatment to all its new patients with probability π_{gt}, regardless of what the average clinical practice norm would suggest. If, on the other hand, $\rho_{gt} = 1$, then the practice's probability of prescribing the recommended treatments increases or decreases according to whether the patient has an index value greater or lesser than that of the average patient treated by his or her practice. And if $\rho_{gt} > 1$ ($\rho_{gt} < 1$) practice g puts more (less) emphasis than what the index suggests.

Having estimated $\hat{\rho}_{gt}$, we now turn to characterise general practices based on the relative prescribing behaviour to respond to a patient's observable information.[19] We generate two mutually exclusive groups of GP practices by taking the median of the distribution of $\hat{\rho}_{gt}$ in every period t to define those most and least responsive. Let the binary variable R_{gt} to take the value of 1 if the sensitivity to patients characteristics of practice g in quarter t is greater than the median in t (i.e. if $\hat{\rho}_{gt} \geq \text{median}_t[\hat{\rho}_{gt}]$, for each t): we call these most responsive, and the opposite group least responsive. Note that this definition allows for the responsiveness to change over time.[20]

With the double characterizations we have defined, in Table 4.4 we show the distribution of patients across the two types of prescribing style. The size-composition of these groups is relatively similar with the 'most compliant–most responsive' group concentrating roughly 22%, and 'most compliant–least responsive' 28% of patients.

TABLE 4.4: Distribution of patients into complying and responsive types

	most responsive	least responsive	Total
most compliant	103,464	133,808	237,272
least compliant	133,548	107,221	240,769
Total	237,012	241,029	478,041

[18] In OLS, the intercept is given by $\hat{\pi}_{gt} = E[RT_{igt}] - \hat{\rho}_{gt}E[z_{igt}]$, and given that the average patient has an index of $z_{igt} = 0$, then $\hat{\pi}_{gt} = E[RT_{igt}]$.

[19] See Appendix Figure 4.9 for the distribution of $\hat{\rho}_{gt}$.

[20] In Appendix Tables 4.6 and 4.7 we show the extent of the variation in responsiveness status, i.e. of R_{gt}, over the full sample period and for each year, across the GP practices in our sample. During the full sample period, 61 quarters between 2003q3 and 2018q3, a GP practice will change 25% of the time between most and least responsive status, with a SD of 6%. Across the years, the variation in responsiveness status remains relatively stable.

Finally, we estimate the following model:

$$Y^j_{igt} = \alpha_0 + \sum_{y=2003}^{T^j} [\beta_y C_{gt} + \gamma_y R_{gt} + \delta_y (C_{gt} \times R_{gt})] \times \mathbb{1}_{\{t \in y(t)\}} + \Phi X_{igt} + \mu Z_{igt} + \eta_g + \theta_{y(t)} + \varphi_{igt}, \tag{4.7}$$

where all variables are defined as above as per model in eq. (4.3). Thus, we will be comparing the outcomes of patients with the same set of observable baseline health characteristics, starting treatment in the same year but treated by GP practices that differ in their levels of compliance with prescription standards and responsiveness to observable patient information. From this model, our interest lies in the following set of coefficients, and combination of coefficients:

$$\begin{aligned}
\beta_y &= E[Y^j_{igt} \mid C_{gt} = 1, R_{gt} = 0] - E[Y^j_{igt} \mid C_{gt} = 0, R_{gt} = 0], \\
\beta_y + \delta_y &= E[Y^j_{igt} \mid C_{gt} = 1, R_{gt} = 1] - E[Y^j_{igt} \mid C_{gt} = 0, R_{gt} = 1], \\
\gamma_y &= E[Y^j_{igt} \mid C_{gt} = 0, R_{gt} = 1] - E[Y^j_{igt} \mid C_{gt} = 0, R_{gt} = 0] \text{ and} \\
\gamma_y + \delta_y &= E[Y^j_{igt} \mid C_{gt} = 1, R_{gt} = 1] - E[Y^j_{igt} \mid C_{gt} = 1, R_{gt} = 0],
\end{aligned}$$

for each year y. The first two, i.e. β_y and $\beta_y + \delta_y$, represent differences in outcomes between most and least compliant practices, within the groups of least and most responsive, respectively. The last two, i.e. γ_y and $\gamma_y + \delta_y$, differences in outcomes between most and least responsive practices, within the groups of least and most compliant, respectively.

4.3.3 Identification assumptions and balancing tests

The identification of the parameters of interest in the outcomes models, i.e. eqs. (4.3) and (4.7), is subject to the assumptions that patients do not register into practices with a certain prescribing style to receive a given treatment, and/or that practices may select patients based on their health characteristics. Both of these assumptions are likely to hold for our sample.

First, patients in our sample have been registered to their GP practice for an average of 12.4 years (median of 8.1 years) prior to the first statin issued to them by their treating practice. This suggests that it is unlikely that patients might have self-select themselves to their physicians while becoming at risk of cardiovascular disease.

Second, all specifications control for the full vector of characteristics that constitute their baseline risk, which may have a direct effect on outcomes themselves (riskier patients are more likely to experience poorer outcomes, everything else equal). The models also include time-invariant practice fixed effect that may capture geographical as well epidemiological characteristics of their registered population under care (every time they

are relatively stable over time), and year fixed effects which may capture unobservables that potentially reflect better management and care of the conditions, the adoption of new information on the safety and side-effect of the treatments, as well as epidemiological improvements or deteriorations in patients underlying health as they arrive to start treatment over time.

Finally, to verify that patients of a given risk are not systematically associated (matched) to and treated by practices of a certain prescribing behaviour, we conduct balancing tests to check that the covariates contained in the baseline risk profile are balanced across the characterization we propose. This is relevant as the initial health status impacts patient prognosis. To test this formally, we follow Pei et al. (2019) and perform a Right Hand Side (RHS) balancing test, in which we run regressions of compliance and responsiveness status of practices, C_{gt} and R_{gt}, on the vector X_{igt} and the patient's index Z_{igt}, and test whether the coefficient vector associated to these explanatory variables is equal to zero.

The models in eqs. (4.3) and (4.7) and the causal interpretation of the parameters of interest in them relies, clearly, on the conditional independence assumption, i.e. conditioning on known and observed covariates that determine the treatment, in this case the complying and responsive status of GP practice, the treatment is independent of potential outcomes. In all our model we control for relevant variables that are arguably related to the health outcomes and the prescribing behaviour of GP practices, such as the patients baseline risk at the moment they start drug treatment, and other set of variables aiming at capturing unobserved components of the health providers and epidemiological trends, through practice and time fixed effects, respectively. However, actual medical practice is plagued by endogeneity, and very likely prescribed behaviour is far from exogenous and depends highly on the very characteristics patients display when seeking care. To the extent that this is relevant and pervasive, the objective of this research is less ambitious and more modest, and we can claim only reasonable associations as we control for relevant but not every conceivable factor that affects clinical practice. Despite this, efforts are conducted to show that, at least statistically, the practice style that we capture appears exogenous from the perspective of the treated patients.

4.4 Results

4.4.1 Balancing tests

We start by discussing the results of the balancing tests. When we assess the balance of patients' baseline risk profile across types of compliance year by year (Appendix Tables 4.8, 4.9 and 4.10), we find that, at standard significance levels, we fail to reject the

null hypothesis of balanced covariates (see F-test and associated p-value reported at the end of the tables) for 14 out of 16 years. For the years 2011 and 2012, we reject such null. For these years, specific coefficient estimates suggest the possibility of matching associated with gender, smoking status and some cardiovascular conditions. When we assess the balance of patients baseline risk profile across types of responsiveness also year by year (Appendix Table 4.11), we fail to reject the null of balanced covariates for 15 out of 16 years and reject only for the first year 2003. However, a t-test on the patient index (included in the models from Tables 4.10 and 4.11) does not show signs of statistically significant correlation with any defined prescribing behaviour. Overall, our characterization of prescribing decision-making seems to result in a distribution of patients relatively similar across them, and therefore they support, to a large extent, our approach of examining outcomes between prescribing styles. In other words, this evidence seems to rule out the possibility that the association between prescribing styles and outcomes are driven by the patient-compositions of practices.

4.4.2 Compliance on outcomes

Figure 4.2 shows the estimated coefficients associated with the most compliant type of practice of model in eq. (4.3) for the four main outcomes, using the three-year cut-off.

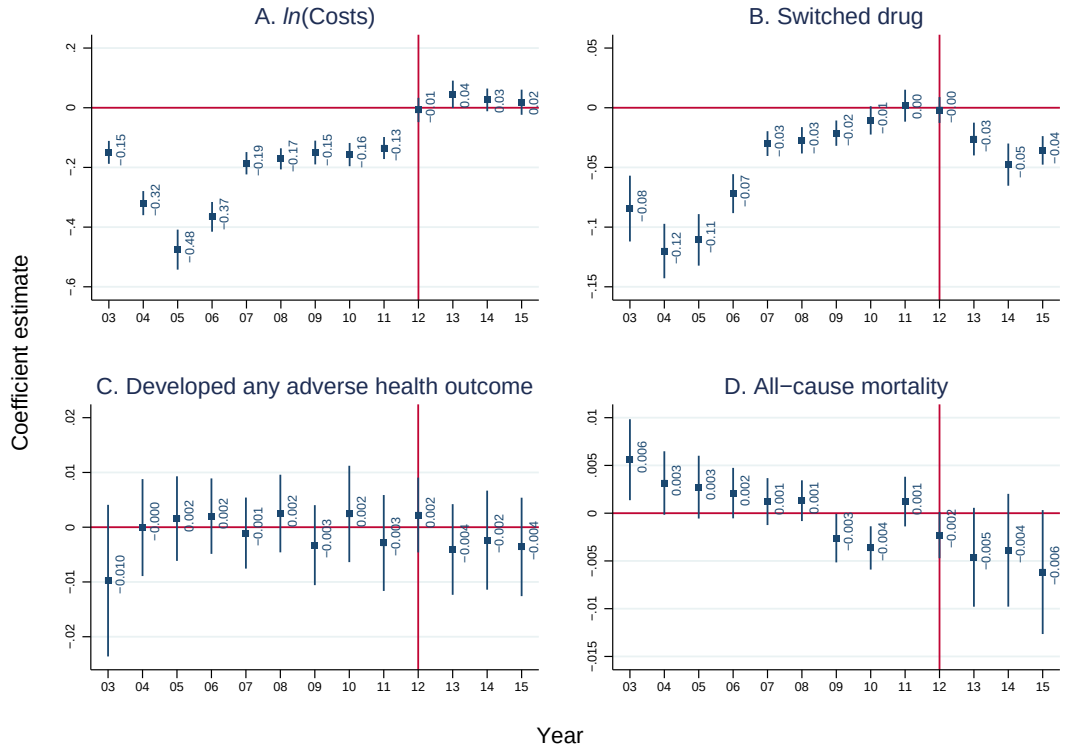


FIGURE 4.2: Outcomes and compliance types

NOTES. The figure plots OLS estimated coefficients β_y for each year y from model in eq. (4.3), along with 90% confidence intervals obtained from robust standard error clustered at the GP practice level, for the four main outcomes measured up to the three year cut-off. panel A uses as dependent variable *ln*(Cost); panel B an indicator for whether a patient had a first switch; panel C an indicator for whether a patient had a first cardiovascular event including hypertension; and panel D an indicator for whether a patient died (from any cause). The sample mean of the dependent variables is shown in Figure 4.1.

With regard to costs of treatment implied to the health service (panel A), not surprisingly we find that most compliant GP practices do exhibit significantly lower costs during the first period (before 2012) reflecting the greater differences in acquisition costs between the two main treatments. As this difference becomes marginal, from 2012 onwards, the difference in costs between complying types becomes insignificant. Specifically, the difference in 3-year treatment cost between least and most compliant providers is 15% in 2003 and 48% in 2005. This difference then decreases to between 19% to 13% for the period 2007-2011, averaging 24% per year before 2012. On average 23.7% per year. From 2012 onwards, no significant differences are found.

The decision to switch drug treatment can be grounded in medical or economic reasons. With the initial prescription patients may experience side-effect, intolerance, low adherence, which may trigger the decision to modify treatments. Alternative, because of cost-consciousness, practitioners might decide to change an initially expensive treatment for the cost-effective one. In terms of the probability that new patients experience

a switching of their originally prescribed treatment (panel B), we find significant differences between complying types, and patterns that are consistent with patent expiration dates. Specifically, the least compliant practices display a greater switching behaviour for their patients later on in the course of treatment. That is, despite them not being initially efficient, their complying behaviour comes with a lag. Further analysis (not presented) shows that more than 85% of switching relates to changes to the recommended treatment. The difference in probability between complying types peaks in 2004 with 12% of new patients starting treatment that year having their drugs switched by least compliant practices, then decrease down to 7% in 2006, to around 2% in 2009, to no significant differences between 2010 and 2012. Then switching is greater again for least compliant, from 2013 to 2015, reaching a probability between 3% to 5%.

Related to the probability of experiencing an adverse health outcome (panel C), we find no significant differences between complying types in any year when using the composite measure. When looking at the individual conditions underlying the composite measure, shown in Figure 4.3, we find isolated significant effects for some conditions in some years,[21] yet they are not systematic over time. Finally, with regard to all-cause mortality (panel D), we find significant effects only for three years of the sample period (2003, 2009 and 2010). Additionally, these results exhibit a clear downward sloping trend, suggesting that this outcome improves over time, for most compliant practices. The trend suggests that patients whose providers are intensive in the use of one particular treatment (i.e. atorvastatin) are associated with lower mortality rates. That is before 2012 most compliant practices are those that prescribe less atorvastatin and more simvastatin, and it is in this period that their patients show a greater probability of any-cause mortality. After 2012, most compliant are the ones prescribing more atorvastatin and less simvastatin, and in this period their patients experience a lower probability of any-cause mortality. Albeit, for most years the effects are not statistically significant, at standard levels.

[21] For example, acute myocardial infarction in 2010, atrial fibrillation in 2015, congestive cardiac failure in 2004, hypertension in 2003, ischaemic stroke in 2009, and peripheral arterial disease in 2008 and 2011.

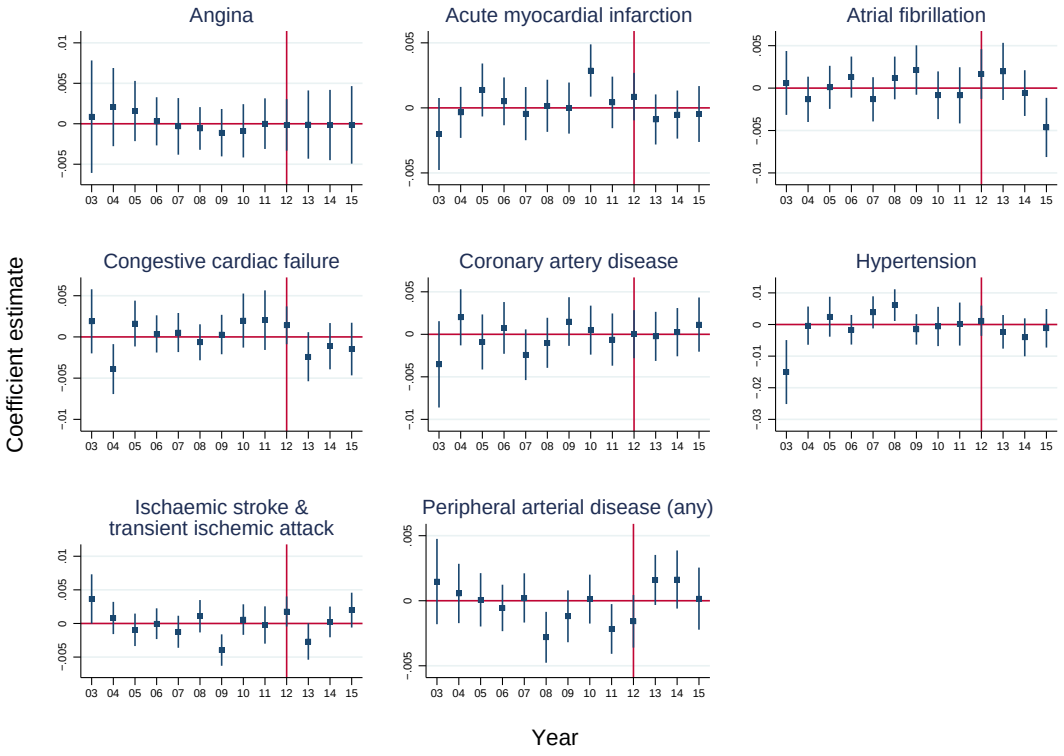


FIGURE 4.3: Specific adverse health outcomes and compliance types

NOTES. The figure plots OLS estimated coefficients β_y for each year y from model in eq. (4.3), along with 90% confidence intervals obtained from robust standard error clustered at the GP practice level, for the eight individual health outcomes measured up to the three year cut-off.

4.4.3 Compliance and responsiveness on outcomes

Empirical clinical practice norm. In section 4.3.2 we use eq. (4.4) to identify those patients who, based on their observable characteristics, are more or less likely to be given a specific treatment, which in turn reflects an empirical clinical practice norm of the general practice sector. Here we discuss these results. Appendix Figures 4.10 and 4.11 show the coefficient estimates of Γ, using the two main treatments as dependent variable. In the model, Γ represents the marginal contribution of each health attribute in X on their likelihood to receive a given treatment. These estimates suggest, on one hand, that some aspects of patient baseline health matter significantly while others do not, and, on the other hand, that the relative importance of these aspects changes over time and is related to the main events in this drug market (especially patent expirations). For example, characteristics that significantly affect the probability of being prescribed either of the two main treatments in the sample period are age, presence of chronic kidney disease, diabetes, hypertension, and CVDs such as angina, acute myocardial infarction (AMI) and coronary artery disease. Other characteristics such as gender, ethnicity,

socio-economic status, and alcohol consumption, do not seem to affect prescription decisions significantly. Regarding the dynamic influence that some characteristics have, it is clear from the sub-figures that it changes according to the main exogenous shocks to prescriptions. Remarkable examples are age, chronic kidney disease, diabetes, angina, AMI, coronary artery disease, and hypertension: the importance of these aspects in affecting the probability of being prescribed the recommended treatments display an inflexion point, especially after the second patent expiration in 2012q2 (i.e. of atorvastatin). Overall, this shows that the same patients, i.e. with the same set of observables, has different probabilities in a different period to be given a specific treatment.

Then, initial drug treatment does select on observable patient characteristics and based on this fact, we can turn to capture responsiveness to patient characteristics and investigate its association with patient outcomes.

Compliance and responsiveness on outcomes. Next, we present the results from the model in eq. (4.7) in which we study outcomes between health providers that differ in compliance and responsiveness. Figure 4.4 shows the results from the costs model. When focusing on differences between most and least compliant, we find that it is within the group of least responsive where the differences are greater (panel A). Within this group, most complaint providers have a cost of treatment, on average, 23.7% lower per year, before 2012, than least compliant ones. Least responsive practices by definition do not place greater emphasis on their patient's health status summarized by the index, and therefore, most compliant providers within this group are more strictly compliant with cost-effective prescribing. Most responsive providers do place more weight on their patient's information, and consequently, they prescribe the more expensive treatment to patients for which the clinical norm would suggest so. This explains that within the most responsive group differences in cost between compliance types are smaller in magnitude (panel C). On the other hand, differences between response types do not imply sizeable differences in the cost of treatment overall. However, there are statistically significant differences within the least compliant group (panel B) suggesting that most responsive are associated with lower costs: of 7% to 18% in the years 2004-2006, and of 6% to 9% in 2011, 2013 and 2014. No meaningful differences are found between responsive types within the most compliant group (panel D).

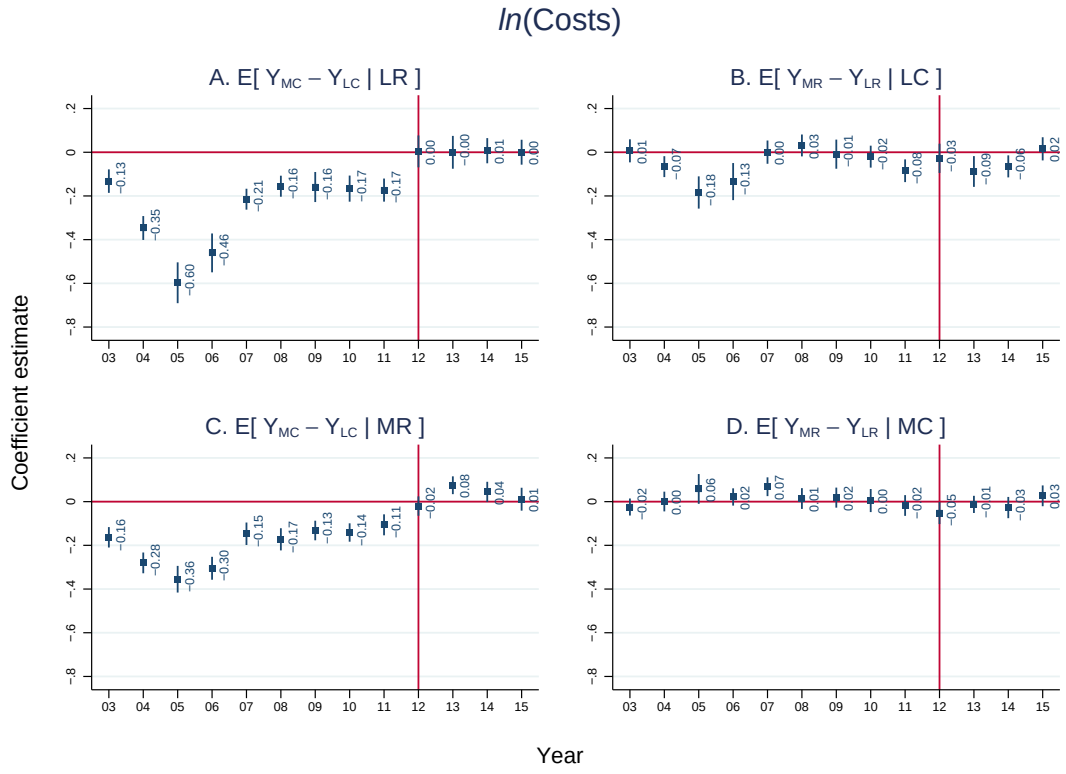


FIGURE 4.4: *ln*(Cost) and compliance and response types

NOTES. The figure plots OLS estimated coefficients associated to types of compliance and responsiveness from model in eq. (4.7), along with 90% confidence intervals obtained from robust standard error clustered at the GP practice level, for *ln*(Cost) measured up to the three year cut-off. Each panel plots the following estimates: panel A, β_y; panel B, γ_y; panel C, $\beta_y + \delta_y$; and panel D, $\gamma_y + \delta_y$, for each year y.

The results for the switching outcome are presented in Figure 4.5. Similarly, with cost, significant differences are found between the most and least compliant groups. In particular, the largest differences persist within the group of least responsive practices (panel A) and smaller in magnitude within most responsive (panel C). No systematic significant differences are found between the most and least responsive, except for a 4% lower probability of switching drugs between most and least responsive within the least compliant group in 2004 (panel B).

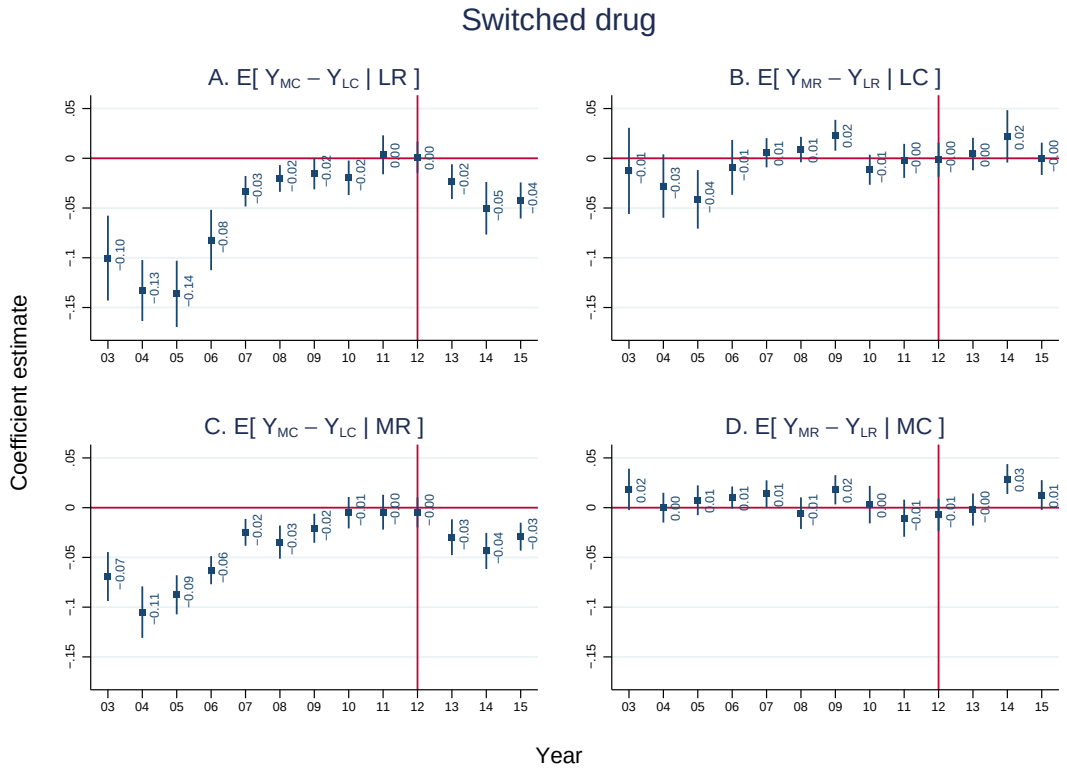


FIGURE 4.5: Switching and compliance and response types

NOTES. The figure plots OLS estimated coefficients associated to types of compliance and responsiveness from model in eq. (4.7), along with 90% confidence intervals obtained from robust standard error clustered at the GP practice level, for the probability of switching drug treatment measured up to the three year cut-off. Each panel plots the following estimates: panel A, β_y; panel B, γ_y; panel C, $\beta_y + \delta_y$; and panel D, $\gamma_y + \delta_y$, for each year y.

Despite significant differences in costs between most and least compliant types, we find no significant differences in health outcomes, measured as the probability of experiencing an adverse cardiovascular event within the first three years of treatment. Although there is some evidence of significant worse outcomes for patients treated by most responsive practices (with respect to least responsive), for example in 2015 within the least compliant group (panel B) and in 2012 and 2014 within most compliant group (panel D) these are not consistent over time. Nevertheless, taken at face value this would suggest that being most responsive to the clinical norm (summarized in the patient index) after 2012, would have been detrimental for health outcomes. Then, simply ignoring patients characteristics, and following cost-effective reasoning would have been less harmful (as in the results from panel A).

Inferring from this evidence that there are no statistically significant differences in health outcomes between any of our defined prescribing types of providers, we conclude that it the most compliant and least responsive type of general practice that behaves more cost-effectively, as they are associated to lower costs.

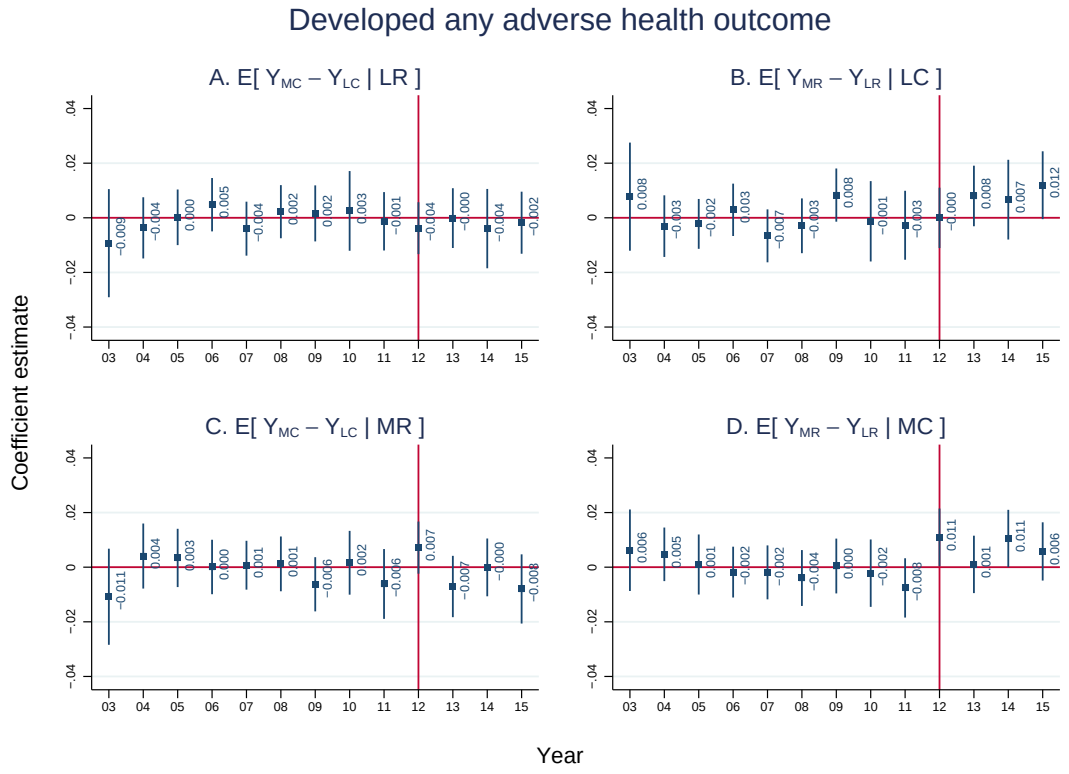


FIGURE 4.6: 3-year risk of adverse health outcomes and compliance and response types

NOTES. The figure plots OLS estimated coefficients associated to types of compliance and responsiveness from model in eq. (4.7), along with 90% confidence intervals obtained from robust standard error clustered at the GP practice level, for the probability of developing an adverse cardiovascular condition measured up to the three year cut-off. Each panel plots the following estimates: panel A, β_y; panel B, γ_y; panel C, $\beta_y + \delta_y$; and panel D, $\gamma_y + \delta_y$, for each year y.

Finally, Figure 4.7 shows the result for all-cause mortality. In terms of differences between most and least compliant providers, we find evidence of significant worse outcomes within the least responsive group early in the sample period before 2012 (panel A), and better outcomes within the most responsive group after 2012 (panel C). Then this suggest strict compliant, those that place less weight to patients characteristics when making drug choices before 2012, i.e. heavy prescribers of simvastatin and light prescribers of atorvasatin, exhibit poorer outcomes, and that compliant also sensitive to their patients characteristics after 2012 perform better. Just as before, this can be consider evidence of a relation between better outcomes (but only specific to all-cause mortality) and the more intense use of one specific treatment: atorvastatin. In terms of responsiveness, we find evidence of significantly worse performance for most responsive within the least complaint group in 2006 and 2015 (panel B) suggesting that, within this group, emphasizing the norm reflected in the index was detrimental for patients. However, within the most compliant group, most responsive practices shows statistically better outcomes later in the sample period, in years 2013 and 2014, which would

suggest that only within high prescribers of atorvastatin, deviations according to norm were beneficial.

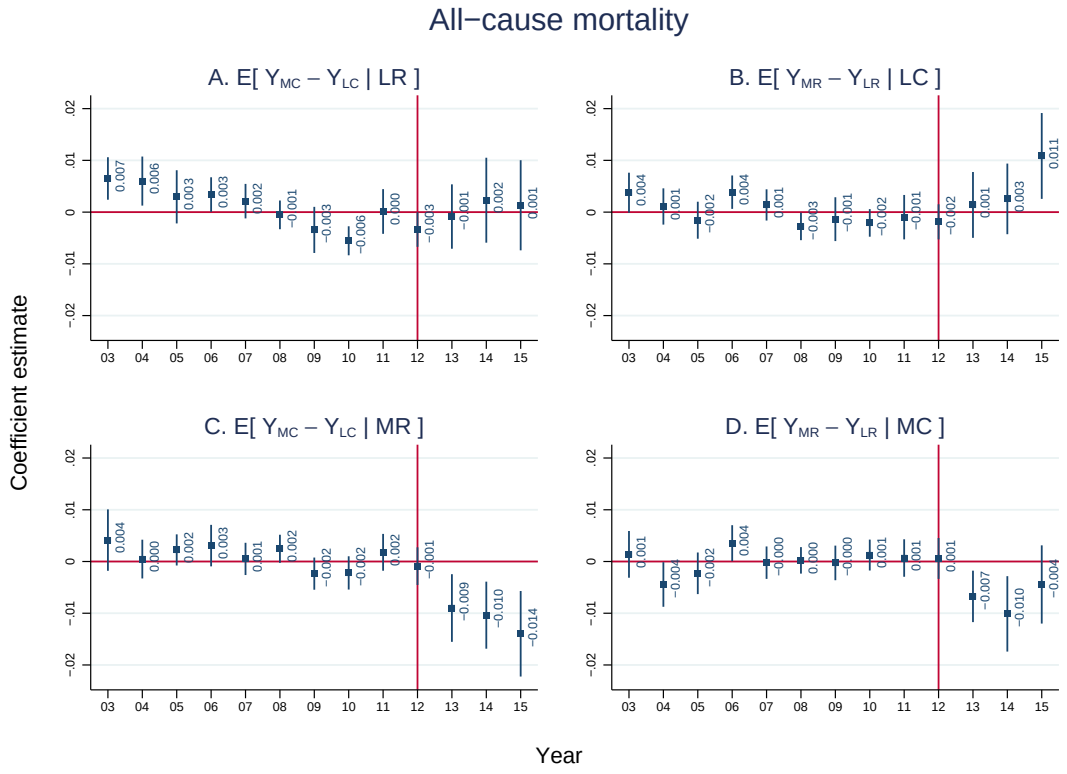


FIGURE 4.7: All-cause mortality and compliance and response types

NOTES. The figure plots OLS estimated coefficients associated to types of compliance and responsiveness from model in eq. (4.7), along with 90% confidence intervals obtained from robust standard error clustered at the GP practice level, for the probability of all-cause mortality measured up to the three year cut-off. Each panel plots the following estimates: panel A, β_y; panel B, γ_y; panel C, $\beta_y + \delta_y$; and panel D, $\gamma_y + \delta_y$, for each year y.

Role of patients in explaining heterogeneity. Finally, we ask what is the role of patients characteristics in explaining the observed variation in drug treatment choices made by the general practices in our sample. For this, we rely again on the model in eq. (4.4). In here the decision to prescribe the recommended treatment has two components: an observable patient characteristics one (given by ΓX_{igt}) and a practice-specific other (λ_{gt}). By estimating the model in eq. (4.4), and similarly for rerunning estimations but including or excluding each of the two components, we can study the extent to which patients observable characteristics help explain variation in initial treatment choices. To do this, we discuss two measures of goodness-of-fit: the standard R^2 and

the area under the receiver operating characteristics (ROC) curve (AUC).[22] These results are presented in Table 4.5. In terms of the R^2, columns (1) and (3) show that patients characteristics alone account for approximately a third of the total variation in recommended treatments that we can account for, as the models including the patient component only (model 1) have a coefficient of determination of around 0.05 while the full model including also the unobservable practice component (model 5) results in a coefficient of around 0.16. Using the AUC statistics we obtain similar findings: although the overall predictability of each model is high (above 0.82), the predictability power of model 1 is the lowest among all specifications. This evidence suggest that the heterogeneity in the allocation of initial treatments across general practices documented in Ortiz de Zarate et al. (2021) is partially explained by the initial health status of the treated patients, but most of it is explained by unobservable characteristics of the health providers themselves.

[22] In linear probability models, the R^2 equals the difference between the average predicted probability in the two groups given by the values of the binary dependent variable; it also measures the fraction of the explained part of the variance due to the difference between the conditional means (see Gronau et al. (1998)). The area under the Receiver Operating Characteristic (ROC) curve, or *c*-statistic, is used to asses the predictability power of the model. This statistic is interpreted as the probability that the models prediction will rank a randomly chosen patient that was prescribed the recommended treatment higher than a randomly chosen patient who was not. Higher values of the AUC indicate a better ability of the model to predict correctly (see Fawcett (2006); Cameron and Trivedi (2005)).

TABLE 4.5: Goodness of fit statistics

	(1)	(2)	(3)	(4)
	panel A: $RT_{igt} = Simvastatin_{igt}$			
	In sample		Out of sample	
Model	R^2	AUC (ROC)	R^2	AUC (ROC)
1. ΓX_{igt}	0.050	0.821	0.051	0.820
2. λ_g	0.087	0.848	0.097	0.839
3. λ_{gt}	0.116	0.856	0.144	0.830
4. $\Gamma X_{igt} + \lambda_g$	0.132	0.861	0.142	0.852
5. $\Gamma X_{igt} + \lambda_{gt}$	0.160	0.868	0.160	0.844
	panel B: $RT_{igt} = Atorvastatin_{igt}$			
	In sample		Out of sample	
Model	R^2	AUC (ROC)	R^2	AUC (ROC)
1. ΓX_{igt}	0.052	0.824	0.053	0.822
2. λ_g	0.080	0.850	0.090	0.840
3. λ_{gt}	0.109	0.858	0.138	0.831
4. $\Gamma X_{igt} + \lambda_g$	0.128	0.863	0.138	0.854
5. $\Gamma X_{igt} + \lambda_{gt}$	0.155	0.870	0.155	0.845

NOTES. The table presents R^2 and the area under the receiver operating characteristics (ROC) curve for different versions of the model in eq. (4.4), under column 'Model'. panel A uses as dependent variable $RT_{igt} = Simvastatin_{igt}$ and panel B $RT_{igt} = Atorvastatin_{igt}$. Columns (1) and (3) show average R^2 for the 61 estimations of each model by OLS (i.e. for the 61 moving 3-quarter periods between 2003q3 and 2018q3)). Columns (2) and (4) show the area under the receiver operating characteristics curve, obtained by performing a non-parametric ROC analyses using the predicted values of each model against the dependent variable. STATA command `roctab` was used. Both statistics are computed in sample (columns (1) and (2)), and out-of-sample (columns (3) and (4)), meaning that estimation was performed for a random 50% sub-sample and the AUC was obtained after predicting out of sample, i.e. on the remaining sub-sample.

4.5 Conclusions

In this paper, we investigate the role of prescribing variation in costs and patient outcomes. To do this we classify our health providers – GP practices – according to two dimensions, namely compliance with cost-effective prescribing and responsiveness to patients characteristics. The characterization we propose is based on the initial drug choice with which a sample of general practices prescribes statins for new patients to prevent CVD risk. Using balancing tests we show that, to a large extent, there is no matching between patients and prescribing types, which seems to rule out that the results are driven by the patient-compositions of practices.

Our analysis is conducted in two stages. In the first stage, we focus only on the complying dimension and its relation with outcomes. Our findings suggest that least compliant practices are associated with higher treatment costs per patient. However, this comes with no additional significant benefit for patients, measured as the 3-year risk of

developing a CVD since starting treatment. We also find that the least compliant practices are more likely to switch their patients' drug treatment, towards the cost-effective alternative later on in the course of treatment, which can attenuate differences in costs that could have been larger otherwise.

In the second stage of the analysis we incorporate the responsive dimension. First, using an index reflecting an empirical clinical norm, we show that the way practitioners weigh patients' health characteristics when allocating treatments changes over time. Second, we find that being most compliant and least responsive (to the norm), i.e. strict compliance, results in having the most cost-effective prescribing behaviour, as this type is associated with the lowest cost levels, while there are no statistically significant differences in health outcomes across all prescribing types. Lastly, we estimate that patients underlying health explains roughly one-third of the variation in initial treatment choices we can account for. Taken all together our findings suggest that, although relevant in explaining variation in treatment choices, responding to patients health information when making prescribing decisions does not produce significant differences in costs or health outcomes.

The conclusions that can be drawn from our analysis are in line with the evidence coming from other health settings: large variation in treatment use across providers, significant differences in costs, at no extra benefit in terms of health outcomes. Then further efforts could have been made to contain heterogeneity without sacrificing outcomes.

This study is not without limitations. First, given data constraints, we can only focus on aggregate decisions made at the general practice level. Further research given proper data availability should focus on heterogeneity in decision making that takes place between physicians within the same practice. Second, primary care is not only about curing but also about caring. In that sense, the characterization and the outcomes we study may not reflect the full scope of the production of health services that are given by general practitioners. Physician behaviour could be further characterized as to reflect, for example, continuity of care (i.e. seeing the same practitioner throughout treatment), engagement with the therapeutic process (e.g. frequency of follow up appointments), stability/volatility of the prescriptions history, referrals to secondary specialized services, etc. And third, we focus on the prescriptions of the single largest therapeutic class prescribed in primary care, cholesterol-lowering drugs, which are used to prevent one of the leading causes of death in the UK. Yet, a more comprehensive study of the variation in the use of medical treatments and interventions in this setting would require to incorporate also the study of drugs used for hypertension, heart failure, stomach ulcers (e.g. proton-pump inhibitor), depressions, asthma, and conditions including diabetes and obesity as all these represents the largest drug classes and prevalent conditions in England, as well as to explore what are the specific factors and events that have influenced the market dynamics of all these relevant therapeutic areas.

4.6 Appendix

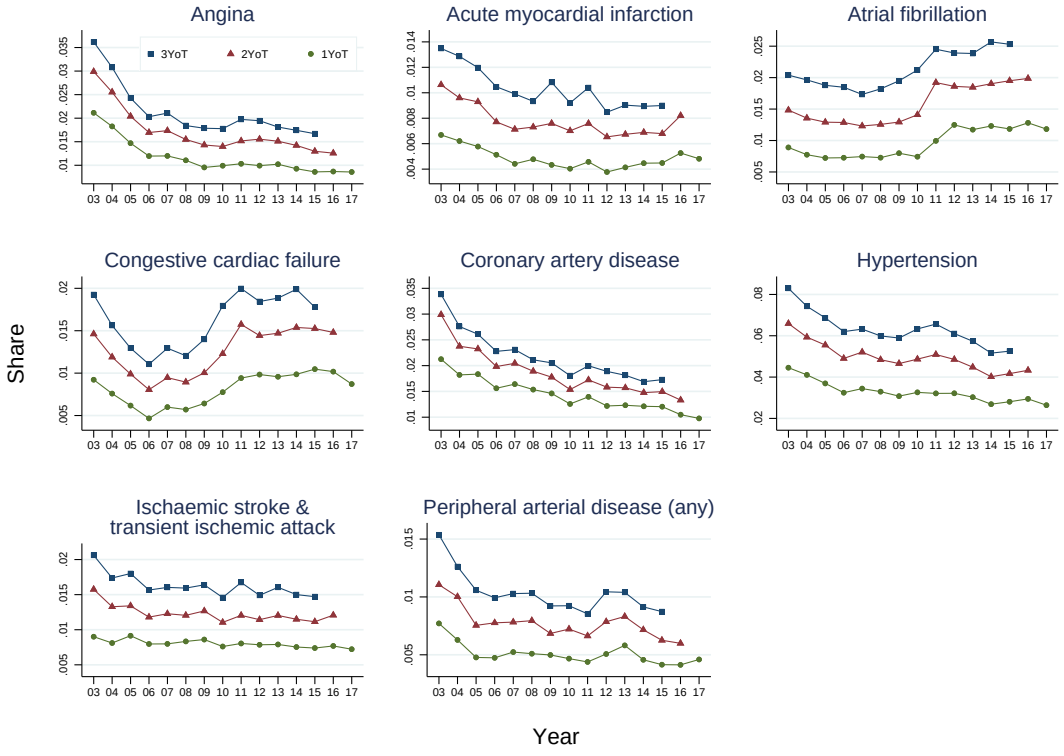


FIGURE 4.8: Sample mean of individual adverse health outcomes

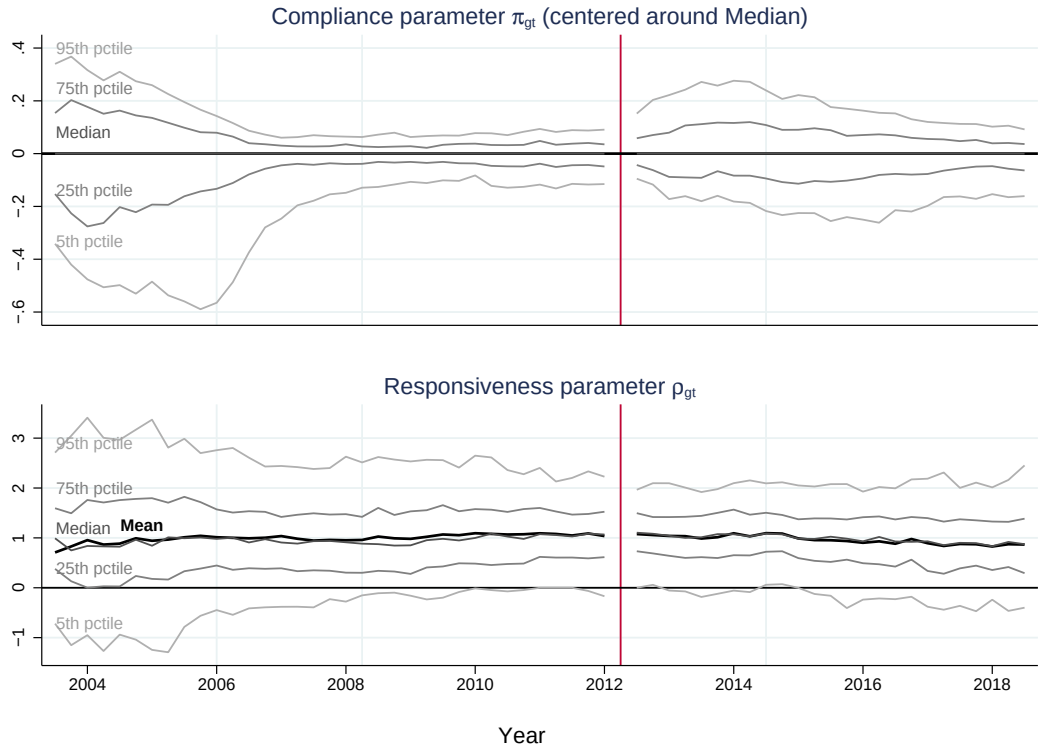


FIGURE 4.9: Distribution of compliance and responsiveness parameters ($\hat{\pi}_{gt}$ and $\hat{\rho}_{gt}$) across practices and time

NOTES. The figure shows the distribution of compliance and responsiveness parameters ($\hat{\pi}_{gt}$ and $\hat{\rho}_{gt}$) across the 243 GP practices and the 61 year-quarters between 2003q3 and 2018q3.

TABLE 4.6: Changes in Compliance and Responsiveness status
Full sample period (Percentage over 61 quarters for the 243 GP practices)

	Mean	SD	Min.	p25	Median	p75	Max.
Compliance	15.87	6.2	0	11.48	16.39	19.67	39.34
Responsiveness	24.58	6.32	6.56	19.67	24.59	27.87	42.62

NOTES. The table presents the percentages (%) of time-quarters a GP practice is observed changing Compliance and Responsiveness status over the full sample period, that is, over the 61 quarters between 2003q3 and 2018q3. These figures are obtained by counting the number of time-quarters a practice changes compliance and responsiveness status, i.e. C_{gt} and R_{gt} in t with respect to $t-1$, for the full sample period.

TABLE 4.7: Changes in Compliance and Responsiveness status
For each year (Percentage over 4 quarters for the 243 GP practices)

	Compliance		Responsiveness	
Year	Mean	SD	Mean	SD
2003	6.90	17.29	18.83	24.28
2004	5.83	11.31	22.22	18.73
2005	9.26	14.79	23.77	20.79
2006	10.89	16.24	18.26	19.73
2007	13.48	17.34	17.80	18.31
2008	13.79	18.19	18.35	19.01
2009	14.09	16.99	15.33	18.01
2010	11.73	17.05	12.55	16.76
2011	14.85	18.81	15.74	19.20
2012	30.35	18.66	17.49	19.22
2013	8.64	15.82	18.11	19.32
2014	11.21	16.24	23.05	20.55
2015	8.74	14.66	19.55	18.98
2016	11.83	17.65	19.55	20.55
2017	11.11	16.23	19.44	17.31
2018	11.25	19.68	13.72	18.54

NOTES. The table presents the percentages (%) of time-quarters a GP practice is observed changing Compliance and Responsiveness status for each year of the sample period. These figures are obtained by counting the number of time-quarters a practice changes compliance and responsiveness status, i.e. C_{gt} and R_{gt} in t with respect to $t-1$, for each year of the data (4 quarter in each year).

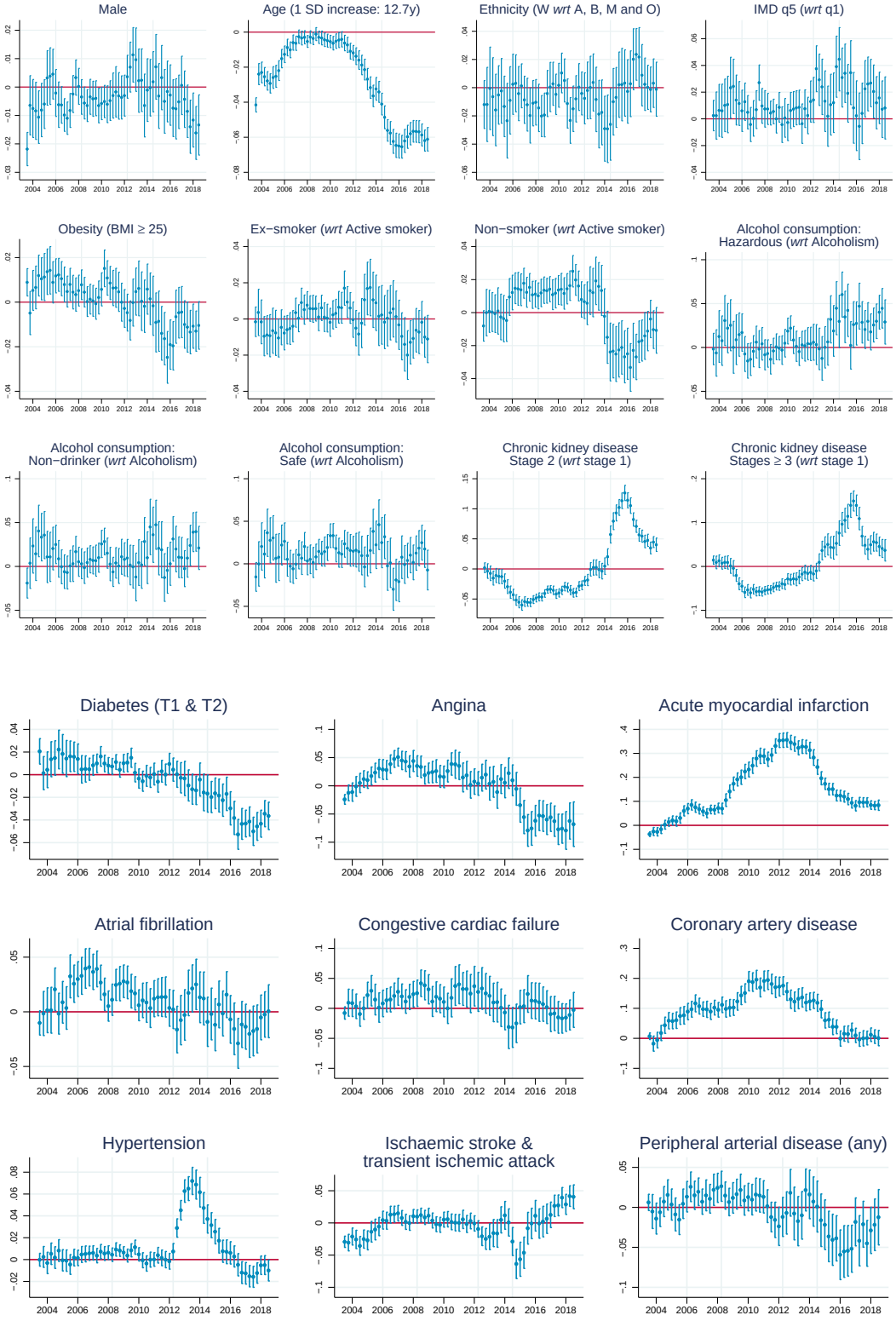


FIGURE 4.10: Estimated coefficients Γ (with $RT_{igt} = Atorvastatin_{igt}$)

NOTES. The figure shows the estimated coefficients Γ from model in eq. (4.4) using $RT_{igt} = Atorvastatin_{igt}$ as dependent variable, for each moving 3-quarter period t. Plotted are the point estimates and 90% confidence intervals, obtained from robust standard errors clustered at the GP practice level. All variables are binary indicators for a patient having the attribute given by the title name each sub-graph, and age is a continuous variable. As age enters in a quadratic form in the model, i.e. $\gamma_1 Age + \gamma_2 Age^2$, in graph 'Age' is plotted $\hat{\gamma}_1 + 2\hat{\gamma}_2 \bar{Age}$, for $\bar{Age} = 63.66$, and in order to express this marginal effect in terms of 1 SD increase in age, is then multiplied by 12.7 years (the sample SD of age).

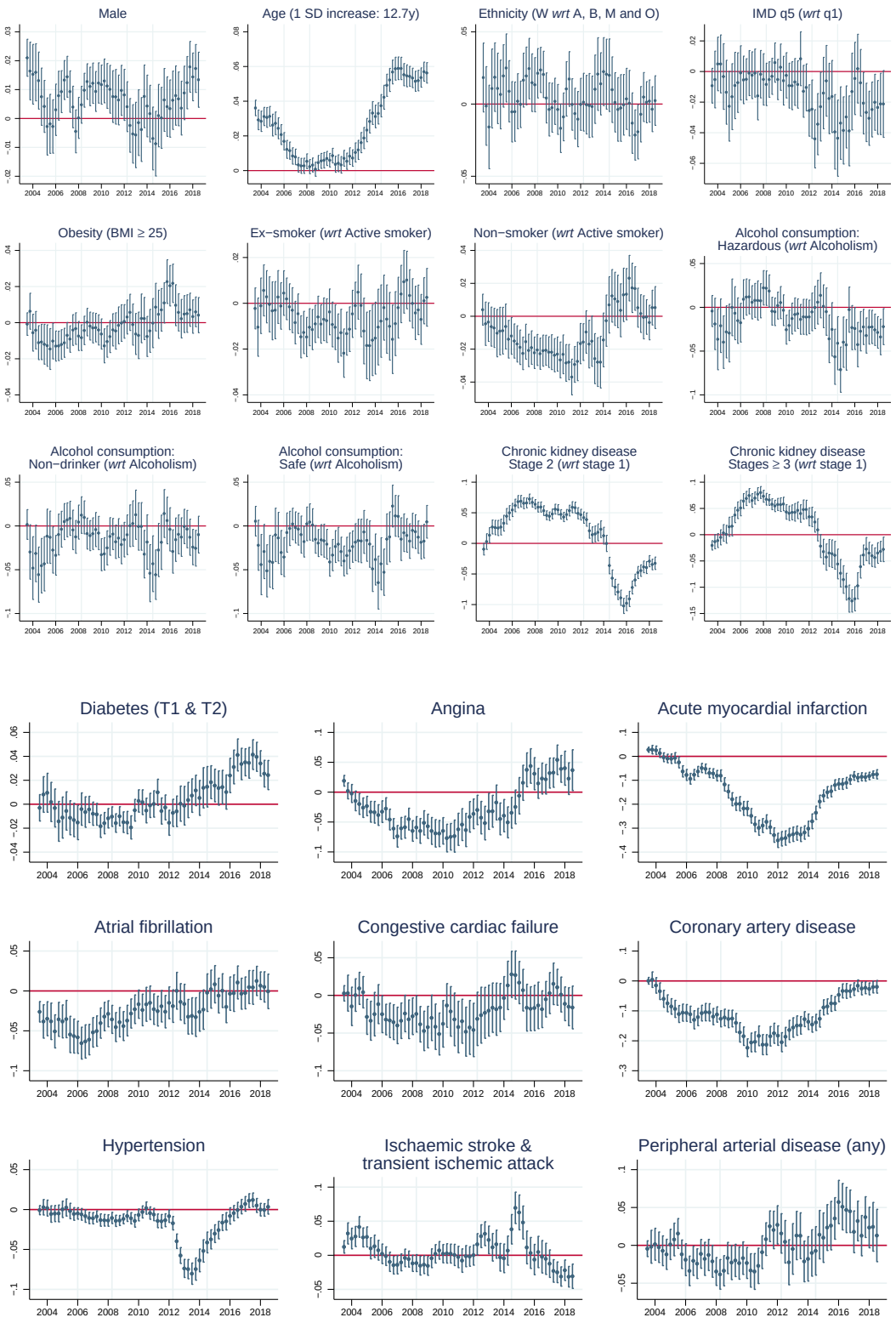


FIGURE 4.11: Estimated coefficients Γ (with $RT_{igt} = Simvastatin_{igt}$)

NOTES. The figure shows the estimated coefficients Γ from model in eq. (4.4) using $RT_{igt} = Simvastatin_{igt}$ as dependent variable, for each moving 3-quarter period t. Plotted are the point estimates and 90% confidence intervals, obtained from robust standard errors clustered at the GP practice level. All variables are binary indicators for a patient having the attribute given by the title name each sub-graph, and age is a continuous variable. As age enters in a quadratic form in the model, i.e. $\gamma_1 Age + \gamma_2 Age^2$, in graph 'Age' is plotted $\hat{\gamma}_1 + 2\hat{\gamma}_2 \bar{Age}$, for $\bar{Age} = 63.66$, and in order to express this marginal effect in terms of 1 SD increase in age, is then multiplied by 12.7 years (the sample SD of age).

TABLE 4.8: Balancing test: Compliance on patient characteristics (PART 1)

	(1) 2003	(2) 2004	(3) 2005	(4) 2006	(5) 2007	(6) 2008	(7) 2009	(8) 2010
Age	0.000	-0.002*	0.000	-0.000	-0.000	0.001	-0.000	0.001
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age^2	0.000	0.000*	-0.000	0.000	0.000	-0.000	0.000	-0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	-0.002	0.002	0.002	0.003	-0.001	0.004	0.008	-0.003
	(0.003)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)	(0.005)	(0.004)
White	0.002	-0.004	-0.006	-0.011	0.001	-0.005	-0.001	-0.011
	(0.005)	(0.005)	(0.005)	(0.007)	(0.008)	(0.008)	(0.009)	(0.007)
IMD Quintile 2	0.003	-0.001	0.002	-0.004	-0.010	0.003	0.005	-0.002
	(0.003)	(0.005)	(0.006)	(0.005)	(0.008)	(0.006)	(0.006)	(0.006)
IMD Quintile 3	0.007	0.004	0.002	-0.001	0.001	0.001	0.010	0.004
	(0.004)	(0.004)	(0.006)	(0.005)	(0.008)	(0.009)	(0.008)	(0.005)
IMD Quintile 4	0.001	0.004	-0.001	-0.005	-0.005	0.008	0.008	-0.001
	(0.004)	(0.004)	(0.006)	(0.006)	(0.008)	(0.008)	(0.009)	(0.006)
IMD Quintile 5	-0.000	0.006	-0.005	-0.004	0.006	-0.003	0.008	-0.001
	(0.004)	(0.005)	(0.006)	(0.005)	(0.009)	(0.010)	(0.009)	(0.006)
CKD Stage 2	0.001	-0.003	-0.000	-0.002	0.008	-0.001	-0.003	0.002
	(0.003)	(0.003)	(0.004)	(0.003)	(0.005)	(0.004)	(0.005)	(0.005)
CKD Stages 3-5	-0.002	-0.002	0.004	-0.002	0.002	0.013	-0.011	-0.004
	(0.004)	(0.003)	(0.004)	(0.005)	(0.007)	(0.007)	(0.010)	(0.006)
Diabetes (T1 & T2)	-0.001	-0.005	-0.002	0.001	-0.001	-0.002	-0.000	-0.002
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.005)	(0.006)	(0.004)
Obesity (BMI≥25)	-0.001	-0.000	-0.000	-0.006	-0.006	-0.002	0.002	-0.001
	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
Ex-smoker	-0.002	-0.001	-0.002	0.002	0.005	-0.003	-0.001	0.002
	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)
Non-smoker	-0.007	-0.001	-0.002	-0.004	-0.001	0.001	-0.004	0.000
	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)	(0.004)
Hazardous	-0.007	-0.007	-0.007	0.006	-0.023*	-0.002	-0.007	-0.017
	(0.008)	(0.006)	(0.008)	(0.009)	(0.009)	(0.008)	(0.011)	(0.009)
Non-drinker	-0.004	-0.007	-0.004	0.011	-0.022*	0.002	0.001	-0.022*
	(0.009)	(0.007)	(0.008)	(0.009)	(0.009)	(0.010)	(0.012)	(0.009)
Safe	-0.007	-0.006	-0.001	0.010	-0.024**	-0.001	0.000	-0.019*
	(0.009)	(0.006)	(0.008)	(0.009)	(0.009)	(0.008)	(0.010)	(0.009)
Angina	0.005	-0.002	-0.002	-0.004	-0.004	0.001	-0.026	-0.017*
	(0.003)	(0.004)	(0.005)	(0.005)	(0.008)	(0.008)	(0.015)	(0.007)
AMI	-0.005	-0.010*	-0.003	-0.002	-0.005	-0.014	-0.003	-0.012
	(0.004)	(0.004)	(0.006)	(0.006)	(0.007)	(0.008)	(0.008)	(0.008)
ATF	0.004	0.002	0.007	0.008	-0.009	-0.010	0.002	-0.010
	(0.006)	(0.004)	(0.005)	(0.005)	(0.008)	(0.009)	(0.009)	(0.007)
CCF	-0.009	-0.000	-0.001	0.009	-0.007	-0.002	0.002	0.022
	(0.007)	(0.005)	(0.007)	(0.008)	(0.010)	(0.012)	(0.012)	(0.013)
CAD	-0.004	-0.001	-0.004	-0.022**	0.002	0.002	-0.008	0.001
	(0.006)	(0.006)	(0.008)	(0.008)	(0.008)	(0.011)	(0.011)	(0.010)
IS & TIA	-0.006	-0.003	-0.004	-0.006	-0.010	-0.009	0.006	-0.002
	(0.004)	(0.003)	(0.005)	(0.005)	(0.007)	(0.007)	(0.006)	(0.006)
PAD	0.015*	0.005	-0.004	-0.011	0.006	0.008	-0.009	0.018
	(0.007)	(0.005)	(0.008)	(0.009)	(0.010)	(0.009)	(0.012)	(0.011)
Hypertension	0.004	-0.001	0.004	0.004	-0.004	0.001	0.001	-0.010**
	(0.003)	(0.002)	(0.003)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)
N	18536	38337	34981	38381	31679	31190	30675	28239
R^2	0.874	0.854	0.764	0.740	0.637	0.656	0.626	0.726
F-stat	1.195	1.088	1.162	1.147	1.085	0.931	0.569	1.476
p-value (F test)	0.245	0.357	0.276	0.291	0.360	0.562	0.953	0.073

NOTES. The table presents OLS estimates from the regression: $C_{gt} = \Omega X_{igt} + \varepsilon_{gt}$, estimated for each year between 2003 and 2018. Each regression controls for GP practice fixed effects. F-stat and p-value correspond to the null hypothesis that coefficient vector $\Omega = 0$, or that patient's baseline risk is balanced across compliance types. Standard errors clustered at the GP practice level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 4.9: Balancing test: Compliance on patient characteristics (PART 2)

	(1) 2011	(2) 2012	(3) 2013	(4) 2014	(5) 2015	(6) 2016	(7) 2017	(8) 2018
Age	0.002	0.001	-0.000	0.000	0.001	-0.001	0.001	-0.001
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Age^2	-0.000	-0.000	0.000	-0.000	-0.000	0.000	-0.000	0.000
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
Male	-0.011**	-0.002	0.000	0.002	-0.000	-0.000	-0.001	-0.001
	(0.004)	(0.005)	(0.002)	(0.003)	(0.003)	(0.004)	(0.003)	(0.003)
With	0.008	-0.007	0.006	-0.001	-0.001	-0.006	-0.001	-0.000
	(0.009)	(0.011)	(0.006)	(0.006)	(0.006)	(0.007)	(0.006)	(0.007)
IMD Quintile 2	-0.015	-0.003	-0.002	-0.001	-0.009	0.000	-0.001	-0.011*
	(0.008)	(0.009)	(0.006)	(0.009)	(0.005)	(0.007)	(0.006)	(0.005)
IMD Quintile 3	-0.004	0.003	0.000	-0.005	-0.002	0.009	-0.002	-0.009
	(0.008)	(0.010)	(0.005)	(0.007)	(0.006)	(0.007)	(0.006)	(0.007)
IMD Quintile 4	-0.002	-0.005	-0.004	-0.014	-0.007	0.006	0.001	-0.010
	(0.008)	(0.010)	(0.005)	(0.008)	(0.007)	(0.008)	(0.007)	(0.008)
IMD Quintile 5	-0.014	0.008	-0.004	-0.008	-0.012*	-0.003	0.003	-0.016
	(0.009)	(0.010)	(0.005)	(0.008)	(0.005)	(0.009)	(0.007)	(0.009)
CKD Stage 2	0.005	0.002	-0.005	-0.001	0.003	0.002	-0.004	0.001
	(0.005)	(0.008)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)
CKD Stages 3-5	0.016	0.008	-0.002	-0.005	-0.005	0.003	-0.009	0.001
	(0.009)	(0.012)	(0.007)	(0.007)	(0.007)	(0.008)	(0.009)	(0.008)
Diabetes (T1 & T2)	0.004	-0.010	0.001	-0.005	-0.007	-0.000	-0.005	0.000
	(0.005)	(0.007)	(0.004)	(0.004)	(0.005)	(0.004)	(0.004)	(0.005)
Obesity (BMI≥25)	-0.001	0.005	-0.003	0.001	-0.003	-0.001	0.000	0.003
	(0.005)	(0.005)	(0.003)	(0.004)	(0.004)	(0.003)	(0.004)	(0.004)
Ex-smoker	0.001	-0.001	0.002	0.006	0.003	-0.001	-0.004	-0.006
	(0.005)	(0.008)	(0.004)	(0.005)	(0.004)	(0.005)	(0.005)	(0.004)
Non-smoker	-0.004	-0.019*	0.005	0.001	0.006	-0.001	-0.005	-0.002
	(0.006)	(0.008)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Hazardous	-0.000	-0.008	-0.004	0.002	0.002	-0.015	0.003	0.007
	(0.009)	(0.013)	(0.007)	(0.008)	(0.010)	(0.009)	(0.006)	(0.008)
Non-drinker	-0.003	-0.002	-0.008	0.008	-0.004	-0.014	0.005	0.003
	(0.009)	(0.015)	(0.007)	(0.007)	(0.010)	(0.010)	(0.007)	(0.007)
Safe	-0.009	-0.005	-0.002	0.010	-0.003	-0.012	0.005	0.008
	(0.008)	(0.014)	(0.007)	(0.007)	(0.009)	(0.009)	(0.006)	(0.007)
Angina	-0.015	-0.003	0.009	0.005	-0.005	-0.015	-0.015	-0.017
	(0.009)	(0.013)	(0.007)	(0.009)	(0.008)	(0.009)	(0.010)	(0.014)
AMI	-0.011	-0.006	0.009	-0.001	-0.009	0.005	-0.012	0.012
	(0.009)	(0.009)	(0.007)	(0.007)	(0.006)	(0.007)	(0.007)	(0.008)
ATF	-0.006	-0.005	0.008	-0.002	-0.000	-0.007	-0.004	-0.000
	(0.008)	(0.011)	(0.008)	(0.006)	(0.005)	(0.006)	(0.007)	(0.008)
CCF	-0.004	0.034*	0.007	-0.006	0.007	0.003	0.004	0.003
	(0.011)	(0.015)	(0.008)	(0.009)	(0.010)	(0.009)	(0.011)	(0.009)
CAD	-0.002	0.025*	-0.006	-0.005	0.002	-0.012	0.021*	-0.001
	(0.010)	(0.012)	(0.007)	(0.006)	(0.008)	(0.008)	(0.009)	(0.008)
IS & TIA	-0.008	0.002	-0.004	-0.005	-0.006	0.000	-0.010	-0.000
	(0.006)	(0.009)	(0.005)	(0.006)	(0.004)	(0.006)	(0.005)	(0.006)
PAD	0.013	-0.029*	-0.000	0.005	-0.009	0.012	-0.001	-0.004
	(0.013)	(0.015)	(0.007)	(0.009)	(0.007)	(0.009)	(0.009)	(0.010)
Hypertension	0.001	0.006	0.002	0.001	0.000	-0.007*	0.004	0.000
	(0.004)	(0.006)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)
N	26230	29388	29235	29317	29884	29263	30549	22157
R^2	0.645	0.319	0.817	0.728	0.778	0.716	0.702	0.772
F-stat	1.620	2.011	1.287	0.905	1.247	1.381	1.072	0.820
p-value (F test)	0.035	0.004	0.169	0.599	0.200	0.113	0.376	0.714

NOTES. The table presents OLS estimates from the regression: $C_{gt} = \Omega X_{igt} + \varepsilon_{gt}$, estimated for each year between 2003 and 2018. Each regression controls for GP practice fixed effects. F-stat and p-value correspond to the null hypothesis that coefficient vector $\Omega = 0$, or that patient's baseline risk is balanced across compliance types. Standard errors clustered at the GP practice level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 4.10: Balancing test: Compliance on patient characteristics and patient index

	(1) 2003	(2) 2004	(3) 2005	(4) 2006	(5) 2007	(6) 2008	(7) 2009	(8) 2010
Patient Index (Z_{igt})	0.015	0.302	-0.425	0.023	0.024	0.580	-0.121	-0.477
	(0.426)	(0.400)	(0.303)	(0.377)	(1.136)	(0.386)	(0.911)	(0.554)
N	18536	38337	34981	38381	31679	31190	30675	28239
R^2	0.874	0.854	0.766	0.740	0.637	0.656	0.626	0.726
F-stat	1.170	1.073	1.130	1.125	1.053	1.032	0.633	1.422
p-value (F test)	0.265	0.374	0.307	0.314	0.398	0.426	0.918	0.090

	2011	2012	2013	2014	2015	2016	2017	2018
Patient Index (Z_{igt})	-0.397	-0.117	0.147	0.070	-0.195	-0.208	-0.617	-1.218
	(0.773)	(0.408)	(0.446)	(0.176)	(0.382)	(0.553)	(1.000)	(1.090)
N	26230	29388	29235	29317	29884	29263	30549	22157
R^2	0.645	0.319	0.817	0.729	0.778	0.716	0.703	0.772
F-stat	1.560	1.935	1.237	0.866	1.193	1.381	1.049	0.826
p-value (F test)	0.045	0.006	0.204	0.657	0.243	0.109	0.404	0.712

NOTES. The table presents OLS estimates from the regression: $C_{gt} = \omega Z_{igt} + \Omega X_{igt} + \varepsilon_{gt}$, estimated for each year between 2003 and 2018. Only ω is presented. Each regression controls for GP practice fixed effects. F-stat and p-value correspond to the null hypothesis that coefficient ω and coefficient vector Ω are jointly equal to 0, or that patient's baseline risk is balanced across compliance types. Standard errors clustered at the GP practice level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

TABLE 4.11: Balancing test: Responsiveness on patient characteristics and patient index

	(1) 2003	(2) 2004	(3) 2005	(4) 2006	(5) 2007	(6) 2008	(7) 2009	(8) 2010
Patient Index (Z_{igt})	0.002	-0.157	0.177	0.242	-0.236	0.133	-0.373	0.712
	(0.766)	(0.730)	(0.424)	(0.472)	(1.183)	(0.470)	(0.978)	(0.590)
N	18536	38337	34981	38381	31679	31190	30675	28239
R^2	0.614	0.517	0.479	0.569	0.515	0.542	0.598	0.682
F-stat	1.755	0.502	0.760	0.910	0.899	0.947	1.184	0.921
p-value (F test)	0.016	0.981	0.795	0.595	0.610	0.543	0.252	0.579

	2011	2012	2013	2014	2015	2016	2017	2018
Patient Index (Z_{igt})	-0.098	0.001	-0.318	0.149	0.054	0.387	0.085	0.287
	(0.722)	(0.389)	(0.617)	(0.214)	(0.522)	(0.673)	(1.278)	(1.298)
N	26230	29388	29235	29317	29884	29263	30549	22157
R^2	0.634	0.582	0.597	0.482	0.520	0.563	0.485	0.677
F-stat	0.837	0.797	0.994	1.374	0.673	1.328	0.940	1.197
p-value (F test)	0.696	0.749	0.477	0.113	0.886	0.139	0.552	0.240

NOTES. The table presents OLS estimates from the regression: $R_{gt} = \omega Z_{igt} + \Omega X_{igt} + \varepsilon_{gt}$, estimated for each year between 2003 and 2018. Only ω is presented. Each regression controls for GP practice fixed effects. F-stat and p-value correspond to the null hypothesis that coefficient ω and coefficient vector Ω are jointly equal to 0, or that patient's baseline risk is balanced across responsiveness types. Standard errors clustered at the GP practice level in parentheses. * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$.

Chapter 5

Conclusions

Using prescription data from general practices, this dissertation explores health-economic questions on the topics of variation in medical care, and populations' health following large-scale shocks. Specifically, first I investigate the heterogeneity in drug treatment choices made by general practitioners in the market for statins, drugs used to prevent cardiovascular conditions, and its relation with costs to the national health service and patients' health outcomes. Second, I study changes in antidepressants prescription in general practice to estimate the impact of the COVID-19 pandemic on mental health in England.

The findings of these investigations highlight the role of promoting evidence-based treatments to reduce unjustified heterogeneity, that is, variation in the use of treatment that does not improve outcomes but may unnecessarily increase costs, and to the trade-off between strictly complying with guidelines, and doing what seems appropriate for the patient (if and only if what is appropriate can be identified). In the case of statins, variation in treatment choice is partially explained by patient demand-side factors, yet our findings rule out that responding to patients characteristics matter for outcomes. However, these results may be specific to this therapeutic market, characterized by a small set of alternative treatments, close substitutability of available treatments, and highly studied and researched illnesses. For a more comprehensive study of variation in primary care, other conditions and drugs classes should be studied and analysed in terms of their own idiosyncrasies (e.g. prevalence, incidence, the number, costs and substitutability of treatments, potential outcomes, market conditions, presence of clinical guidance, etc.)

The interest in mental health, both as a social-health concern and as a topic of academic inquiry, is growing. While most of the existing studies rely on a self-reported assessment of health, in this dissertation I explore mental health outcomes that reflect not only an individual's subjective evaluation of symptoms but one where patients are assessed

by a professional practitioner, with whom they decide to start (or not) pharmacological treatment, depending on the assessed severity of the condition. However, there are several other measures to capture depression, anxiety, and other forms of deteriorated mental health, such as prescription drugs used to treat substance dependence, anxiety, and psychosis; referrals to psychological talking therapies or related interventions; psychiatric interventions; suicide statistics; and even self-reported health from surveys that may capture true symptoms of poorer mental health that simply do not require clinical attention. A more comprehensive study of the effects of COVID-19 on mental health would have to follow a more inclusive approach acknowledging all such possible outcomes and how they are internalized and registered by the health care system.

References

Alberto Abadie and Javier Gardeazabal. The economic costs of conflict: A case study of the basque country. *American economic review*, 93(1):113–132, 2003.

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Synthetic control methods for comparative case studies: Estimating the effect of californias tobacco control program. *Journal of the American statistical Association*, 105(490):493–505, 2010.

Alberto Abadie, Alexis Diamond, and Jens Hainmueller. Comparative politics and the synthetic control method. *American Journal of Political Science*, 59(2):495–510, 2015.

Abi Adams-Prassl, Teodora Boneva, Marta Golin, and Christopher Rauh. The impact of the coronavirus lockdown on mental health: evidence from the us. 2020.

Pierre Amarenco and Julien Labreuche. Lipid management in the prevention of stroke: review and updated meta-analysis of statins for stroke prevention. *The Lancet Neurology*, 8(5):453–463, 2009.

Stephanie Armbruster and Valentin Klotzbücher. Lost in lockdown? covid-19, social distancing, and mental health in germany. Technical report, Diskussionsbeiträge, 2020.

Jane Armitage. The safety of statins in clinical practice. *The Lancet*, 370(9601):1781–1790, 2007.

Richard Armitage. Antidepressants, primary care, and adult mental health services in england during covid-19. *The Lancet Psychiatry*, 8(2):e3, 2021.

Anthony A Avery, Nick Barber, Maisoon Ghaleb, Bryony Dean Franklin, Sarah Armstrong, Sarah Crowe, Soraya Dhillon, Anette Freyer, Rachel Howard, Cinzia Pezzolesi, et al. Investigating the prevalence and causes of prescribing errors in general practice: the practice study. 2012.

James Banks and Xiaowei Xu. The mental health effects of the first two months of lockdown and social distancing during the covid-19 pandemic in the uk. Technical report, IFS Working Papers, 2020.

Eliana Barrenho, Marisa Miraldo, Carol Propper, Christiern Rose, et al. Peer and network effects in medical innovation: the case of laparoscopic surgery in the english nhs. Technical report, HEDG, c/o Department of Economics, University of York, 2019.

J Bedson, J Belcher, OI Martino, M Ndlovu, T Rathod, K Walters, KM Dunn, and KP Jordan. The effectiveness of national guidance in changing analgesic prescribing in primary care from 2002 to 2009: an observational database study. *European journal of pain*, 17(3):434–443, 2013.

Seth A Berkowitz and Sanjay Basu. Unemployment insurance, health-related social needs, health care access, and mental health during the covid-19 pandemic. *JAMA internal medicine*, 181(5):699–702, 2021.

Prachi Bhatnagar, Kremlin Wickramasinghe, Elizabeth Wilkins, and Nick Townsend. Trends in the epidemiology of cardiovascular disease in the uk. *Heart*, 102(24):1945–1952, 2016.

John Bound et al. Self-reported versus objective measures of health in retirement models. *Journal of Human Resources*, 26(1):106–138, 1991.

Abel Brodeur, Andrew E Clark, Sarah Fleche, and Nattavudh Powdthavee. Covid-19, lockdowns and well-being: Evidence from google trends. *Journal of public economics*, 193:104346, 2021.

Samantha K Brooks, Rebecca K Webster, Louise E Smith, Lisa Woodland, Simon Wessely, Neil Greenberg, and Gideon James Rubin. The psychological impact of quarantine and how to reduce it: rapid review of the evidence. *The lancet*, 395(10227): 912–920, 2020.

Joseph S Butler, Richard V Burkhauser, Jean M Mitchell, and Theodore P Pincus. Measurement error in self-reported health variables. *The Review of Economics and Statistics*, pages 644–650, 1987.

A Colin Cameron and Pravin K Trivedi. *Microeconometrics: methods and applications*. Cambridge university press, 2005.

Colleen Carey, Ethan MJ Lieber, and Sarah Miller. Drug firms' payments and physicians' prescribing behavior in medicare part d. Technical report, National Bureau of Economic Research, 2020.

Matthew J Carr, Sarah Steeg, Roger T Webb, Nav Kapur, Carolyn A Chew-Graham, Kathryn M Abel, Holly Hope, Matthias Pierce, and Darren M Ashcroft. Effects of the covid-19 pandemic on primary care-recorded mental illness and self-harm episodes in the uk: a population-based cohort study. *The Lancet Public Health*, 6(2):e124–e135, 2021.

Amitabh Chandra and Douglas O Staiger. Productivity spillovers in health care: evidence from the treatment of heart attacks. *Journal of Political Economy*, 115(1):103–140, 2007.

Amitabh Chandra, David Cutler, and Zirui Song. Who ordered that? the economics of treatment choices in medical care. *Handbook of health economics*, 2:397–432, 2012.

Zhiming Cheng, Silvia Mendolia, Alfredo R Paloyo, David A Savage, and Massimiliano Tani. Working parents, financial insecurity, and childcare: mental health in the time of covid-19 in the uk. *Review of Economics of the Household*, 19(1):123–144, 2021.

Paris Cleanthous. Patient welfare implications of innovation in the us antidepressant market. *Unpublished working paper, New York University*, 2002.

Ana Correa, William Hinton, Andrew McGovern, Jeremy van Vlymen, Ivelina Yonova, Simon Jones, and Simon de Lusignan. Royal college of general practitioners research and surveillance centre (rcgp rsc) sentinel network: a cohort profile. *BMJ open*, 6(4): e011092, 2016.

Andrea Coscelli. The importance of doctors and patients preferences in the prescription decision. *The Journal of Industrial Economics*, 48(3):349–369, 2000.

Janet Currie, Valentina Duque, and Irwin Garfinkel. The great recession and mothers health. *The Economic Journal*, 125(588):F311–F346, 2015.

Janet Currie, W Bentley MacLeod, and Jessica Van Parys. Provider practice style and patient health outcomes: the case of heart attacks. *Journal of health economics*, 47:64–80, 2016.

Janet M Currie and W Bentley MacLeod. Understanding doctor decision making: The case of depression treatment. *Econometrica*, 88(3):847–878, 2020.

Timothy M Dall, Paul D Gallo, Ritasree Chakrabarti, Terry West, April P Semilla, and Michael V Storm. An aging population and growing disease burden will require alarge and specialized health care workforce by 2025. *Health affairs*, 32(11):2013–2020, 2013.

Michael Daly, Angelina R Sutin, and Eric Robinson. Longitudinal changes in mental health and the covid-19 pandemic: Evidence from the uk household longitudinal study. *Psychological medicine*, pages 1–10, 2020.

Apostolos Davillas and Andrew M Jones. The covid-19 pandemic and its impact on inequality of opportunity in psychological distress in the uk. *Available at SSRN 3614940*, 2020.

DHSC and ONS. Direct and indirect health impacts of covid-19 in england. https://www.gov.uk/government/publications/

dhsc-direct-and-indirect-health-impacts-of-covid-19-in-england-long-paper-9-september-202
September 2021.

Ben Etheridge and Lisa Spantig. The gender gap in mental well-being during the covid-19 outbreak: evidence from the uk. Technical report, ISER Working paper series, 2020.

Tom Fawcett. An introduction to roc analysis. *Pattern recognition letters*, 27(8):861–874, 2006.

John Fitzgerald, Peter Gottschalk, and Robert Moffitt. An analysis of sample attrition in panel data: The michigan panel study of income dynamics. *The Journal of Human Resources*, 33(2):251–299, 1998. ISSN 0022166X. URL http://www.jstor.org/stable/146433.

Osea Giuntella, Kelly Hyde, Silvia Saccardo, and Sally Sadoff. Lifestyle and mental health disruptions during covid-19. *Proceedings of the National Academy of Sciences*, 118(9), 2021.

J Alison Glover. The incidence of tonsillectomy in school children, 1938.

Ben Goldacre, Carl Reynolds, Anna Powell-Smith, Alex J Walker, Tom A Yates, Richard Croker, and Liam Smeeth. Do doctors in dispensing practices with a financial conflict of interest prescribe more expensive drugs? a cross-sectional analysis of english primary care prescribing data. *BMJ open*, 9(2):e026886, 2019.

Google. Google covid-19 community mobility reports. https://www.google.com/covid19/mobility/, 2021. Accessed: 2021-02-23.

Reuben Gronau et al. A useful interpretation of r2; in binary choice models (or, have we dismissed the good old r2; prematurely). Technical report, 1998.

Mirella P Hage and Sami T Azar. The link between thyroid function and depression. *Journal of thyroid research*, 2012, 2012.

Stephen Harrison, George Dowswell, John Wright, and Ian Russell. General practitioners' uptake of clinical practice guidelines: a qualitative study. *Journal of health services research & policy*, 8(3):149–153, 2003.

William Hinton, Andrew McGovern, Rachel Coyle, Thang S Han, Pankaj Sharma, Ana Correa, Filipa Ferreira, and Simon de Lusignan. Incidence and prevalence of cardiovascular disease in english primary care: a cross-sectional and follow-up study of the royal college of general practitioners (rcgp) research and surveillance centre (rsc). *BMJ open*, 8(8):e020282, 2018.

Julia Hippisley-Cox, Carol Coupland, Yana Vinogradova, John Robson, Rubin Minhas, Aziz Sheikh, and Peter Brindle. Predicting cardiovascular risk in england and wales: prospective derivation and validation of qrisk2. *Bmj*, 336(7659):1475–1482, 2008.

Carly Hire and Bruno Rushforth. General practitioners views on using a prescribing substitution application (scriptswitch®). *Journal of Innovation in Health Informatics*, 21(1):1–11, 2013.

Katharina Janke, Kevin Lee, Carol Propper, Kalvinder Shields, and Michael Shields. Macroeconomic conditions and health in britain: aggregation, dynamics and local area heterogeneity. 2020.

Angela Ju, Camilla S Hanson, Emily Banks, Rosemary Korda, Jonathan C Craig, Tim Usherwood, Peter MacDonald, and Allison Tong. Patient beliefs and attitudes to taking statins: systematic review of qualitative studies. *British Journal of General Practice*, 68(671):e408–e419, 2018.

Andres M Kanner. Depression in epilepsy: a neurobiologic perspective. *Epilepsy currents*, 5(1):21–27, 2005.

John A Kaufman, Melvin D Livingston, and Kelli A Komro. Unemployment insurance program accessibility and suicide rates in the united states. *Preventive Medicine*, 141: 106318, 2020.

Elira Kuka. Quantifying the benefits of social insurance: unemployment insurance and health. *Review of Economics and Statistics*, 102(3):490–505, 2020.

Margaret Kyle and Heidi Williams. Is american health care uniquely inefficient? evidence from prescription drugs. *American Economic Review*, 107(5):486–90, 2017.

Malcolm R Law, Nicholas J Wald, and AR Rudnicka. Quantifying effect of statins on low density lipoprotein cholesterol, ischaemic heart disease, and stroke: systematic review and meta-analysis. *Bmj*, 326(7404):1423, 2003.

Emily Leslie and Riley Wilson. Sheltering in place and domestic violence: Evidence from calls for service during covid-19. *Journal of Public Economics*, 189:104241, 2020.

Haiyang Lu, Peng Nie, and Long Qian. Do quarantine experiences and attitudes towards covid-19 affect the distribution of mental health in china? a quantile regression analysis. *Applied Research in Quality of Life*, 16(5):1925–1942, 2021.

Ramon Luengo-Fernandez, Jose Leal, Alastair Gray, Sophie Petersen, and Mike Rayner. Cost of cardiovascular diseases in the united kingdom. *Heart*, 92(10):1384–1389, 2006.

Anne Mason. New medicines in primary care: a review of influences on general practitioner prescribing. *Journal of clinical pharmacy and therapeutics*, 33(1):1–10, 2008.

Anthony Matthews, Emily Herrett, Antonio Gasparrini, Tjeerd Van Staa, Ben Goldacre, Liam Smeeth, and Krishnan Bhaskaran. Impact of statin related media coverage on use of statins: interrupted time series analysis with uk primary care data. *bmj*, 353, 2016.

James C Moon and Richard G Bogle. Switching statins. *BMJ*, 332(7554): 1344–1345, 2006. ISSN 0959-8138. . URL https://www.bmj.com/content/332/7554/1344. Publisher: BMJ Publishing Group Ltd _eprint: https://www.bmj.com/content/332/7554/1344.full.pdf.

Peter Moss, Gavin Barlow, Nicholas Easom, Patrick Lillie, and Anda Samson. Lessons for managing high-consequence infections from first covid-19 cases in the uk. *The Lancet*, 395(10227):e46, 2020.

NICE. Technology appraisal 94. Statins for the prevention of cardiovascular events, 2006. URL https://www.nice.org.uk/guidance/ta94. Publisher: NICE.

NICE. Clinical guideline 67. Lipid modification: Cardiovascular risk assessment and the modification of blood lipids for the primary and secondary prevention of cardiovascular disease, 2008. URL https://www.nice.org.uk/guidance/cg67. Publisher: NICE.

NICE. Clinical guideline 181. Cardiovascular disease: risk assessment and reduction, including lipid modification, 2014. URL https://www.nice.org.uk/guidance/cg181/chapter/appendix-a-grouping-of-statins. Publisher: NICE.

NICE. The NICE strategy 2021 to 2026, 2021. URL https://www.nice.org.uk/about/who-we-are/corporate-publications/the-nice-strategy-2021-to-2026.

OECD. Tackling the mental health impact of the covid-19 crisis: An integrated, whole-of-society response. 2021.

Matias Ortiz de Zarate, Emmanouil Mentzakis, Simon DS Fraser, Paul Roderick, Paul Rutter, and Carmine Ornaghi. Price versus clinical guidelines in primary care statin prescribing: a retrospective cohort study and cost simulation model. *Journal of the Royal Society of Medicine*, 2021.

Sara Parker-Lue. The impact of reducing pharmaceutical industry payments on physician prescribing. *Health economics*, 29(3):382–390, 2020.

Zhuan Pei, Jörn-Steffen Pischke, and Hannes Schwandt. Poorly measured confounders are more useful on the left than on the right. *Journal of Business & Economic Statistics*, 37(2):205–216, 2019.

Charles E Phelps. Information diffusion and best practice adoption. In *Handbook of health economics*, volume 1, pages 223–264. Elsevier, 2000.

Berkeley Phillips, Craig Roberts, Amy E Rudolph, Steve Morant, Fayaz Aziz, and Christopher P ORegan. Switching statins: the impact on patient outcomes. *Br J Cardiol*, 14(5):280–5, 2007.

Eugenio Proto and Climent Quintana-Domeque. Covid-19 and mental health deterioration by ethnicity and gender in the uk. *PloS one*, 16(1):e0244419, 2021.

Arndt R Reichert and Harald Tauchmann. The causal impact of fear of unemployment on psychological health. *Ruhr Economic Paper*, (266), 2011.

Meredith B Rosenthal, Ernst R Berndt, Julie M Donohue, Richard G Frank, and Arnold M Epstein. Promotion of prescription drugs to consumers. *New England Journal of Medicine*, 346(7):498–505, 2002.

Rita Santos, Hugh Gravelle, and Carol Propper. Does quality affect patients choice of doctor? evidence from england. *The Economic Journal*, 127(600):445–494, 2017.

Anthony Scott. Economics of general practice. *Handbook of health economics*, 1:1175–1200, 2000.

Bettina Siflinger, Michaela Paffenholz, Sebastian Seitz, Moritz Mendel, and Hans-Martin von Gaudecker. The covid-19 pandemic and mental health: Disentangling crucial channels. 2021.

Adan Silverio-Murillo, Lauren Hoehn-Velasco, Jose Roberto Balmori de la Miyar, and Abel Rodríguez. Covid-19 and womens health: Examining changes in mental health and fertility. *Economics letters*, 199:109729, 2021.

Barbara Starfield, Leiyu Shi, and James Macinko. Contribution of primary care to health systems and health. *The milbank quarterly*, 83(3):457–502, 2005.

Maxime Taquet, Emily A Holmes, and Paul J Harrison. Depression and anxiety disorders during the covid-19 pandemic: knowns and unknowns. *The Lancet*, 398(10312): 1665–1666, 2021.

Annie Tubadji, Frederic Boy, and D Webber. Narrative economics, public policy and mental health. *Center for Economic Policy Research*, 20:109–131, 2020.

UK Health Security Agency. Health Matters: Preventing cardiovascular disease, 2019. URL https://ukhsa.blog.gov.uk/2019/02/14/health-matters-preventing-cardiovascular-disease/.

Epidemiological Studies Unit. Efficacy and safety of cholesterol-lowering treatment: prospective meta-analysis of data from 90 056 participants in 14 randomised trials of statins. *Lancet*, 366(9493):1267–1278, 2005.

Alex J Walker, Felix Pretis, Anna Powell-Smith, and Ben Goldacre. Variation in responsiveness to warranted behaviour change among nhs clinicians: novel implementation of change detection methods in longitudinal prescribing data. *bmj*, 367, 2019.

Bernard Wathen and Tara Dean. An evaluation of the impact of nice guidance on gp prescribing. *British Journal of General Practice*, 54(499):103–107, 2004.

Richard Williams, David A Jenkins, Darren M Ashcroft, Ben Brown, Stephen Campbell, Matthew J Carr, Sudeh Cheraghi-Sohi, Navneet Kapur, Owain Thomas, Roger T Webb, et al. Diagnosis of physical and mental health conditions in primary care during the covid-19 pandemic: a retrospective cohort study. *The Lancet Public Health*, 5 (10):e543–e550, 2020.

Peter WF Wilson, Ralph B DAgostino, Daniel Levy, Albert M Belanger, Halit Silbershatz, and William B Kannel. Prediction of coronary heart disease using risk factor categories. *Circulation*, 97(18):1837–1847, 1998.

World Health Organization. Global spending on health: a world in transition. Technical Report WHO/HIS/HGF/HFWorkingPaper/19.4, World Health Organization, 2019. URL https://apps.who.int/iris/handle/10665/330357. number-of-pages: 49.