UNIVERSITY of York

RESEARCH

CHE

Centre For Health Economics

Financial Incentives and Prescribing Behaviour in Primary Care

Olivia Bodnar, Hugh Gravelle,
Nils Gutacker, Annika Herr

CHE Research Paper 181

Financial incentives and prescribing behaviour in primary care

[a]Olivia Bodnar
[b]Hugh Gravelle
[b]Nils Gutacker
[a,c]Annika Herr

[a]DICE, Heinrich-Heine-University Düsseldorf, Germany
[b]Centre for Health Economics, University of York, UK
[c]Institute of Health Economics, Leibniz University Hannover, Germany

April 2021

Background to series

CHE Discussion Papers (DPs) began publication in 1983 as a means of making current research material more widely available to health economists and other potential users. So as to speed up the dissemination process, papers were originally published by CHE and distributed by post to a worldwide readership.

The CHE Research Paper series takes over that function and provides access to current research output via web-based publication, although hard copy will continue to be available (but subject to charge).

Acknowledgements

We are grateful to participants at the German Health Econometrics Workshop (2020), the German Health Economics Association Meeting (Augsburg 2019), the EuHEA PhD Student-Supervisor conference (Porto 2019), and the CINCH Academy (Essen, 2019) as well as seminars at Düsseldorf, Hannover, York, Monash, City, Manchester, Essen and Turin for valuable comments and suggestions. OB received funding from the DFG Research Training Group Competition Economics (project 235577387/GRK1974). Declaration of interest: none. No ethical approval was needed.

Further copies

Only the latest electronic copy of our reports should be cited. Copies of this paper are freely available to download from the CHE website www.york.ac.uk/che/publications/. Access to downloaded material is provided on the understanding that it is intended for personal use. Copies of downloaded papers may be distributed to third parties subject to the proviso that the CHE publication source is properly acknowledged and that such distribution is not subject to any payment.

Printed copies are available on request at a charge of £5.00 per copy. Please contact the CHE Publications Office, email che-pub@york.ac.uk for further details.

Centre for Health Economics
Alcuin College
University of York
York,
YO10 5DD, UK
www.york.ac.uk/che

Abstract

Many healthcare systems prohibit primary care physicians from dispensing the drugs they prescribe due to concerns that this encourages excessive, ineffective or unnecessarily costly prescribing. Using data from the English National Health Service for 2011 to 2018, we estimate the impact of physician dispensing rights on prescribing behaviour at the extensive margin (comparing practices that dispense and those that do not) and the intensive margin (comparing practices with different proportions of patients to whom they dispense). Our empirical strategy controls for practices selecting into dispensing based on observable (OLS, entropy balancing) and unobservable practice characteristics (2SLS). We show that physician dispensing raises drug costs per patient by 4.2%, which reflects more and more expensive drugs being prescribed, including potentially inappropriate substances such as opioids. Dispensing practices also prescribe smaller packages as reimbursement is partly based on a fixed fee per prescription dispensed. Similar effects are observed at the intensive margin.

JEL: I11, I18, L10

Keywords: Physician dispensing, primary care, drug expenditure, financial incentives, physician agency

1. Introduction

Pharmaceutical drugs are a cornerstone of modern medicine and a major contributor to population health (Buxbaum et al. 2020; Lichtenberg 2012). Most drugs are prescribed in primary care, where they are used to treat a range of common health problems such as hypertension, diabetes or bacterial infections. However, patients may not benefit from pharmaceutical prescribing to the full extent if they struggle to have their medicines dispensed, for example because they live far away from the closest community pharmacy. In order to facilitate patients' access to drugs, some countries permit physicians to dispense the medication they prescribe through their own co-located pharmacies (*physician dispensing* (PD)).[1] This form of vertical integration benefits patients, who incur lower travel costs. However, it may also harm the interests of patients and funders if physicians exploit their information advantage to extract dispensing rents through excessive, ineffective or unnecessarily costly prescribing.

In this paper, we analyse the effect of dispensing rights on the prescribing behaviour of primary care physicians ('general practitioners' (GPs)) in the English National Health Service (NHS) using detailed quarterly data from 2011 to 2018 on the cost, volume, and pack size (drug amount per prescription) of all types of drugs prescribed in every general practice in England. We exploit the unique regulatory environment in the UK that permits general practices to dispense medicines only for patients who live more than 1 mile (1.6km) away from the nearest community pharmacy and who have asked their practice to dispense to them. These regulations create variation in dispensing status and intensity across practices, which we use to estimate the effects of PD at the extensive margin (comparing practices which do and do not dispense) and at the intensive margin (comparing practices with different proportions of patients to whom they dispense), thus providing estimates of the effects of banning dispensing entirely and of differences in demand for dispensing. We control for the potential endogeneity of PD status in two ways. First, we use detailed information on the characteristics of practice patients (i.e. age/sex composition, prevalence rates for 12 chronic conditions, and deprivation profile) and practice organisation (e.g. list size, number of GPs, proportion of GPs who are profit sharing partners, proportion trained in the UK, GP age) to create entropy balanced (EB) samples of dispensing and non-dispensing practices ('selection on observables'). Second, we capitalise on the eligibility rule to construct an instrument for practice dispensing based on the number of potentially eligible patients in an area around a practice and the travel distance they would save if the practice dispensed ('selection on unobservables').

Our models (OLS, EB, and 2SLS) yield the same pattern of results. Test statistics from the 2SLS model fail to reject the null hypothesis of no endogenous selection conditional on observed practice characteristics. Based on our preferred approach, EB, we find that the prescribed drug cost per patient is 4.2% higher in dispensing than in non-dispensing practices, which results from both prescribing more often (2.4% more prescriptions per patient), and more expensive drugs (1.6% higher cost per prescription; 0.7% more branded drugs). English dispensing practices receive a fee for each item they dispense and we find that they respond to this incentive by reducing the volume of drugs in each prescription by 16.1% compared with non-dispensing practices. There are also differences in the pattern of prescribing: dispensing practices prescribe, per patient, 3.7% more opioids, 2.9% more anti-depressants and 3.7% more over-the-counter (OTC) medicines.

[1] For example, PD is permitted in Canada, parts of Switzerland, most US states, and many Asian countries but is banned in Germany and much of Scandinavia. In some countries, such as Austria, France and the United Kingdom, physicians are only allowed to dispense in areas where patients have greater difficulty in accessing community pharmacies. See Eggleston (2012) for a review of the history of PD and an economic analysis of factors that give rise to integration or segregation of prescribing and dispensing in different healthcare systems.

At the intensive margin, we show that practices which have a larger proportion of patients to whom they dispense have higher drug costs, more prescriptions per patient, smaller pack sizes, and a smaller proportion of generic prescriptions.

Our study contributes to the literature on PD and the more general literature on physician agency and supplier-induced demand (e.g. McGuire 2000; Clemens and Gottlieb 2014). Previous studies find both positive (Chou et al. 2003; Kaiser and Schmid 2016; Burkhard et al. 2019) and negative effects (Trottmann et al. 2016; Ahammer and Schober 2020) of PD on drug expenditure.[2] Possible channels through which PD may increase expenditure include volume effects (e.g. Burkhard et al. 2019), including increased prescribing of antibiotics with potentially negative externalities (Trap 2002; Park et al. 2005; Filippini et al. 2014), or choice of drugs with higher profit margins (e.g Liu et al. 2009; Iizuka 2012). The existing economic literature has mostly focused on Switzerland (where dispensing is regulated at canton level) and Asia. For England, Morton-Jones and Pringle (1993) analysed data from 108 practices in Lincolnshire and found prescribing expenditures to be 13% higher in dispensing practices. More recently, Goldacre et al. (2019) compared the probability of prescribing four categories of high cost drugs versus therapeutically equivalent low cost drugs. They found that dispensing practices were up to five times more likely to choose high cost drugs, and that this effect intensifies as the share of eligible dispensing patients in the practice increases.

Our study expands on these analyses by conducting a comprehensive assessment of PD effects on *all* pharmaceutical prescribing in English GP practices. Furthermore, we investigate how the incentives created through a per-item dispensing fee, a possible channel that has received very limited attention in the literature (Rischatsch 2014), affects GPs' choice of pack size and prescribing frequency. Whereas the existing international literature has either relied on before-and-after studies (e.g. due to bans on PD) or regulatory differences across large, heterogeneous jurisdictions (e.g. cantons in the Swiss context), we exploit variation in dispensing status and intensity across practices in the same geographic areas to examine behaviour at the intensive and extensive margins.

The remainder of the paper is organised as follows: Section 2 provides the institutional background and the dispensing regulation in the NHS. Sections 3 and 4 describe the data and empirical methods. Section 5 presents the results and Section 6 concludes.

[2] See Lim et al. (2009) and Eggleston (2012) for systematic reviews of earlier literature.

2. Institutional setting

The English NHS has a list system in primary care: patients register with a single general practice that acts as the gatekeeper to most other NHS services, including non-emergency hospital care. Almost all general practices are small businesses owned and run by partnerships of GPs who share profits and losses. In 2018, there were 7,148 general practices in England with an average list size of 8,279 patients and 3.37 full time equivalent GPs.

The English NHS is funded almost entirely by taxation. There is a small patient charge (£8.80 in 2018/19) when a primary care prescription is dispensed, be it by a pharmacy or an on-site GP dispensary. Around half of the population are exempt from this charge on grounds of age (under 16, in full time education and under 18, or over 60), current or recent pregnancy, specified medical conditions, and low income (House of Commons Library 2020). As a result, approximately 90% of prescriptions are dispensed without charge.[3]

General practices have contracts with the NHS under which they are paid by a mixture of capitation payments, quality incentives, and items of service such as vaccinations. They are reimbursed for the costs of their premises and information technology but meet all other costs, such as hiring practice staff, including salaried non-partner GPs, from their revenue. There are two main type of contracts. The General Medical Services (GMS) contract is negotiated centrally by the Department of Health & Social Care and the British Medical Association, the doctors' trade union. The most common non-GMS contract is the Primary Medical Services (PMS) contract, which is negotiated between individual practices and local healthcare purchasers, known as clinical commissioning groups (CCGs).[4] Under PMS contracts, the practice receives a lump sum for providing a set of services similar to those required by the GMS contract plus additional services for particular groups of patients.

2.1. General practice dispensing

Most patients who receive a drug prescription from their GP must take it to a pharmacy in contract with the NHS to have it dispensed. Patients who would have serious difficulty in accessing a pharmacy or who live in an area which has been designated as rural in character and who are more than 1 mile (1.6km) away from a pharmacy, can ask their general practice to dispense drugs to them.[5] The practice decision on whether to dispense is all or nothing: if it agrees to dispense for one eligible patient it must dispense to any eligible patient who requests it.

In 2018, 2.9m (5.5%) of 53.8m general practice patients were in the 916 dispensing practices (16.3% of all practices) which had agreed to dispense to their eligible patients. Most dispensing practices are in rural areas, as shown in Figure 1, though there are some in urban areas where a small number of patients have claimed eligibility on grounds of serious difficulty in accessing community pharmacies.

Like pharmacies, practices receive two main types of payments for dispensing:

[3] This reduces concerns that affordability may affect GPs' choice of medication (e.g. Lundin 2000; Crea et al. 2019).
[4] Prior to 2012/13, healthcare was purchased by Primary Care Trusts. We will use the term CCG for both types of purchasers.
[5] There are regulations restricting the entry of new pharmacies into rural areas (Department of Health 2012) and attempts to enter are strongly resisted by local dispensing general practices.

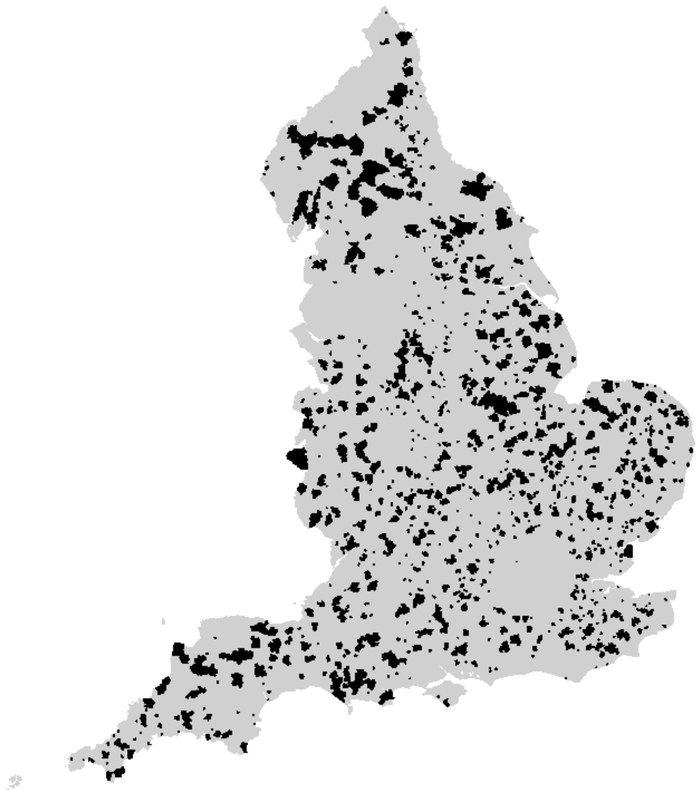


Figure 1: Areas in England with at least one dispensing GP practice

(a) a fee per prescription they dispense that is independent of the type, cost, or quantity of the drug.[6] Thus, a practice is paid more for dispensing two separate prescriptions on separate occasions each for a month's supply of a drug than for a single prescription for two months' supply of the same drug. The dispensing fee declines with the total number of prescriptions dispensed in the financial year. The maximum dispensing fee per prescription is currently £1.99 and the minimum is £1.76.

(b) reimbursement for the dispensed drugs bought by the practice. The reimbursement is based on the manufacturers' list price, known as the Net Ingredient Cost (NIC), minus an adjustment for discounts from suppliers. Dispensing practices can usually buy drugs from wholesalers at a discount, which depends on a number of factors such as volume purchased or temporary promotions. In recognition of these discounts, the NHS reduces the reimbursement by a fraction of the NIC, known as the "clawback". The clawback increases with the total NIC of all drugs dispensed and ranges from around 3% to 11%, with most dispensing practices facing the full clawback of 11%. The maximum clawback is less than the discount that practices receive on the NIC.

In addition, dispensing practices can earn up to £2.58 per dispensing patient per year if they meet various process requirements of the Dispensary Services Quality Scheme. The scheme aims to ensure minimum competency standards for dispensing staff by mandating training requirements and standard operating procedures for dispensaries. It also requires practices to review prescriptions for 10% of their dispensing patients each year.

[6] We define a prescription as a specified amount (pack size) of one drug. The prescription *form* given to the patient by a GP could contain more than one prescription, e.g. for two different drugs.

Dispensing is often profitable for practices. In 2017/18, partner GPs in dispensing practices with GMS contracts had a mean pre-tax income of £121,300 compared with £104,800 for partner GPs in non-dispensing GMS practices (NHS Digital 2019).

2.2. Physician incentives

We assume that GPs in dispensing practices, like those in non-dispensing practices, are partially altruistic (McGuire 2000) and care about the effect of their decisions on their income and on the well-being of their patients. Practices must decide whether to dispense to all eligible patients requesting it or to none of them. GPs also decide what they prescribe and dispense to dispensing patients and what they prescribe to non-dispensing patients.[7] Utility for patient h in practice i in period t is b^d_{hit} if patient h is on the dispensing list and b^{nd}_{hit} if h is not on the dispensing list (either because the practice has chosen not to dispense, or, if the practice dispenses, h is not eligible or has not asked the practice to be put on the dispensing list). Patient utility depends not just on the prescriptions that they are given but also on the time, travel, and other costs of having their prescriptions dispensed in the practice or in a pharmacy. If their practice has eligible patients and decides to dispense, the gain in utility for an eligible patient who requests that the practice dispense to them is $g^d_{hit} = b^d_{hit} - b^{nd}_{hit} \geq 0$.

Given the prescribing decision, the utility gain G to a practice i that has eligible patients and decides to dispense is

$$G_{it}(L^d_{it}) = \pi_{it}(L^d_{it}) - F_{it} + \alpha \sum\nolimits_{h \in DL_{it}} g^d_{hit} \tag{1}$$

where $\pi_{it}(L^d_{it})$ is the total profit from dispensing to the L^d_{it} patients in the set DL_{it} of patients on the dispensing list, F_{it} is the fixed cost of running a dispensary, g_{hit} is the patient's utility increase from receiving the drug at the practice rather than at a community pharmacy and $\alpha > 0$ is a parameter reflecting GP altruism. A dispensing practice cannot control the size (L^d_{it}) of its dispensing list. Once it has decided to dispense, it must dispense to any eligible patient who requests it.

Total profit from dispensing is

$$\pi_{it} = \sum\nolimits_k \sum\nolimits_\ell p_{it\ell k} n_{it\ell k} + r(n_{it}) + mL^d_{it} - c(n_{it}) \tag{2}$$

where $p_{it\ell k}$ is the net reimbursement from the NHS for drug k dispensed in pack size ℓ (defined as reimbursement price minus the purchase price and clawback). $n_{it\ell k}$ is the number of prescriptions of drug type k and pack size ℓ. $r(n_{it})$ is dispensing fee income from dispensing $n_{it} = \sum_k \sum_\ell n_{it\ell k}$ prescriptions. m is payment per dispensing patient if the practice meets the quality standards of the Dispensing Services Quality Scheme. $c(n_{it})$ is the variable cost, such as additional staff, of dispensing n_{it} prescriptions.

A practice with eligible patients will choose to dispense if and only if $G_{it}(L^d_{it}) \geq 0$. The total utility gain to dispensing patients will increase with L^d_{it}. There are also likely to be economies of scale affecting total profit since practices can achieve bigger discounts on larger drug purchases and use their specialist dispensary staff more productively. Thus, it is plausible that practices with more eligible patients are more likely to choose to dispense.

[7] We assume that all patients who are given a prescription have it dispensed, either by their practice or by a pharmacy. Thus, we do not distinguish between prescribed and dispensed drugs. A proportion of prescriptions are never dispensed: the patient may subsequently decide that they have recovered and do not need the drug or that it is not worth paying the prescription charge if they are not exempt (Beardon et al. 1993). We assume that GPs can predict the probability that the prescription is not dispensed and allow for this in their decisions on prescribing and whether to open a dispensary.

Practices can increase revenue from dispensing in a number of ways. They can prescribe more expensive prescriptions with a higher NIC and hence a higher payment net of the clawback $(p_{it\ell k})$. For example, they can prescribe proprietary drugs rather than cheaper equivalent generics. Second, they can increase the number of prescriptions $(n_{it\ell k})$ of a given drug of a given pack size, for example by increased prescribing of antibiotics for upper respiratory tract infections. This will increase their reimbursement $p_{it\ell k}n_{it\ell k}$ and also their dispensing fee income $r(n_{it})$. Third, since dispensing fee income $r(n_{it})$ varies with the number of prescriptions, not the quantities of drugs prescribed, practices will increase their dispensing fee income if they prescribe several smaller packages rather than a single larger package with the same total quantity of the drug.

In some of these decisions GPs will be trading off patient utility against greater income. For example, they may choose drug treatment over non-pharmacological alternatives even if the latter are at least as effective. Other responses may increase patient utility as well as practice profit. For example, patients may regard receiving a prescription as a validation of their decision to consult the GP and so will be more satisfied (Zgierska et al. 2012; Ashworth et al. 2016).

If they respond to these incentives dispensing practices will have different prescribing patterns compared to non-dispensing practices: they will prescribe more expensive drugs, they will prescribe a greater total quantity of drugs, and, on average, each prescription will be for a smaller amount. In the next section, we explain how we measure practice prescribing patterns and describe the data.

3. Data

We link administrative data to construct a quarterly panel of 7,979 practices for 2011 to 2018. The panel covers all practices in England, but is unbalanced because of practice entries, exits and mergers. For each practice, we have data on organisational structure, characteristics of the patient population, and detailed information on prescribing. Appendix Table A1 lists the data sources and reporting frequencies.

Our initial sample contains 233,048 practice-quarter observations. We exclude 40 very small practices (less than 1,000 patients) since these are likely to be in the process of opening or closing and this may affect their prescribing behaviour.[8] Furthermore, we exclude practices in CCGs where all or none of the practices dispense to ensure within CCG variation in dispensing status. The final sample has 130,113 observations.

3.1. Prescribing measures

The prescribing data covers all medicines, dressings and appliances prescribed by English practices and dispensed to patients anywhere in the United Kingdom. We use the shorthand "drug" to cover all types of prescriptions.

Drugs are labelled with a 15 digit British National Formulary (BNF) code that identifies the name of the drug, whether it is generic or branded, its formulation (e.g. capsule, tablet, liquid), its strength, and the quantity (e.g. number of pills, volume of liquid). For each drug type (BNF code) the data reports, for each month for each practice, the total number of prescriptions dispensed, the total quantity, and the total NIC. The latter is based on the list price for the drug excluding VAT, and does not take account of discounts, dispensing costs, or prescription charge income. We aggregate the data at quarterly level at which we observe other practice characteristics.

The practice prescribing data do not differentiate between prescriptions for dispensing and non-dispensing patients, nor by whether the prescription was dispensed in a community pharmacy or practice on-site dispensary. Our prescribing variables are therefore measured at practice level and are a weighted average of prescribing for dispensing and non-dispensing patients. Since we do observe the number of dispensing and non-dispensing patients in each practice we use this to test our hypotheses about the effects of having dispensing patients (see the discussion of Methods in Section 4).

Our discussion in Section 2.2 suggests that practices with on-site dispensaries are likely to have greater drug costs per patient because they gain financially by prescribing more expensive drugs and by prescribing a greater amount of drugs. Because they also receive a fee per prescription dispensed they have an incentive to write more prescriptions with smaller drug quantities. We construct nine measures of practice prescribing to test these hypotheses. Details are in Appendix Table A2. Four measures are based on overall practice prescribing: total NIC per patient, relative pack size, total NIC per prescription, and the number of prescription per patient. Relative pack size compares the average pack size (quantity of drugs) per prescription of a given type with the modal pack size across all prescriptions of this type across all dispensing and non-dispensing practices in England. We expect relative pack size to be smaller in dispensing practices because the fee per prescription dispensed creates an incentive to prescribe a given quantity of a given quantity of a drug in several small packs rather than in a single large pack. To

[8] We also exclude two practices with unusually high prevalence rates for chronic diseases.

allow for this when examining total NIC per prescription and the number of prescriptions, we adjust the number of prescriptions by the relative pack size so that for a practice with small relative pack sizes the adjusted number of prescriptions is smaller than the number observed.

We also test whether dispensing status affects prescribing of particular types of drugs. Generic drugs are usually cheaper than equivalent patented proprietary versions and GPs are encouraged by the NHS to prescribe generic drugs whenever possible. However, prescribing cheaper drugs is likely to reduce dispensing practice income and so we compare the percentage of generic prescriptions in dispensing and non-dispensing practices. Over the counter (OTC) drugs can be bought by patients without the need for a prescription from a GP and usually treat minor or short-term conditions. They are widely available in supermarkets, petrol stations and community pharmacies. To reduce costs, the NHS discourages GPs from prescribing OTCs. Dispensing practices, however, will lose financially if they reduce the number of OTC prescriptions that are dispensed in their on-site pharmacies. We compare the number of OTC prescriptions per patient in dispensing and non-dispensing practices. We adjust the number of prescriptions by relative pack sizes for both generic and OTC prescriptions to allow for the possibility that dispensing practices may prescribe more but smaller packages.

Overuse of some drugs, such as opioids, antidepressants and antibiotics, can be harmful to patients and to wider population health (e.g. through increased antimicrobial drug resistance). The NHS, like many other healthcare systems, discourages GPs from prescribing these drugs if there is no clear medical reason to do so. We measure the number of prescriptions per patient, adjusted for relative pack size, for each of these three potentially harmful drugs to test if dispensing practices also take financial incentives into account when prescribing them.

3.2. Practice and patient characteristics

We have quarterly data on practice list size and the number of their dispensing patients. We measure practice organisational structure using annual data on the number of full time equivalent GPs, the proportion of them who are partners (and so residual claimants on practice profit), rather than salaried, their age, gender and whether they qualified in the UK or elsewhere. We also know the type of practice contract with the NHS (GMS vs all other types).

To control for differences in patient case-mix across practices, we use data on patient demographics (fourteen age by gender categories). We attribute a measure of average patient deprivation to each practice by using the proportions of practice patients resident in each Lower Super Output Area (LSOA)[9] and the Index of Multiple Deprivation (IMD) for each LSOA. Finally, we use annual prevalence data for twelve chronic conditions treated in primary care, which are reported for each practice as part of a national pay for performance programme - the Quality and Outcomes Framework (QOF).

3.3. Potential miles saved through dispensing

We construct a proxy measure for the potential demand for dispensing services at each GP practice based on the total travel distance that would be saved across the eligible population if the practice had a dispensary (and the eligible population would be enrolled with the practice). To do so, we first

[9] There are 34.753 LSOAs in England with a mean population of 1500 (ONS 2012).

calculate the straight-line distance from the centroid of every Output Area (OA)[10] in England to the nearest community pharmacy. The median area covered by OAs in England is 0.03 square miles (7 hectares), which suggests that OA centroids are a reasonable approximation to the location of individual residents. For each eligible OA (with centroid more than 1 mile from the nearest pharmacy) within 3 miles (4.8 km) of a practice[11] we then calculate the difference between twice the straight line distance from the centroid of the OA to the practice (i.e. the minimum distance a patient would have to travel if the practice dispenses) and the sum of the distances from the OA centroid to the practice, from the practice to the pharmacy nearest to the patient, and from the nearest pharmacy to the OA centroid (i.e. the minimum distance a patient would have to travel to obtain a prescription and to have it dispensed in a community pharmacy). We then multiply this quantity by the population of the OA and sum over all OAs within 3 miles of the practice to give an estimate of the potential miles saved by the local population if the practice had an on-site dispensary.

[10] There are 181,408 OAs in England and Wales with an average population of 309 (ONS 2012).

[11] Santos et al. (2017) report that patients are registered with general practices that are, on average, 1.2 miles (1.9km) away from their home. We use a slightly larger radius to allow for the wider catchment area of practices in rural areas.

4. Methods

We study the effect of GP dispensing on our prescribing measures both at the extensive margin (i.e. whether a practice has any patients to whom it dispenses) and at the intensive margin (i.e. the proportion of a practice's patient list to whom it has agreed to dispense, if any).

4.1. Extensive margin

Our baseline regression model is

$$y_{ijt} = \beta_0 + \beta x_{it} + \delta D_i + \omega_t + \gamma_j + \varepsilon_{ijt} \tag{3}$$

where y_{ijt} is the prescribing measure for practice i in CCG j in quarter t, x_{it} is a row vector of practice and patient characteristics, D_i is an indicator for practice dispensing status, ω_t are quarter t fixed effects, γ_j are CCG fixed effects, and ε_{ijt} is a zero mean error. The coefficient of interest is δ, which measures the average difference in prescribing behaviour between dispensing and non-dispensing practices.

A practice's dispensing status may vary over time due to changes in their patient population or local market entry of community pharmacies. However, during our study period, dispensing status was essentially time-invariant, with only 95 out of 7,850 (1.2%) practices in our sample changing status over time. We therefore model dispensing status as a time-invariant practice characteristic and exclude practices that switch dispensing status throughout the study period. This precludes the use of practice fixed effects to control for time-invariant unobserved practice factors not picked up in x_{it}. However, with very few practices changing dispensing status the effect of dispensing status would be extremely imprecisely estimated. Instead we use CCG fixed effects as dispensing status varies sufficiently across the, on average, 40 practices within each CCG. The CCG fixed effects are expected to absorb a large part of unobserved heterogeneity. CCGs can influence the prescribing of its practices via its clinical governance procedures and local prescribing incentive schemes. Local hospital provision and policies can also affect practice prescribing. For example, practices in areas where patients wait longer for hip replacements may prescribe more painkillers. The CCG effect will also allow for local area characteristics, such as air quality, availability of outside spaces or housing stock, which may affect patient morbidity and, hence, prescribing. Finally, it will control for the overall rurality of an area, which will influence patient eligibility for dispensing and, thus, practice decisions about dispensing status.

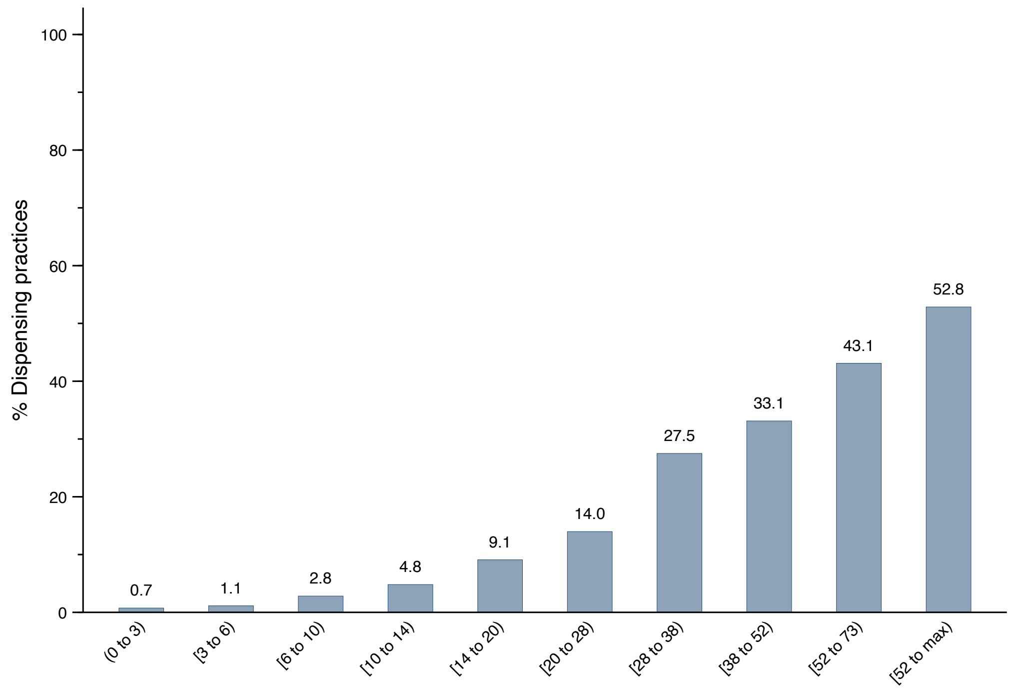
The model includes a rich set of GP and patient characteristics in x_{it} to reduce concerns over differences in patient or GP characteristics that affect prescribing and the practices' decisions to dispense. However, there may still be residual differences in observed and unobserved characteristics between dispensing and non-dispensing practices that would bias our estimates. We use two separate strategies to deal with this.

First, we pre-process our data using Entropy Balancing (EB) (Hainmueller 2012) to reduce imbalance in observable characteristics that may determine selection into treatment.[12] EB is a re-weighting approach in which the weights for each observation are chosen to approximately equalise pre-specified moments of the covariate distributions between the treatment and the control group. The pre-processed data are then analysed using weighted least squares controlling for observable characteristics. The resulting estimates have been shown to have doubly-robust properties (Zhao and Percival 2016), which reduces model

[12] We use the user-written Stata command `ebalance` (Hainmueller and Xu 2013) to estimate EB weights.

dependence and misspecification bias (Robins 1994). Assuming that observable and unobservable practice characteristics are highly correlated, EB yields unbiased estimates of the effect of treatment (dispensing status). We require the weighted data to exhibit balance in both the means and the variances of all control variables, including CCG membership. By definition, no set of weights can satisfy these requirements for practices in CCGs without any (or with only) dispensing practices, and practices in these CCGs are therefore dropped from the sample. Analysis of the re-weighted data recovers the average treatment effect on the treated (ATT), i.e. the consequences of a policy banning dispensing by practices.

Second, we use an instrumental variable (IV) strategy to control for potential selection into dispensing status based on unobservable patient or practice characteristics. Note that the regulatory framework of the NHS prohibits practices that do not treat eligible patients (e.g. those located in urban areas with close proximity to community pharmacies) from opting into dispensing services. Hence, we analyse a situation with potential one-sided non-compliance, which would result in a downward biased estimate of δ. To address endogenous selection, we estimate two-stage least squares (2SLS) models on our unweighted data using the potential local demand for dispensing services that each GP practice faces as an instrument for observed dispensing status. Specifically, our IV Z_i is defined as the total miles saved by the local resident population eligible for dispensing services who live within a 3 mile radius of practice i (see Section 3.3). Assuming that GPs are at least partially altruistic, we expect practices that can achieve a larger reduction in the travel burden of their patient population will be more likely to select into dispensing, all else equal. Figure 2 demonstrates a strong positive, monotonic relationship between miles saved by the local eligible population Z_i and dispensing status D_i. The resulting estimates are local average treatment effects (LATE) and, therefore, are informative for the policy question of what would happen if all English GP practices (that respond to the IV) were allowed to dispense.



Notes: Categories reflect deciles of the distribution. Excludes practices without eligible patients within 3 mile radius.

Figure 2: Share of dispensing practices by potential miles saved (3 mile radius around the practice)
Note: Potential miles saved is the difference between the straight line distance from the patient's OA centroid to the practice and back and the sum of the distances via the nearest pharmacy, which is then multiplied with the eligible population in this OA and summed over all OAs within a 3 mile radius.

The validity of our IV may be compromised if the potential number of eligible patients in the local vicinity of the practice is in itself correlated with its prescribing behaviour. This may be the case if practices compete for patients and seek to attract demand by prescribing over-generously. However, GP practices located in rural areas, where dispensing status is potentially endogenous, face little competition for patients and, therefore, have less need to attract patients than their urban counterparts. This is consistent with Schaumans (2015), who finds that Belgian GPs write fewer prescriptions per patient in areas with less competition. Furthermore, prescribing behaviour is difficult to ascertain for patients at the stage of practice selection since these data are not reported publicly, e.g. on websites such as nhs.net. This reduces the scope for competition based on prescribing. We, therefore, argue that our IV is both strong and valid.[13]

All standard errors are clustered at practice level to account for serial correlation. In the context of the EB analysis, inference is conditional on the estimated weights.

4.2. Intensive margin

Our weighted prescribing measures are averages at practice level where the weights are the proportions of dispensing and non-dispensing patients. Thus,

$$y_{ijt} \equiv s_{ijt} y^d_{ijt} + (1 - s_{ijt}) y^{nd}_{ijt} = y^{nd}_{ijt} + s_{ijt}(y^d_{ijt} - y^{nd}_{ijt}) = y^{nd}_{ijt} + s_{ijt}\Delta_{ijt} \tag{4}$$

where y^d_{ijt} and y^{nd}_{ijt} are the prescribing measures for dispensing and non-dispensing patients, Δ_{ijt} denotes the difference in prescribing measures between both patient groups, and $s_{ijt} = L^d_{ijt}/L_{ijt}$ is the proportion of patients to whom the practice dispenses.

The analysis at the extensive margin recovers Δ_{ijt} at $\bar{s}|s > 0$, i.e. at the average value of s_{ijt} over all dispensing practices. We now seek to establish whether Δ_{ijt} is in itself a function of s_{ijt} and thus varies across the range $s_{ijt} \in (0, 1]$. This would indicate that GP practices change their relative prescribing behaviour as the proportion of dispensing patients in the practice changes, for example because dispensing income becomes a more important source of overall practice income or, conversely, because the fixed costs of operating a dispensary are distributed over more patients. For simplicity, we assume that Δ_{ijt} is approximately linear in s_{ijt} so that

$$\Delta_{ijt} = \phi_1 + \phi_2 s_{ijt} \tag{5}$$

where ϕ_1 is the constant difference in prescribing measures between dispensing and non-dispensing patients at any level of s_{ijt}, and ϕ_2 is a behavioural parameter that reflects the responsiveness of GPs' prescribing behaviours to the share of dispensing patients in their practices.

We estimate an intensive margin model for the sub-sample of practices with a positive number of dispensing patients to recover ϕ_1 and ϕ_2. We exploit variation within practices in terms of the number of patients to whom they can dispense drugs assuming that this varies exogenously within and between

[13] Another concern is that the instrument might be correlated with the outcome variable through its correlation with the *share* of dispensing patients s_{ijy} (see Section 4.2). However, we find that the share of dispensing patients in a practice and our instrumental variable ('total miles saved if dispensing') are only weakly correlated (ρ=0.198). People living in rural areas are unlikely to be equally distributed around practices, which might result in a small share of dispensing patients but a high amount of miles saved.

practices once a dispensary has been established. This is a reasonable assumption since dispensing practices cannot exclude eligible patients from drug dispensing services without terminating dispensing entirely. GP practice dispensing status is largely time-invariant and so any inter-temporal variation in the number of patients receiving dispensing services is solely the effect of variation in local demand and not an effect of selection.

We use similar covariates to the extensive margin model in Eq. (3) but replace the binary dispensing status variable with the share of dispensing patients s_{ijt} and its squared value so that

$$y_{ijt} = \alpha_i + \beta x_{it} + \phi_1 s_{ijt} + \theta s_{ijt}^2 + \omega_t + \varepsilon_{ijt} \tag{6}$$

where α_i is a practice fixed effect and $\theta = \frac{1}{2} \cdot \phi_2$. Taking the first derivative of Eq (6) with respect to s_{ijt} yields the expression in Eq. (5). Note that if $\theta = 0$ then Δ_{ijt} is constant in s_{ijt} and the model collapses to the extensive margin model in Eq. (3). Since we have limited within GP practice variation in s_{ijt}, we also estimate a model with CCG fixed effects instead of practice fixed effects. Standard errors are clustered at practice level.

5. Results

5.1. Descriptive statistics

Table 1 presents the descriptive statistics by practice dispensing status for the prescription measures, practice and patient characteristics as well as the distance measure. Figure A1 in the Appendix presents kernel density plots for the prescribing indicators. Thirteen percent (1,023) of practices in the full sample and 29 percent (988) in the final sample have dispensing patients.[14] These practices prescribe, on average, more and smaller packages[15] than non-dispensing practices. They also have higher net ingredient costs (NIC) per patient. Dispensing practices also appear to be more homogeneous in their prescribing behaviour as indicated by the lower variance of the prescribing indicators. Dispensing practices have slightly smaller list sizes, more GPs per patient, and a higher proportion who qualified in the UK. They also have somewhat older patients. Dispensing practices have higher rates of disease prevalence, except for chronic obstructive pulmonary disease and mental health problems but their patients are less deprived than those in non-dispensing practices.

The average dispensing practice has dispensing rights for approximately 3,200 patients, or 49% of their list in the final sample (Figure 3). There is substantial variation in the number of dispensing patients across practices, with some practices being allowed to dispense for up to 10,000 patients. The share of dispensing patients in dispensing practices ranges from nearly 0% to 100% and has a bimodal distribution.

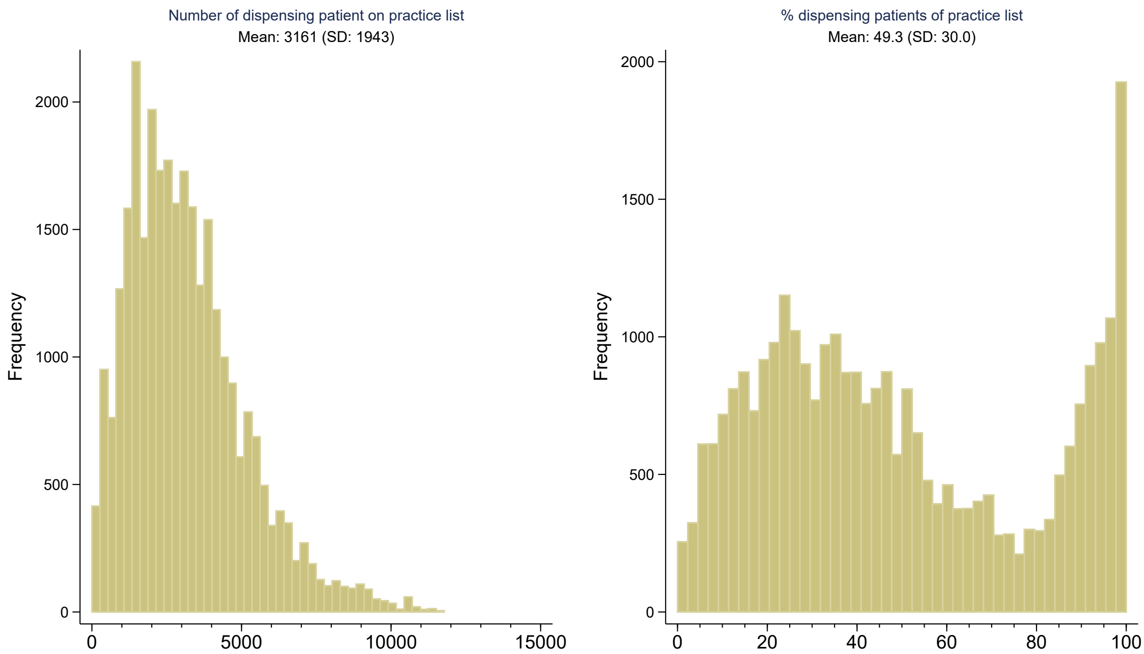


Figure 3: Number and share of dispensing patients in dispensing practices in the final sample

[14] For the descriptive statistics and the number and share of dispensing patients in dispensing practices in the full sample see Appendix Table A3 and Figure A2.
[15] Note that relative pack size is defined with respect to the mode, not the average, pack size.

Table 1: Descriptive statistics by practice dispensing status for the final sample

	Dispensing practices		Non-dispensing practices	
	Mean	SD	Mean	SD
Prescribing measures				
Cost per patient	41.11	7.65	38.55	9.02
Cost per prescription	6.73	0.71	6.70	0.75
Prescriptions per patient	6.16	1.31	5.82	1.47
OTC prescriptions per patient	0.92	0.26	0.89	0.28
Antibiotic prescriptions per patient	0.18	0.04	0.18	0.09
Opioid prescriptions per patient	0.23	0.08	0.23	0.10
Antidepressant prescriptions per patient	0.35	0.09	0.35	0.11
Relative pack size	1.11	0.16	1.28	0.23
Proportion generic prescriptions	0.80	0.06	0.79	0.06
Organisational structure of practice				
List size	7812.58	4687.23	8222.02	4678.23
Full-time equivalent GPs per 1,000 patients	0.53	0.30	0.46	0.27
GP partners (%)	0.69	0.24	0.68	0.29
UK-trained GPs (%)	0.65	0.36	0.53	0.38
Age structure of GPs (proportion)				
Age <40	0.27	0.22	0.28	0.24
Age 40 to 59	0.66	0.25	0.60	0.28
Age 60+	0.07	0.16	0.12	0.24
Contract type				
GMS	0.75	0.43	0.62	0.49
other (incl. PMS)	0.25	0.43	0.38	0.49
Patient characteristics				
Age-sex proportions				
Male - 0 to 4	0.02	0.01	0.03	0.01
Male - 5 to 19	0.08	0.01	0.09	0.01
Male - 20 to 44	0.13	0.02	0.17	0.04
Male - 45 to 59	0.11	0.01	0.10	0.01
Male - 60 to 74	0.10	0.02	0.08	0.02
Male - 75 to 84	0.03	0.01	0.03	0.01
Male - 85+	0.01	0.00	0.01	0.00
Female - 0 to 4	0.02	0.01	0.03	0.01
Female - 5 to 19	0.08	0.01	0.08	0.01
Female - 20 to 44	0.13	0.02	0.16	0.03
Female - 45 to 59	0.11	0.01	0.10	0.02
Female - 60 to 74	0.10	0.02	0.08	0.02
Female - 75 to 84	0.04	0.01	0.03	0.01
Female - 85+	0.02	0.01	0.01	0.01
Prevalence of chronic conditions (%)				
Coronary heart disease	3.78	0.91	3.46	1.06
Stroke	2.10	0.55	1.85	0.62
Hypertension	15.99	2.82	14.32	3.42
Chronic obstructive pulmonary disease	1.81	0.62	1.96	0.84
Cancer	2.92	0.77	2.28	0.83
Mental health problems	0.64	0.22	0.85	0.35
Asthma	6.45	0.96	6.15	1.15
Heart failure	0.85	0.34	0.78	0.35
Palliative care	0.35	0.37	0.30	0.31
Dementia	0.74	0.36	0.71	0.42
Atrial fibrillation	2.20	0.59	1.73	0.67
Cardiovascular disease (aged 30-74)	1.92	1.10	1.77	1.03
Index of Multiple Deprivation (2015)	0.09	0.03	0.14	0.07
IV for dispensing status				
Miles saved if dispense (1000s)	56.40	31.57	27.07	27.84
Practices	988		3,410	
Practice-quarter observations	30,203		99,910	

Note: Data: quarter-practice level 2011Q1 to 2018Q4. See Appendix Table A1 for sources. All prescribing measures, except cost per patient, are adjusted by relative pack size.

5.2. Extensive margin

Table 2 presents the estimated effects of dispensing status on our prescribing measures. The OLS results are consistent with the predictions in Section 2.1: dispensing practices prescribe more drugs, drugs that are more expensive, and in smaller pack sizes. Dispensing practices have 0.1 more prescriptions per quarter, even after allowing for the fact that they, on average, have 17% (100*(-0.22/1.28)) smaller packages compared to non-dispensing practices. The average NIC per prescription is £0.09 higher than in non-dispensing practices and they prescribe fewer generic drugs that cost less than branded alternatives. Finally, dispensing practices prescribe more OTC drugs, antibiotics and opioids.

Overall, the differences in prescribing behaviour are associated with an additional expenditure of £1.23 per patient per quarter, or (100*(1.23/38.55))=3.19% of the mean quarterly expenditure per patient for non-dispensing practices. For a dispensing practice of average list size, this amounts to £38,677 of additional prescribing expenditure per year or £38M per year over all dispensing practices.

Table 2: Effect of dispensing status at the extensive margin

	Pooled OLS		EB + WLS		2SLS	
	Est	SE	Est	SE	Est	SE
Cost per patient	1.228***	0.159	1.605***	0.189	1.950**	0.731
Cost per prescription	0.093***	0.016	0.108***	0.021	0.170*	0.080
Prescriptions per patient	0.099***	0.023	0.138***	0.030	0.278*	0.117
OTC prescriptions per patient	0.060***	0.006	0.049***	0.006	0.042	0.026
Antibiotic prescriptions per patient	0.007***	0.002	0.002	0.002	0.031*	0.013
Opioid prescriptions per patient	0.007***	0.002	0.009**	0.003	0.027**	0.009
Antidepressant prescriptions per patient	0.001	0.002	0.010***	0.003	0.033**	0.011
Relative pack size	-0.219***	0.007	-0.207***	0.011	-0.258***	0.034
% generic prescriptions	-0.004**	0.001	-0.005**	0.002	-0.001	0.007
Practice-quarter observations	130,113		130,113		130,113	
Partial F-test of excluded instrument					132.7	
Test of endogeneity (p-value)					0.706	

*** $p<0.01$, ** $p<0.05$, * $p<0.1$
Est = Coefficient estimate; SE = Standard error; OTC = Over-the-counter; OLS = Ordinary Least Squares; EB = Entropy balancing; WLS = Weighted Least Squares; 2SLS = Two-stage Least Squares.
Note: All models control characteristics of the patient population and the organisational structure of the practice. Quarterly data 2011Q1 to 2018Q4 (Sources see Table A1).

Some of the differences in observable characteristics between dispensing and non-dispensing practices may not be fully accounted for by OLS regression adjustment. Entropy balancing equalises the first two moments of the covariate distributions (see Appendix Table A4 for descriptive statistics). The resulting WLS estimates are generally in line with the OLS estimates, although there are two noteworthy differences. First, the effect of dispensing status on both the prescription costs and the number of prescriptions issued is 30% and 39% larger than under OLS, respectively. The increase in the latter appears to be driven mainly by increased prescribing of non-OTC medications, which account for 64.5% (=1 - 0.049/0.138) of the additional prescribing volume associated with dispensing status. Second, we now find dispensing status to be linked to increased prescribing of antidepressants, whereas there is no longer a statistical significant effect on antibiotic prescribing.

The 2SLS estimates allow for selection into treatment due to unobservable practice characteristics and identify LATEs. The IV is the total potential travel distance saved for the resident population within a 3-mile radius around the practice if the practice operated an on-site dispensary and all residents were registered with this practice. (Semi-)altruistic practice owners are expected to consider possible reductions in travel burden for their patient population when deciding whether to opt into dispensing (see Section 2.2). However, the reduction in travel distance if the practice chooses to dispense is unlikely to affect prescribing behaviour conditional on the included patient and practice characteristics. The IV is a strong predictor of dispensing status with a partial F-statistic of 132.7 (see Appendix Table A5 for first stage results). The resulting point estimates are broadly consistent with the OLS and EB+WLS estimates although they tend to be a bit larger. However, the 2SLS estimates do not support the notion that dispensing practices prescribe statistically significantly more OTCs. Since the robust score test of endogeneity (Wooldridge 1995) fails to reject the null hypothesis of conditional exogeneity (p=0.706) we prefer the more efficient OLS/WLS estimates to judge the ATE of dispensing status.

Table 3 shows that our results are robust to alternative adjustments for selection on observables, namely 1:1 nearest neighbour matching (NNM) without replacement and inverse probability weighting with regression adjustment (IPWRA) for the observations on the common support (129,163). Both methods allow calculation of the ATT and ATEs under the assumption that practices do not select into dispensing based on unobservable characteristics.

Table 3: Extensive margin - alternative adjustments for selection on observables for the final sample

	NNM (ATT)		IPWRA (ATT)		NNM (ATE)		IPWRA (ATE)	
	Est	SE	Est	SE	Est	SE	Est	SE
Cost per patient	1.832***	0.052	1.615***	0.067	1.382***	0.044	1.819***	0.109
Cost per prescription	0.172***	0.005	0.109***	0.006	0.130***	0.004	0.016	0.012
Prescriptions per patient	0.116***	0.009	0.132***	0.013	0.088***	0.007	0.291***	0.026
OTC prescriptions per patient	0.05***	0.002	0.051***	0.002	0.024***	0.001	0.063***	0.005
Antibiotic prescriptions per patient	0.004***	0.0003	0.003***	0.0005	-0.004***	0.0004	0.005***	0.0006
Opioid prescriptions per patient	0.009***	0.001	0.01***	0.0007	0.19***	0.0005	0.028***	0.0008
Antidepressant prescriptions per patient	0.001**	0.001	0.011***	0.0008	0.002***	0.0001	0.017***	0.001
Relative pack size	-0.237***	0.002	-0.218***	0.003	-0.158***	0.002	-0.103***	0.004
Proportion generic prescriptions	-0.008***	0.0005	-0.006***	0.0006	-0.007***	0.0004	-0.007***	0.001
Practice-quarter observations	130,113		129,163		130,113		129,163	

*** p<0.01, ** p<0.05, * p<0.1
Est = Coefficient estimate; SE = Standard error; OTC = Over-the-counter; NNM = 1:1 Nearest-neighbour matching without replacement with Abadie-Imbens SEs (Abadie and Imbens 2006); IPWRA = Inverse probability weighting with regression adjustment. ATE = Average treatment effect; ATT = Average treatment effect on the treated.
Note: All models control for a full set of control variables for characteristics of the patient population and the organisational structure of the practice. Quarterly data 2011Q1 to 2018Q4 (Sources see Table A1). All estimations were performed using Stata's `teffects` routine.

5.3. Intensive margin

Table 4 presents the results of our analysis of the intensive margin. For each prescribing measure we present two coefficient estimates: the difference in prescribing between dispensing and non-dispensing patients that is independent of the actual share of dispensing patients in the practice (ϕ_1), and the estimated marginal effect of increasing the share of dispensing patients in the practice by 100 percentage points (θ). For ease of interpretation, the share variable s_{ijt} is mean centered so that θ denotes deviations away from the sample mean of approximately 49%.

Most (97.5%) of the variation in the share of dispensing patients in our sample is between GP practices so that point estimates are poorly identified once GP fixed effects are introduced. We therefore prefer to focus on the model with CCG fixed effects that capture unobserved differences at the regional level.[16] In line with the pooled OLS and EB+WLS results at the extensive margin (Table 2), we find that dispensing status increases prescribing costs, prescriptions per patient (both OTC and non-OTC), and the cost per prescription but reduces pack size and % generic prescribing. In addition, our estimates of θ (and

[16] On average CCGs in our sample have 40 GP practices.

thus ϕ_2) suggest a diminishing marginal effect of dispensing share on most prescribing measures as evidenced by the opposing signs of ϕ_1 and θ. Put differently, GPs in dispensing practices prescribe more similarly to their peers in non-dispensing practices when the share of patients to whom they can dispense is high. However, this diminishing effect is often small in magnitude.[17]

Figure 4 plots the predicted values of parametric regression models assuming either a linear or parabolic relationship between s_{ijt} and the prescribing measures of interest. In addition, a spline model with 20 equally spaced knots serves as a non-parametric approximation of these relationships.

Table 4: Effect at intensive margin

	Pooled OLS		CCG fixed effects		GP fixed effects	
	Est	SE	Est	SE	Est	SE
Cost per patient						
ϕ_1	2.261***	0.593	1.784***	0.518	1.263	1.154
θ	-5.007***	1.691	-3.194**	1.403	2.632	2.629
Cost per prescription						
ϕ_1	0.200***	0.068	0.115**	0.058	0.295*	0.156
θ	-0.639***	0.188	-0.336**	0.146	0.328	0.372
Prescriptions per patient						
ϕ_1	0.239***	0.092	0.246***	0.083	0.063	0.212
θ	0.010	0.219	-0.049	0.209	-0.189	0.567
OTC prescriptions per patient						
ϕ_1	0.047**	0.023	0.068***	0.018	0.029	0.031
θ	-0.078	0.049	-0.124***	0.044	0.005	0.104
Antibiotic prescriptions per patient						
ϕ_1	0.007*	0.004	0.009*	0.004	0.000	0.008
θ	-0.022*	0.012	-0.017	0.012	-0.004	0.019
Opioid prescriptions per patient						
ϕ_1	0.004	0.006	0.005	0.006	0.009	0.009
θ	-0.011	0.017	-0.001	0.016	0.028	0.023
Antidepressant prescriptions per patient						
ϕ_1	0.005	0.008	-0.005	0.007	-0.002	0.013
θ	-0.018	0.023	-0.001	0.020	0.047	0.031
Relative pack size						
ϕ_1	-0.156***	0.020	-0.158***	0.018	-0.038	0.035
θ	0.362***	0.051	0.302***	0.044	0.124	0.105
Proportion generic prescriptions						
ϕ_1	-0.014***	0.005	-0.012**	0.005	-0.006	0.019
θ	-0.048***	0.014	-0.036***	0.014	-0.074	0.047
Practice-quarter observations	30,203		30,203		30,203	

*** p<0.01, ** p<0.05, * p<0.1
Est = Coefficient estimate; SE = Standard error.
All models control for a full set of control variables for characteristics of the patient population and the organisational structure of the practice.
Note: ϕ_1 and $\theta = \frac{1}{2}\phi_2$ denote the regression coefficients on the share of dispensing patients s_{ijt} and s_{ijt}^2, respectively. See Section 4.2 for further details.

[17] Table A6 in the Appendix reports the regression coefficients of a model assuming that prescribing behaviour is constant in s_{ijt}.

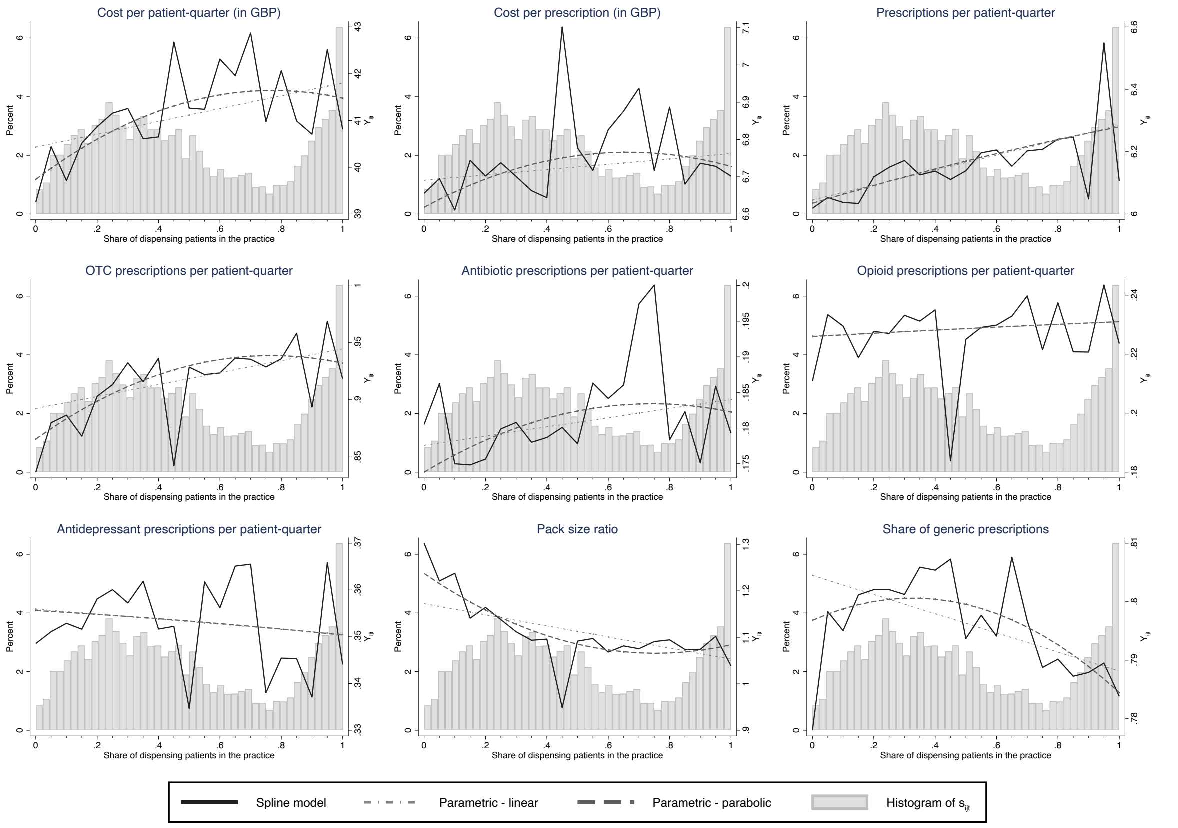


Figure 4: Prescribing outcomes at different proportions of dispensing patients in dispensing practices

6. Conclusion

The English NHS is one of several healthcare systems that permit GPs to prescribe *and* dispense medicines under specific circumstances (Eggleston 2012). GP practices are allowed to dispense prescriptions to patients who live more than 1 mile away from the nearest community pharmacy or who otherwise struggle to access a pharmacy. We present evidence that English GPs respond to the financial incentives created by geographic, exogenous variation in dispensing rights by prescribing, on average, more and more costly medications that are provided in smaller pack sizes. These results are consistent with a rent extraction motive in which GP practices prescribe to increase dispensing income.

We obtain broadly similar results from the analysis of the intensive margin, with drugs costs per patient, drug costs per prescription, and prescriptions per patient higher, and pack size and share of generic prescribing lower, in dispensing practices with a higher proportion of dispensing patients. However, the effects of the share of dispensing patients falls as the share increases. While our data do not allow us to unpick this finding further, we note that it is consistent with a more altruistic interpretation of practice motivation, in which practices provide dispensing services in the interests of patients while seeking to recover the fixed costs of running a dispensary.

Our prescribing indicators are averages across all patients within practices since the data do not distinguish between prescriptions for dispensing and non-dispensing patients. One could argue that GPs make no distinction between dispensing and non-dispensing patients and prescribe more and more expensively to all patients when operating a dispensary. However, it is also possible that doctors distinguish patient types in their practice. If so, our analyses provide a conservative assessment of the impact of PD on GPs' prescribing behaviour towards patients to whom they can dispense.

We can use our estimates to provide a back-of-the-envelope calculation of the additional expenditure due to PD in the English NHS. Based on our EB estimates, we calculate that a dispensing GP practice of average size (i.e. approx. 7,800 patients) has additional revenue of £57,650 per year, on average, of which 87% are due to additional prescribing expenditure and the remaining 13% reflect dispensing fees linked to additional prescribing (assuming all prescriptions are dispensed on-site). Aggregated over the 918 dispensing practices in the English NHS in 2018, this amounts to an additional expenditure of approximately £53m per year, or the equivalent of nearly 3,900 quality-adjusted life years that could have been produced with this money if invested otherwise (Claxton et al. 2015). However, our analysis only provides a partial picture of the full consequences of PD. We do not know whether the additional expenditure due to PD generates health benefits (e.g. due to improved adherence to pharmacological therapy), nor how much PD reduces travel time for patients and how this is valued. Furthermore, we do not observe the cost of operating an on-site dispensary to GP practices and the degree of rent extraction. We, therefore, caution against drawing strong inference about the welfare implications of the English PD policy from our results.

In summary, our analysis provides evidence that English GPs modify their prescribing behaviour when permitted to dispense medications in ways that are consistent with a profit motive. These behavioural differences are unlikely to be explained by differences in the health care needs of their local patient populations. However, our results do not permit us to exclude the possibility that PD increases overall welfare. Future studies should explore the consequence of PD on population health and access to care to permit a more comprehensive assessment of costs and benefits.

References

Abadie, A. and G. W. Imbens (2006). 'Large sample properties of matching estimators for average treatment effects'. *Econometrica*, 74: 235–267.

Ahammer, A. and T. Schober (2020). 'Exploring variations in health-care expenditures. What is the role of practice styles?' *Health Economics*, 29: (6), 683–699.

Ashworth, M., P. White, H. Jongsma, P. Schofield and D. Armstrong (2016). 'Antibiotic prescribing and patient satisfaction in primary care in England: cross-sectional analysis of national patient survey data and prescribing data'. *British Journal of General Practice*, 66: e40–e46.

Beardon, P., M. McGilchrist, A. McKendrick, D. McDevitt and T. MacDonald (1993). 'Primary non-compliance with prescribed medication in primary care.' *British Medical Journal*, 307: 846–848.

Burkhard, D., C. P. R. Schmid and K. Wüthrich (2019). 'Financial incentives and physician prescription behavior: Evidence from dispensing regulations'. *Health Economics*, 1114–1129.

Buxbaum, J., M. Chernew, A. Fendrick and D. Cutler (2020). 'Contributions Of Public Health, Pharmaceuticals, And Other Medical Care To US Life Expectancy Changes, 1990-2015'. *Health Affairs*, 39: 1546–1556.

Chou, Y., W. C. Yip, C.-H. Lee, N. Huang, Y.-P. Sun and H.-J. Chang (2003). 'Impact of separating drug prescribing and dispensing on provider behaviour: Taiwan's experience'. *Health Policy and Planning*, 18: 316–329.

Claxton, K., S. Martin, M. Soares, N. Rice, E. Spackman, S. Hinde, N. Devlin, P. C. Smith and M. Sculpher (2015). 'Methods for the estimation of the National Institute for Health and Care Excellence cost-effectiveness threshold'. *Health Technology Assessment*, 19: 1–503.

Clemens, J. and J. D. Gottlieb (2014). 'Do Physicians' Financial Incentives Affect Medical Treatment and Patient Health?' *American Economic Review*, 104: 1320–1349.

Crea, G., M. Galizzi, I. Linnosmaa and M. Miraldo (2019). 'Physician altruism and moral hazard: (no) Evidence from Finnish national prescriptions data'. *Journal of Health Economics*, 65: 153–169.

Department of Health (2012). *Regulations under the Health Act 2009: Market entry by means of Pharmaceutical Needs Assessments: Information for Primary Care Trusts Executive Summary and Chapters 1-4*.

Eggleston, K. (2012). 'Prescribing institutions: Explaining the evolution of physician dispensing'. *Journal of Institutional Economics*, 8: 247–270.

Filippini, M., F. Heimsch and G. Masiero (2014). 'Antibiotic consumption and the role of dispensing physicians'. *Regional Science and Urban Economics*, 49: 242–251.

Goldacre, B., C. Reynolds, A. Powell-Smith, A. J. Walker, T. A. Yates, R. Croker and L. Smeeth (2019). '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: e026886.

Hainmueller, J. (2012). 'Entropy balancing for causal effects: A multivariate reweighting method to produce balanced samples in observational studies'. *Political Analysis*, 25–46.

Hainmueller, J. and Y. Xu (2013). 'Ebalance: A Stata package for entropy balancing'. *Journal of Statistical Software*, 54:

House of Commons Library (2020). *House of Commons Library briefing on NHS charges.* Briefing paper number 07227.

Iizuka, T. (2012). 'Physician Agency and Adoption of Generic Pharmaceuticals'. *American Economic Review*, 102: 2826–2858.

Kaiser, B. and C. Schmid (2016). 'Does Physician Dispensing Increase Drug Expenditures? Empirical Evidence from Switzerland'. *Health Economics*, 25: 71–90.

Lichtenberg, F. R. (2012). 'Contribution of Pharmaceutical Innovation to Longevity Growth in Germany and France, 2001-7'. *PharmacoEconomics*, 30: 197–211.

Lim, D., J. Emery, J. Lewis and V. B. Sunderland (2009). 'A systematic review of the literature comparing the practices of dispensing and non-dispensing doctors'. *Health Policy*, 92: 1–9.

Liu, Y.-M., Y.-H. K. Yang and C.-R. Hsieh (2009). 'Financial incentives and physicians' prescription decisions on the choice between brand-name and generic drugs: Evidence from Taiwan'. *Journal of Health Economics*, 28: 341–349.

Lundin, D. (2000). 'Moral hazard in physician prescribing behaviour'. *Journal of Health Economics*, 19: 639–662.

McGuire, T. G. (2000). 'Chapter 9 Physician agency'. In: *Handbook of Health Economics*. Vol. 1, 461–536.

Morton-Jones, T. J. and M. A. L. Pringle (1993). 'Prescribing costs in dispensing practices'. *BMJ*, 306: 1244–1246.

NHS Digital (2019). *GP Earnings and Expenses Estimates 2017/18*.

ONS (2012). *2011 Census: Population and Household Estimates for Small Areas in England and Wales*.

Park, S., S. B. Soumerai, A. S. Adams, J. A. Finkelstein, S. Jang and D. Ross-Degnan (2005). 'Antibiotic use following a Korean national policy to prohibit medication dispensing by physicians'. *Health Policy and Planning*, 20: 302–309.

Rischatsch, M. (2014). 'Lead me not into temptation: drug price regulation and dispensing physicians in Switzerland'. *European Journal of Health Economics*, 15: 697–708.

Robins, J. M. (1994). 'Correcting for non-compliance in randomized trials using structural nested mean models'. *Communications in Statistics-Theory and methods*, 23: 2379–2412.

Santos, R., H. Gravelle and C. Propper (2017). 'Does Quality Affect Patients' Choice of Doctor? Evidence from England'. *Economic Journal*, 127: 445–494.

Schaumans, C. (2015). 'Prescribing behavior of general practitioners: competition matters'. *Health Policy*, 119: 456–463.

Trap, B. (2002). 'Prescription habits of dispensing and non-dispensing doctors in Zimbabwe'. *Health Policy and Planning*, 17: 288–295.

Trottmann, M., M. Frueh, H. Telser and O. Reich (2016). 'Physician drug dispensing in Switzerland: association on health care expenditures and utilization'. *BMC Health Services Research*, 16: 238.

Wooldridge, J. M. (1995). 'Score diagnostics for linear models estimated by two stage least squares'. *Advances in econometrics and quantitative economics: Essays in honor of Professor CR Rao*, 66–87.

Zgierska, A., M. Miller and D. Rabago (2012). 'Patient satisfaction, prescription drug abuse, and potential unintended consequences'. *JAMA*, 307: 1377–1378.

Zhao, Q. and D. Percival (2016). 'Entropy balancing is doubly robust'. *Journal of Causal Inference*, 5: 1.

Appendix

Table A1: Data sources and reporting frequencies

Data	Frequency	Source
Prescriptions per GP practice	Monthly	NHSBSA
Number of dispensing patients per practice	Quarterly	NHSBSA
Patient demographics	Annual	NHS Digital
GP characteristics	Annual	NHS Digital
QOF disease prevalences	Annual	NHS Digital
Population density	Annual	ONS

Abbreviations: NHSBSA = NHS Business Services Authority; ONS = Office for National Statistics; QOF = Quality & Outcomes Framework.

Table A2: Definition of practice prescribing measures

Prescribing measure	Definition
Cost per patient	$\left(\sum_k \sum_\ell N_{k\ell it} C_{k\ell t}\right) / L_{it} = C_{it}/L_{it}$
Cost per prescription	C_{it} / N_{it}^A
Relative pack size	$RPS_{it} = \sum_k \left(\frac{N_{kit}}{N_{it}}\right) \left(\frac{1}{M_k} \frac{\sum_\ell N_{k\ell it} Q_{k\ell t}}{\sum_\ell N_{kit}}\right)$ $= \sum_k \left(\frac{N_{kit}}{N_{it}}\right) RPS_{itk}$
Prescriptions per patient	$\left(\sum_k N_{itk} RPS_{itk}\right) / L_{it}$ $= \left(\sum_k N_{ikt}^A\right) / L_{it} = N_{it}^A / L_{it}$
Proportion generic prescriptions	$\sum_{k \in generic} N_{kit}^A \Big/ N_{it}^A$
Variable	**Definition**
k	drug name and formulation
ℓ	pack size category
i	practice
t	period (quarter)
L_{it}	list size of practice i in period t
$N_{k\ell it}$	number of prescriptions (items) of drug k pack size ℓ
N_{it}	total prescriptions
$C_{k\ell t}$	net ingredient cost drug k in pack size ℓ
C_{it}	total net ingredient cost
L_{it}	list size of practice i in period t
$Q_{k\ell t}$	quantity drug k supplied in pack size ℓ
M_k	Modal pack size drug k (over all periods, practices)
$RPS_{itk} = \left(\frac{1}{M_k} \frac{\sum_\ell N_{k\ell it} Q_{k\ell t}}{\sum_\ell N_{kit}}\right)$	Relative pack size drug k, practice i, period t
$N_{itk}^A = N_{itk} RPS_{itk}$	Quantity adjusted number of items drug k
$N_{it}^A = \sum_k N_{itk} RPS_{itk}$	Total quantity adjusted items

Note: Prescribing volume of specific medications (i.e. antibiotics, opioids, and antidepressants) and OTCs are defined in the same way as prescriptions per patient.

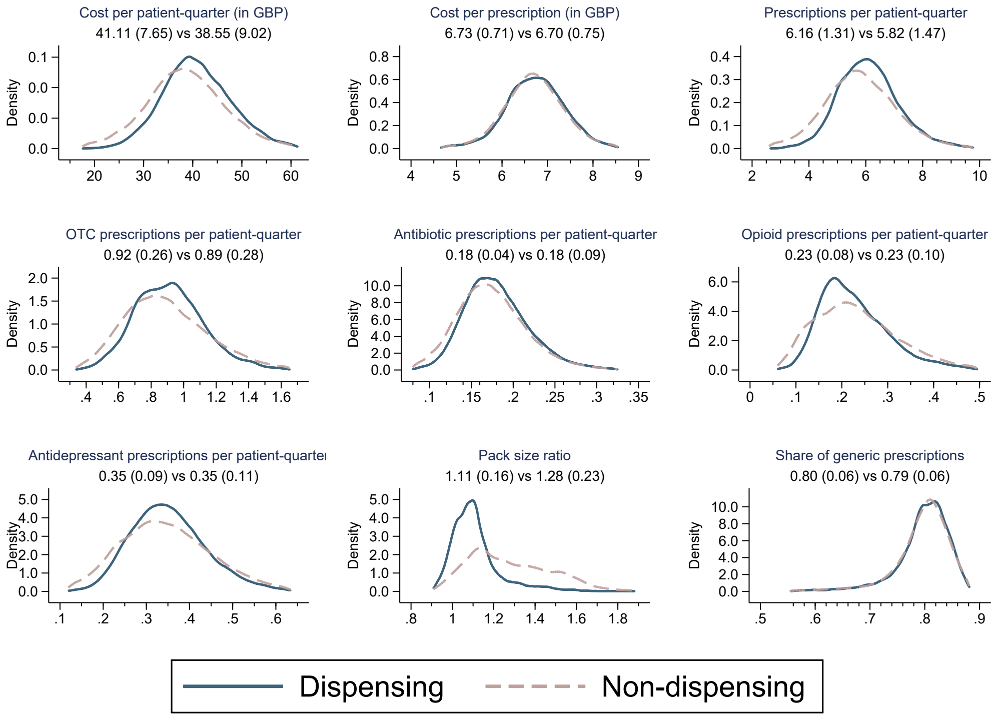


Figure A1: Kernel densities for prescribing measures for the final sample

Note: Shown densities exclude top and bottom percentiles of sample values. Mean and SD are reported for the full sample.

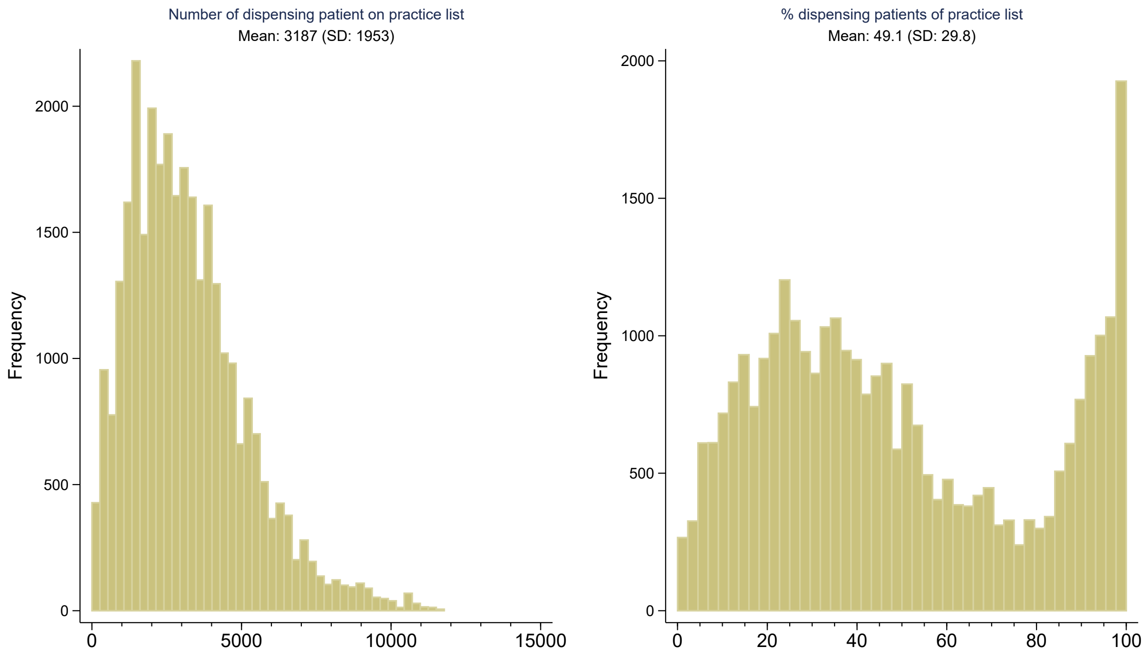


Figure A2: Number and share of dispensing patients in dispensing practices in the full sample

Table A3: Descriptive statistics - initial sample including practices in CCGs with 0 or 100% dispensing practices

	Dispensing practices		Non-dispensing practices	
	Mean	SD	Mean	SD
Prescribing measures				
Cost per patient	41.34	7.78	37.01	10.21
Cost per prescription	6.73	0.70	6.74	0.80
Prescriptions per patient	6.19	1.32	5.55	1.60
OTC prescriptions per patient	0.92	0.25	0.90	0.30
Antibiotic prescriptions per patient	0.18	0.04	0.17	0.08
Opioid prescriptions per patient	0.23	0.08	0.21	0.11
Antidepressant prescriptions per patient	0.36	0.09	0.31	0.13
Relative pack size	1.11	0.16	1.28	0.23
% generic prescriptions	0.80	0.06	0.79	0.06
Organisational structure of practice				
List size	7873.76	4707.96	7377.11	4466.22
Full-time equivalent GPs per 1,000 patient	0.52	0.30	0.46	0.27
GP partners (%)	0.69	0.24	0.67	0.30
UK-trained GPs (%)	0.65	0.36	0.51	0.38
Contract type				
GMS	0.74	0.44	0.60	0.49
other (incl. PMS)	0.26	0.44	0.40	0.49
Age structure of GPs (%)				
Age <40	0.27	0.22	0.29	0.26
Age 40 to 59	0.67	0.24	0.56	0.31
Age 60+	0.07	0.16	0.16	0.27
Patient characteristics				
Age-sex proportions				
Male - 0 to 4	0.02	0.01	0.03	0.01
Male - 5 to 19	0.08	0.01	0.09	0.02
Male - 20 to 44	0.13	0.02	0.18	0.05
Male - 45 to 59	0.11	0.01	0.10	0.02
Male - 60 to 74	0.10	0.02	0.07	0.02
Male - 75 to 84	0.03	0.01	0.02	0.01
Male - 85+	0.01	0.00	0.01	0.00
Female - 0 to 4	0.02	0.01	0.03	0.01
Female - 5 to 19	0.08	0.01	0.08	0.02
Female - 20 to 44	0.13	0.02	0.18	0.04
Female - 45 to 59	0.11	0.01	0.10	0.02
Female - 60 to 74	0.10	0.02	0.07	0.02
Female - 75 to 84	0.04	0.01	0.03	0.01
Female - 85+	0.02	0.01	0.01	0.01
Prevalence of chronic conditions or major health shocks (per 1000)				
Coronary heart disease	3.80	0.91	3.21	1.16
Stroke	2.11	0.55	1.63	0.65
Hypertension	16.05	2.83	13.62	3.63
Chronic obstructive pulmonary disease	1.82	0.63	1.86	0.94
Cancer	2.93	0.78	2.02	0.85
Mental health problems	0.65	0.22	0.94	0.48
Asthma	6.47	0.98	5.84	1.34
Heart failure	0.85	0.34	0.73	0.35
Palliative care	0.35	0.37	0.28	0.31
Dementia	0.75	0.37	0.63	0.43
Atrial fibrillation	2.21	0.60	1.49	0.71
Cardiovascular disease (aged 30-74)	1.92	1.11	1.75	1.02
Index of Multiple Deprivation (2015)	0.09	0.03	0.17	0.08
Potential miles saved if dispensing (in 1000)				
3 mile radius	57.17	32.08	16.27	23.73
Practice-quarter observations	31,263		197,915	

Note: Data: quarter-practice level 2011Q1 to 2018Q4 including practices in CCGs with 0 or 100% dispensing practices. See Appendix Table A1 for sources. All prescribing measures, except cost per patient, are adjusted by relative pack size.

Table A4: Descriptive statistics after EB+WLS for the final sample

	Unweighted				Weighted	
	Treated		Control		Control	
	Mean	SD	Mean	SD	Mean	SD
Organisational structure of practice						
List size	8.80	0.59	8.85	0.61	8.80	0.59
Full-time equivalent GPs per 1,000 patients	0.53	0.30	0.46	0.27	0.53	0.30
GP partners (%)	0.69	0.24	0.68	0.29	0.69	0.24
UK-trained GPs (%)	0.65	0.36	0.53	0.38	0.65	0.36
Age structure of GPs (proportion)						
Age $<$40	0.27	0.22	0.28	0.24	0.27	0.22
Age 40 to 59	0.66	0.25	0.60	0.28	0.66	0.25
Age 60+	0.07	0.16	0.12	0.24	0.07	0.16
Contract type						
GMS	0.75	0.43	0.62	0.49	0.75	0.43
other (incl. PMS)	0.25	0.43	0.38	0.49	0.25	0.43
Patient characteristics						
Age-sex proportions						
Male - 0 to 4	0.02	0.01	0.03	0.01	0.02	0.01
Male - 5 to 19	0.08	0.01	0.09	0.01	0.08	0.01
Male - 20 to 44	0.13	0.02	0.17	0.04	0.13	0.02
Male - 45 to 59	0.11	0.01	0.10	0.01	0.11	0.01
Male - 60 to 74	0.10	0.02	0.08	0.02	0.10	0.02
Male - 75 to 84	0.03	0.01	0.03	0.01	0.03	0.01
Male - 85+	0.01	0.00	0.01	0.00	0.01	0.00
Female - 0 to 4	0.02	0.01	0.03	0.01	0.02	0.01
Female - 5 to 19	0.08	0.01	0.08	0.01	0.08	0.01
Female - 20 to 44	0.13	0.02	0.16	0.03	0.13	0.02
Female - 45 to 59	0.11	0.01	0.10	0.02	0.11	0.01
Female - 60 to 74	0.10	0.02	0.08	0.02	0.10	0.02
Female - 75 to 84	0.04	0.01	0.03	0.01	0.04	0.01
Female - 85+	0.02	0.01	0.01	0.01	0.02	0.01
Prevalence of chronic conditions (%)						
Coronary heart disease	3.78	0.91	3.46	1.06	3.78	0.91
Stroke	2.10	0.55	1.85	0.62	2.10	0.55
Hypertension	15.99	2.82	14.32	3.42	15.99	2.82
Chronic obstructive pulmonary disease	1.81	0.62	1.96	0.84	1.81	0.62
Cancer	2.92	0.77	2.28	0.83	2.92	0.77
Mental health problems	0.64	0.22	0.85	0.35	0.64	0.22
Asthma	6.45	0.96	6.15	1.15	6.45	0.96
Heart failure	0.85	0.34	0.78	0.35	0.85	0.34
Palliative care	0.35	0.37	0.30	0.31	0.35	0.37
Dementia	0.74	0.36	0.71	0.42	0.74	0.36
Atrial fibrillation	2.20	0.59	1.73	0.67	2.20	0.59
Cardiovascular disease (aged 30-74)	1.92	1.10	1.77	1.03	1.92	1.10
Index of Multiple Deprivation (2015)	0.09	0.03	0.14	0.07	0.09	0.03
Practices	988		3,410		3,410	
Practice-quarter observations	30,203		99,910		99,910	

Note: CCG membership is also balanced after weighting but is not reported here. Data based on quarter-practice level 2011Q1 to 2018Q4. Data sources can be found in the Appendix in Table A1.

Table A5: Predictors of observed dispensing status (2SLS first-stage) for the final sample

	Est	SE
Potential miles saved if dispensing (in 1000)	0.003***	0.0002
log(list size)	-0.015	0.010
Full-time equivalent GPs per 1	0.097***	0.018
GP partners (%)	-0.048**	0.015
UK-trained GPs (%)	0.040***	0.012
PMS contract	-0.014	0.011
Age structure of GPs (%)		
Age 40 to 59	0.027	0.017
Age 60+	-0.030	0.022
Demographic composition by age-sex band (%)		
Male - 0 to 4	-0.982	1.066
Male - 5 to 19	-0.161	0.634
Male - 45 to 59	-1.658*	0.664
Male - 60 to 74	10.269***	0.862
Male - 75 to 84	9.638***	1.507
Male - 85+	6.415*	2.690
Female - 0 to 4	-1.214	1.104
Female - 5 to 19	2.301***	0.655
Female - 20 to 44	-0.819	0.505
Female - 45 to 59	1.293*	0.578
Female - 60 to 74	-2.819***	0.833
Female - 75 to 84	-11.168***	1.310
Female - 85+	-10.374***	1.467
Prevalence of chronic conditions or major health shocks (per 1000)		
Coronary heart disease	-0.056***	0.012
Stroke	0.001	0.018
Hypertension	-0.005	0.003
Chronic obstructive pulmonary disease	-0.021*	0.009
Cancer	0.020	0.012
Mental health problems	-0.059***	0.016
Asthma	0.012**	0.005
Heart failure	-0.040*	0.019
Palliative care	0.037**	0.013
Dementia	0.029*	0.014
Atrial fibrillation	0.064***	0.017
Cardiovascular disease (aged 30-74)	0.008**	0.005
Index of Multiple Deprivation (2015)	-0.399	0.142
Practice-quarter observations	130,113	

Table A6: Effect at intensive margin - assumed linear relationship

	Pooled OLS		GP fixed effects		CCG fixed effects	
	Est	SE	Est	SE	Est	SE
Cost per patient	1.674***	0.587	1.482	1.121	1.372***	0.507
Cost per prescription	0.125*	0.065	0.323**	0.152	0.071	0.058
Prescriptions per patient	0.210**	0.094	0.047	0.191	0.240***	0.087
OTC prescriptions per patient	0.037*	0.022	0.029	0.032	0.052***	0.017
Antibiotic prescriptions per patient	0.005	0.004	-0.001	0.008	0.006	0.004
Opioid prescriptions per patient	0.003	0.006	0.011	0.008	0.005	0.006
Antidepressant prescriptions per patient	0.003	0.008	0.002	0.013	-0.005	0.007
Relative pack size	-0.113***	0.018	-0.027	0.033	-0.119***	0.017
Proportion generic prescriptions	-0.020***	0.005	-0.012	0.018	-0.016***	0.005
Practice-quarter observations	30,203		30,203		30,203	

*** p<0.01, ** p<0.05, * p<0.1
Est = Coefficient estimate; SE = Standard error.
All models control for a full set of control variables for characteristics of the patient population and the organisational structure of the practice.