Patient Choice, Payment Systems and Multidimensional Quality

Enrico Camarda* Sebastian Fleitas †

Preliminary and incomplete

January 2022

Abstract

Competition and empowering patient choice have been popular policy tools to incentivize quality in health care markets. Market mechanisms work if consumers follow high quality providers. However, quality can have different dimensions with very little correlation across them. In this paper, we explore the role of quality multi-dimensionality in the market for public General Practitioner services in England. In equilibrium, multi-dimensional quality affects not only patients' allocation and choices, but also practices quality provision choices. In counterfactual simulations we compare how different reforms in payment systems, such as the introduction of risk adjustment or increasing quality rewards, perform in terms of welfare. We find that the introduction of risk-adjustment yields the highest increases in consumer welfare for a unit increase in government expenditure. This happens because of mechanisms that cannot materialize in absence of multi-dimensional quality.

JEL-code: L15, L44, L51, I11, I18
Keywords: Regulation, Health Economics, Quality of care, Supply-side incentives

*KU Leuven and FWO. Email: enricomaria.camarda@kuleuven.be
†KU Leuven and CEPR. Email: sebastian.fleitas@kuleuven.be

1 Introduction

In many health care systems patients' choice and managed competition have played an important role in improving the allocation of resources and welfare. Recent studies (Chandra, et al., 2016, Gaynor et al., 2016 and Santos et al., 2017) show that consumers have a preference for quality, creating incentives for providers to compete in quality. However, quality is "multi-dimensional" and the regulator and consumers may care about different types of qualities. For example, while consumers may care more about patient experience, the regulator may care more about the quality captured by some specific clinical process measures. Concerns about the performance of different payment systems in equilibrium may arise if the correlation between the different components of quality is low. In fact, in that case, the impact of payment regimes is affected by the presence of multidimensional quality. This paper analyzes the equilibrium welfare effects of different payment systems under the presence of multi-dimensional quality. We find that payments that directly reward process measures may not significantly change reallocation and might not be efficient, if consumers do not have strong preferences for this type of quality. The introduction of risk-adjustment, instead, or paying more for more complex patients, can have different welfare implications compared to a scenario where quality is uni-dimensional.

We focus on the market for General Practitioners ("GP") services, also known as family doctors, in the English National Health Service. GP practices are for profit organizations and are the sole providers of primary care for patients within the National Health Service ("NHS"). Virtually all primary care services in the United Kingdom are covered by the NHS, the socialized health care system prevalent in the U.K..[1] All the payments to the practices are from the regulator (NHS), which uses a combination of capitation and a quality rewarding system called Quality of Outcomes Framework (QOF). The English setting is characterized by free patient choice of general practices (with few restrictions) and the consequent competition among practices to attract patients. Patients have to enroll with GP practices to receive primary health care and referral to hospital and specialists. Crucially, patients cannot rely on experts in their choice of practices as they can when they choose their provider of hospital care (in that case, the expert is the general practitioner himself).

In this paper, we analyze the equilibrium effects of three different payments systems on patients' welfare and on practice quality decisions in the context of the

[1] Research by the King's Fund into private health care counted that around 97% of all GP consultations happen within the NHS.

English general practitioners. We study i) increases to the capitation payments, the yearly amount set and paid by the NHS for each patient enrolled with a practice; ii) changes to the capitation payments to better reflect patients needs based on their medical conditions, so called "risk-adjustment"; iii) changes to the QOF system where specific yearly payment are made by the NHS to practices based on their performance in some metrics (process measures) capturing the quality of care of patients with chronic conditions.

In our study we make an important distinction: patients with no chronic conditions (hereafter, "healthy patients") and patients with different chronic conditions. Our focus on patients with conditions is important because there has been a shift in their care towards the community and family doctors. This is due to the fact that patients with chronic conditions use up the most of the expenditure in health care systems. Providing good GP and preventive care in the management of these conditions can have a very positive effect on government expenditure as well as on their health (reducing complications and hospital admissions).[2] Then, preferences of patients become crucial in determining an allocation of patients to physicians that would maximize the benefits of preventive care. Healthy and chronic conditions patients care about different dimensions of quality. Healthy patients care about patient experience (or ability to manage the health of all patients). For chronic patients, instead, quality is both general patient experience as well as their ability to manage patients with certain conditions, captured by specific process measures.

These dimensions of quality are captured by metrics and data that we use in our analysis. For patient experience, we use the results from surveys of GP patients that are published on a website set up by the NHS to help patient choices. In particular, we use a measure that indicates the percentage of patients of a practice who would recommend such practice to their new neighbors. This measure has been used prominently by the NHS on the website.[3] For condition specific quality we use process measures from the reward payment system (QOF). QOF rewards the achievements of each practice along some metrics suggested by an independent body of medical experts for each of the conditions. One example is the percentage in each practice of patients with Chronic Obstructive Pulmonary Disease performing a spirometry in the last twelve months.

We start the analysis by documenting the low correlation between the different dimensions of quality that are reported and used in QOF and the patient

[2] In England, they use the majority of the resources. Roughly 70% of all hospital bed days for just approximately 20% of patients.

[3] Santos et al. (2017) also use this measure in part of their analysis.

experience quality indicator. We observe that doctors abilities may vary significantly across these two different dimensions and that there is little correlation between them. This makes non-trivial the choice of chronic patients: practices may be characterized by good patient experience, but not manage well patients with certain conditions. This highlights potential issues of allocation of chronic patients depending on their preferences for the two dimensions.

Second, we exploit a reform in the available choice sets of patients to better understand patients preferences for the different dimensions of quality. We analyze two choice reforms (in 2012 and 2015) that allowed patients to choose GP practices farther away from their place of residence (and not just among practices close to home as previously devised). If patients with conditions preferred more one quality dimension to the other we would observe that practices with the preferred quality would attract more patients post-reform. We adopt a difference-in-differences strategy to uncover the impact of this reform and we find interesting evidence pointing to the fact patients with a condition may follow more patient experience.

Third, we develop and estimate a partial equilibrium model of supply and demand for GP services. This model allows us to confirm the results of the reduced form analysis of the reform and recover the structural parameters that guide consumers and firm behavior. On the demand side, we develop and estimate a model of demand similar to the one in Ellickson et al. (2020) and Holmes (2011). We estimate the preference parameters of patients for the different quality dimensions as well as for distance and other practice characteristics. As expected, we find that patients are not willing to travel far away from home. Additionally, the estimation results confirm the evidence from the reforms: patients with chronic conditions have a stronger preference for patient experience than for the process measures relevant for their condition. On the supply side, instead, we endogeneize doctors' decisions about quality which determine the attractiveness of practices to patients and indirectly their level of revenues. Given some assumptions about practice choices and objective functions we back out costs parameters from the first order condition of the practices (that do not set prices).

Finally, leveraging on the model of the market we developed, we simulate changes to the payment system. We highlight how patients' preferences and different price regimes lead doctors to modify their quality provision in different ways. In particular, we look at the effects of an increase in the uniform capitation price, the introduction of risk adjustment in the capitation system (higher payments for chronic patients) and an increase in rewards for achievements in

condition specific process measures. Finally, we perform an exercise where we explore the role of information, by equalizing the preference parameters for patient experience and process measures for chronic patients. The rationale is to understand how the different payments regimes would perform if chronic patients cared more about process measures, for example because they are better informed. The size of the different reforms we analyze are in line with changes discussed and implemented by the NHS in the past. While the potential reforms are realistic, they are characterized by largely different changes in government expenditure. We use the change in consumer surplus per additional pound spent by the government as a metric to compare the different reforms.

Summary of counterfactual results. Here we present a small summary of the main results. First, we look at the case of an increase in uniform capitation price. We find that doctors would improve patient experience to attract more patients. However, this move would attract both healthy patients and chronic patients for which practices have lower or even negative margins. Practices may not increase/may decrease condition specific quality in an attempt to select away chronic patients because they cannot reject them, according to law. Second, we analyze the introduction of risk adjustment in the capitation system. In this case we find that practices would increase both patient experience and condition specific quality (at least for some chronic patients). This would benefit all patients, even if the government would only increase expenditure for chronic patients. Finally, in the case of an increase in rewards for condition specific quality we find that this reform is the most cost effective to make practices increase specific quality, but it would not affect significantly patient experience. We end with a discussion of the welfare effects of these different measures considering that patients preferences may not be in line with socially optimal quality provision.

Finally, as a final step in our analysis, we study the potential role of information. One of the reasons why chronic patients may care less about process measures is that they do not have easy access to this information. To analyze what would happen if they had better access, we equalize the preferences parameters of chronic patients for patient experience and their condition specific process measures. The aim is also to understand the welfare effects in the different payment scenarios when patients preferences are more aligned with an hypothetical social planner that care more about process measures. We find that the results are not very different. For example, higher uniform prices still present an incentive to select away less profitable patients by reducing process measures. Interestingly, risk-adjustment leads to a relatively higher increase in process measures than patient experience compared to the previous scenarios where the parameters are

not equalized. For this reason, risk-adjustment benefits less the other patients without conditions and leads to a lower welfare increase in comparison.

Relation to the literature. The main contribution of this paper is to highlight the importance and the role of quality multi-dimensionality especially in how it can affect the welfare impact of different payment regimes. First, we find that patient choice can have negative effects on patient allocation in presence of multi-dimensional quality. This is the case when there is little correlation between different quality components and when what patients prefer may not be socially optimal. We add to the analysis of Gaynor et al. (2016), Santos et al. (2017) and Gravelle et al. (2019) about patient preference for quality and the impact of competition on quality. Second, we provide a comparison of the welfare effects of different payment systems in presence of multi-dimensional quality. We show how competition and price regulation may lead to different welfare outcomes in our context and how the choice of price regulation is important in increasing quality dimensions and welfare. In the study of this relationship between price regulation and quality provision, our paper adds to the work of Hackmann (2019), Camarda (2022) as well as Kolstad et al. (2021). Our findings are particularly important for the care of chronic patients who receive little guidance in their choice of GP even if their choices may have important consequences for both the patients and the financial sustainability of the entire health care system.

A second contribution is to the modelling of GP and doctors supply choices. Primary care decisions are under-explored and we provide a framework to think about doctors trade-offs in the context of a publicly funded capitation system. This is an addition to some emerging work on privately insured primary care settings as in Shurtz et al. (2019) and Kolstad et al. (2021). Additionally, our analysis contributes to the study of the quality choices of doctors in the care of patients with different levels of needs, for example in Chan (2018) and Chan and Gruber (2020).

The paper is structured as follows. Section 2 and 3 provide a description of the institutional details of the market and the data. Section 4 presents the analysis and the evidence of the choice reform of 2012 and 2015. Section 5 develops a model of demand and supply and presents the results of the estimation of these models. Section 6 performs various policy counterfactuals showing the demand and supply responses to these policy changes. A final sub-section explores the role of information in the welfare effects of different price regimes by simulating equilibrium outcomes in presence of better information availability. Section 7 concludes.

2 Health care in the United Kingdom

In the UK, the NHS is responsible for around 90% of healthcare expenditure in curative and rehabilitative care in both hospital and primary care.[4] The government funds 99.9% of this expenditure and regulates the functioning of the NHS. The government has an interest in providing good and affordable care, as voters react to both taxes and the quality of care received. The two other important players in the GP market are the GP practices that provide the services and the patients who register with one of these practice to receive treatment.

2.1 Main players in the family doctor setting: GP practices and patients

In the NHS healthcare services are funded through tax revenues and social security contributions and they are free at the point of delivery, meaning no out-of-pocket expenses. This is true for both secondary care (e.g. hospitals) and primary care (General Practitioners or "GPs"). While hospitals are NHS bodies, General Practices are usually organized in the form of private limited liability partnerships that enter into contracts with the NHS. Even if the amount of money spent in general practice is dwarfed by the amount spent on hospital care[5] the role of GPs is crucial. Firstly, GPs are the gatekeepers of the system, in the sense that access to specialized and hospital non-emergency care can only be granted through a GP referral, a way to control expenditure even if there are no out-of-pocket expenses. Secondly, proper primary care has the potential of reducing the need for (expensive) hospital care, because primary care plays an important role in preventive care and in the management of long term conditions, e.g. diabetes and asthma. Today there are approximately 7,000 practices with an average of 4.2 doctors per practice and 8,500 patients on average.

Patients register with only one practice where they receive care. Patients choose their GP using online platforms like the website nhs.uk (former nhschoices) where they can see which practices are closer to their address as well as additional information from patient surveys and information about the language spoken by the doctors, their gender, the additional services offered, the opening hours and some performance indicator from QOF (this information has not been consist-

[4] https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthcaresystem/bulletins/ukhealthaccounts/2017

[5] Hospital care expenditure is more than 5 times greater than GP expenditure
https://www.ons.gov.uk/peoplepopulationandcommunity/healthandsocialcare/healthcaresystem/bulletins/ukhealthaccounts/2017

ently available). Patients cannot be refused registration based on their gender, religion, age or medical condition, except if they live too far from the practice, in which case they may be denied registration. Santos et al. (2017) have shown that patients take in consideration distance from home, but also the reported patient experience of the practices when choosing a practice.

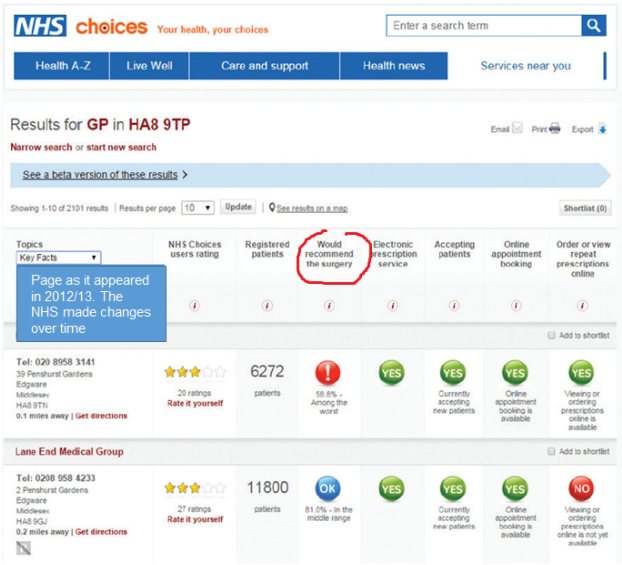


Figure 1: NHS choices search result page example

Of particular interest for our study are patients with long term conditions. In 2017 they were approximately 26% of all patients, but they are responsible for about 50% of all GP consultations and use up more than 70% of inpatient bed days.[6] Interestingly patients with multiple (more than 2) long term conditions are on the rise, but they represented 18% of all patients with long term conditions and only 4.9% of the total. Instead, patients with only one condition were around 63% of patients with long term conditions and patients with only two conditions were around 20% of this group.[7]

2.2 GP practices revenues

The NHS has two main form of contracting with General Practices: i) more than half practices use the General Medical Services ("GMS") contract, which is set nationally, ii) less than half use the Primary Medical Services ("PMS") contract that is negotiated with the local health authority called Clinical Commissioning

[6] $https://www.kingsfund.org.uk/projects/time-think-differently$ on long-term conditions and multi-morbidity
[7] $https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216528/dh_134486.pdf$

Group ("CCG"). The two contracts are very similar and a higher number of patients leads to higher revenues. In particular, around 75% of revenues is directly linked to number of patients, with around 60% that comes from a capitation system where practices receive a fixed amount of money per year per patient, where the amount varies with age, gender and morbidity. Around 15% of the revenues instead comes from the Quality and Outcomes Framework ("QOF") where practice receive bonuses for achieving specific objectives. The reminder of the revenues comes from transfers for rents and mortgage payments as well as additional services that the practices may offer both through a contract with the CCG or directly to patients.

Capitation system. The capitation system consists in paying practices a certain amount of money (£65 per year in 2011/12) for each patient registered with the practice. The number of patients, however, is weighted by the expected number of consultations per category of patients. For example, a male patient aged over 85 has a weight more than 6 times higher than a male patient aged 15-44. These weights cover four main categories: 1) age and gender, 2) morbidities and mortality, 3) being a patient living at a nursing home, 4) being a newly registered patient (they typically go more often) and 5) rurality of the practice. While 1), 3) and 4) depend on the actual patients registered with a practice 2) and 5) are based on the location of the practice irrespective of the actual patients a practice may receive. In particular, for 2) the practice would weight all patients by a higher or lower factor depending on the characteristics of the geographical area of where a practice is (electoral wards, smaller than the CCG area).

Importantly, the weights are standardized and scaled down to reflect the relative difference in patients composition across practices. So, even if a practice has more patients over 85, the weight for age and gender the weight varies only little, because other practices also have patients over 85. In this sense the weights are only reflecting the relative prevalent age of one practice with respect to the average within the health authority. The same reasoning applies to all other categories mentioned before. In practice, these weights vary very little and also mostly across geographical areas and we do not use them in our subsequent counterfactual analysis, as we focus on one specific city.

QOF. In our study we will rely on several quality measures that come from QOF. QOF is a voluntary program for practices (with a more than 98% coverage) that has been created to incentivize the good management of patients with long term conditions and, to a lesser extent, reducing obesity, smoking or improve

management practices. Currently, QOF allows practices to receive up to 559 QOF points, each valued at the base value of £187.74. QOF is composed by three "domains" or indicator categories: i) clinical which covers 77% of the points, ii) public health which covers 17% of the points and iii) public health – additional services that covers 6% of the points. The clinical domain is composed by quality indicators for 19 chronic conditions, for example asthma, diabetes, chronic kidney disease and dementia.

As anticipated, the inclusion of these indicators is meant to incentivize the good care of patients with these conditions (See Appendix for detailed calculations). In particular, these indicators have been designed by the National Institute for Health and Care Excellence that uses the best experts in the field to guide the adoption of new treatments, best practices and medicines in the NHS. The idea is to create a battery of targets that reflect the best practice in the management of patients with these conditions. To better understand what "management" means, it is important to know that while patients with these conditions may need a specialist for certain extraordinary interventions they require continuous care by medical professionals to keep their condition under control and this task has been given to the GPs. Such task consists of making sure that patients receive the checks[8] and the medications they need and to avoid expensive complications that may require hospitalization. Finally, the other two QOF domains include a smaller number of indicators that cover recording the number of smokers and obese patients, prevention of cardiovascular diseases, prescription of contraceptives and cervical screening.[9]

3 Data

The data used in this paper come from QOF and from GP patients surveys. We combine and match these two sources and few others using practice unique identifiers provided by the NHS.

The data in the QOF series includes the number of patients enrolled in each practice, by gender, age and condition and it is collected and published on a

[8]For example, for a patient with diabetes this may be checking the long run sugar level or checking the state of the feet and hands. These checks are not urgent, but need to be performed once or twice a year, and cannot be done by the patients. They are very important, so that, for example, they may prevent amputation or kidney failure.

[9]Other domains regarding organization, management and adoption of IT system have been retired in 2013/14 and 2014/15.

	2010/11	2011/12	2012/13
No. Practices			
Total	8,245	8,123	8,020
No. of patients per practice			
Average	6,691	6,835	6,984
Median	5,871	5,987	6,142
Standard deviation	4,210	4,274	4,311
Patient experience (% recommend)			
Average	81.9	80.7	79.0
Median	84	82.8	80.9
Standard deviation	10.7	11.2	11.7
QOF (COPD)			
Average	91.2	91.5	90.9
Median	92.9	92.9	92.5
Standard deviation	10.7	9.3	9.4
QOF (Hearth failure)			
Average	90.3	90.3	90.8
Median	90.9	90.9	91.3
Standard deviation	12.1	11.2	10.5
QOF (Mental health)			
Average	89.4	88.5	87.7
Median	92.6	91.2	90.0
Standard deviation	11.8	10.9	10.6

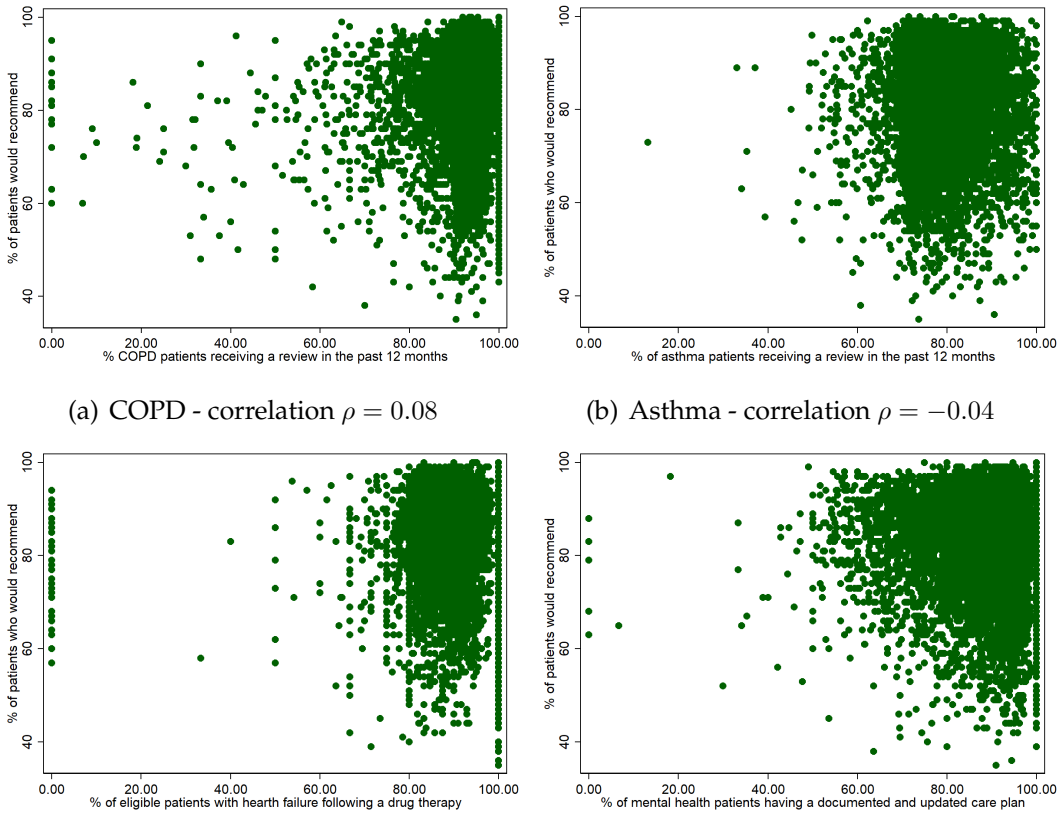
Table 1: Summary statistics

yearly basis. For example, it is possible to know how many patients are diagnozed with asthma in each practice in each year. From the data at the practice level it is also possible to construct prevalence rates of different chronic diseases at different levels of geography, assuming all patients with a certain condition in that geography are diagnozed correctly.

Additionally, the QOF series includes the levels of levels of achievements of the practices with respect to the different targets set by the QOF system. For each practice, as an example, it is possible to know how many patients with COPD (Chronic Obstructive Pulmonary Disease) received a spirometry test in the previous 12 months. The level of achievements vary across practices (see Appendix A for discussion).

The GP survey series contains answers from the patients who participated to such surveys. The data is available at the practice level on a yearly basis. There are questions about different aspects of the GP experience, from easiness to get secretaries on the phone to the speed at which it is possible to have an appointment. The particular information we use is the percentage of respondents who would recommend the practice to someone who moved in the neighborhood. This information is used as a measure of the practice patient experience.

We emphasize that the correlation between these two measures is low. This is exemplified in the graph below that shows a low correlation between patient experience and condition specific quality for Chronic Obstructive Pulmonary Disease ("COPD"). This rises concerns as patient need to choose which quality to prioritize and their priority may not be the socially optimal one.



(a) COPD - correlation $\rho = 0.08$

(b) Asthma - correlation $\rho = -0.04$

(c) Heart failure - correlation $\rho = -0.05$

(d) Mental health - correlation $\rho = -0.03$

Figure 2: Correlations between patient experience and process measures

4 Evidence from choice reforms

To collect evidence about patients preferences with respect to general patient experience and condition specific quality we exploited two reforms in GP choices that have implemented in the NHS in the last years.

Historically since the establishment of the NHS in 1948 patients could registered with a GP practice if they lived in the so-called catchment area of that practice. Each practice defines an area around the practice and only patients living in this area can register. In 2012 this paradigm changed as the government asked the practices to provide outer boundaries of the catchment area. People living outside the original boundaries but inside the outer boundaries had the option to apply for the registration with the practice in question. If the GPs at the practice found that there was no clinical reason to reject the application they had to register the patients. In 2015 the same process was extended to patients living anywhere in England.

The change in the choice set of patients allows us to try to identify patients preferences for patient experience and condition specific quality. If after the reform more patients with conditions were moving to practices with higher patient experience than places with higher condition specific process measures we could infer the relative importance of these types of quality. For each condition (k) we specify a diff-in-diff type strategy:

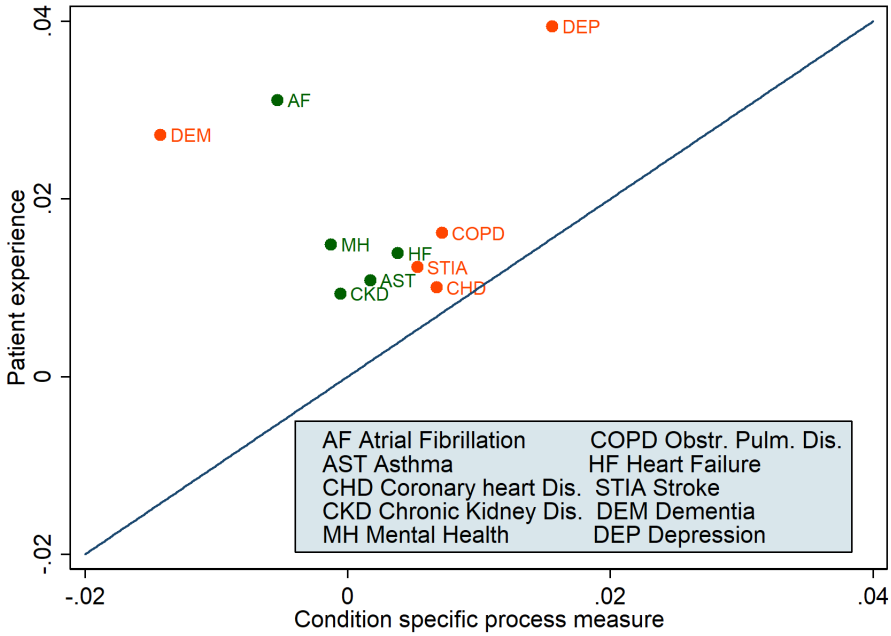
$$Register^a_{jt} = \gamma^a_0 z_{j2011} * Reform_{2015,t} + \gamma^a_1 z^a_{j2011} * Reform_{2015,t} + \tau^a_t + \eta^a_j + \epsilon^a_{jt}$$

Where we have: quality z, practice j, time t, time-area fixed effects τ^k_t. We restricted our analysis to cardiac conditions, respiratory conditions and kidney-related conditions.

In the graph below the points have coordinates: $\frac{\sigma^{z^a}\gamma^a_1}{\overline{Register_a}}, \frac{\sigma^z\gamma^a_0}{\overline{Register_a}}$, where σ stands for standard deviation and $\overline{Register_a}$ for average list size for patients with condition a (σ^{z^a} is a one standard deviation increase in z^a and σ^z is a one standard deviation increase in z). The results of our analysis show that with both reforms patients with different conditions seem to show preferences for patient experience, but smaller or non-significant preferences for condition specific quality. However, we want to highlight some caveats. The effect of the reform does not really capture preferences of patients with a certain condition, but rather an indic-

ation of such preferences . The coefficients γ_a^0, γ_a^1 capture the net flows of patients across the different practices, rather than preference parameters. Additionally, the reform would only show the effect of differences in quality for patients willing to travel farther away from home. Finally, practices could use a certain discretion at rejecting patients for clinical reasons and this could also interfere in the interpretation of the results. For example, practices with higher condition specific quality may prefer to focus more on existing patients rather than taking on additional ones. For these reasons, the evidence from the reforms is not conclusive about patients preferences and we adopted and estimated a structural model of demand for GP practices.

General v. condition specific quality post 2015 reform



Notes: All effects of patient experience are significant, non significant effects of process measures in green. Results from 2012 reform are very similar to the ones displayed.

5 Model and estimation

5.1 Demand

In this section we describe the model and the strategy we adopted to estimate the primitives of demand. Consider a geographic market n, in our case the equivalent

of a US Census tract (called Lower Layer Super Output Areas), where patients residing in the tract n choose the GP practice that maximizes their individual utility (patient i in the healthy patients group h and patient p with condition $a \in K$, where K is the set of conditions) :

$$U^h_{ijt} = \delta^h_{ijt} + \epsilon^h_{ijt} \text{ where } \delta^h_{ijt} = \alpha^h_0 log(z_{jt}) + \beta^h dist_{ij} + X^c_{jt}\beta^h_c \quad (1)$$

$$U^a_{pjt} = \delta^a_{pjt} + \epsilon^a_{pjt} \text{ where } \delta^a_{pjt} = \alpha^a_0 log(z_{jt}) + \alpha^a_1 log(z^a_{jt}) + \beta^a dist_{pj} + X^c_{jt}\beta^a_c \quad (2)$$

In the equations above z_{jt} is the patient experience of practice j at time t, z^a_{jt} is the condition specific process measure provided at practice j at time t, $dist_{jn}$ is the distance between practice j and the population-weighted centroid of Census tract n and X^c_{jt} indicates observable characteristics that may affect patients' utility. Idiosyncratic errors $\epsilon^h_{ijt}, \epsilon^a_{pjt}$ are i.i.d. following an EVT1 distribution. In our model healthy patients do not have utility from condition specific process measures and patients with conditions only care about patient experience and the quality related to their condition, not quality related to other conditions.

If we had data on practice market shares for people living in each Census tract we would use a Berry logit model (Berry (1994)) to retrieve the parameters of interest. Unfortunately, we do not have this data, but we do have data on the number of people living in each tract and on the number of patients going to each practice j, disaggregated at the condition level a (e.g., for patients with asthma). We can use this information to construct the equation below (as Elickson et al. (2020) and Holmes (2011)). Then, GP practice demand for all types of patients $k = h, a$ (both healthy h and with condition a), aggregated from demand in Census tract n:

$$q^k_{jt} = \sum_{n \in M_j} \psi^k pop_{nt} s^k_{njt} + \eta^k_{jt} \quad where \quad s^k_{njt} = \frac{e^{\delta^k_{njt}}}{1 + \sum_{u \in J_n} e^{\delta^k_{nut}}}$$

In the equation we have the following additional elements. J_n is the set of practices in the choice set of individuals living in Census tract n (3km radius),[10]. M_j is the set of Census tracts included in the catchment area of practice j, this

[10]We perform robustness checks with 5km and 10km radiuses.

set is also defined using a 3km radius around the practice based on the the average size of catchment areas.[11] ψ_k is instead a regional prevalence rate for patient type k which is used as an approximation for the actual tract-specific prevalence rate. pop_{nt} is the population of tract n at time t. η_{jt}^k captures firm specific error term which can be interpreted as measurement error or as an unexpected demand shock.[12] Finally, we can write the formula for the market share of practice j in tract n in the way spelled out above based on Berry (1994).

	Dependent variable: number of patients per category					
	Patient exper.	Distance	Distance^2	Process measure	# GPs	GP exper.
Healthy	0.47	-0.61	-0.17		0.06	0.00
	(0.03)	(0.27)	(0.10)		(0.00)	(0.00)
Respiratory group	0.79	-1.14	0.03	0.04	0.05	0.00
	(0.04)	(0.25)	(0.08)	(0.02)	(0.00)	(0.00)
Cardiac group	0.44	-1.02	-0.06	0.41	0.05	0.00
	(0.06)	(0.27)	(0.10)	(0.06)	(0.00)	(0.00)
Kidney Disease	0.66	-0.45	-0.34	0.15	0.06	0.00
	(0.05)	(0.43)	(0.16)	(0.03)	(0.00)	(0.00)
Mental Health Group	0.85	-1.91	0.25	0.09	0.06	0.00
	(0.05)	(0.34)	(0.12)	(0.02)	(0.00)	(0.00)
Cancer	0.62	-0.32	-0.33	0.17	0.04	0.00
	(0.04)	(0.30)	(0.11)	(0.03)	(0.00)	(0.00)

Note: standard errors in parentheses. All English cities >250k inhabitants in 2010-2012.
"GP experience" is average GPs' time in the practice (in months), "#GPs" is no. of GPs in the practice

Table 2: Estimates from demand model 5km radius

Now we can use the formula in (3) instead of the usual Berry logit formula to retrieve the parameters of interest. In particular, we use non-linear least squares to estimate equation (3) minimizing the difference between the observed q_{jt}^k and the predicted equivalent. As outside option we considered that 2.5% of healthy patients and 1% of patients with conditions are not registered with a GP.[13] The data used for this estimation is from all large English cities above 250,000 inhabitants in the period 2010-2012, this covers a large part of the English population and allows to consider a smaller radius around the practices (and therefore a more numerically tractable amount of Census tracts).

Some of the practices may be at the outskirts of the city and in our model they may be able to collect patients only from a smaller number of tracts even if in

[11] We actually differentiate between urban catchment areas and rural catchment areas.
[12] This shock will be kept constant in the subsequent counterfactual analysis.
[13] For example, patients with conditions may change address after the diagnose and neglect their health.

reality would be able to attract patients from outside the city. For this reason, we excluded those practices that have only a small number of tracts in their catchment area.[14]

As shown in Table 1 and Table 2, patients have significant, but low preferences for quality. Additionally, patients with chronic conditions value more patient experience than healthy patients, but they value less condition specific quality than patient experience.

		Healthy	Resp.	Cardiac	Kidney	MH	Cancer
Willingness to travel	Patient experience	44.5	68.5	35.8	50.7	59.3	54.6
	Process measure	-	5.7	21.5	19.9	16.4	13.3

Note: Willingness to travel for 1 std. dev. increase in quality measure (from the average), based on specification w/out $(distance)^2$

Table 3: Willingness to travel in response to quality (meters)

5.2 Supply

We formulate the following revenue function for practice j in a market shared with other practices, indicated with $-j$:

$$Revenues_j = \underbrace{\bar{p}q_j^h(z_j^h, z_{-j}^h)}_{\text{Healthy patients}} + \sum_{a=1}^{K} \underbrace{\bar{p}q_j^a(z_j^h, z_{-j}^h, z^a, z_{-j}^a)}_{\text{Patients specific condition } a} + \sum_{a=1}^{K} \underbrace{QOF(z_j^a)}_{\text{Rewards spec. quality}} + \underbrace{T}_{\text{Lump-sum}}$$

As discussed earlier, the main sources are i) revenues directly linked to the number of patients (healthy and chronic ones), ii) QOF payments and iii) lump-sum transfers (for premises and seniority). The three together cover around 97% of revenues. The $\bar{p_j}$'s are the prices from the capitated system and they are responsible for the majority of the revenue per patient (approximately 65%). QOF payments, instead, represent around 25% of the revenues. In our model, we drop the subscript j from $\bar{p_j}$ because price do not vary based on the conditions of patients for practices within the same area.[15]

We model the costs of the practices in the following way:

[14] We performed robustness check around the chose threshold.

[15] Only small variations in price based on conditions prevalence exist across different geographical areas. Other sources in variation are given by rurality and age and gender of patients. We consider only practices from a city area and the effect of demographics alone is small and we are not modeling it in our analysis.

$$C_{jt} = \mu_j^{qh}(q_{jt}^h(z_{jt}^h))^3 + \sum_{a=1}^{K} \mu_j^{qa}(q_{jt}^a(z_{jt}^h, z_{jt}^a))^3 + \mu^{zh}(z_{jt}^h)^3 + \sum_{a=1}^{K} \mu^{za}(z_{jt}^a)^3 - F$$

We do not explicitly model a production function in these two qualities and number of patients (quantity). However, we assume quality is a function of time spent with a specific category of patients. In particular, practices face a trade off between spending time with one category of patient or another categories of patients, as well as between patient time and idle time at work that reduces their work-related stress. This is in line also with Gaynor and Gertler (1995) and Shurtz et al. (2020). This translates into the fact that as the number of patients per doctor increases, the more difficult it becomes to increase quality for all types of patients as the time available shrinks. We model this aspect of the cost function of practices introducing convexity in number of patients (quantity) and quality provided.[16]

We have several indications from decisions taken by the NHS that support our view on cost and quality. For example, higher amount of QOF money is given for more time consuming quality targets (and related condition specific quality) or to practices with higher levels of disease prevalence (and therefore higher workload).[17] Another example indicating that quality is costly is that in few districts the local health authorities have suspended QOF to face a spike in flu during the winter.[18] Interestingly, this would indicate that financial incentives are important, as practices have to choose to spend their time between QOF related activity and non-QOF related activites. Then, it is not surprising that, answering to a survey by the British Medical Association,[19] around 25% of doctors answered that the best way to reduce the workload is to scrap the QOF system.

Finally, differences of quality across practices is due both to different market conditions affecting marginal revenues and differences in marginal costs. Practices are more or less able to achieve a certain level of quality and may be better at a certain specific condition. This would be reflected in differences in the slope of the marginal cost curve as well as in their intercept. This view is confirm by the

[16] The time constraint could be relaxed if the practice was hiring more doctors, but we abstract away from this decision as this is costly for practices to hire new doctors and high levels of practice exits from the market in recent years indicate that practices are typically under-funded to cover these additional fixed costs.

[17] Additionally, we found several other indications in this sense. For example, this trade-off is recognized by the documents of the negotiations between the association representing the GPs (the British Medical Association General Practitioners Committee BMA GPC) and NHS England.

[18] https://www.gponline.com/qof-suspended-across-leeds-ease-incredible-strain-gp-practices/article/1422925

[19] https://www.slideshare.net/citiustech/quality-outcomes-framework-qof-81647848

general guidance given by the NHS in using QOF points, they do not produce rankings of practices based on the totality of QOF points, because the heterogeneity in achievements needs to be taken into account.

5.3 Marginal costs

We adopted an OLS strategy to recover the marginal costs and the marginal cost of quality (where K is the set of chronic conditions). To this purpose we use the first order conditions of practices objective function, given that we know prices and estimates of marginal residual demand (where $\omega^h_{jt}, \omega^a_{jt}$ are measurement errors in marginal costs orthogonal to the rest).[20]

$$\bar{p}\frac{\partial q^h_j(z^h_j)}{\partial z^h_j} + \sum_{a=1}^{K} \bar{p}\frac{\partial q^a_j(z^h_j, z^a_j)}{\partial z^h_j} = \frac{\partial C_j}{\partial z^h_j} + \frac{\partial C_j}{\partial q^h_j}\frac{\partial q^h_j(z^h_j)}{\partial z^h_j} + \sum_{a=1}^{K} \frac{\partial C_j}{\partial q^a_j}\frac{\partial q^a_j(z^h_j, z^a_j)}{\partial z^h_j} + \omega^h_{jt} \tag{3}$$

$$\bar{p}\frac{\partial q^a_j(z^h_j, z^a_j)}{\partial z^a_j} + \frac{\partial QOF(z^a_j)}{\partial z^a_j} = \frac{\partial C_j}{\partial z^a_j} + \frac{\partial C_j}{\partial q^a_j}\frac{\partial q^a_j(z^h_j, z^a_j)}{\partial z^a_j} + \omega^a_{jt} \qquad \forall a \tag{4}$$

We back out the parameters of the cost function making assumptions about the parametrization of the marginal costs:

$$\frac{\partial C_j}{\partial q^a_j} = 3\mu^{qa}_j (q^a_{jt})^2, \frac{\partial C_j}{\partial q^h_j} = 3\mu^{qh}_j (q^h_{jt})^2, \frac{\partial C_j}{\partial z^h_j} = 3\mu^{zh}_j (z^h_{jt})^2, \frac{\partial C_j}{\partial z^a_j} = 3\mu^{za} (z^a_{jt})^2 \qquad \forall a$$

We assume that marginal costs are convex in both quantity and quality. The assumption about convexity in quality comes from the fact that it is increasingly difficult to achieve higher levels of quality. Additionally the strong convexity in both dimensions is a reduced form approach to model the presence of capacity constraints. As quality increases and more patient enroll, the time available to doctors and nurses decreases considerably making it hard to, for example, improve patient experience. Both the time spent with patients and other components of patient experience, e.g. unobserved waiting time or easiness of booking an appointment, are affected by capacity constraints.

The marginal cost of quality for patients with conditions is based on the marginal QOF which is determined by the NHS considering the different levels of workload required to reach a specific QOF target. However, this type of information is not available for the marginal cost of quality for healthy patients (patient experience). Therefore, the marginal cost of quality for healthy patients has been normalized to the average marginal cost for the different conditions times

[20] Proving the existence and uniqueness of equilibria in the different markets in this context is beyond the scope of this paper.

a factor capturing the ratio of healthy patients over conditions specific patients. The reason for this assumption is that the cost of quality is proportional to the number of patients treated: an increase in quality for a category of patients would affect all those patients.

In the regression analysis we first retrieved the marginal costs for the patients with conditions and then we used equation (3) to retrieve the marginal costs for healthy patients.

		Mean	90 perc.	75 perc.	50 perc.	25 perc.	10 perc.
Panel A: Healthy patients	Marginal cost	58.3	64.7	63.7	62.5	60.7	45.6
	MC of quality	1291	1470	1385	1282	1206	1140
Panel B: Respiratory group	Marginal cost	70.3	64.6	64.4	64.3	64.1	63.7
	MC of quality	2.2	2.6	2.4	2.3	2.0	1.7
Panel C: Cardiac group	Marginal cost	62.5	68.4	65.4	62.7	59.7	55.4
	MC of quality	123.6	136.3	132.2	125.9	116.6	109.0
Panel D: Mental Health group	Marginal cost	57.4	99.6	54.1	50.0	38.0	23.7
	MC of quality	130.2	201.4	157.6	126.0	105.5	82.2

Note: Estimating marginal cost regressing FOC conditions and MC of quality based on QOF achievements
MC of quality are for an increase of 1 standard deviation in quality

Table 4: Estimates of the Marginal Costs and Marginal Costs of Quality

From Table 3 note that marginal costs of patients varies across different types of patients. Additionally, some practices have negative margins $\bar{p} - \frac{\partial C}{\partial q_j^a}$ ($\bar{p} = 65$) when in equation (4) the marginal revenues for the QOF are larger than $\frac{\partial C}{\partial z_j^a}$. Typically, marginal costs for healthy patients is lower than for patients with conditions.

6 Counterfactuals

In this section we cover three different counterfactuals where we modify the payments to the practices. The main objective is to compare the effects on quality and welfare of alternative payment system to assess which one is better for the different categories of patients and for total welfare. The counterfactuals are: an increase in the capitation price, an increase in price only for patients with conditions (risk-adjustment) and an increase in the points of QOF for each condition. In the appendix we also explored a combination of risk-adjustment and the possibility of practices to reject patients.

We assume competition is Nash-in-quality and we simulate what would hap-

pen under the different scenarios. For this analysis we considered 4 groups: healthy patients (79.7%) and patients with respiratory, cardiac and mental health conditions (20.3%). As anticipated, the analysis is limited to the city of Bristol for the years 2010-2012 to reduce the computational burden of the counterfactuals. Additionally, Bristol with a population of 0.4 million people and 60 practices is representative of large cities in England.

The changes in prices are based on what could be reasonable changes implemented by the NHS. Given the different number of patients directly involved in each change to the payment system, the amount of government expenditure is different in each simulation. To draw a comparison between the effects of the different proposed reforms we consider their impact on quality as well as their relative efficiency in terms of $\frac{\Delta \text{Consumer Surplus}}{\Delta \text{Government Expenditure}}$.

Higher uniform price. In the first counterfactual we look at what happens when the price paid with the capitation system is increased by 10% uniformly for all types of patients. To understand the mechanisms at play consider the first order conditions (3) and (4) above. The revenues from the capitation system can be divided in different revenue streams: one for each type of patient. For each of them demand responses are different because different types of patients have different preferences for quality and they are more or less numerous. Additionally, the variable profits for each type of patient are also different, because there are differences in marginal costs across patients, even if prices are the same.

The level of margins is determined by the first order conditions. As explained in the previous section, the choice of condition specific qualities from (4) can lead to either positive or negative margins. Similarly, at the optimal choice of patient experience from (3), some of margins from different types of patients may be positive and some may be negative as long as the equation is satisfied. This means that the presence of some negative margins for some type of patient in (3) may reduce the incentives of practices to provide additional quality (compared to a case where the patients with negative margins are excluded).

In this context with different margins, an increase in the regulated prices may have unintended consequences. In particular, a higher uniform price can lead practices to discriminate against cardiac patients and lower their condition-specific quality. The practices would receive the same higher uniform price regardless of the condition of the patient, however, some patients with chronic conditions are more costly, because they require more time. From equation (3) there would be an incentive to increase patient experience, especially because the majority of patients are healthy and have lower marginal costs. However, at the

same time, given the convexity of marginal costs, an increase in general patient experience would lead to higher marginal costs, especially for chronic patients. The higher level of marginal costs for chronic patients would also be relevant in equation (4) determining condition specific quality. For this reason the level of specific quality may be unchanged even if the prices $\bar{p}$ increase. The level may even decrease to compensate the increase in costs from (3).

Intuitively, practices facing higher prices may have the incentive to select patients. They could try to turn away some chronic patients by lowering condition specific quality. So, an increase in prices which is not accompanied by additional measures, like an increase in QOF compensation, may affect negatively patients with chronic conditions.[21]

Notably, while this effect holds on average, there is heterogeneity in the responses of practices, because of heterogeneity in costs (margins) and marginal residual demand. There are also differences in the effects on cardiac and respiratory quality. Given the higher marginal costs for cardiac conditions and the higher preference parameter for cardiac specific quality, the practices would react on average by decreasing their cardiac quality more significantly.

Risk adjustment. Consider now an increase in prices only for patients with conditions, in a way to compensate for the the higher marginal costs they cause. As anticipated, this would depart from the current situation where the price does not vary with the type of patients practices have. In particular, practices receive the same price regardless of the conditions of the patients, even if practices from different geographical areas can receive different prices based on differences in average diseases prevalence across geographical areas (based on electoral districts).

This reform would affect patient experience. On the one side, given that chronic patients have preferences for both patient experience and their specific process measures, practices will improve patient experience and condition specific quality to attract chronic patients. This change would benefit all patients. On the other side, the increase in patient experience is smaller on average than with higher uniform prices because practices do not earn higher margins on healthy patients.

[21] This would be particularly true for areas with higher prevalence of patients with conditions (typically poorer areas with a larger minority population, be it of African or Asian origin). This is happening because the prices are already higher in areas with higher prevalence of chronic conditions, making the trade-off even stronger. This also points out the potential problems that may arise from a phasing out of QOF (as it happened in Scotland) keeping prices as they are and not having specific prices for patients with certain conditions.

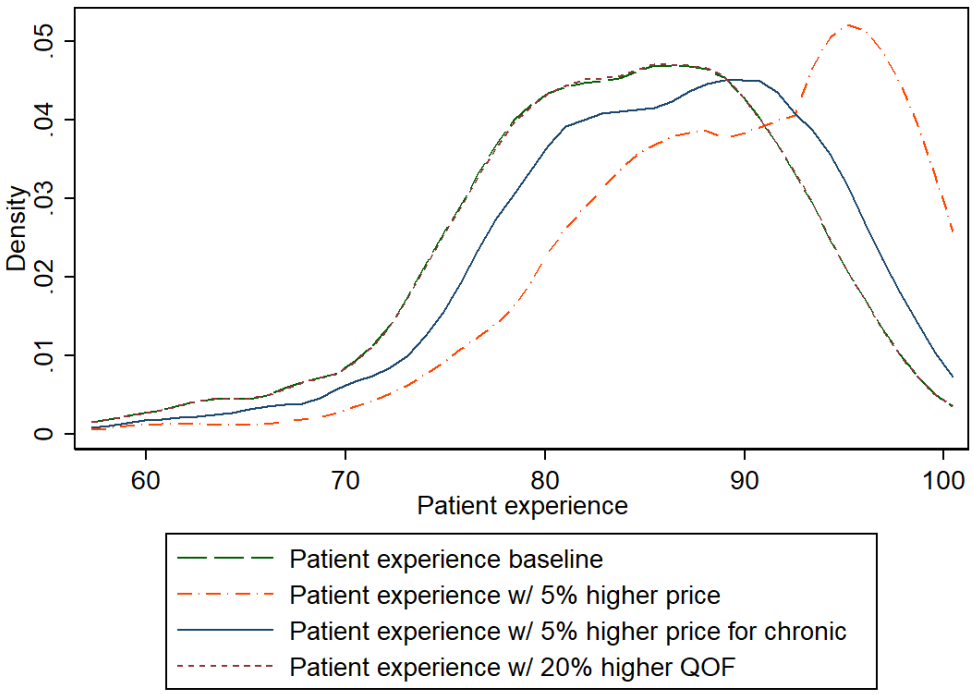


Figure 3: Patient experience distribution under different reform scenarios

The reform would also affect condition specific process measures. Specific quality is increased when risk adjustment is introduced, however, the impact is larger for cardiac quality. This is happening because cardiac patients have a preference parameter for their condition specific quality which is relatively larger than the preference parameter for patient experience, compared to mental health patients or respiratory patients. For example, in (4) practices would not see much of an effect of a raise in price in the case of respiratory quality.

Increase in rewards for quality (QOF). Finally, consider a scenarios specifically tailored for patients with conditions: an increase in the amount of money paid in for one QOF point (or the number of point available per condition). This change to the payment system affects only specific quality and benefits only chronic patients. The effect varies depending if the quality already provided by the practices is above or below the maximum threshold for payments. In the case where the majority of the practices already provide quality above the threshold we observe almost no effect (as for respiratory patients, maximum threshold is 75%). When, instead, the quality provided is below the maximum threshold the effect of the change is larger, as in the cardiac case (maximum is 100%).

Additionally, for some practices patient experience is even negatively affected by a small margin. This is due to the fact that by decreasing patient experience they can select for the most profitable patients (for these practices, in this case,

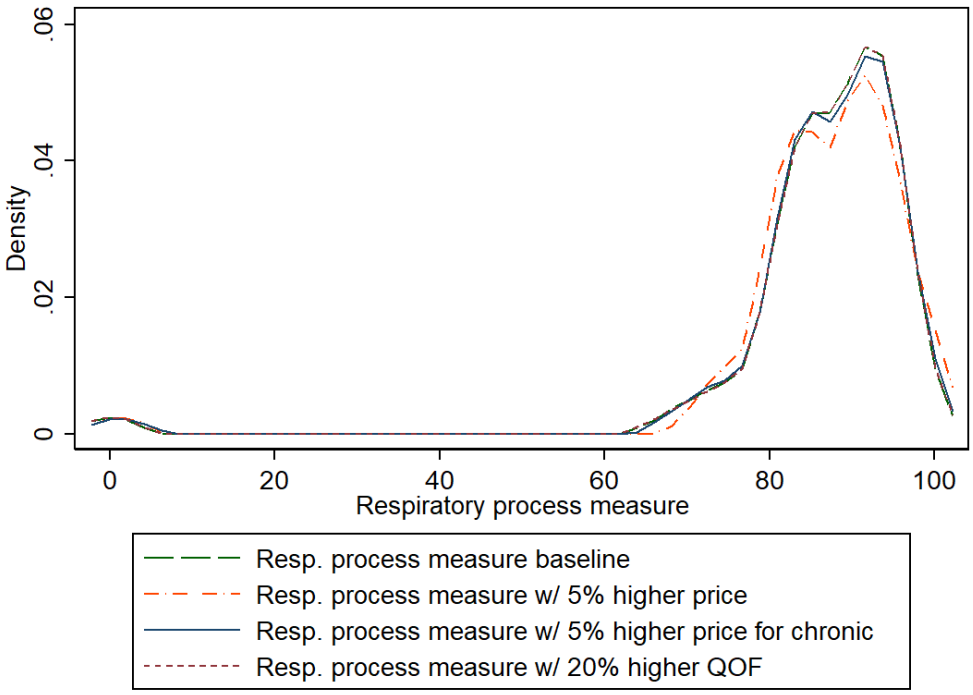


Figure 4: Respiratory quality distribution under different reform scenarios

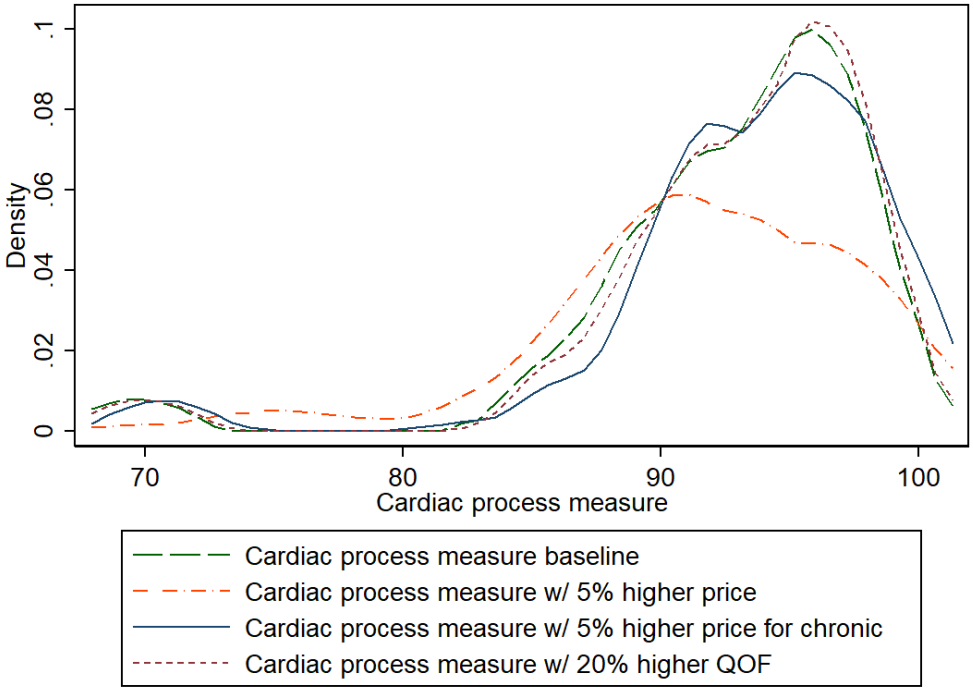


Figure 5: Cardiac quality distribution under different reform scenarios

chronic patients). This may or may not be desirable because patients with conditions also value patient experience. The desirability depends on what the regulator thinks is the most appropriate practice for the patients with conditions: i.e.

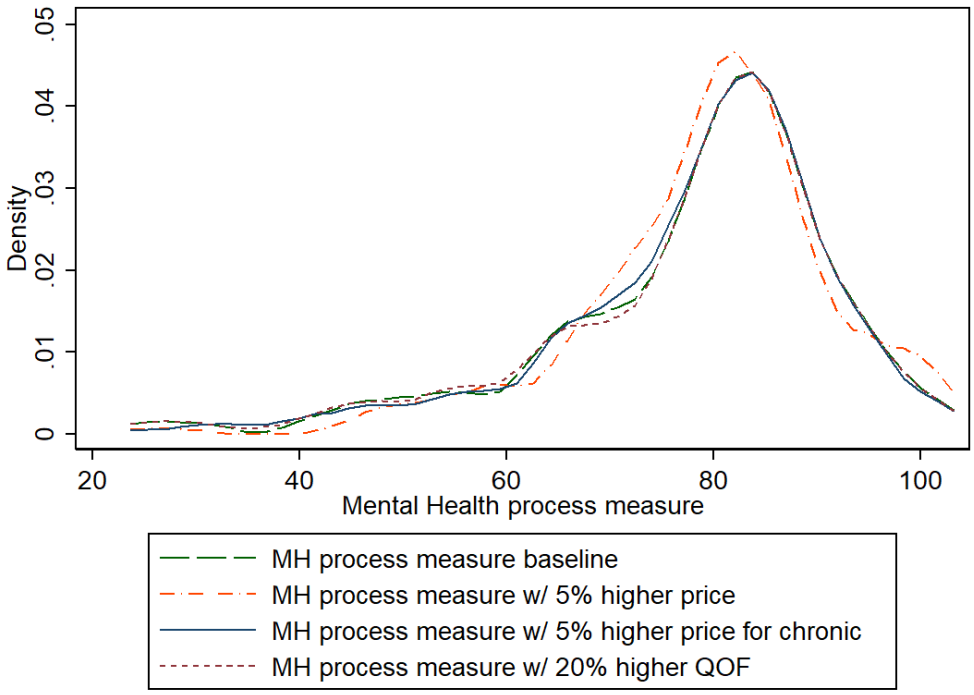


Figure 6: Mental health quality distribution under different reform scenarios

where they can manage better their long term condition.

Welfare effects. Using results from Small and Rosen (1981) we can calculate the consumer surplus and the welfare effects of the policies, making an assumption on patient disutility of spending (money).[22]

We can observe that higher uniform prices increase more consumer surplus than the other two measures. However, risk-adjustment is the most welfare efficient delivering the highest amount of consumer surplus per pound of government expenditure. Finally, the increase in QOF point is the least efficient.

	Δwelfare for Bristol in 2010-2012 in GBP						
	ΔCS_H	$\sum_K \Delta CS_K$	$\Delta V.Profits$	$\Delta Costs$	$\Delta Gov.Exp.$	$\Delta Welfare$	$\frac{\Delta CS_H + \Delta CS_K}{\Delta Gov.Exp.}$
Panel A: Uniform price increase							
$+5\%\bar{p}$ for all patients	3.6m	1.8m	4.8m	0.3m	5.2m	3.4m	1.0
Panel B: Risk Adjustment							
$Higher\bar{p_a}$ Resp., $\bar{p_a}$ Cardiac & $\bar{p_a}$ MH	1.7m	0.9m	1.2m	0.1m	1.3m	2.1m	2.1
Panel C: Increase in QOF point value							
$+20\%QOF$ point	0.01m	0.02m	0.06m	0.00m	0.06m	0.01m	0.5

Note: Welfare = Consumer Surplus CS_H + $\sum_K$Consumer Surplus (condition K) CS_K - practice costs - $\lambda Gov.Exp.$
Note: The cost of raising public funds $\lambda = 0.3$. Marginal utility of income for consumers ≈0.011

Table 5: Welfare change after reforms

The comparison with the quality rewards may be incomplete. The different effects are driven by patients preferences for the different qualities. For example, given that a uniform prices increase leads to a larger increase in patient experience, it has the largest positive consumer welfare effect even if it does not affect/affect negatively some condition specific quality. However, in this analysis we are assuming that patients display the "right" preferences, in the sense of having access to complete information and having no behavioral bias in their decisions. This may not be the case. The preferences we estimated may not reflect a correct valuation of the benefits arising from condition specific process measures. For this reason the results from the welfare analysis may be misleading about the actual effects of the different changes to the payment system. Taking into consideration the changes in quality under the different scenarios helps interpreting the results.

[22]This assumption is based on monetary estimates of disutility from travelling.

The role of information. There multiple explanations for the difference in preference parameters for patient experience and process measures. One explanation that we contemplate here is that patients are not sufficiently informed about the process measures. In fact, the set of process measures has not been easily publicly accessible for many years on NHS websites. The measures have been collected for regulatory purposes for the QOF payments, but have been only recently briefly available on NHS comparison websites. We can try to understand the impact of better availability and accessibility of information with a counterfactual simulation. In this exercise, we equalize the preference parameters of chronic patients for patient experience and for process measures. This would reflect not only increased awareness, but also an increase valuation of process measures in the social welfare function.

To understand the impact of the different reforms analyzed in the previous section we need to simulate a new baseline where the preference parameters are equalized. The new baseline quality distributions are different as higher parameters translate to increased importance of margins. On average mental health patients are characterized by lower marginal costs than other conditions, for this reason when the parameters are equalized the quality level increases. For the other conditions the effects is much smaller and slightly negative, due to smaller or negative margins. This is also reflected in the effect of the equalized preference parameters on patient experience. There is a small negative effect on this quality dimension driven by the higher marginal costs and smaller margins for chronic conditions.

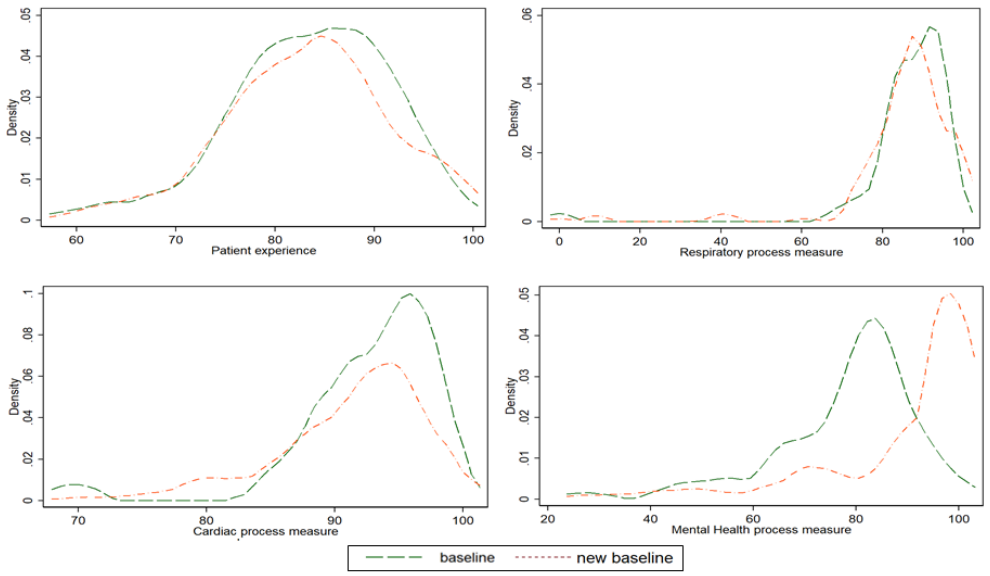


Figure 7: Baseline quality dimensions distributions when preferences are equalized

The results of the different reforms are similar to what we described in the previous section. Nevertheless, there are noticeable differences. All reforms have a larger impact on chronic conditions specific quality measures. A uniform price increase still leads to a decrease in condition specific quality measures, albeit small. The difference is margins still drives these results with mental health quality measure being affected the least. Additionally, risk-adjustment has a larger impact on these measures than on patient experience, contrary to what was described in the previous section. In this scenario practices can increase condition specific quality to easily attract patients.

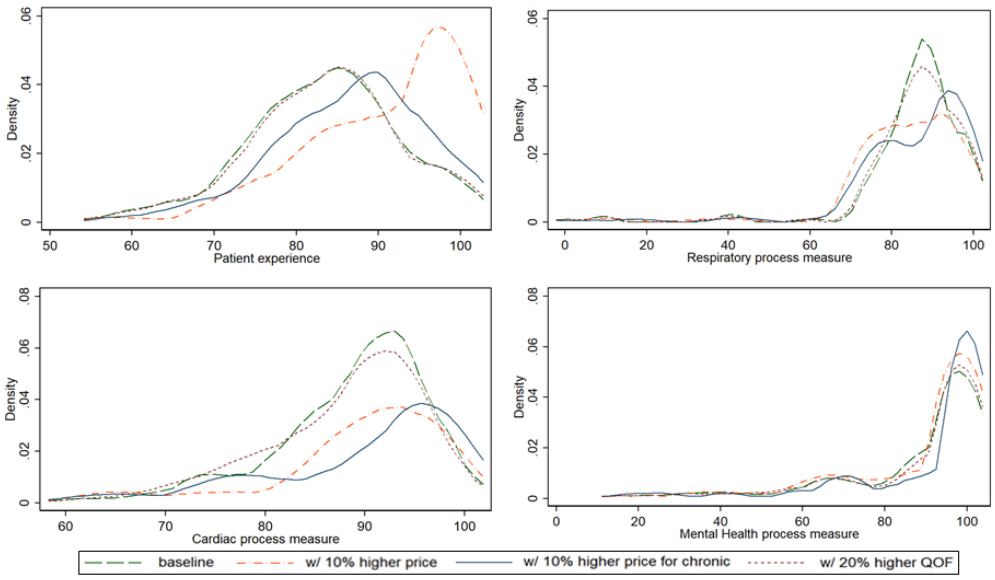


Figure 8: Quality dimensions distributions when preferences are equalized - reforms

	Δwelfare for Bristol in 2010-2012 in GBP						
	ΔCS_H	$\sum_K \Delta CS_K$	$\Delta V.Profits$	$\Delta Costs$	$\Delta Gov.Exp.$	$\Delta Welfare$	$\frac{\Delta CS_H + \Delta CS_K}{\Delta Gov.Exp.}$
Panel A: Uniform price increase							
$+5\%\bar{p}$ for all patients	4.0m	2.0m	4.7m	0.4m	5.2m	4.0m	1.2
Panel B: Risk Adjustment							
$Higher\bar{p}_a$ Resp., $\bar{p}_a$ Cardiac & $\bar{p}_a$ MH	1.2m	0.7m	1.2m	0.1m	1.3m	1.5m	1.5
Panel C: Increase in QOF point value							
$+20\%QOF$ point	0.02m	0.01m	0.08m	-0.02m	0.05m	0.03m	0.5

Note: Welfare = Consumer Surplus CS_H + $\sum_K$Consumer Surplus (condition K) CS_K - practice costs - $\lambda Gov.Exp.$
Note: The cost of raising public funds $\lambda = 0.3$. Marginal utility of income for consumers ≈0.011

Table 6: Welfare change after reforms - with better access to information

7 Final remarks

Competition is a tool that has been used to help efficient allocation of patients and incentivize quality. However, this relies on the assumptions that patients have a preference for quality and that quality can be approximated to a single dimension. In this paper we highlight the importance of multi-dimensional quality for patient allocation through demand and for supply incentives for quality provision. In particular, mis-allocation of chronic patients may arise if they attend providers inappropriate for their conditions following other dimensions of quality not as relevant to them from a social welfare perspective. From the supply side perspective, we focus on the role of price regulation in this context. We highlight how uniform prices can lead to perverse incentives when patients display differences in profitability. Practices may decide to use some dimensions of quality to select away unprofitable patients. Instead, quality payments emerge as a useful option to incentivize condition specific process measures when patient display little preference for condition specific quality.[23]

An additional aspect that emerged from our analysis is that risk adjustment can have more social welfare benefits than in a setting where quality is unidimensional. In particular, we observe in our counterfactuals that welfare can be boosted by risk-adjustment when chronic patients care also about patient experience, an attribute also valued by healthy patients. Increasing payments for chronic patients requires less government expenditure than an uniform price increase for all patients. However, this increase can lead to an improvement in patient experience given that chronic patients care about this dimension of quality enjoyed by all patients. This leads to an efficient increase in consumer welfare for all patients.

As a final consideration we want to point out the following. Family doctor care is becoming more important and even the care of chronic patients is moved to the community and family doctors. The aim is to reduce government expenditure by reducing expensive hospital care and promote preventive care and good management of chronic conditions which is beneficial to both patients and government coffers. We think that our analysis should be taken into account during this process in order to maximize the benefits from this move and avoid potential unexpected market distortions. These could arise in presence of multidimensional quality as discussed in our paper.

[23]However, there is a risk that doctors may end up focusing on specific metrics and not on the overall patient health. For this reason the metric should capture some aspects of care that is positively correlated with other aspects of care. We abstract from this discussion in this paper.

References

Ackerberg, Daniel A, Caves, Kevin, & Frazer, Garth. 2015. Identification properties of recent production function estimators. *Econometrica*, **83**(6), 2411–2451.

Alkalay, Adi, Eizenberg, Alon, Lahad, Amnon, Shurtz, Ity, *et al.* 2018. Physician workload and treatment choice: the case of primary care.

Berry, Steven T. 1994. Estimating Discrete-Choice Models of Product Differentiation. *The RAND Journal of Economics*, **25**(2), 242–262.

Bloom, Nicholas, Propper, Carol, Seiler, Stephan, & Van Reenen, John. 2015. The impact of competition on management quality: evidence from public hospitals. *The Review of Economic Studies*, **82**(2), 457–489.

Camarda, Enrico. 2022. Welfare Effects of Outcome-based prices in the English National Health Service. *Working paper*.

Chan, David C. 2018. THE EFFICIENCY OF SLACKING OFF: EVIDENCE FROM THE EMERGENCY DEPARTMENT. *Econometrica*, **86**(3), 997–1030.

Chan, David C., & Gruber, Jonathan. 2020. Provider Discretion and Variation in Resource Allocation: The Case of Triage Decisions. *AEA Papers and Proceedings*, **110**(May), 279–83.

Chandra, Amitabh, & Staiger, Douglas. 2016. Sources of Inefficiency in Healthcare and Education. *American Economic Review*, **106**(5), 383–387.

Chandra, Amitabh, Finkelstein, Amy, Sacarny, Adam, & Syverson, Chad. 2016. Health Care Exceptionalism? Performance and Allocation in the US Health Care Sector. *American Economic Review*, **106**(8), 2110–44.

Crawford, Gregory S, Shcherbakov, Oleksandr, & Shum, Matthew. 2019. Quality Overprovision in Cable Television Markets. *American Economic Review*, **109**(3), 956–995.

Dalton, Christina M, Gowrisankaran, Gautam, & Town, Robert. 2019. *Salience, myopia, and complex dynamic incentives: Evidence from Medicare Part D*. Tech. rept. National Bureau of Economic Research.

Einav, Liran, Finkelstein, Amy, & Mahoney, Neale. 2018. Provider Incentives and Healthcare Costs: Evidence From Long-Term Care Hospitals. *Econometrica*, **86**(6), 2161–2219.

Eliason, Paul J, Grieco, Paul L E, McDevitt, Ryan C, & Roberts, James W. 2018. Strategic patient discharge: The case of long-term care hospitals. *American Economic Review*, **108**(11), 3232–3265.

Fan, Ying. 2013. Ownership Consolidation and Product Characteristics: A Study of the US Daily Newspaper Market. *American Economic Review*, **103**(5), 1598–1628.

Fleitas, Sebastián. 2018. Who benefits when inertia is reduced? Competition, quality and returns to skill in health care markets.

Gaynor, Martin. 2006. *What Do We Know About Competition and Quality in Health Care Markets?* Working Paper 12301. National Bureau of Economic Research.

Gaynor, Martin, & Town, Robert J. 2011. *Competition in Health Care Markets*. Working Paper 17208. National Bureau of Economic Research.

Gaynor, Martin, Moreno-Serra, Rodrigo, & Propper, Carol. 2013. Death by Market Power: Reform, Competition, and Patient Outcomes in the National Health Service. *American Economic Journal: Economic Policy*, **5**(4), 134–166.

Gaynor, Martin, Ho, Kate, & Town, Robert. 2014 (jan). *The Industrial Organization of Health Care Markets*. Working Paper 19800. National Bureau of Economic Research.

Gaynor, Martin, Propper, Carol, & Seiler, Stephan. 2016. Free to Choose? Reform, Choice, and Consideration Sets in the English National Health Service. *American Economic Review*, **106**(11), 3521–3557.

Gravelle, Hugh, Santos, Rita, & Siciliani, Luigi. 2014. Does a hospital's quality depend on the quality of other hospitals? A spatial econometrics approach. *Regional Science and Urban Economics*, **49**, 203–216.

Gravelle, Hugh, Liu, Dan, Propper, Carol, & Santos, Rita. 2019. Spatial competition and quality: Evidence from the English family doctor market. *Journal of Health Economics*, **68**, 102249.

Grieco, Paul L E, & McDevitt, Ryan C. 2016. Productivity and quality in health care: Evidence from the dialysis industry. *The Review of Economic Studies*, **84**(3), 1071–1105.

Hackmann, Martin B. 2019. Incentivizing Better Quality of Care: The Role of Medicaid and Competition in the Nursing Home Industry. *American Economic Review*, **109**(5), 1684–1716.

Handel, Benjamin R, Kolstad Jonathan T Root Allyson B & Whiston Michael D. 2021. Outcomes-Based Payments and Physician Productivity: Evidence from Diabetes Care in Hawaii.

Ho, Kate, Hogan, Joseph, & Scott Morton, Fiona. 2017. The impact of consumer inattention on insurer pricing in the Medicare Part D program. *The RAND Journal of Economics*, **48**(4), 877–905.

Holmes, Thomas J. 2011. The Diffusion of Wal-Mart and Economies of Density. *Econometrica*, **79**(1), 253–302.

Lancsar, Emily, Gu, Yuanyuan, Gyrd-Hansen, Dorte, Butler, Jim, Ratcliffe, Julie, Bulfone, Liliana, & Donaldson, Cam. 2020a. The relative value of different QALY types. *Journal of Health Economics*, 102303.

Lancsar, Emily, Gu, Yuanyuan, Gyrd-Hansen, Dorte, Butler, Jim, Ratcliffe, Julie, Bulfone, Liliana, & Donaldson, Cam. 2020b. The relative value of different QALY types. *Journal of Health Economics*, 102303.

Oakley, Ben, Nightingale, Jessica, Moran, Christopher G, & Moppett, Iain K. 2017. Does achieving the best practice tariff improve outcomes in hip fracture patients? An observational cohort study. *BMJ open*, **7**(2), e014190.

Pennington, Mark, Baker, Rachel, Brouwer, Werner, Mason, Helen, Hansen, Dorte Gyrd, Robinson, Angela, Donaldson, Cam, & Team, EuroVaQ. 2015. Comparing WTP values of different types of QALY gain elicited from the general public. *Health economics*, **24**(3), 280–293.

Propper, Carol. 1995. The disutility of time spent on the United Kingdom's National Health Service waiting lists. *Journal of Human Resources*, 677–700.

Santos, Rita, Gravelle, Hugh, & Propper, Carol. 2017. Does quality affect patients' choice of doctor? Evidence from England. *The Economic Journal*, **127**(600), 445–494.

Small, Kenneth A, & Rosen, Harvey S. 1981. Applied welfare economics with discrete choice models. *Econometrica: Journal of the Econometric Society*, 105–130.

Spence, Michael;. 1975. Monopoly, quality, and regulation.

White, Lawrence. 1972. Quality Variation When Prices Are Regulated. *Bell Journal of Economics*, **3**(2), 425–436.

8 Appendix A - Additional details on QOF payments and achievements

QOF payment calculation for different indicators. Of the indicators included in QOF, some are related to achieving specific targets, for example establishing a list of patients with diabetes (so called Boolean indicators). Others are more related to actual interventions or level of quality outcomes of a certain practice. Examples are checking the limbs of patients with diabetes at least once a year or having diabetes patients with long term blood sugar levels under a certain threshold. These indicators capture the percentage of patients receiving such interventions or achieving such targets. Typically, a practice would receive a certain amount of money for achieving a Boolean indicator target, in particular for creating a register of patients with diabetes a practice will receive 6 QOF points. For the other type of indicators the practices would attract a linearly increasing amount of money starting from 20% of patients up to an upper threshold, usually 80% or 90% of patients. However, for these types of indicators the payment is calculated in a slightly more complicated way, adjusting the payments depending on the total amount of patients in the practice and the total number of patients with a certain condition. The objective is to make sure that payments reflect the workload needed for achieving a specific target. We explain the details below.

Clinical indicators typically fall in the category of non-Boolean indicators and the payment calculation, as anticipated, is more complicated and involve several steps. The first step is calculating achievements. Achievements are the number of desired interventions (e.g., a treatment, but could also be having a certain level of long term sugar in the blood) on patients with a specific condition over the total number of patient with such condition. Crucially, this measure is net of "exceptions", which means that in calculating this ratio they exclude from the denominator (total number of patients with a certain condition, e.g. diabetes) some patients. These patients are excepted mainly for three reasons: i) they do not show up for the treatment after being invited three times to the practice for this purpose ii) they did not want to receive treatment (giving written disagreement), iii) it is not clinically appropriate for them to receive that specific treatment (taking a certain medicine is incompatible with taking another medication).

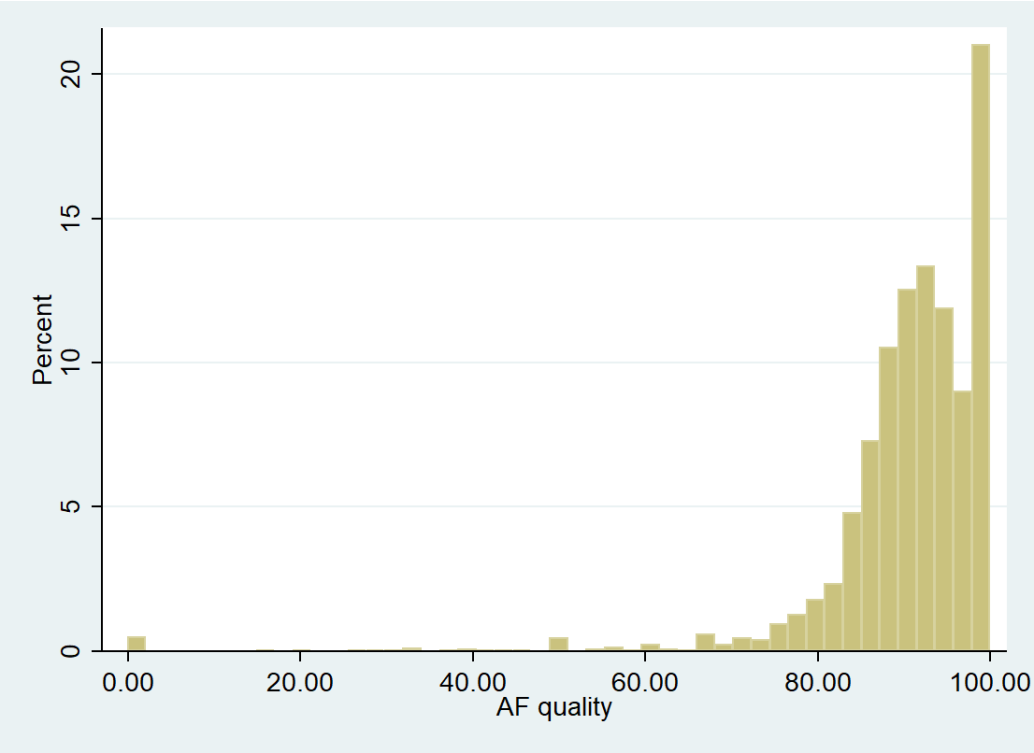
$$\text{Achievement for indicator 2 for condition C} = \frac{\text{No. of patients w/ condition C receiving an intervention}}{\text{No. of patients w/ condition C} - \text{No. of patients ``excepted'' w/ condition C}}$$

The second step is to calculate the final payment for each of these non-Boolean indicators. The amount paid out will increase linearly with the percentage of achievement from the minimum threshold (e.g. 20%) to the maximum (e.g. 80%). The amount will depend on the total number of QOF points (1 QOF point is £187.74) available for each indicator. Additionally there are two adjustments factors that modify the final payment depending on the size of the practice (in terms of total number of patients) and the percentage of patients with the condition relevant for each indicator.
In particular, the Adjusted Practice Disease Factor ("APDF") is measuring how

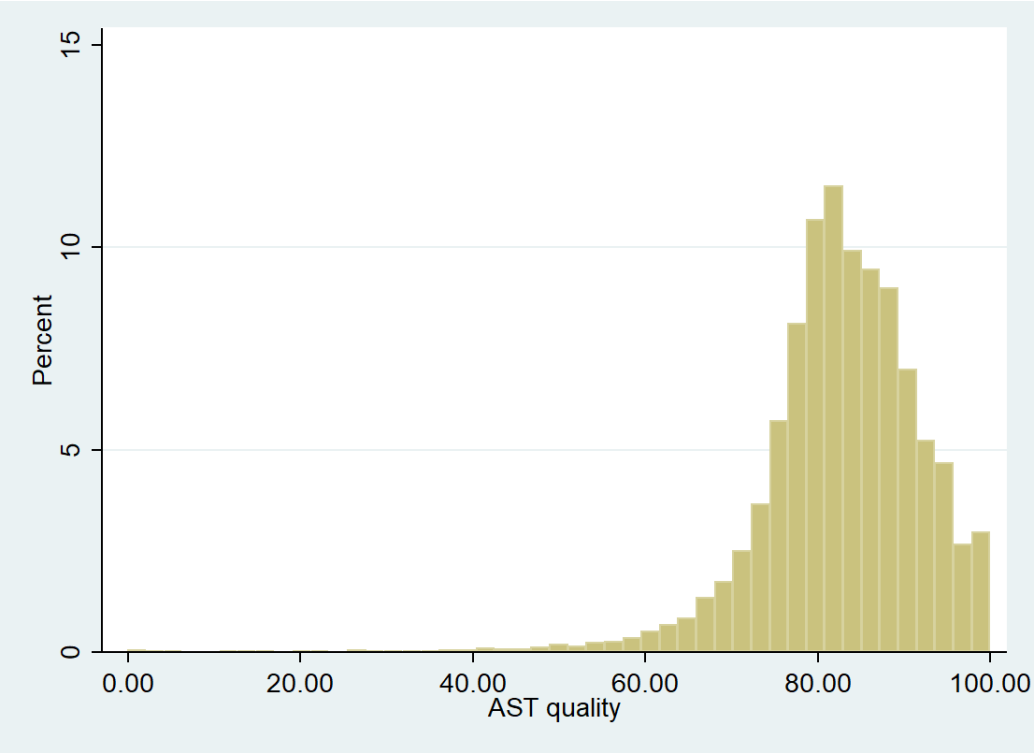
far the prevalence of a certain disease in a practice is from the national prevalence rate. For example, if the national prevalence rate for diabetes is 6.7% and the prevalence rate of a practice is 13.4% then the APDF is 2, indicating that the prevalence rate is twice as much. The idea is that it is going to be more difficult/time consuming/resource intensive to provide interventions in that practice.
Following a similar logic the Contractor Population Index ("CPI") indicates how big a practice is compared to the average size of the practice in England. The average size of GP practices in England in terms of registered patients is around 8,500 if a practice has 17,000 patients then the CPI will be equal to 2.

Final payment for indicator 2 for condition C = Achievement for indicator 2 for condition C * No. of point for indicator 2 * APDF for condition C * CPI

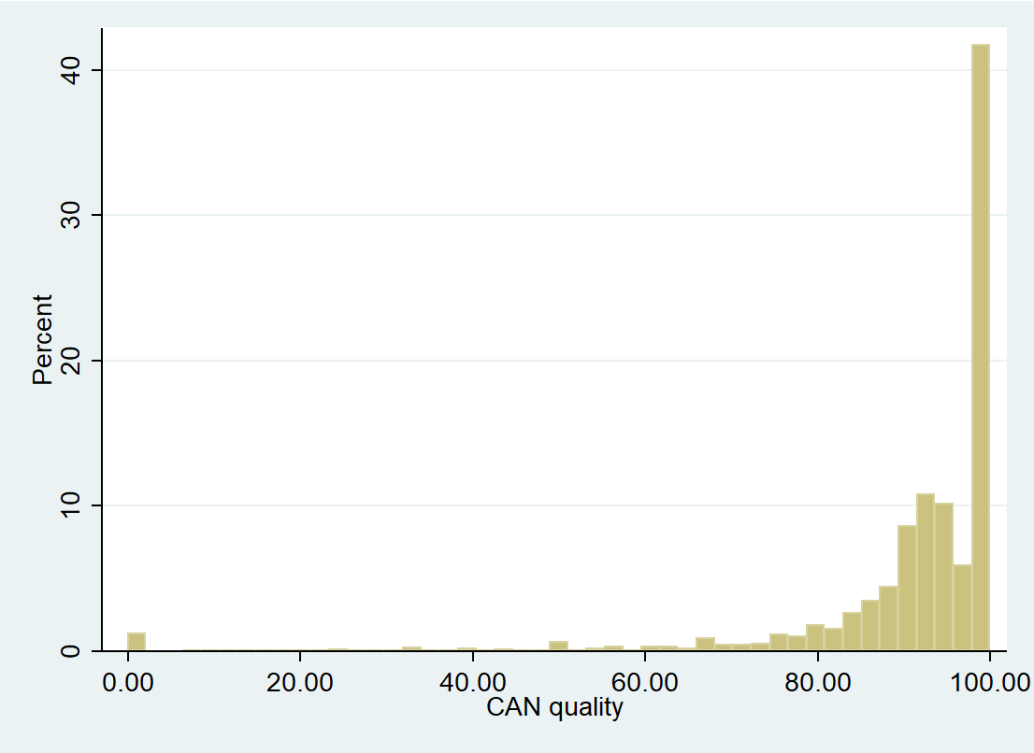
Note on variation in QOF achievements. As a final note, we find variation in the achievements of these targets. Differences in achievements are can be due to a multiplicity of factors. We can make sense of these factors considering that practices may have different levels of workload. This can happen because they have many patients with different conditions, patients with multiple conditions, patients with other diseases not covered by QOF or patients who are more frail than patients in the average practice on top of having one of the conditions covered by QOF. This can happen because the demographics of the area where the practice operates is different from the norm. For example, a practice may be operating where a more elderly population lives or in a more economically deprived area. The important point to highlight here is that the risk-adjustment of QOF and of the capitation system is not taking into account all possible aspects of patients health. This has raised questions about the effect of QOF on inequality of healthcare provision, especially considering that poorer patients are more likely to be sick and have long term conditions. Additionally, even if practices may not face a particularly difficult set of patients, these practices may be having difficult in attracting good staff or staff in general given the geographic location or simply because there is a shortage of nurses and GP. In light of these considerations it will be easier to understand why achieving QOF targets is not compulsory, they are time consuming and the money received for them is an important source of funding; not achieving the targets is punished by withholding this source of funding. In fact, following the same line of reasoning, the NHS discourages ranking practices by QOF points, as it may be inherently more difficult for some practices to achieve certain targets, because of higher prevalence, different socio-demographic characteristics and higher co-morbidity. Nevertheless they are informative for the quality of care provided to a particular type of patient and for patients with a certain condition seeking to manage it better.



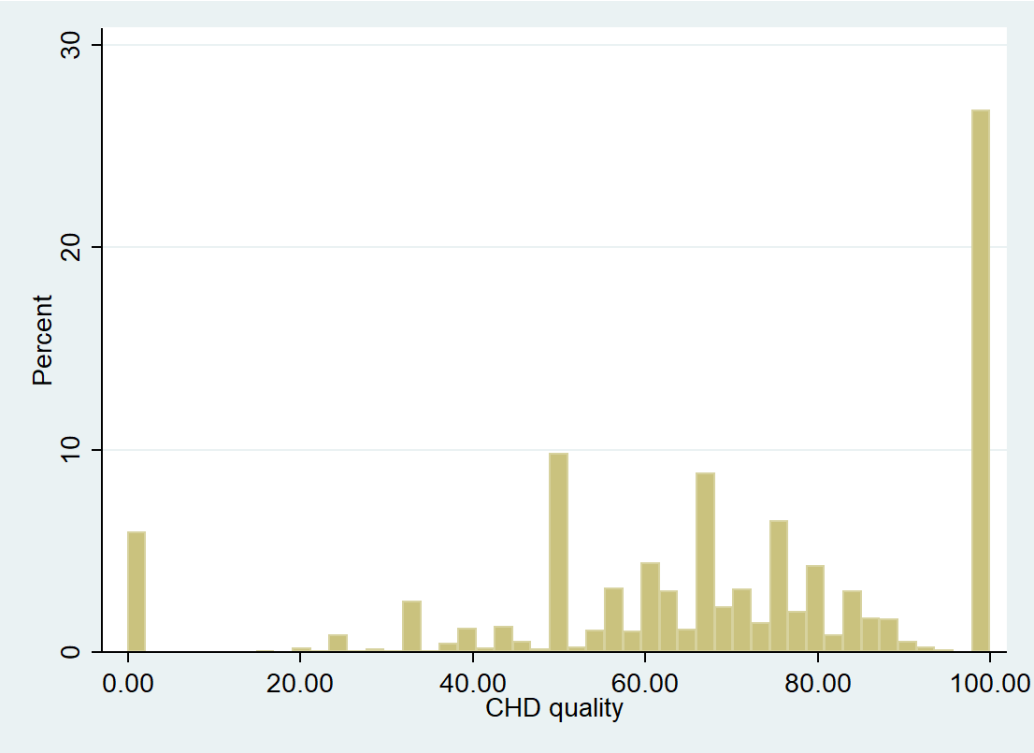
Notes: Quality distribution as percentage of practices with that level of treatment



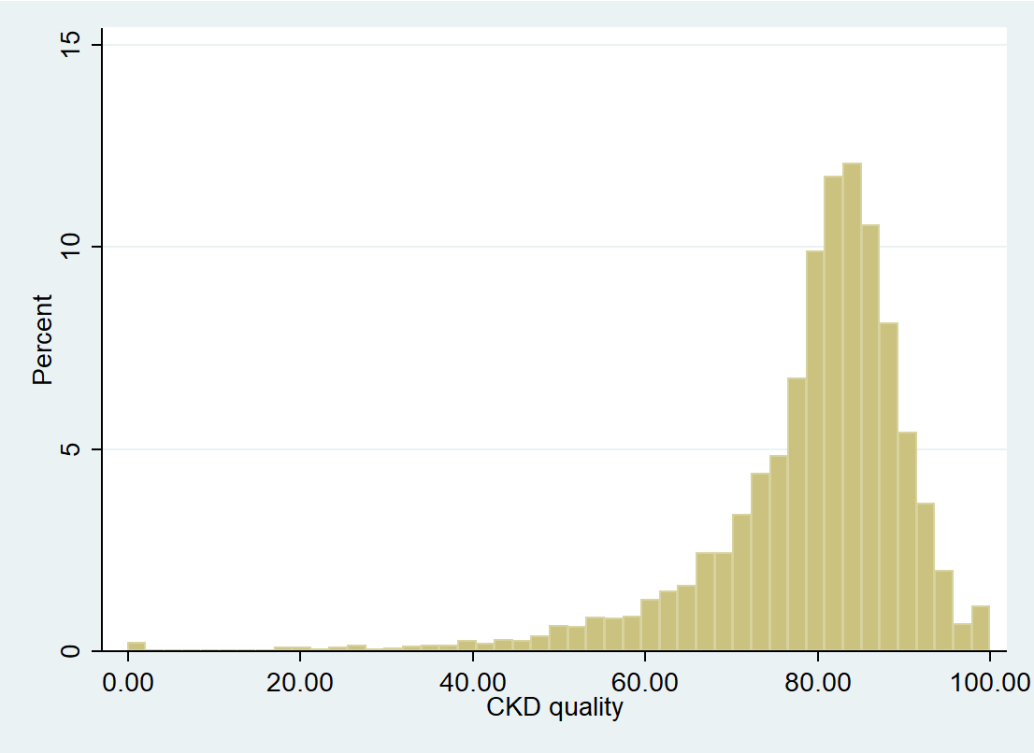
Notes: Quality distribution as percentage of practices with that level of treatment



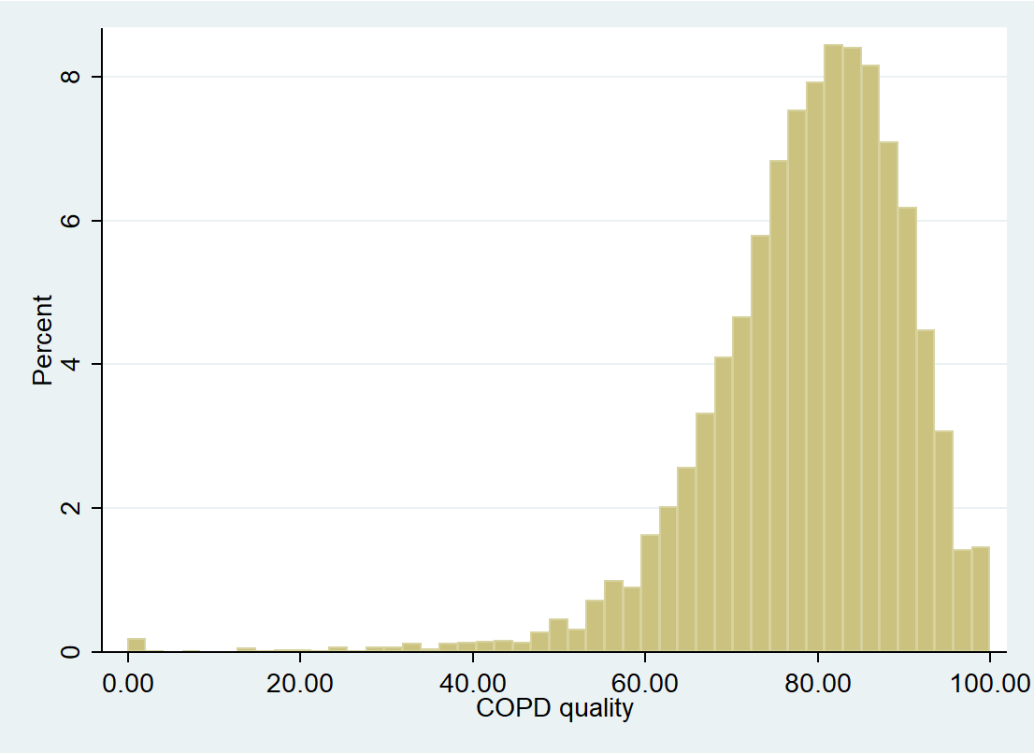
Notes: Quality distribution as percentage of practices with that level of treatment



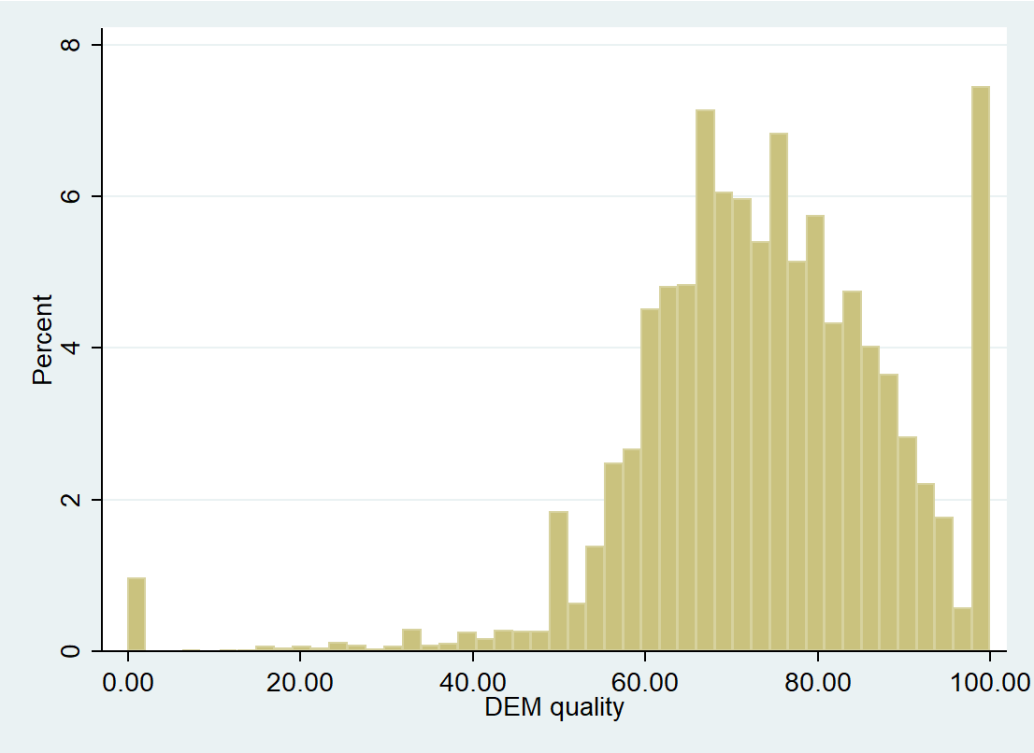
Notes: Quality distribution as percentage of practices with that level of treatment



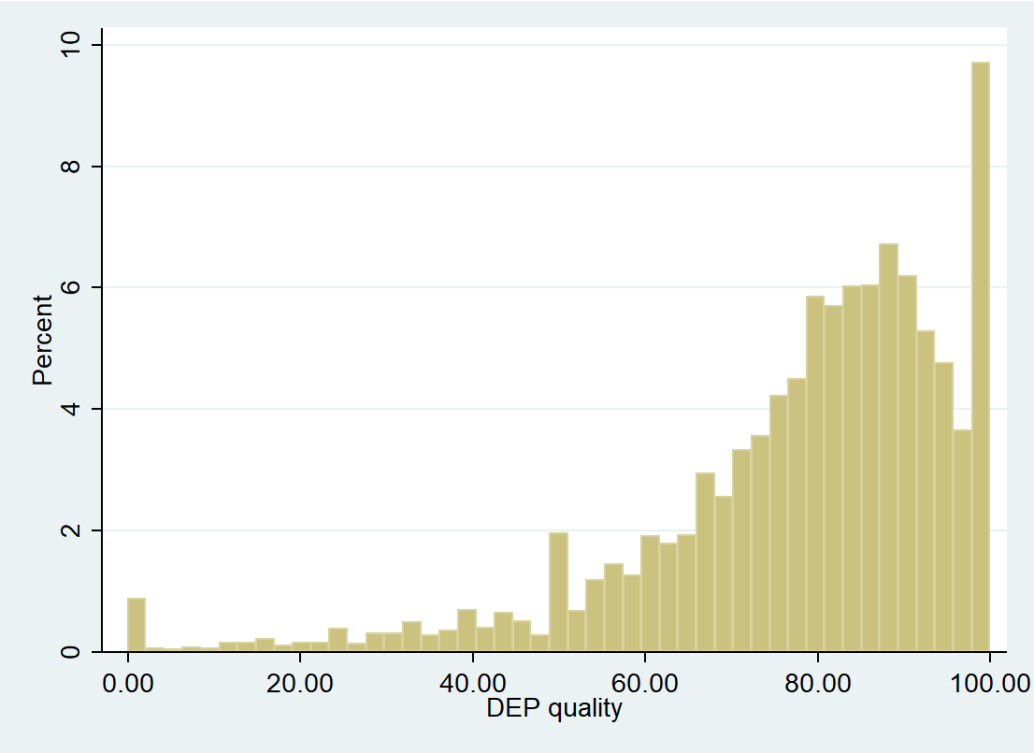
Notes: Quality distribution as percentage of practices with that level of treatment



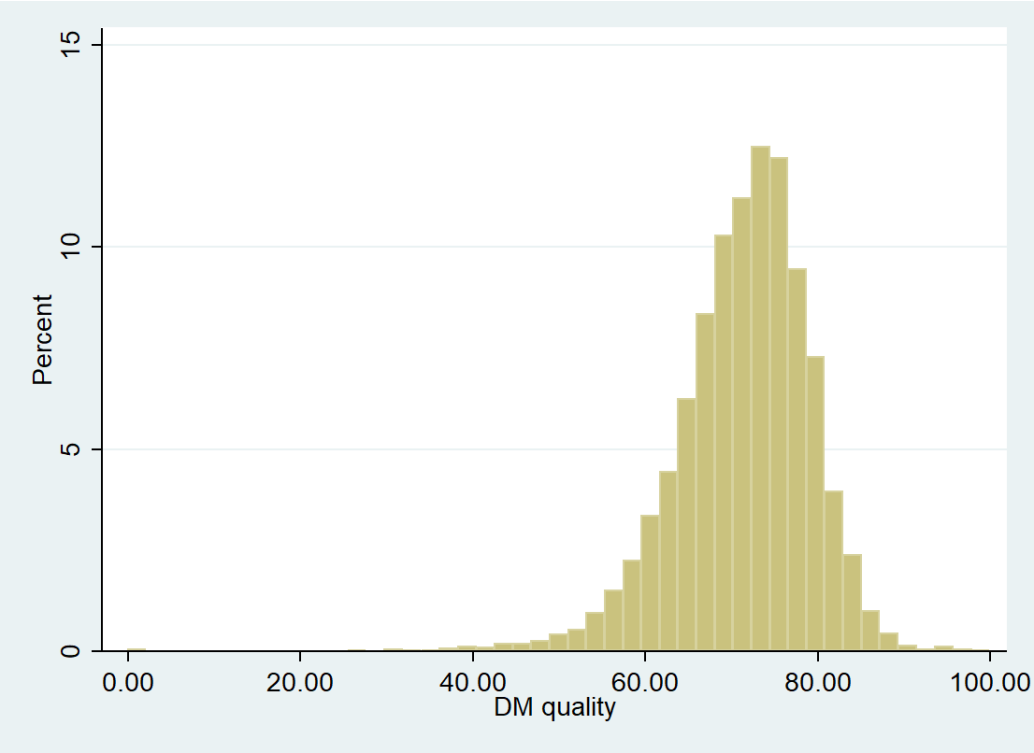
Notes: Quality distribution as percentage of practices with that level of treatment



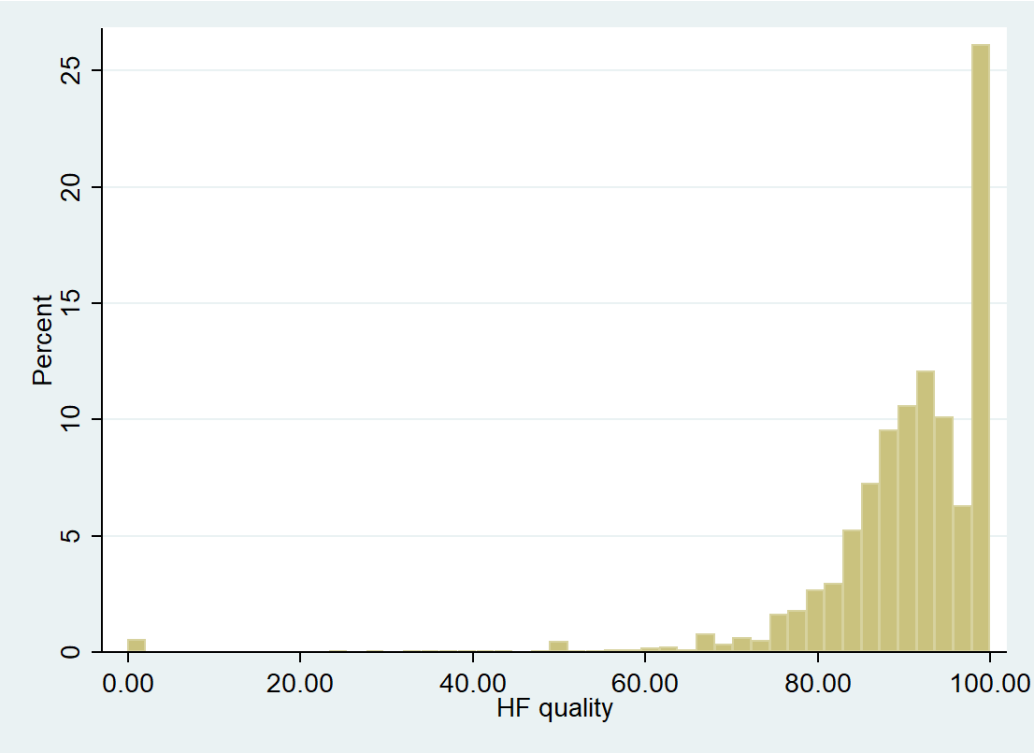
Notes: Quality distribution as percentage of practices with that level of treatment



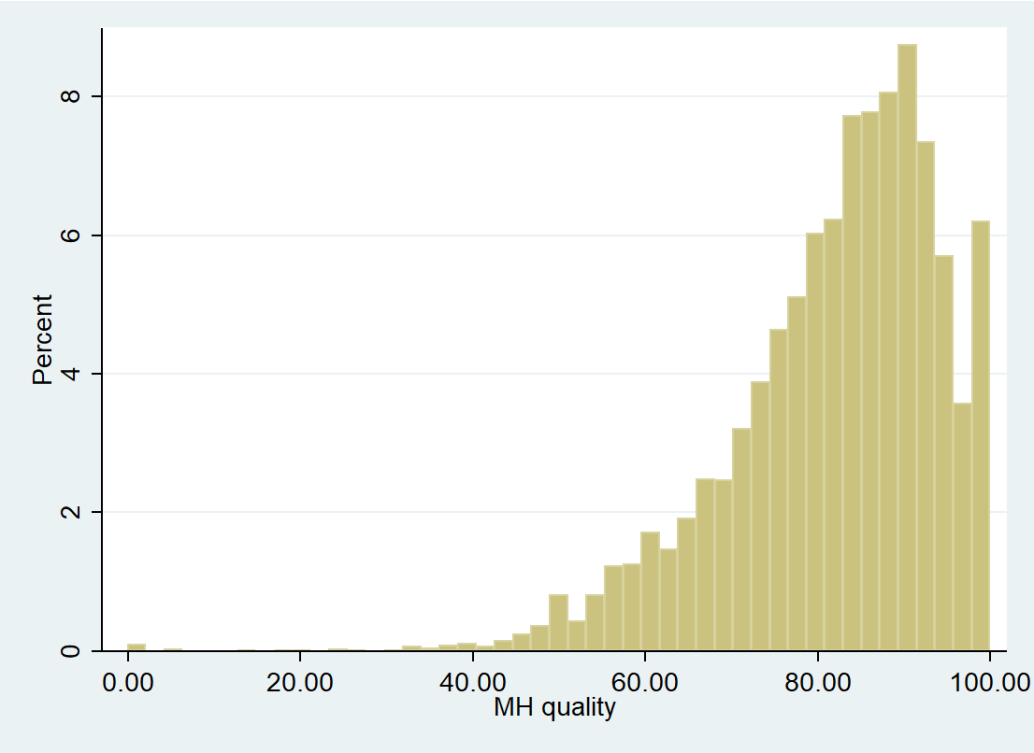
Notes: Quality distribution as percentage of practices with that level of treatment



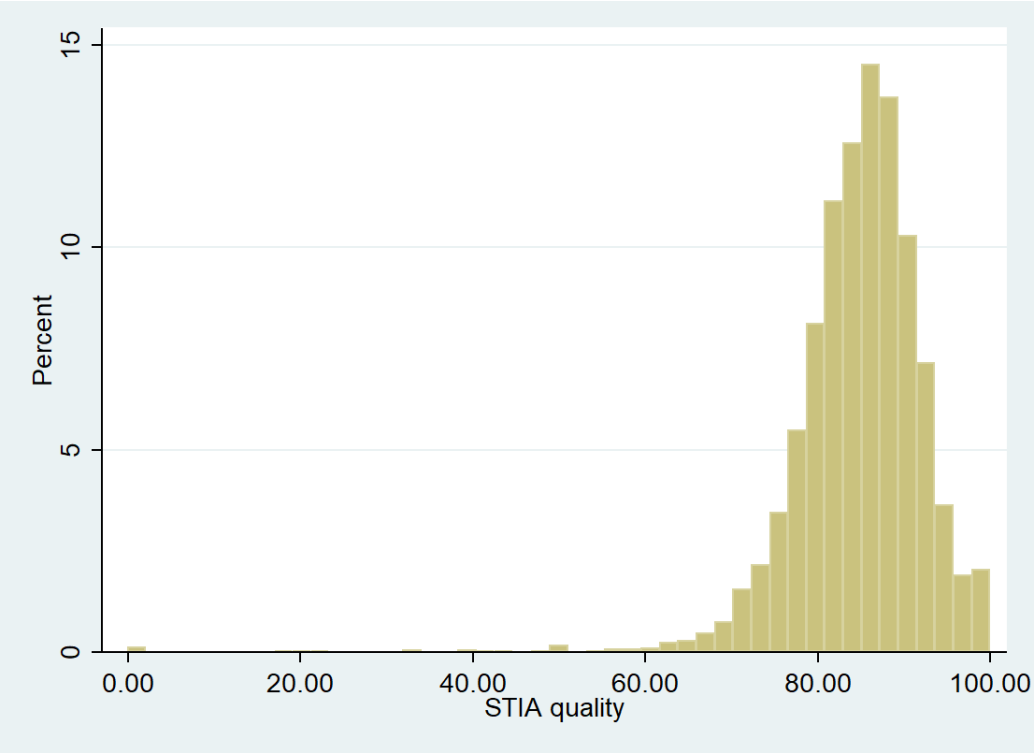
Notes: Quality distribution as percentage of practices with that level of treatment



Notes: Quality distribution as percentage of practices with that level of treatment



Notes: Quality distribution as percentage of practices with that level of treatment



Notes: Quality distribution as percentage of practices with that level of treatment

9 Appendix B - Rejections

Rejections and risk adjustments. In this scenario a partial risk adjustment would be implemented and practices would be allowed to reject patients who are not profitable after the price increase. Crucially, the partial risk adjustment would be in the form of a price increase for chronic patients only for new patients. The combination of the two policies is meant to ensure that all patients would find a practice where to receive treatment. It also meant to improve quality and patient allocation for a small increase in government expenditure. This also assumes that practices with a lower cost to provide care to a certain group of patients are better in the care of these patients.

As discussed earlier, practices can have negative or positive margins with respect to different categories of patients. Additionally, given that marginal costs are increasing, for every condition k there exist a quantity $\bar{q^k} = \sqrt{\frac{\bar{p}}{3\mu_j^{qk}}}$ above which $\bar{p} - \frac{\partial C}{\partial q^k} < 0$. Practices would reject patients when they are unprofitable (negative margins), if they were given the possibility. In that case their objective functions would become the following (where is only one condition a to simplify notation).

$$\max_{z_j^h, z_j^a} U_j = \underbrace{\bar{p} q_j^h(z_j^h, z_{-j}^h)}_{\text{Healthy patients}} + \underbrace{\bar{p} q_j^a(z_j^h, z_{-j}^h, z^a, z_{-j}^a)}_{\text{Patients spec. condition } a} + \underbrace{QOF(z_j^a)}_{\text{Rewards spec. quality}} - \underbrace{C(z^h, z^a, q_j^h, q_j^a)}_{\text{Total cost}} \quad \text{if } z^h, z^a : \bar{p} - \tfrac{\partial C}{\partial q^a} > 0$$

$$\max_{z_j^h, z_j^a} U_j = \underbrace{\bar{p} q_j^h(z_j^h, z_{-j}^h)}_{\text{Healthy patients}} + \underbrace{\bar{p} \bar{q_j^a}}_{\text{Patients spec. condition } a} + \underbrace{QOF(z_j^a)}_{\text{Rewards spec. quality}} - \underbrace{C(z^h, z^a, q_j^h, \bar{q_j^a})}_{\text{Total cost}} \quad \text{if } z^h, z^a : \bar{p} - \tfrac{\partial C}{\partial q^a} <= 0$$

Their first order conditions would become:

$$\begin{aligned}
\bar{p} \frac{\partial q_j^h(z_j^h)}{\partial z_j^h} + \sum_{a=1}^{K} \bar{p} \frac{\partial q_j^a(z_j^h, z_j^a)}{\partial z_j^h} &= \frac{\partial C}{\partial z_j^h} + \frac{\partial C}{\partial q_j^h} \frac{\partial q_j^h(z_j^h)}{\partial z_j^h} + \sum_{a=1}^{K} \frac{\partial C}{\partial q_j^a} \frac{\partial q_j^a(z_j^h, z_j^a)}{\partial z_j^h} \\
\bar{p} \frac{\partial q_j^a(z_j^h, z_j^a)}{\partial z_j^a} \mathbb{1}(\bar{p} - \frac{\partial C}{\partial q_a} > 0) + \frac{\partial QOF(z_j^a)}{\partial z_j^a} &= \frac{\partial C}{\partial z_j^a} + \frac{\partial C}{\partial q_j^a} \frac{\partial q_j^a(z_j^h, z_j^a)}{\partial z_j^a} \mathbb{1}(\bar{p} - \frac{\partial C}{\partial q_a} > 0)
\end{aligned}$$

In our simulation we start from the current levels of patients in each practice such that practices would not be able to reject patients already registered with them, so that $\bar{q^a} = max\{\sqrt{\frac{\bar{p}}{3\mu_j^{qa}}}, q_{baseline}^a\}$, where $q_{baseline}^a$ is the baseline equilibrium quantity of patients with condition a.

We formulate an iterative procedure in finding an equilibrium with rejections. In a first stage all patients apply for a practice and practices decide how many patients to accept based on their profitability. In a second stage chronic patients who are rejected would be looking for a new practice among the practices that have not rejected in the previous stage.[24] The procedure is repeated until all patients are registered with a practice. Importantly, during all stages all practices still modify

[24]Note that $(\frac{\partial q_j^a(z_j^h, z_j^a)}{\partial z_j^h}, \frac{\partial q_j^a(z_j^h, z_j^a)}{\partial z_j^a})$ for non-rejecting practices can include rejected patients.

their quality offering and the allocation of healthy patients may change over the different stages. In this procedure we also make the assumption that patients are not forward-looking and cannot anticipate if they are gonna be rejected.

The benefits of this potential reform are twofold. On the one side, it involves a smaller amount of government expenditure compared to the introduction of risk-adjustment discussed in the main body of the paper. On the other side, it incentives a better allocation of patients in terms of cost-efficiency. Only practices that have a cost advantage in treating chronic patients would accept them.

Interestingly, competition in patient experience and condition specific process measures is intensified under this scenario as practices that have a cost advantage in treating chronic patients would try to attract them using both types of quality. In particular, it would lead to higher levels of patient experience for all practices than simple risk-adjustment benefiting all patients. This reform would be the most efficient in terms of consumer welfare because of this mechanism and because it would not require a large increase in government expenditure or in practices costs. Profits, however, would increase the least under this scenario, given the higher level of competition and the smaller increase in revenues than in other scenarios.

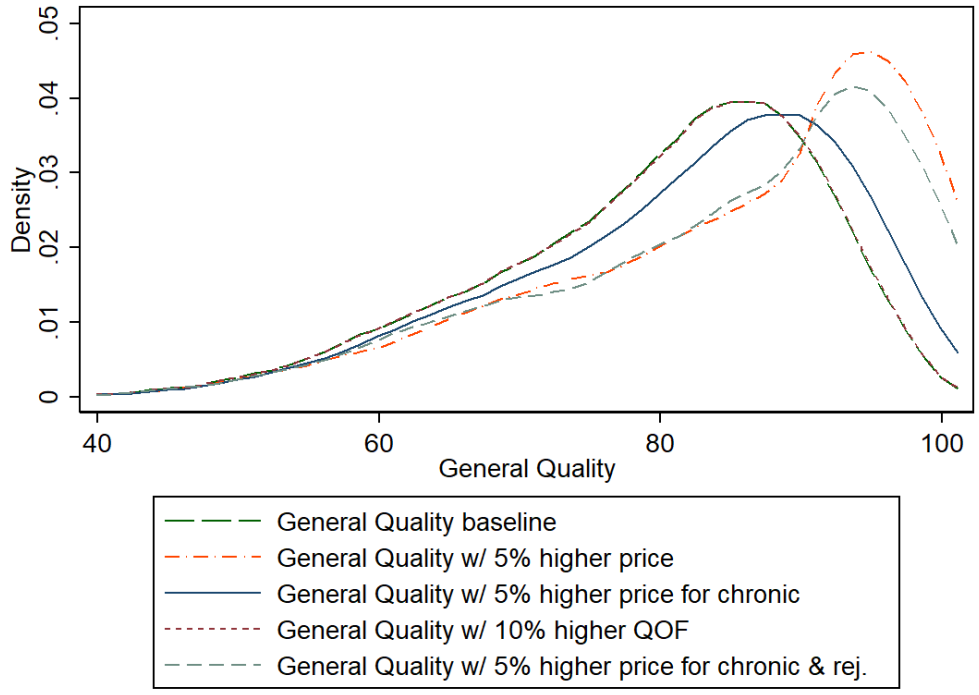


Figure 9: Patient experience distribution under different reform scenarios

In the welfare analysis similar considerations about patients preferences would apply as for the other counterfactuals.

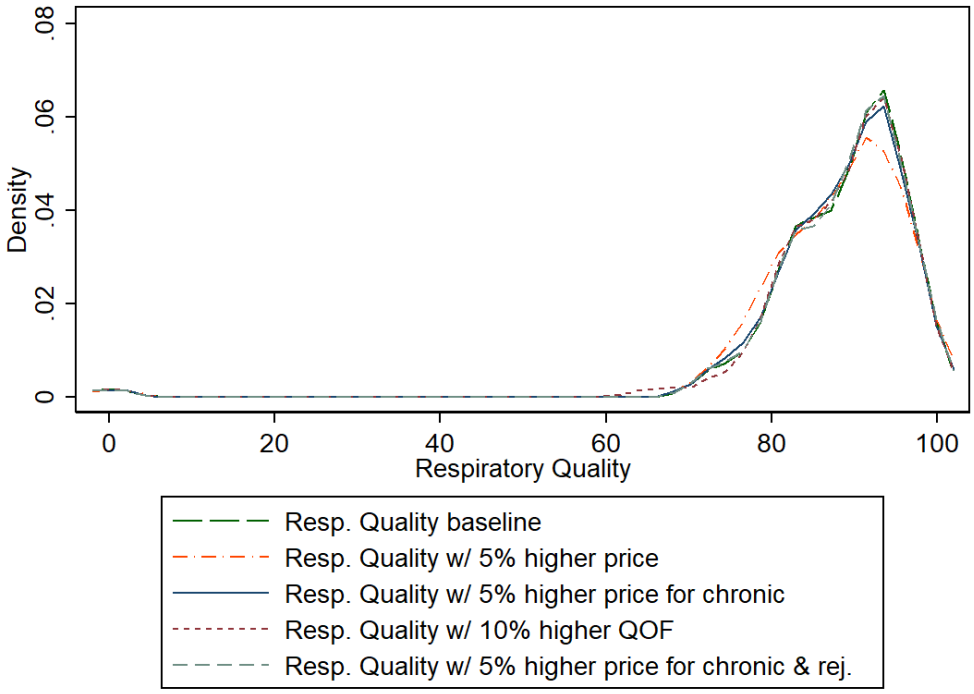


Figure 10: Respiratory quality distribution under different reform scenarios

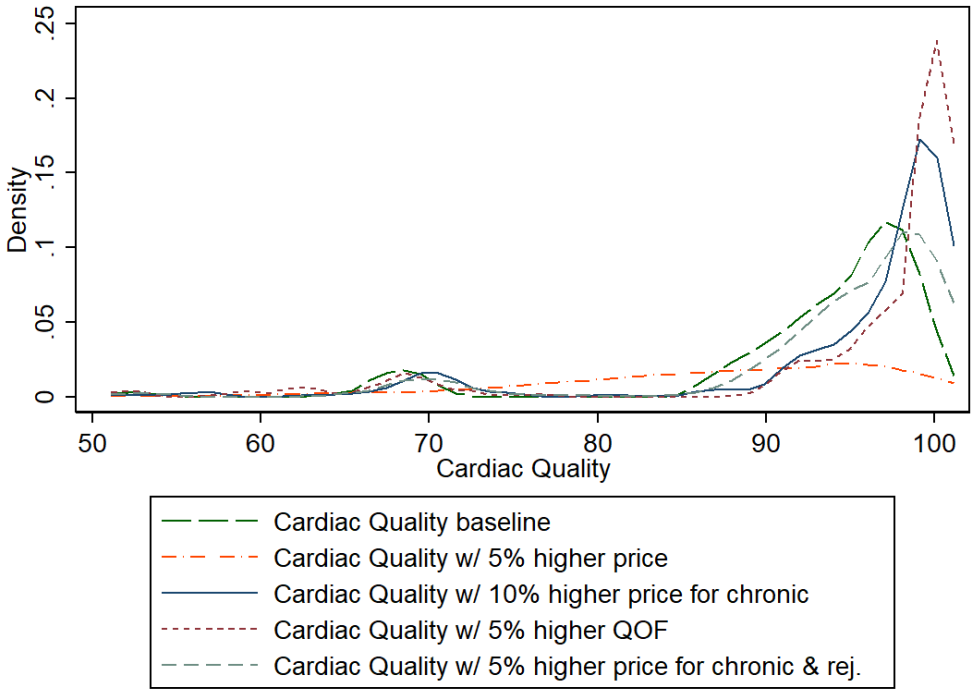


Figure 11: Cardiac quality distribution under different reform scenarios

	Δwelfare for Leeds in 2010-2012 in GBP						
	ΔCS_H	ΔCS_K	$\Delta Profits$	$\Delta Costs$	$\Delta Gov.Exp.$	$\Delta Welfare$	$\frac{\Delta CS_H + \Delta CS_K}{\Delta Gov.Exp.}$
Panel A: Uniform price increase							
$+5\%\bar{p}$ for all patients	4.7m	1.8m	6.9m	0.4m	7.3m	3.9m	0.9
Panel B: Risk Adjustment							
Higher $\bar{p_a}$ Resp. $\bar{p_a}$ Cardiac	2.1m	1.0m	0.9m	0.2m	1.1m	2.6m	1.2
Panel C: Increase in QOF point value							
$+10\% QOF$ point	-0.01m	0.05m	0.05m	0.005m	0.06m	0.2m	0.7
Panel D: Risk Adjustment and rejections							
$+10\% QOF$ point	4.1m	1.9m	0.3m	0.4m	0.6m	5.4m	10

Note: Welfare = Consumer Surplus CS_H + Consumer Surplus CS_K - practice costs - $\lambda Gov.Exp.$
Note: The cost of raising public funds $\lambda = 0.3$. Marginal utility of income for consumers ≈0.011