



Universidade do Minho

Documentos de Trabalho
Working Paper Series

“Competition and Equity in Health Care Markets”

Luigi Siciliani
Odd Rune Straume
NIPE WP 01/ 2018

NÚCLEO DE INVESTIGAÇÃO EM POLÍTICAS ECONÓMICAS
UNIVERSIDADE DO MINHO

“Competition and Equity in Health Care Markets”

Luigi Siciliani
Odd Rune Straume

NIPE* WP 01/2018

URL:

<http://www.nipe.eeg.uminho.pt/>



«This work was carried out within the funding with COMPETE reference nº POCI-01-0145-FEDER-006683 (UID/ECO/03182/2013), with the FCT/MEC's (Fundação para a Ciência e a Tecnologia, I.P.) financial support through national funding and by the ERDF through the Operational Programme on "Competitiveness and Internationalization – COMPETE 2020 under the PT2020 Partnership Agreement»

Competition and Equity in Health Care Markets*

Luigi Siciliani[†]

Odd Rune Straume[‡]

January 2018

Abstract

We provide a model where hospitals compete on quality under a fixed price regime to investigate (i) whether hospital competition, as measured by an increase in fixed prices or increased patient choice, increases or reduces the gap in quality between high- and low-quality hospitals, and as a result, (ii) whether competition increases or reduces (pure) health inequalities across hospitals and patient severities. The answer to the first question is generally ambiguous, but we find that the scope for competition to result in quality convergence across hospitals is larger when the marginal patient health gains from quality decrease at a faster rate. Whether competition increases health inequalities depends on the type and measure of inequality. If the patient health benefit function is not too concave in quality, health inequalities due to postcode lottery will increase (decrease) whenever competition induces quality dispersion (convergence). Competition reduces health inequalities between high- and low-severity patients if patient composition effects, due to high-severity patients being more likely to exercise choice, are small. We also investigate the effect of competition on health inequalities as measured by the Gini and the Generalised Gini coefficients, and highlight differences compared to the simpler dispersion measures.

Keywords: Hospital competition; quality; health inequalities; Gini coefficient.

JEL Classification: I11, I14, L13

*Odd Rune Straume acknowledges funding from COMPETE (ref. no. POCI-01-0145-FEDER-006683), with the FCT/MEC's (Fundação para a Ciência e a Tecnologia, I.P.) financial support through national funding and by the ERDF through the Operational Programme on Competitiveness and Internationalization – COMPETE 2020 under the PT2020 Partnership Agreement.

[†]Department of Economics and Related Studies, University of York, Heslington, York YO10 5DD, UK; E-mail: luigi.siciliani@york.ac.uk

[‡]Corresponding author. Department of Economics/NIPE, University of Minho, Campus de Gualtar, 4710-057 Braga, Portugal; and Department of Economics/HEB, University of Bergen. E-mail: o.r.straume@eeg.uminho.pt

1 Introduction

Recent and ongoing reforms in several OECD countries aim at stimulating competition and patient choice among publicly-funded hospitals in order to improve quality of care (EXPH, 2015; OECD, 2012). In the U.S. Medicare and Medicaid programmes, hospitals are paid by Diagnosis Related Group (DRG) since 1983. Medicare and Medicaid cover respectively individuals older than 65 years old and poor patients. The DRG system involves paying a fixed tariff for every patient treated. In the United Kingdom, under a policy commonly known as ‘Payment by Results’, hospitals are also paid a tariff for every patient treated, and patients are free to choose the hospital. Hospital competition is also present in other countries such as Denmark, France, Germany, Italy, and Norway. The idea is that hospitals ‘compete’ on quality to attract patients and are rewarded financially for doing so.

Opponents of hospital competition argue that these policies will harm equity. For example, high-quality hospitals will respond to competition by improving even more, while low-quality hospitals will be left behind. A recent report by the European Commission highlights that despite the extensive literature investigating the effect of competition in the health sector, there is very limited literature focusing on its equity implications (EXPH, 2015). Reduction in health inequalities are an ubiquitous policy objective, and it is surprising that it has received little attention in relation to competition. We contribute to fill this gap in knowledge.

In this study we extend the received theoretical literature by investigating (i) whether competition increases or reduces the gap in quality between high- and low-quality hospitals, and (ii) whether, as a result, competition increases or reduces health inequalities. We focus on two dimensions of (pure) health inequalities (Wagstaff and van Doorslaer, 2000, section 5). The first type of health inequalities is what is commonly known, in the hospital context, as inequalities due to ‘postcode lottery’: a patient living close to a given hospital might receive much poorer quality compared to a patient living close to a good hospital (Dalton, 2014, p.4). The second type of health inequalities relates to disparities in health across patients with different severity: if high-severity patients benefit less from competition than low-severity patients, health inequalities will worsen. The equity concern across severity groups is regularly reflected in sub-group analysis (by severity type) in cost-effectiveness analysis (Sculpher and Gafni, 2001). Given that we have two sources of health inequalities, we also investigate how competition affects the Gini coefficient, a commonly used measure to empirically assess health inequalities within or across countries (Wagstaff and van

Dooerslaer, 2000).

Our choice of theoretical framework is a Hotelling model with two hospitals competing on quality and located at the extremes of a unit line. In this respect we follow the existing theoretical literature, where quality competition is typically analysed within a spatial competition framework. We allow one hospital to have a comparative advantage so that hospitals provide different qualities in equilibrium. We focus on two measures of competition: (i) an increase in the fixed price, and (ii) a reduction in transportation costs. A higher price increases the profitability of attracting more patients, and therefore stimulates the hospitals to compete on quality. A reduction in transportation costs can be interpreted as an increase in patient choice, where patients are encouraged to choose hospitals based on quality and therefore stimulates competition. Patient choice can be enhanced by the introduction or the enhancement of public reporting of quality indicators (Siciliani, Chalkley and Gravelle, 2017).

Our key findings are as follows. Whether competition increases or reduces quality differences across hospitals is generally ambiguous, and depends on two key factors related to the demand for health care and the cost of health care provision, namely (i) the marginal health gains from quality and (ii) the extent to which quality affects marginal treatment costs. The answer also depends, to some extent, on how we measure competition, whether by an increase in the fixed price or by an increase in the degree of patient choice. Our most clear-cut result is that quality convergence across hospitals is a more likely effect of increased competition, regardless of how it is measured, if marginal health gains decrease with quality at a faster rate.

Whether competition increases health inequalities depends on the type of inequality and the effect does not necessarily have the same sign as the change in hospital quality differences. If health gains are linear or not too concave in quality, health inequalities due to postcode lottery go hand in hand with health inequalities: they will increase (decrease) whenever competition induces quality dispersion (convergence). But if health gains are concave in quality to a sufficient degree, then health inequalities can reduce even if competition induces quality dispersion, and they will always reduce if competition induces quality convergence. Competition generally reduces health inequalities between high- and low-severity patients, because high-severity patients benefit more from higher quality than do low-severity patients. However, this reduction can be strengthened or weakened by what we refer to as ‘composition effects’, which relate to competition inducing high-severity patients to exercise choice more than low-severity patients by selecting hospitals with

higher quality.

We then derive the effect of competition on aggregate measures of absolute and relative inequality, namely the Generalised Gini and Gini coefficients, respectively. These measures are conceptually distinct from the above-mentioned measures of dispersion across hospitals and severity groups. Consider for example the case with just one severity group. Even if competition increases differences in health outcomes across hospitals (an increase in inequalities due to postcode lottery), the Generalised Gini coefficient may still reduce if competition induces more patients to go to the high-quality hospital. Similarly, if competition has no effect on differences on health outcomes, the Gini coefficient will still reduce as a result of the overall increase in quality. With two severity groups, numerical simulations based on two different parameterisations of the model suggest that competition (whether measured by price or patient choice) tends to reduce both absolute and relative inequality when the shares of high- and low-severity patients are not too different. One of the main driving forces is that competition tends to reduce inequalities between high- and low-severity patients, regardless of how competition is measured and regardless of whether competition leads to quality dispersion or quality convergence.

In line with the existing literature, our theoretical model assumes that hospitals are profit maximisers and suggests that an increase in competition increases quality (Ma and Burgess, 1993; Wolinsky, 1997; Gravelle, 1999; Beitia, 2003; Nuscheler, 2003; Brekke, Nuscheler and Straume, 2006, 2007; Gaynor, 2006; Karlsson, 2007). This result also holds with altruistic providers but only if the degree of altruism is not too high (Brekke, Siciliani and Straume, 2011, 2012; see also Barigozzi and Burani, 2016).

The seminal empirical study by Kessler and McClellan (2000) suggests that competition increases quality. This result is also confirmed by Tay (2003), but only partially by Shen (2003) while Gowrinsankaran and Town (2003) find a negative effect. The latest evidence from England suggests that competition, as measured by the introduction of patient choice policies, increases quality under different empirical approaches (Cooper et al., 2011; Gaynor et al., 2013; Bloom et al., 2015). There is only one empirical study which directly tests the effect of competition on equity. Cookson et al. (2013) find that competition did not harm equity, as measured by difference in hip replacement utilisation across socioeconomic status in England. This study is not directly relevant for us given the focus on utilisation as opposed to quality and health outcomes, and the focus on socioeconomic inequalities as opposed to pure health inequalities. Although not focussing on equity,

Kessler and Geppert (2005) find that competition improved health for high-severity patients but not for low-severity patients, therefore providing indirect evidence that health inequalities across severity groups reduced. Some empirical studies (Dafny, 2005; Farrar et al., 2009) also test the effect of price changes on quality, but none of them focuses on equity implications. Our approach is positive rather than normative. Although we could derive the optimal pricing rule set by a welfare maximising regulator, in reality hospital prices are fixed and are set to reflect average treatment costs. We therefore prefer to investigate how competition affects health inequalities under current common financial arrangements.

The study is organised as follows. In Section 2, we present the model and derive equilibrium quality. In Section 3, we investigate how competition affects quality differences across hospitals, and in Section 4, how competition affects health inequalities. Section 5 draws implications for empirical analyses. Section 6 concludes the study.

2 Model

Consider a market for a healthcare treatment (e.g., a coronary bypass or a hip replacement) offered by two different providers (hospitals), located at opposite endpoints of a Hotelling line of length 1. Demand comes from a unit mass of patients who are uniformly distributed on the line. At each point of the line there is a share λ of high-severity patients, denoted by h . The remaining patients have lower severity and are denoted by l . A patient of type k who is treated at Hospital i has the following utility:

$$U_i^k(q_i) = B^k(q_i) - td, \quad k = h, l; \quad i = 1, 2, \quad (1)$$

where $B^k(\cdot)$ is the (expected) health status of a patient with severity k following healthcare treatment; $q_i \geq \underline{q}$ is the quality of treatment at Hospital i ; d is the distance travelled by the patient, and t is the marginal cost of travelling. The lower bound \underline{q} on quality represents the minimum treatment quality that the hospitals are allowed to offer, and we can interpret the case of $q_i < \underline{q}$ as malpractice. We assume that: (i) for a given level of treatment quality, the patient with higher severity is in worse health, even after treatment, $B^h(q) < B^l(q)$; and (ii) the patient with higher severity benefits more from a marginal increase in treatment quality, i.e. $\partial B^h / \partial q > \partial B^l / \partial q > 0$ for all q . Thus, for a given level of treatment quality, the difference in health status across high- and low-severity patients is smaller the higher the quality of treatment.

Under the assumption of unit demand and full market coverage, utility-maximising behaviour leads to the following demand functions for high- and low-severity patients, respectively, at Hospital i :

$$D_i^h := \lambda \left(\frac{1}{2} + \frac{B^h(q_i) - B^h(q_j)}{2t} \right), \quad (2)$$

$$D_i^l := (1 - \lambda) \left(\frac{1}{2} + \frac{B^l(q_i) - B^l(q_j)}{2t} \right), \quad (3)$$

where $i = 1, 2$, $j = 1, 2$, and $i \neq j$. Total demand for Hospital i is then

$$D_i = D_i^h + D_i^l, \quad (4)$$

while total demand for Hospital j is $D_j = 1 - D_i$.

Each hospital is assumed to maximise profits. Under the assumption that the (regulated) price p is the same for both types of patients (e.g., DRG tariff for a coronary bypass), profits of Hospital i are given by

$$\pi_i = \left(p - c_i^h(q_i) \right) D_i^h + \left(p - c_i^l(q_i) \right) D_i^l - C(q_i), \quad (5)$$

where $c_i^k(q_i)$ is the marginal cost of treating a patient with severity k , and $C(q_i)$ is the fixed cost of quality (e.g., MRI machines). We assume that the fixed cost of quality increases with quality at an increasing rate, $\partial C / \partial q_i > 0$ and $\partial^2 C / \partial q_i^2 > 0$, that the marginal treatment cost increases (weakly) with quality, $\partial c_i^k(q_i) / \partial q_i \geq 0$, and that the cost of treating a high-severity patient is higher than the cost of treating a low-severity patient, $c_i^h(q_i) > c_i^l(q_i)$ for all q_i . We also assume that hospitals differ in marginal treatment costs, with Hospital 1 having a cost advantage: $c_1^k(q_1) < c_2^k(q_2)$ and $\partial c_1^k(q_1) / \partial q_1 \leq \partial c_2^k(q_2) / \partial q_2$ for $q_1 = q_2$.

The hospitals simultaneously choose qualities in a non-cooperative one-shot game. We consider an interior-solution Nash equilibrium in which both hospitals choose treatment quality above the minimum level. This Nash equilibrium is implicitly characterised by a pair of first-order conditions, given by¹

$$\frac{\partial \pi_i}{\partial q_i} = \sum_k \left(p - c_i^k(q_i^*) \right) \frac{\partial D_i^k(q_i^*, q_j^*)}{\partial q_i} - \sum_k \frac{\partial c_i^k(q_i^*)}{\partial q_i} D_i^k(q_i^*, q_j^*) - \frac{\partial C(q_i^*)}{\partial q_i} = 0, \quad (6)$$

¹Second-order and stability conditions are given in the Appendix.

where

$$\frac{\partial D_i^h}{\partial q_i} = \frac{\lambda}{2t} \frac{\partial B^h}{\partial q_i}; \quad \frac{\partial D_i^l}{\partial q_i} = \frac{1 - \lambda}{2t} \frac{\partial B^l}{\partial q_i}. \quad (7)$$

Given our assumptions on the hospitals' cost functions, the Nash equilibrium is asymmetric and the hospital with a cost advantage provides a higher quality, $q_1^* > q_2^*$.

3 Competition and quality differences

What is the effect of competition on quality provision? In particular, does fiercer competition reduce or amplify quality dispersion between the hospitals? Our modelling framework allows us to consider two different policy measures that stimulate competition: (i) more high-powered financial incentives in the form of a higher treatment price, p , and (ii) increased patient choice, which is captured by a reduction in the transportation cost parameter, t . For ease of exposition, we refer to the degree of patient choice (i.e., a positive measure of competition, as for an increase in price) as $r := -t$. Higher r could for example be due to policies which implement public reporting of quality measures in the public domain. The former policy makes it more profitable to attract patients, whereas the latter policy makes demand more responsive to quality changes. In both cases, incentives for competition are intensified.

3.1 Higher treatment price p

It is possible to show (see Appendix) that a higher price leads to higher quality for both hospitals in equilibrium: $\partial q_i^* / \partial p > 0$, $i = 1, 2$. The main effect is that a higher price increases the price-cost margin and therefore makes it more profitable for each hospital to attract more patients by providing higher quality. If the marginal treatment costs increase with quality (i.e., if $\partial c_i^k(q_i) / \partial q_i > 0$), this effect will be reinforced by competition due to qualities being strategic complements.²

These effects are well known from previous literature. In this study we are interested in investigating whether the price increase amplifies or reduces equilibrium quality differences, defined by $\Delta := q_1^* - q_2^*$. Using (A5)-(A6) in the Appendix, this effect is given by

$$\frac{\partial \Delta}{\partial p} = \frac{1}{H} \left[\frac{\partial D_2}{\partial q_2} \left(\frac{\partial^2 \pi_1}{\partial q_1^2} + \frac{\partial^2 \pi_1}{\partial q_2 \partial q_1} \right) - \frac{\partial D_1}{\partial q_1} \left(\frac{\partial^2 \pi_2}{\partial q_2^2} + \frac{\partial^2 \pi_2}{\partial q_1 \partial q_2} \right) \right], \quad (8)$$

²If marginal treatment costs depend positively on quality, higher quality by Hospital i will reduce the marginal cost of quality provision for Hospital j through lower demand. As a result, Hospital j will respond by increasing its quality as well.

where the expressions for $H > 0$, $\partial^2\pi_i/\partial q_j\partial q_i \geq 0$ and $\partial^2\pi_i/\partial q_i^2 < 0$ are given in the Appendix. The sign of $\partial\Delta/\partial p$ is generally ambiguous. It depends on the difference between the demand responsiveness to quality of Hospital 2, weighted by the sum of the degree of concavity of the profit function of Hospital 1 and its degree of profit complementarity in qualities, and the demand responsiveness of Hospital 1, similarly weighted.

The condition for whether a higher price leads to *quality convergence* or *quality dispersion* can be more extensively stated as follows:

$$\frac{\partial\Delta}{\partial p} < (>)0 \quad \text{if}$$

$$\bar{B}_2 \left(\frac{\partial^2 C}{\partial q_1^2} + \bar{c}_1 + \psi_1 + \bar{p}_1 \right) > (<) \bar{B}_1 \left(\frac{\partial^2 C}{\partial q_2^2} + \bar{c}_2 + \psi_2 + \bar{p}_2 \right), \quad (9)$$

where

$$\bar{B}_1 := \lambda \frac{\partial B^h}{\partial q_1} + (1 - \lambda) \frac{\partial B^l}{\partial q_1} > 0, \quad \bar{B}_2 := \lambda \frac{\partial B^h}{\partial q_2} + (1 - \lambda) \frac{\partial B^l}{\partial q_2} > 0, \quad (10)$$

$$\bar{p}_1 := - \left[(p - c_1^h) \frac{\lambda}{2t} \frac{\partial^2 B^h}{\partial q_1^2} + (p - c_1^l) \frac{1 - \lambda}{2t} \frac{\partial^2 B^l}{\partial q_1^2} \right] > 0, \quad (11)$$

$$\bar{p}_2 := - \left[(p - c_2^h) \frac{\lambda}{2t} \frac{\partial^2 B^h}{\partial q_2^2} + (p - c_2^l) \frac{1 - \lambda}{2t} \frac{\partial^2 B^l}{\partial q_2^2} \right] > 0, \quad (12)$$

$$\bar{c}_1 := \frac{\partial^2 c_1^h}{\partial q_1^2} D_1^h + \frac{\partial^2 c_1^l}{\partial q_1^2} D_1^l \geq 0, \quad \bar{c}_2 := \frac{\partial^2 c_2^h}{\partial q_2^2} D_2^h + \frac{\partial^2 c_2^l}{\partial q_2^2} D_2^l \geq 0, \quad (13)$$

$$\psi_1 := \frac{\lambda}{2t} \frac{\partial c_1^h}{\partial q_1} \left(2 \frac{\partial B^h}{\partial q_1} - \frac{\partial B^h}{\partial q_2} \right) + \frac{1 - \lambda}{2t} \frac{\partial c_1^l}{\partial q_1} \left(2 \frac{\partial B^l}{\partial q_1} - \frac{\partial B^l}{\partial q_2} \right) \geq 0, \quad (14)$$

$$\psi_2 := \frac{\lambda}{2t} \frac{\partial c_2^h}{\partial q_2} \left(2 \frac{\partial B^h}{\partial q_2} - \frac{\partial B^h}{\partial q_1} \right) + \frac{1 - \lambda}{2t} \frac{\partial c_2^l}{\partial q_2} \left(2 \frac{\partial B^l}{\partial q_2} - \frac{\partial B^l}{\partial q_1} \right) > 0. \quad (15)$$

We can interpret \bar{B}_i as the expected marginal health gain from quality in Hospital i across both severity types; \bar{c}_i as the degree of convexity of marginal treatment costs with respect to quality in Hospital i , and \bar{p}_i as Hospital i 's profit margin, weighted by the degree of concavity of the demand function.

Under the condition of equilibrium stability, both sides of the inequality in (9) are positive. Since by assumption Hospital 1 provides a higher quality, the marginal health gain from quality, and therefore demand responsiveness, is (weakly) higher in Hospital 2 relative to Hospital 1, $\bar{B}_2 \geq \bar{B}_1$ (with a strict inequality if $B^k(\cdot)$ is strictly concave). In turn, this tends to give a stronger incentive for Hospital 2 to increase quality, relative to Hospital 1, therefore reducing dispersion in qualities

across hospitals.

However, changes in quality also affect the marginal profitability of quality (i.e., the degree of concavity of the profit function) and the degree of complementarity in qualities across hospitals, which are captured by the terms in brackets in (8). For example, if the marginal profitability of quality is higher for Hospital 1 as a result of the price and quality increases, then Hospital 1 may increase quality more than Hospital 2. In order to further characterise the other relevant terms, we introduce some additional assumptions:

A1 The weighted profit margin is higher for the provider with a competitive advantage, $\bar{p}_1 \geq \bar{p}_2$.

A2 The fixed-quality-cost function $C(\cdot)$ is quadratic.

A3 The marginal treatment cost, $c_i^k(q_i)$, is linear in quality.

A1 holds if the equilibrium profit margin of the high-quality hospital is sufficiently large relative to the low-quality hospital. If $\partial^3 B^k / \partial q_i^3 \leq 0$, A1 holds as long as the marginal treatment cost advantage of Hospital 1 is not overturned in equilibrium (i.e., $c_1^k(q_1^*) \leq c_1^k(q_2^*)$). Furthermore, A2 essentially rules out the possibility that the sign of (9) is determined by potential mechanisms related to the sign of the third-order derivative of the fixed-quality-cost function. Finally, A3 ensures that the degree of convexity of marginal treatment costs with respect to quality is the same across the two hospitals, which implies $\bar{c}_1 = \bar{c}_2 = 0$ and $\psi_1 \leq \psi_2$. After deriving results that hold under A1-A3, we will briefly discuss the implications of relaxing A2.

Applying A1-A3, the condition in (9) reduces to

$$\frac{\partial \Delta}{\partial p} < (>) 0 \quad \text{if} \quad \frac{\bar{B}_2}{\bar{B}_1} > (<) \frac{k + \psi_2 + \bar{p}_2}{k + \psi_1 + \bar{p}_1}, \quad (16)$$

where $k := \partial^2 C / \partial q_1^2 = \partial^2 C / \partial q_2^2$. The general ambiguity remains, but we can now more precisely characterise each of the two possibilities by identifying the following set of sufficient conditions:³

Proposition 1 *Given assumptions A1-A3:*

(i) *If the marginal health gain from quality is strictly decreasing, $\partial^2 B^k / \partial q_i^2 < 0$, and if the effect of quality on marginal treatment costs, $\partial c_i^k / \partial q_i$, is sufficiently small, a higher price leads to quality convergence in equilibrium, $\partial \Delta / \partial p < 0$.*

³Proofs of Propositions 1 and 2 are given in the in Appendix.

(ii) If the marginal health gain from quality is constant or decreases slowly with quality, and if the high-quality hospital has a cost advantage in the provision of quality in equilibrium, $\partial c_1^k(q_1^*)/\partial q_1 < \partial c_2^k(q_2^*)/\partial q_2$, a higher price leads to quality dispersion in equilibrium, $\partial\Delta/\partial p > 0$.

(iii) If the marginal health gain from quality is constant, and if marginal treatment costs are constant, $\partial c_i^k/\partial q_i = 0$, a higher price does not affect quality differences in equilibrium, $\partial\Delta/\partial p = 0$.

The proposition highlights the two main mechanisms at work: (i) the concavity of the health benefit function, which determines the relative magnitudes of \bar{B}_1 and \bar{B}_2 , and (ii) the differences in the effect of quality on marginal treatment costs, which determines the relative magnitudes of ψ_1 and ψ_2 .

The former mechanism contributes to quality convergence in response to a price increase. A strictly concave health benefit function implies that the marginal health benefit of quality is higher for patients in the low-quality hospital, which in turn implies that demand responds more strongly to quality for this hospital. A price increase will therefore lead to a larger increase in the marginal revenue of quality for Hospital 2 than for Hospital 1, contributing, all else equal, to quality convergence between the two hospitals. However, this effect is counteracted by differences in the effects of quality on marginal treatment costs, which relate to the degree of profit concavity and of profit complementarity in qualities. If quality has a smaller effect on marginal treatment costs in Hospital 1, this hospital has a stronger incentive to increase quality in response to a price increase, leading, all else equal, to higher quality dispersion between the two hospitals.

In addition to these two counteracting effects, which are given by the relative magnitudes of \bar{B}_2/\bar{B}_1 and ψ_2/ψ_1 , respectively, there is the effect of $\bar{p}_1 \geq \bar{p}_2$. This effect, which is also related to differences in the degree of concavity of the profit functions of the two hospitals, works in the direction of quality convergence in response to a price increase. Thus, in broad terms, demand effects (mainly through demand responsiveness) tend to induce quality convergence, whereas cost advantages in quality provision tend to induce quality dispersion.

Assumption A2 eliminates potential effects due to the convexity of the fixed costs of quality. To highlight and isolate the role played by the fixed costs of quality, suppose that health benefits are linear in quality and treatment costs are constant (but lower for Hospital 1 than for Hospital 2). This implies $\bar{B}_1 = \bar{B}_2$ and $\bar{p}_i = \bar{c}_i = \psi_i = 0$, which means that (9) reduces to

$$\frac{\partial\Delta}{\partial p} > (<)0 \quad \text{if} \quad \frac{\partial^2 C}{\partial q_2^2} > (<)\frac{\partial^2 C}{\partial q_1^2}. \quad (17)$$

Whether qualities diverge or converge as a result of a price increase depends on the difference in the degree of convexity of the cost function across hospitals (evaluated at equilibrium qualities). Since $q_1^* > q_2^*$, the condition in (17) is equivalent to

$$\frac{\partial \Delta}{\partial p} > (<) 0 \quad \text{if} \quad \frac{\partial^3 C}{\partial q^3} < (>) 0.$$

Thus, if the degree of convexity in fixed costs increases (reduces) with quality, a price increase will induce quality convergence (dispersion).

3.2 Increased patient choice r

In a symmetric model with profit-maximising providers and regulated prices, there is a well-established positive relationship between increased patient choice, measured as a reduction of transportation costs, and equilibrium quality provision. In our asymmetric setting, lower transportation costs have however additional effects on unilateral quality provision incentives. On the one hand, as in a symmetric model, increased patient choice makes demand more quality elastic, which gives both hospitals an incentive to increase quality. On the other hand, for given quality levels, increased patient choice implies that a larger share of each patient type chooses the high-quality hospital. If marginal treatment costs increase with quality, $\partial c_i^k / \partial q_i > 0$, such a reallocation of demand implies higher (lower) marginal cost of quality provision, and therefore weaker (stronger) incentives for quality provision, for the high-quality (low-quality) hospital. However, by applying the first-order conditions, (6), it can be shown (see Appendix) that the former effect dominates the latter, implying that the results from a symmetric model also carry over to an asymmetric one. Increased patient choice leads to higher quality provision in equilibrium for both hospitals: $\partial q_i^* / \partial r > 0$, $i = 1, 2$.

Whether patient choice leads to quality dispersion or quality convergence is *a priori* ambiguous. Using (A10)-(A11) in the Appendix, the exact condition is given by

$$\frac{\partial \Delta}{\partial r} > (<) 0 \quad \text{if} \quad \xi_2 \left(\psi_1 + \bar{p}_1 + \bar{c}_1 + \frac{\partial^2 C_1}{\partial q_1^2} \right) < (>) \xi_1 \left(\psi_2 + \bar{p}_2 + \bar{c}_2 + \frac{\partial^2 C_2}{\partial q_2^2} \right), \quad (18)$$

where \bar{p}_i , \bar{c}_i and ψ_i are defined by (11)-(15), and

$$\xi_1 := \frac{1}{2t} \left(\lambda \frac{\partial c_1^h}{\partial q_1} + (1 - \lambda) \frac{\partial c_1^l}{\partial q_1} + 2 \frac{\partial C}{\partial q_1} \right) > 0, \quad (19)$$

$$\xi_2 := \frac{1}{2t} \left(\lambda \frac{\partial c_2^h}{\partial q_2} + (1 - \lambda) \frac{\partial c_2^l}{\partial q_2} + 2 \frac{\partial C}{\partial q_2} \right) > 0. \quad (20)$$

Compared to the case of a higher price, the criterion for quality convergence or dispersion as a result of lower transportation costs depends somewhat less on the concavity of the health benefit function and somewhat more on the characteristics of the treatment cost function, in particular the relationship between marginal treatment costs and quality. Once more applying A1-A3, the condition in (18) reduces to

$$\frac{\partial \Delta}{\partial r} > (<) 0 \quad \text{if} \quad \frac{\xi_2}{\xi_1} < (>) \frac{k + \psi_2 + \bar{p}_2}{k + \psi_1 + \bar{p}_1}. \quad (21)$$

Notice that, whereas $\psi_2 \geq \psi_1$ and $\bar{p}_1 \geq \bar{p}_2$, the relative magnitudes of ξ_1 and ξ_2 are *a priori* ambiguous. The next proposition establishes a sufficient condition for one of the two possible outcomes:

Proposition 2 *Given assumptions A1-A3, if the marginal health gain from quality is constant or decreases slowly with quality, and if the effect of quality on marginal treatment cost, $\partial c_i^k / \partial q_i$, is sufficiently small, increased patient choice leads to quality dispersion in equilibrium, $\partial \Delta / \partial r > 0$.*

To gain some intuition for this condition, notice that the right-hand sides of (16) and (21) are equal, whereas the left-hand sides are different. Thus, whether considering a price increase or an increase in patient choice, the effects of $\psi_2 \geq \psi_1$ and $\bar{p}_1 \geq \bar{p}_2$ are similar in both cases and work in the directions of quality dispersion and quality convergence, respectively. However, the left-hand side of (21) introduces two new effects that are specific to the case of increased patient choice. Both effects are related to treatment cost advantages but work in opposite directions.

(i) Increased patient choice r implies that demand becomes more responsive to quality, which increases the marginal revenue of quality and gives both hospitals an incentive to increase quality. If $c_1(q_1^*) < c_2(q_2^*)$, the profit margin is higher for Hospital 1, which implies that the increase in marginal revenue of quality, due to more quality-responsive demand, is also higher for Hospital 1, which gives this hospital a stronger incentive to increase quality. This effect contributes, all else

equal, to quality dispersion.

(ii) The increase in demand responsiveness due increased patient choice also implies that, for given qualities, demand is shifted towards the high-quality hospital (i.e., $\partial D_1/\partial r > 0$ and $\partial D_2/\partial r < 0$). If marginal treatment costs depend on quality, the demand increase (decrease) for the high-quality (low-quality) hospital implies that the marginal cost of quality provision increases (decreases) for the high-quality (low-quality) hospital. All else equal, this gives the high-quality (low-quality) hospital an incentive to reduce (increase) quality. This effect works in the opposite direction of (i) and contributes, all else equal, to quality convergence.

The opposite natures of (i) and (ii) contribute to the general ambiguity of (21). However, if the effect of quality on marginal treatment costs is sufficiently small, (ii) becomes irrelevant and so does the effect related to $\psi_2 \geq \psi_1$. Furthermore, if the marginal health gain from quality decreases at a sufficiently slow rate, the effect related to $\bar{p}_1 \geq \bar{p}_2$ also becomes irrelevant. In this case, which is identified by Proposition 2, the only relevant effect is (i). Notice that, if $\partial c_i^k/\partial q_i$ is sufficiently small, the basic assumptions of our model ensures that $c_1(q_1^*) < c_2(q_2^*)$.

Again, assumption A2 eliminates potential effects due to the convexity of the fixed costs of quality. In the special case of constant marginal treatment costs ($c_1 < c_2$) and linear health benefit functions, which implies $\psi_1 = \psi_2 = \bar{p}_1 = \bar{p}_2 = 0$, the condition in (21) reduces to

$$\frac{\partial \Delta}{\partial r} > 0 \quad \text{if} \quad \frac{\partial C}{\partial q_1} > \frac{\partial C}{\partial q_2}, \quad (22)$$

which, due to the strict convexity of fixed costs of quality C , is always true for $q_1^* > q_2^*$.⁴ Differently from an increase in competition through an increase in price, increased patient choice always induces quality dispersion under constant marginal treatment costs and linear health benefit functions, regardless of the degree of convexity in the fixed costs of quality.

3.3 Parametric examples

In order to illustrate the general results stated in Propositions 1 and 2, and to gain some additional insights regarding the main mechanisms of the model, we proceed by exploring some parametric examples.

⁴Notice that the expressions for ξ_1 and ξ_2 have been obtained after substitutions using the first-order conditions of the hospitals' maximisation problems (see Appendix). Thus, the second effect identified above, which relies on differences in profit margins across hospitals, is captured by the differences in marginal costs of quality in the expressions for ξ_1 and ξ_2 .

The previous analysis has revealed two key determinants of whether increased competition leads to quality convergence or quality dispersion, namely (i) the degree of concavity of the health benefit function, and (ii) the degree to which quality affects marginal treatment costs. In the following two parametric examples, we will consider each of these two dimensions separately, which also allows us to obtain closed-form solutions. In both examples, we parameterise the fixed costs (of quality) function as follows: $C(q_i) = (k/2)q_i^2$. Furthermore, for simplicity we disregard patient heterogeneity (with respect to severity) by setting $\lambda = 1$. This is without loss of generality, since the share of high-severity versus low-severity patients does not qualitatively affect the relationship between competition and quality dispersion.

Example 1 *Decreasing marginal health gain from quality and constant marginal treatment costs.* Suppose that marginal treatment costs are constant and given by $c_1 < c_2$. We will consider two different parameterisations of the health benefit functions. (i) Suppose that $B(\cdot)$ is quadratic and given by

$$B(q_i) = \alpha + \beta q_i - \frac{\gamma}{2} q_i^2.$$

In this case, equilibrium qualities are given by

$$q_i^* = \frac{\beta(p - c_i)}{\gamma(p - c_i) + 2kt}.$$

The effect of increased competition on quality differences is given by

$$\frac{\partial \Delta}{\partial p} = -\frac{2kt\beta\gamma(c_2 - c_1)(\gamma(2p - c_1 - c_2) + 4kt)}{(\gamma(p - c_1) + 2kt)^2(\gamma(p - c_2) + 2kt)^2} < 0$$

and

$$\frac{\partial \Delta}{\partial r} = \frac{2k\beta(c_2 - c_1)(4k^2t^2 - \gamma^2(p - c_1)(p - c_2))}{(\gamma(p - c_1) + 2kt)^2(\gamma(p - c_2) + 2kt)^2} < (>) 0$$

if $\gamma > (<) \frac{2kt}{\sqrt{(p - c_1)(p - c_2)}}$.

(ii) Suppose that $B(\cdot)$ is logarithmic and given by

$$B(q_i) = \alpha + \beta \ln q_i.$$

In this case equilibrium qualities are given by

$$q_i^* = \frac{\sqrt{2}}{2kt} \sqrt{kt\beta(p - c_i)}.$$

The effect of increased competition on quality dispersion is in this case given by

$$\frac{\partial \Delta}{\partial p} = - \frac{\sqrt{2} \sqrt{kt\beta(p - c_1)} \sqrt{kt\beta(p - c_2)} \left(\sqrt{kt\beta(p - c_1)} - \sqrt{kt\beta(p - c_2)} \right)}{4k^2 t^2 \beta (p - c_1) (p - c_2)} < 0$$

and

$$\frac{\partial \Delta}{\partial r} = \frac{\sqrt{2} \left(\sqrt{kt\beta(p - c_1)} - \sqrt{kt\beta(p - c_2)} \right)}{4kt^2} > 0.$$

In the benchmark case of constant marginal health gains of quality and constant marginal treatment costs, we know from Propositions 1 and 2 that a higher price has no effect on quality dispersion whereas increased patient choice lead to increased quality dispersion. Example 1 illustrates that this is changed by the introduction of a concave health benefit function. Whether the health benefit function is quadratic or logarithmic, a higher price always leads to quality convergence, which is consistent with Proposition 1. In the case of a quadratic health benefit function, increased patient choice also leads to quality convergence if the degree of concavity (measured by the parameter γ) is sufficiently large, which is consistent with Proposition 2. In this case, different competition measures have the same effect on the sign of quality differences.

However, if the health benefit function is logarithmic (or quadratic with a low degree of concavity), increased patient choice always leads to quality dispersion, and different competition measures have opposite effects on hospital quality differences. This is explained by the fact that $\bar{p}_1 = \bar{p}_2$ in this parametric example, which implies that the only mechanism in play is the effect related to the higher profit margin of Hospital 1 (which implies $\xi_1 > \xi_2$).

Overall, Example 1 illustrates that the presence of decreasing marginal health gains of quality increases the scope for quality convergence as a result of more competition, but also that different competition measures can have different effect on quality dispersion across hospitals. It is also worth noticing that A1 holds for both quadratic and logarithmic health benefits, even though the logarithmic form implies $\partial^3 B / \partial q_i^3 \geq 0$, which suggests that this assumption is not overly restrictive.

Example 2 *Constant marginal health gains of quality and quality-dependent marginal*

treatment costs. Suppose that the health benefit function is given by

$$B(q_i) = \alpha + \beta q_i,$$

and that marginal treatment costs are given by

$$c(q_i) = c_i q_i,$$

where $c_1 < c_2$. In this case, equilibrium qualities are given by

$$q_i^* = \frac{2kt(p\beta - tc_i) + \beta(p\beta(c_i + 2c_j) - 3tc_i c_j)}{3\beta^2 c_i c_j + 4kt(\beta(c_i + c_j) + kt)}.$$

The effect of increased competition on quality dispersion is given by

$$\frac{\partial \Delta}{\partial p} = \frac{(c_2 - c_1)\beta^2}{3\beta^2 c_1 c_2 + 4kt(\beta(c_1 + c_2) + kt)} > 0$$

and

$$\frac{\partial \Delta}{\partial r} = \frac{4k\beta(c_2 - c_1)(p\beta(\beta(c_1 + c_2) + 2kt) - t(2kt(c_1 + c_2) + 3\beta c_1 c_2))}{(3\beta^2 c_1 c_2 + 4kt(\beta(c_1 + c_2) + kt))^2} > (<) 0$$

if $p > (<) \frac{t(2kt(c_1 + c_2) + 3\beta c_1 c_2)}{\beta(\beta(c_1 + c_2) + 2kt)}$.

Consistent with Proposition 1, the combination of constant marginal health gains and quality-dependent marginal treatment costs imply that a higher price leads to quality dispersion. Increased patient choice yields the same outcome if the price is sufficiently high relative to marginal treatment costs. Thus, even if marginal treatment costs are strongly affected by quality (i.e., c_i is large), increased patient choice will nevertheless lead to higher quality dispersion if the price is sufficiently large. In this case, the dominating mechanism is the one that is caused by the profit margin being higher for the high-quality hospital, as explained in Section 3.3.

The main insights from the above examples are summarised in Table 1. The most clear-cut conclusion that can be drawn is that the scope for increased competition to instigate quality convergence increases with the concavity of the health benefit function. This applies in particular to the case of increased competition induced by price increases. However, there is still a relatively

wide range of parameter configurations for which price increases and increased patient choice have opposite effects on quality differences.

Table 1. The effects of competition on quality differences

<i>Health gain</i>	<i>Marginal costs</i>	<i>Increase in price</i>	<i>Increase in patient choice</i>
Linear	Constant	No effect	Dispersion
Linear	Linear in quality	Dispersion	Dispersion (convergence) if price is high (low)
Logarithmic	Constant	Convergence	Dispersion
Quadratic	Constant	Convergence	Convergence (dispersion) if γ is high (low)*

*The parameter γ denotes the degree of concavity of the health benefit function.

4 Competition and health inequalities

In the previous section we have characterised the conditions under which competition induces a reduction or an increase in inequalities in the level of *quality* across hospitals, which we have referred to as quality convergence and quality dispersion, respectively.

In this section we investigate how competition affects *health* inequalities. In our model we have four groups of patients who differ in severity and the provider from which they receive the treatment, and we answer this question in three steps. First, we look at inequalities in health outcomes across hospitals. These can be thought of as inequalities arising from the ‘postcode lottery’: some patients will have worse health outcomes than others simply because they live closer to a low-quality hospital. Second, we look at inequalities in health outcomes between patients with high and low severity, and check whether competition increases or reduces the health gap between the two patient groups. Third, we look at aggregate measures of (relative and absolute) health inequality based on the Gini and Generalised Gini coefficients since these have been commonly used in the health economics empirical literature to measure health inequalities.

4.1 Absolute health inequalities across hospitals (postcode lottery)

When considering health inequalities across hospitals, we restrict attention to inequalities within each patient type. As long as health outcomes (e.g., mortality rates) are risk adjusted, the analysis would be similar in the presence of patients with different severity. For a given level of severity, the difference in health outcomes of patients being treated at Hospital 1 and 2, respectively, is given by

$$\Omega^k := B^k(q_1^*) - B^k(q_2^*), \quad k = h, l, \text{ with } \Omega^h > \Omega^l. \quad (23)$$

The effect of competition on health inequalities is consequently given by

$$\frac{\partial \Omega^k}{\partial p} = \frac{\partial B^k}{\partial q_1} \frac{\partial q_1^*}{\partial p} - \frac{\partial B^k}{\partial q_2} \frac{\partial q_2^*}{\partial p}, \quad (24)$$

$$\frac{\partial \Omega^k}{\partial r} = \frac{\partial B^k}{\partial q_1} \frac{\partial q_1^*}{\partial r} - \frac{\partial B^k}{\partial q_2} \frac{\partial q_2^*}{\partial r}. \quad (25)$$

If competition induces quality *convergence*, i.e., if it reduces inequalities in quality across hospitals in addition to raising quality in both hospitals, then competition also reduces health inequalities. If the marginal health gain from quality is constant, inequalities are driven by differences in quality. This effect is reinforced if the marginal health gain from quality is decreasing and therefore smaller in the hospital with higher quality. Reductions in inequalities in quality always reduce health inequalities.

If competition induces quality *dispersion*, i.e., if it increases inequalities in quality across hospitals, then the effect of competition on health inequalities is instead ambiguous. It is only when the health gain from quality is linear or not too concave that inequalities in levels of quality go hand-in-hand with health inequalities, so that quality dispersion increases health inequalities. If the marginal health gain from quality is decreasing, the larger quality increase in Hospital 1 arising from competition can be dampened or even offset by the smaller marginal health gain of quality, and quality dispersion can therefore reduce health inequalities.

Proposition 3 (i) *If competition induces quality convergence, then it reduces health inequalities across hospitals for each severity type. (ii) If competition induces quality dispersion, it increases health inequalities when the health gain from quality is not too concave in quality; it reduces health inequalities when the health gain from quality is sufficiently concave. (iii) If competition has no effect on quality differences across hospitals, it reduces health inequalities if the health gain from*

quality is concave.

The second part of the Proposition 3 can be illustrated by considering the parameterisations used in Example 1 of the previous section. For both types of health benefit functions (quadratic and logarithmic), a higher price leads to quality convergence and therefore (by Proposition 3) to reduced health inequality between patients within each severity type. On the other hand, the effect of increased patient choice on health inequalities is *a priori* ambiguous. With a *quadratic* health benefit function, the effect is given by

$$\frac{\partial \Omega}{\partial r} = \frac{4k^2 t \beta^2 (c_2 - c_1) [8k^3 t^3 - \gamma^2 (p - c_2) (p - c_1) (\gamma (2p - c_1 - c_2) + 6kt)]}{(\gamma (p - c_1) + 2kt)^3 (\gamma (p - c_2) + 2kt)^3}.$$

It is relatively straightforward to see that the sign of this expression is positive (negative) if γ is sufficiently low (high). There are two different forces at play here, both of which work in the same direction. A higher value of γ (which implies a more concave benefit function), increases the scope for increased patient choice leading to less inequality even if differences in quality increase. In addition, a more concave benefit function also increases the scope for quality convergence as a result of increased patient choice, as shown by Example 1.

Consider an illustrative numerical example, with $p = k = t = 2$, $\beta = c_2 = 1$ and $c_1 = 0.5$, which yields the following effects of increased patient choice on quality dispersion and health inequality:

- (i) $\gamma < 3.26$: $\frac{\partial \Delta}{\partial r} > 0$ and $\frac{\partial \Omega}{\partial r} > 0$.
- (ii) $3.26 < \gamma < 6.53$: $\frac{\partial \Delta}{\partial r} > 0$ and $\frac{\partial \Omega}{\partial r} < 0$.
- (iii) $\gamma > 6.53$: $\frac{\partial \Delta}{\partial r} < 0$ and $\frac{\partial \Omega}{\partial r} < 0$.

The interesting case is (ii). When the degree of concavity is in an intermediate range, increased patient choice leads to quality dispersion but simultaneously reduces health inequalities within each severity group, because marginal health gains from quality is decreasing at a sufficiently high rate.

Using instead the logarithmic health benefit function in Example 1, it is straightforward to show that $\partial \Omega / \partial r = 0$. In this case, and regardless of the degree of concavity of the health benefit function, the increase in quality difference across hospitals due to increased patient choice is exactly offset by the counteracting effect of decreasing marginal health gains, leaving health inequalities unaffected.

4.2 Absolute health inequalities between high- and low-severity patients

In this sub-section we investigate how competition affects health inequalities across patient severity. These could be due to patients differing in severity within the same condition or across conditions. For example, for patients who had a heart attack (within the same health condition), high severity patients have a history of heart conditions or other comorbidities. Across conditions, we could think of high-severity patients as patients with cancer as opposed to patients in need of a cataract surgery (low-severity patients).⁵

The average (or expected) health outcome for a high-severity patient is given by

$$\bar{B}^h = \frac{1}{\lambda} \left(D_1^h B^h(q_1^*) + (\lambda - D_1^h) B^h(q_2^*) \right), \quad (26)$$

which can be re-written as

$$\bar{B}^h = \frac{B^h(q_1^*) + B^h(q_2^*)}{2} + \frac{(B^h(q_1^*) - B^h(q_2^*))^2}{2t}. \quad (27)$$

The similar expression for a low-severity patient is

$$\bar{B}^l = \frac{B^l(q_1^*) + B^l(q_2^*)}{2} + \frac{(B^l(q_1^*) - B^l(q_2^*))^2}{2t}. \quad (28)$$

Health inequalities between patient types can then be defined as $\Phi := \bar{B}^l - \bar{B}^h$.

Higher treatment price The effect of a price increase can be expressed as

$$\begin{aligned} \frac{\partial \Phi}{\partial p} &= -\frac{1}{2} \sum_{i=1}^2 \left(\frac{\partial B^h}{\partial q_i} - \frac{\partial B^l}{\partial q_i} \right) \frac{\partial q_i^*}{\partial p} - \frac{1}{t} \left(B^h(q_1^*) - B^h(q_2^*) \right) \left(\frac{\partial B^h}{\partial q_1} \frac{\partial q_1^*}{\partial p} - \frac{\partial B^h}{\partial q_2} \frac{\partial q_2^*}{\partial p} \right) \\ &\quad + \frac{1}{t} \left(B^l(q_1^*) - B^l(q_2^*) \right) \left(\frac{\partial B^l}{\partial q_1} \frac{\partial q_1^*}{\partial p} - \frac{\partial B^l}{\partial q_2} \frac{\partial q_2^*}{\partial p} \right). \end{aligned} \quad (29)$$

The first term is the effect on health inequality for given patient allocations. A higher price leads to higher quality provision at both hospitals. Since the health gain of higher quality is larger for high-severity than for low-severity patients, the inequality in health outcomes between the two patient groups is reduced. Therefore, the first effect is unambiguously negative, and this is regardless of

⁵Although our model has only one price, and therefore implicitly considers only one condition, the effects of competition on health inequalities would be similar in a model with more than one condition as long as the price differences across conditions remain constant.

whether a price increase induces quality convergence or quality dispersion. The last two terms capture the effects of changes in *patient composition* as a result of the price increase, and the sum of these (second-order) effects is *a priori* indeterminate.

If the marginal benefit of quality is decreasing at a sufficiently low rate, the direction of the patient composition effect (i.e., the second and third terms in (29)) is uniquely determined by whether or not competition leads to quality dispersion. To see this, consider the extreme case of linear health benefits, which implies $\partial B^k/\partial q_1 = \partial B^k/\partial q_2 = \partial B^k/\partial q$. The expression in (29) can then be rewritten as

$$\frac{\partial \Phi}{\partial p} = -\frac{1}{2} \sum_{i=1}^2 \left(\frac{\partial B^h}{\partial q} - \frac{\partial B^l}{\partial q} \right) \frac{\partial q_i^*}{\partial p} - \frac{1}{t} \left[\left(B^h(q_1^*) - B^h(q_2^*) \right) \frac{\partial B^h}{\partial q} - \left(B^l(q_1^*) - B^l(q_2^*) \right) \frac{\partial B^l}{\partial q} \right] \frac{\partial \Delta}{\partial p}. \quad (30)$$

By the assumption $\partial B^h/\partial q > \partial B^l/\partial q$, the expression in square brackets is positive. This implies that, if competition leads to quality *dispersion*, i.e., if $\partial \Delta/\partial p > 0$, the first- and second-order effects go in the same direction, and a higher price always leads to less health inequality (on average) between high- and low-severity patients. Thus, in the case of $\partial \Delta/\partial p > 0$, the first-order effect is reinforced by the following second-order effect: Since a higher price increases quality dispersion between the hospitals, and since high-severity patients are more responsive than low-severity patients to changes in quality dispersion, the share of high-severity patients in the high-quality hospital will increase, which further reduces the health inequality between these two groups of patients. This illustrates how increased disparities in quality across hospitals do not necessarily imply increased disparities in health outcomes across patient types. In the above example, the opposite holds. Since it is the most disadvantaged group, i.e., the high-severity patients, who benefit most from differences in qualities across hospitals, health inequalities are actually reduced.

Increased patient choice The effect of an increase in patient choice is given by

$$\begin{aligned} \frac{\partial \Phi}{\partial r} = & -\frac{1}{2} \sum_{i=1}^2 \left(\frac{\partial B^h}{\partial q_i} - \frac{\partial B^l}{\partial q_i} \right) \frac{\partial q_i^*}{\partial r} - \frac{1}{t} \left(B^h(q_1^*) - B^h(q_2^*) \right) \left(\frac{\partial B^h}{\partial q_1} \frac{\partial q_1^*}{\partial r} - \frac{\partial B^h}{\partial q_2} \frac{\partial q_2^*}{\partial r} \right) \\ & + \frac{1}{t} \left(B^l(q_1^*) - B^l(q_2^*) \right) \left(\frac{\partial B^l}{\partial q_1} \frac{\partial q_1^*}{\partial r} - \frac{\partial B^l}{\partial q_2} \frac{\partial q_2^*}{\partial r} \right) \\ & - \frac{1}{2t^2} \left(\left(B^h(q_1^*) - B^h(q_2^*) \right)^2 - \left(B^l(q_1^*) - B^l(q_2^*) \right)^2 \right). \end{aligned} \quad (31)$$

The first three terms are completely equivalent to the previously explained effects of a price increase. However, a change in patient choice also has an additional effect, given by the last term in (31). As previously explained, an increase in patient choice makes demand more sensitive to quality differences, which implies that a relatively larger share of high-severity patients will choose the high-quality hospital. All else equal, this extra effect contributes in the direction of competition leading to less inequality in health outcomes between high- and low-severity patients.

Whether we consider an increase in price or in patient choice, notice that the effect via changes in quality differences, represented by the second and third terms in (29) and (31), respectively, is identical in both cases. If the marginal health benefit of quality is constant, the sum of these two terms are negative, thereby contributing to lower health inequality, if competition leads to quality dispersion. The reason, as previously explained, is that it is the most disadvantaged patient group who benefits more from quality dispersion. By continuity, this holds also for health benefit functions with a sufficiently low degree of concavity, which allows us to summarise the above derived results as follows:

Proposition 4 *(i) An increase in competition (whether by a higher price or by increased patient choice) reduces inequalities across patients with different severity if the subsequent changes in patient composition at each hospital are sufficiently small. (ii) If the marginal health gain from quality is constant or decreases slowly with quality, a sufficient condition for increased competition to reduce inequalities across severity types is that competition leads to quality dispersion.*

When seeing Proposition 4 in conjunction with Propositions 1 and 2, we can further pin down the cases in which competition reduces health inequalities across severity types:

Corollary 1 *Suppose that A1-A3 hold, and suppose that the marginal health gain from quality is constant or decreases slowly with quality. In this case, a price increase will always reduce inequality between patient types, whereas increased patient choice will reduce inequality if the effect of quality on marginal treatment cost is sufficiently small.*

The following Table 2 summarises and illustrate some insights that can be drawn from the model. It shows that (i) changes in inequalities in quality across hospitals go hand-in-hand with changes in health inequalities if the health benefit function is not too concave, (ii) health inequalities across different severity levels reduce if composition effects are small, and (iii) health inequalities

due to postcode lottery can go in the opposite direction of health inequalities between high- and low-severity patients. The latter is exemplified by the case of linear health gain and competition inducing quality dispersion. As a result, while health inequalities due to postcode lottery are increased, health inequalities between high- and low-severity patients are reduced.

Table 2. The effect of competition on health inequalities

<i>Health gain</i>	<i>Quality difference</i>	<i>Health inequalities due to postcode lottery</i>	<i>Health inequalities between high- and low-severity patients</i>
Linear	No effect	Unchanged	Reduced
Linear	Dispersion	Increased	Reduced
Linear	Convergence	Reduced	Reduced**
Concave	No effect	Reduced	Reduced**
Concave	Dispersion	Increased*	Reduced**
Concave	Convergence	Reduced	Reduced**

*If the health benefit function is not too concave. **If composition effects are sufficiently small.

4.3 Aggregate measures of (absolute and relative) health inequality

In the previous subsections we have studied the effect of competition on health inequalities along two different dimensions: (i) inequalities between patients treated at different hospitals (arising from the postcode lottery) and (ii) inequalities between high- and low-severity patients. An aggregate measure of inequality which allows to trade off inequalities along different dimensions is the Gini coefficient, which is also a function of the share of (high/low severity) patients who receive high and low quality. To illustrate the role of the latter we start with a simplified framework with only one severity level, and then extend to two severity levels.

4.3.1 One severity level

With only one severity level, there are only two patient groups, those receiving high quality (at Hospital 1) and those receiving low quality (at Hospital 2). Using the notational short-hand $B_i :=$

$B(q_i^*)$, the Lorenz curve is given by

$$L(x) = \begin{cases} \frac{B_2}{\bar{B}}x & \text{if } 0 \leq x \leq 1 - D_1 \\ -\frac{(B_1 - B_2)(1 - D_1)}{\bar{B}} + \frac{B_1}{\bar{B}}x & \text{if } 1 - D_1 < x \leq 1 \end{cases}, \quad (32)$$

where $\bar{B} := D_1 B_1 + (1 - D_1) B_2$ is average health outcome. The Gini coefficient is then given by

$$G = 1 - 2 \int_0^1 L(x) dx = 1 - \frac{[B_1 - (B_1 - B_2)(1 - D_1)(1 + D_1)]}{\bar{B}}, \quad (33)$$

where

$$\frac{\partial G}{\partial B_1} = \frac{(1 - D_1) D_1 B_2}{\bar{B}^2} > 0, \quad (34)$$

$$\frac{\partial G}{\partial B_2} = -\frac{(1 - D_1) D_1 B_1}{\bar{B}^2} < 0 \quad (35)$$

and

$$\frac{\partial G}{\partial D_1} = -(B_1 - B_2) \frac{B_2(2D_1 - 1) + D_1^2(B_1 - B_2)}{\bar{B}^2} < 0. \quad (36)$$

All else equal, a marginal increase in the health outcome of patients at the high-quality (low-quality) hospital will increase (reduce) the Gini coefficient. Furthermore, an increase in the market share of the high-quality hospital – which initially has the larger market share – will reduce the Gini coefficient. Notice also that

$$\frac{\partial G}{\partial B_1} + \frac{\partial G}{\partial B_2} = -\frac{(1 - D_1) D_1}{\bar{B}^2} (B_1 - B_2) < 0. \quad (37)$$

Thus, a marginal increase in health outcome for all patients will, all else equal, reduce the Gini coefficient. This is a reflection of the Gini coefficient being a *relative* measure of inequality, which is reduced when all patients experience an equal absolute increase in health status.

We can convert the Gini coefficient to a measure of *absolute* inequality by multiplying G with the average health outcome, which yields the Generalised Gini coefficient:

$$\tilde{G} := \bar{B}(q_1^*, q_2^*) G = D_1(1 - D_1)(B_1 - B_2), \quad (38)$$

where

$$\frac{\partial \tilde{G}}{\partial B_1} = -\frac{\partial \tilde{G}}{\partial B_2} = D_1(1 - D_1) > 0 \quad (39)$$

and

$$\frac{\partial \tilde{G}}{\partial D_1} = -(B_1 - B_2)(2D_1 - 1) < 0. \quad (40)$$

As for the Gini coefficient, a higher market share for the high-quality hospital will also reduce absolute inequality, whereas a marginal improvement in the health status of patients at the high-quality (low-quality) hospital will increase (reduce) absolute inequality, as measured by the Generalised Gini coefficient. However, for given patient allocations between the two hospitals, an equal absolute increase in the health status of all patients has no effect on absolute inequality (i.e., $\partial \tilde{G}/\partial B_1 + \partial \tilde{G}/\partial B_2 = 0$).

Higher treatment price The effect of a higher treatment price on *absolute inequality*, as given by the Generalised Gini coefficient, is given by

$$\frac{\partial \tilde{G}}{\partial p} = D_1(1 - D_1) \left(\frac{\partial B_1}{\partial p} - \frac{\partial B_2}{\partial p} \right) - (B_1 - B_2)(2D_1 - 1) \frac{\partial D_1}{\partial p}, \quad (41)$$

which can be re-written as

$$\frac{\partial \tilde{G}}{\partial p} = \left[D_1(1 - D_1) - \frac{1}{2t}(2D_1 - 1)(B_1 - B_2) \right] \frac{\partial \Omega}{\partial p}, \quad (42)$$

where $\Omega := B_1 - B_2$, and where $\partial \Omega/\partial p$ is given by (24).

A price increase affects absolute inequality only if it affects inequality due to the postcode lottery (given by $\partial \Omega/\partial p$). Suppose that a higher price leads to increased inequality between the hospitals ($\partial \Omega/\partial p > 0$), which implies a reallocation of patients towards the high-quality hospital ($\partial D_1/\partial p > 0$). This has two counteracting effects on the Generalised Gini coefficient, given by the two terms in square brackets on the right-hand side of (42). On the one hand, for given market shares, absolute inequality increases because of increased inequality in health outcomes. However, the reallocation of patients towards the high-quality hospital implies that a lower share of patients experience low quality, which reduces the Generalised Gini coefficient.

The relative strength of these two effects depends on the initial quality difference. If the quality difference is small, so that D_1 is close to $\frac{1}{2}$ and B_1 close to B_2 , then the first effect dominates and a dispersion in health outcomes increases absolute inequality. On the other hand, if the quality difference is very large, so that D_1 is close to 1, the second effect dominates and further dispersion in health outcomes actually reduces absolute inequality.

The effect of a higher treatment price on *relative inequality* can be expressed as

$$\begin{aligned} \frac{\partial G}{\partial p} = & \frac{1}{\bar{B}^2} \left[(1 - D_1) D_1 B_2 - \frac{1}{2t} (B_1 - B_2) (B_2 (2D_1 - 1) + D_1^2 (B_1 - B_2)) \right] \frac{\partial \Omega}{\partial p} \\ & - \frac{(1 - D_1) D_1 B_1}{\bar{B}^2} \frac{\partial B_2}{\partial q_2} \frac{\partial q_2}{\partial p}. \end{aligned} \quad (43)$$

The first term in (43) is completely equivalent to (42) and contains the two counteracting effects described above. The second term, which is negative, is specific to the Gini coefficient and reflects the fact that G measures *relative* inequality. Even if $\partial \Omega / \partial p = 0$, a higher price leads to a reduction in G . This is a pure *level effect*. Even if a price increase does not lead to any patient reallocations, the resulting higher quality at both hospitals reduces the relative health inequality between the two patient groups.

Increased patient choice The effects of increased patient choice on absolute and relative inequality, respectively, are given by

$$\frac{\partial \tilde{G}}{\partial r} = \left[D_1 (1 - D_1) - \frac{1}{2t} (2D_1 - 1) (B_1 - B_2) \right] \frac{\partial \Omega}{\partial r} - \frac{(B_1 - B_2)^2 (2D_1 - 1)}{2t^2} \quad (44)$$

and

$$\begin{aligned} \frac{\partial G}{\partial r} = & \frac{1}{\bar{B}^2} \left[(1 - D_1) D_1 B_2 - \frac{1}{2t} (B_1 - B_2) (B_2 (2D_1 - 1) + D_1^2 (B_1 - B_2)) \right] \frac{\partial \Omega}{\partial r} \\ & - \frac{(1 - D_1) D_1 B_1}{\bar{B}^2} \frac{\partial B_2}{\partial q_2} \frac{\partial q_2}{\partial r} - (B_1 - B_2)^2 \frac{B_2 (2D_1 - 1) + D_1^2 (B_1 - B_2)}{2t^2 \bar{B}^2}. \end{aligned} \quad (45)$$

Comparing (44)-(45) with (42)-(43), we see that the effects are similar to the effects of a price increase, but with one additional sub-effect, represented by the last term in (44) and in (45). For given quality levels, increased patient choice implies a reallocation of patients towards the high-quality hospital. This effect contributes to lower relative and absolute inequality. Thus, if a price increase and an increase in patient choice have the exact same effect on quality at both hospitals, the scope for a subsequent reduction in (absolute and relative) inequality is larger in the case of increased patient choice.

We summarise the above analysis as follows:

Proposition 5 *Suppose that all patients have the same severity level.*

(i) If competition leads to a dispersion (convergence) of health outcomes between the two hospitals, this will, all else equal, contribute towards an increase (reduction) in absolute and relative inequality if the initial quality difference is sufficiently small, and towards a reduction (increase) in absolute and relative inequality if the initial quality difference is sufficiently large.

(ii) If the difference in health outcomes is unaffected by the degree of competition, a price increase has no effect on absolute inequality but reduces relative inequality, whereas an increase in patient choice reduces both absolute and relative inequality.

4.3.2 Two severity levels

The previous analysis with one severity level can be seen as an approximation of the case where severity differences are small relative to quality differences between the hospitals, such that a patient treated at the high-quality hospital always has a better health outcome than a patient treated at the low-quality hospital, regardless of severity.

Consider now the opposite, that severity differences are large relative to quality differences, in the sense that the health outcome is always better for a low-severity patient than for a high-severity patient, regardless of which hospital the patient is treated at. Thus, and using again the notational short-hand $B_i^k := B^k(q_i^*)$, suppose that $B_1^l > B_2^l > B_1^h > B_2^h$. In this case, the Lorenz curve is given by

$$L(x) = \begin{cases} \frac{B_2^h}{\bar{B}}x & \text{if } 0 \leq x \leq \lambda - D_1^h \\ -\frac{(B_1^h - B_2^h)x_1}{\bar{B}} + \frac{B_2^h}{\bar{B}}x & \text{if } \lambda - D_1^h < x \leq \lambda \\ -\frac{(B_1^h - B_2^h)x_1 + (B_2^l - B_1^l)x_2}{\bar{B}} + \frac{B_2^l}{\bar{B}}x & \text{if } \lambda < x \leq 1 - D_1^l \\ -\frac{(B_1^h - B_2^h)x_1 + (B_2^l - B_1^l)x_2 + (B_1^l - B_2^l)x_3}{\bar{B}} + \frac{B_1^l}{\bar{B}}x & \text{if } 1 - D_1^l < x \leq 1 \end{cases}, \quad (46)$$

where $\bar{B} := (\lambda - D_1^h) B_2^h + D_1^h B_1^h + (1 - \lambda - D_1^l) B_2^l + D_1^l B_1^l$ is average health outcome. The Gini coefficient is given by

$$G = 1 - \frac{1}{\bar{B}} \left[B_2^l + (D_1^l)^2 (B_1^l - B_2^l) - \lambda(2 - \lambda) (B_2^l - B_2^h) + D_1^h (B_1^h - B_2^h) (D_1^h + 2(1 - \lambda)) \right], \quad (47)$$

while the Generalised Gini coefficient is given by

$$\tilde{G} = D_1^l (1 - D_1^l) (B_1^l - B_2^l) + D_1^h (2\lambda - 1 - D_1^h) (B_1^h - B_2^h) + \lambda(1 - \lambda) (B_2^l - B_2^h). \quad (48)$$

Competition can affect absolute and relative inequality along three main dimensions:

1. Competition can affect inequalities due to the postcode lottery. For given patient allocations, this effect is described in Proposition 3.
2. Competition can affect inequalities between high- and low-severity patients. This effect is described in Proposition 4.
3. Competition can affect the relative shares of different patient groups, as highlighted by the analysis in the previous subsection, which is summarised in Proposition 5.

For the case of one severity level, the effects along the third dimension listed above are straightforward. If competition leads to patient reallocation towards the high-quality (low-quality) hospital, this will – all else equal – contribute to lower (higher) inequality. For the case of two severity types, which implies four different patient groups, the effects along this dimension are somewhat more complicated. To illustrate this, consider the effect on absolute inequality of patient reallocation towards the high-quality hospital. From (48) we derive

$$\frac{\partial \tilde{G}}{\partial D_1^h} = - \left(B_1^h - B_2^h \right) \left(2 \left(D_1^h - \lambda \right) + 1 \right) < 0.$$

and

$$\frac{\partial \tilde{G}}{\partial D_1^l} = \left(B_1^l - B_2^l \right) \left(1 - 2D_1^l \right) < (>) 0 \quad \text{if} \quad D_1^l > (<) \frac{1}{2}.$$

A reallocation of *high-severity* patients towards the high-quality hospital implies a reallocation of patients from the group with the worst health outcome to the group with the second-worst outcome. This will always reduce inequality. However, a reallocation of *low-severity* patients towards the high-quality hospital, which implies a reallocation of patients from the group with the second-best health outcome to the group with the best health outcome, will reduce inequality only if the latter group constitutes more than half of all patients, which requires that the share of high-severity patients (λ) is very low.

The effects of increased competition on absolute and relative inequality are analytically given by some very involved expressions that yield limited additional insights. It is therefore more illustrative to display the effects by numerical simulations based on our previous parameterisations. Table 3 shows the effects of increased competition (higher price or increased patient choice) based on

the parameterisations in Example 1, with a quadratic health benefit function.⁶ We consider two different cases: $\lambda = \frac{1}{2}$ and $\lambda = 1$. The latter case implies only one severity level and therefore removes effects related to inequalities between high- and low-severity patients.

Table 3: Quadratic health function and constant marginal treatment costs (Example 1)

	$\lambda = \frac{1}{2}$						$\lambda = 1$			
	Δ	Ω^h	Ω^l	Φ	G	\tilde{G}	Δ	Ω	G	\tilde{G}
$p = \frac{1}{2}$	0.167	0.319	0.153	0.877	0.150	0.247	0.222	0.420	0.067	0.086
$\mathbf{p} = \mathbf{1}$	0.133	0.236	0.102	0.744	0.112	0.206	0.178	0.300	0.042	0.068
$p = 2$	0.091	0.141	0.050	0.537	0.070	0.146	0.121	0.169	0.020	0.041
$t = 2$	0.078	0.147	0.068	0.868	0.137	0.230	0.105	0.191	0.036	0.047
$\mathbf{t} = \mathbf{1}$	0.133	0.236	0.102	0.744	0.112	0.206	0.178	0.300	0.042	0.068
$t = \frac{1}{2}$	0.200	0.320	0.120	0.512	0.071	0.147	0.267	0.391	0.018	0.038

Remaining parameter values: $c_1 = 0$, $c_2 = \frac{1}{2}$, $\alpha_h = \beta_l = \gamma = 1$, $k = \beta_h = \alpha_l = 2$.

In the example shown in Table 3, a higher price leads to quality convergence whereas increased patient choice leads to quality dispersion. Despite decreasing marginal health gains, quality dispersion (convergence) also implies dispersion (convergence) in health outcomes for each severity type.

Consider first the case of $\lambda = 1$. The effect of competition on absolute inequality (as measured by \tilde{G}) is then determined by changes in inequality along two different dimensions. On the one hand, higher (lower) inequalities due to the postcode lottery contributes to higher (lower) absolute inequality, whereas, on the other hand, increased (reduced) market share of the high-quality hospital contributes to lower (higher) absolute inequality. These two effects are always counteracting, as discussed in the previous sub-section. For the case of a price increase, the former effect dominates. The reduction in postcode inequality is sufficiently strong to reduce the Generalised Gini coefficient. However, for the case of increased patient choice, the overall effect is non-monotonic. Absolute inequality increases as t is reduced from 2 to 1, but decreases as t reduced from 1 to $\frac{1}{2}$. This

⁶Analytical expressions for the Nash equilibrium, on which these numerical examples are based, are given in the Appendix.

illustrates the first part of Proposition 6, which states that dispersion in health outcomes contributes towards more (less) inequality if the initial quality difference is sufficiently small (large).

The effects of competition on relative inequality (as measured by G) are qualitatively identical to the effect on absolute quality, even if the former measure is sensitive to a *level effect*, whereby higher quality in itself reduces relative inequality. However, this effect is not strong enough to prevent an increase in the Gini coefficient when t is reduced from 2 to 1.

Consider now the case of $\lambda = \frac{1}{2}$. The effect of competition on absolute and relative inequality is now determined also by changes in inequalities along a third dimension, namely inequalities between high- and low-severity types, as measured by Φ . We see that competition always reduces inequality along this dimension, regardless of whether competition leads to quality convergence or quality dispersion. The reason is that high-severity patients benefit more from higher quality than low-severity patients. The reduction of inequality along this dimension implies that increased competition always reduces both absolute and relative inequality for all the numerical values considered in this example.

In Table 4 we show an equivalent numerical analysis based on the parameterisation in Example 2.

Table 4: Linear health function and quality-dependent treatment costs (Example 2)

	$\lambda = \frac{1}{2}$						$\lambda = 1$			
	Δ	Ω^h	Ω^l	Φ	G	\tilde{G}	Δ	Ω	G	\tilde{G}
$p = \frac{1}{2}$	0.087	0.174	0.087	0.879	0.141	0.236	0.092	0.183	0.033	0.044
$\mathbf{p} = \mathbf{1}$	0.106	0.212	0.106	0.703	0.101	0.195	0.122	0.244	0.032	0.057
$p = 2$	0.144	0.289	0.144	0.347	0.046	0.112	0.183	0.366	0.029	0.079
$t = 2$	0.093	0.185	0.093	0.882	0.142	0.238	0.095	0.191	0.035	0.047
$\mathbf{t} = \mathbf{1}$	0.106	0.212	0.106	0.703	0.101	0.195	0.122	0.244	0.032	0.057
$t = \frac{1}{2}$	0.164	0.328	0.164	0.338	0.043	0.105	0.213	0.426	0.011	0.029

Parameter values: $c_1 = 0.1$, $c_2 = 0.5$, $\alpha_h = \beta_l = 1$, $k = \beta_h = \alpha_l = 2$.

In the numerical examples shown in Table 4, more competition (either by a higher price or by increased patient choice) always leads to quality dispersion. Because of the linearity of the health

benefit function, this also implies that competition leads to increased postcode inequality. The implications for absolute and relative inequality are very similar to those shown in Table 3, though. The only difference is that, for the case of $\lambda = 1$, the *level effect* is sufficiently strong to ensure that more competition leads to a monotonic reduction in relative inequality, even if the postcode inequality increases.

For the case of $\lambda = \frac{1}{2}$, more competition always leads to a reduction in both absolute and relative inequality. As for the examples shown in Table 3, the driving force is the reduction in inequality between high- and low-severity patients. The strong negative effect of competition on inequality along this dimension, for different parameterisations and different numerical parameter values, suggests that this is a fairly general result. Furthermore, our numerical simulations suggest that competition will lead to a reduction in both absolute and relative inequality for a wide range of parameters. In fact, for the two different parameterisations explored here, it is hard to find examples of the opposite as long as the shares of high- and low-severity patients are not too uneven.

5 Implications for empirical analyses

In this section we discuss possible approaches which could be pursued to test empirically how competition affects health inequalities. First, to test for the effect of competition on health inequalities due to postcode lottery, researchers could compute measures of dispersion of health outcomes, such as the standard deviation or the coefficient of variation, within a given hospital catchment area and relate them to the degree of patient choice and market structure. For example, future empirical work could test whether in more competitive areas the introduction of patient choice policies lead to an increase or a reduction in AMI mortality dispersion across hospitals.

Second, to test for the effect of competition on health inequalities a sub-group analysis by degree of severity may be appropriate. In line with Geppert and Kessler (2005), high severity could be measured based on the number of previous hospital admissions preceding a health shock (such as AMI). By comparing the effect of competition on mortality for high- and low-severity patients, we can infer the effect on health inequalities across severity groups.

Third, the two types of inequality could be brought together by developing a Generalised Gini or Gini index in a given market area, where patients are ordered by their level of health, i.e., starting with patients with highest severity and lowest hospital quality and ending with patients with lowest severity and highest quality.

Our analysis also illustrates the importance of patient ‘composition effects’ when measuring the effect of competition due to patients with high and low severity exercising choice to a different degree. Competition affects differentially the health gains for patients with differing severity but also changes the number of patients receiving high and low quality through the composition effect. These will affect both the Gini coefficients and the simple measures of dispersion of health outcomes across hospitals.

The empirical literature which estimates patient choice models as a function of quality and severity tends to confirm that high-severity patients are more likely to choose high-quality hospitals. The elasticity of hospital demand to quality are however generally low and so are the interactions between quality and severity (see Brekke et al., 2014, Section 3.1, for a review of the evidence). We therefore conjecture that overall composition effects are likely to be small in empirical analyses.

Finally, our analysis highlights the importance of distinguishing empirically between quality and health outcomes. Although health outcomes are often used as a proxy of hospital quality, our study highlights how inequalities in qualities do not necessarily go hand-in-hand with inequalities in health outcomes. In relation to inequalities due to postcode lottery, an increase in inequalities in *quality* across hospitals is compatible with a reduction in *health* inequalities across hospitals if the marginal health gain is decreasing, so that patients in high-quality hospitals benefit less from a given quality increase than do patients in low-quality hospitals. Similarly, an increase in quality differences across hospitals is compatible with a reduction in health inequalities across severity types and this is due to patients with higher severity benefiting more from the increase in quality compared to patients with lower severity.

6 Concluding remarks

Several OECD countries have introduced pro-market policy interventions in the health sector with the aim of stimulating quality of care. Such policies are generally contentious and the subject of an intense political debate. The existing literature has extensively investigated, both theoretically and empirically, the effect of competition on quality but there is very little work on its impact on equity. This is surprising given that reduction in health inequalities is an ubiquitous policy objective. Our study has contributed to fill this gap in knowledge by carefully characterising the conditions under which competition (i) increases or reduces the gap between high-quality and low-quality hospitals (due to postcode lottery), and, as a result, (ii) contributes to an increase or reduction in health

inequalities.

Our first key finding is that the effect of competition on hospital quality gap depends on demand and supply factors affecting health care provision, more precisely captured by (i) the marginal health gains from quality – a demand parameter – and (ii) the extent to which quality affects marginal treatment costs – a supply parameter. Our most clear-cut result is that competition, regardless of how it is measured, is more likely to lead to quality convergence across hospitals if marginal health gains decrease with quality at a faster rate. The answer also depends, to some extent, on how we measure competition, whether by an increase in the fixed price or by an increase in the degree of patient choice. Cost factors increase the scope for quality dispersion when competition is measured by an increase in price, but not necessarily when measured by patient choice. Such factors will vary by medical condition, diagnosis and treatment. For example, standardised treatments such as cataract surgery will have treatment costs mildly increasing with quality. This may not be the case for more serious treatments, such as a coronary bypass, where costs will increase more rapidly with quality.

Our second key finding is that health inequalities due to *postcode lottery* go hand in hand with health inequalities but only if health gains are not too concave in quality. Instead, we find that competition generally reduces health inequalities across patients with different *severity*, because high-severity patients benefit more from higher quality than do low-severity patients. This reduction can be strengthened or weakened by what we refer to as ‘composition effects’, which relate to competition inducing more high-severity patients to exercise choice and to select hospitals with higher quality. Reductions in inequalities across severity types also drive reductions in the Gini and Generalised Gini coefficients, which aggregate different sources of health inequalities both across hospitals and severity types.

Finally, we highlight that measuring the effect of competition on health inequalities through simple measures of dispersion or through the Gini coefficient is important since different measures can lead to different conclusions. If competition increases differences in health outcomes across hospitals, the Generalised Gini coefficient may still reduce due to the composition effects, and the Gini coefficient will reduce further as a result of the overall increase in quality.

In terms of policy implications, our analysis highlights that whether competition induces an equity-efficiency trade-off depends on the particular dimension of equity which policy makers focus on. For example, if policy makers focus on equity due to postcode lottery, then an equity-efficiency

trade-off may arise though it is less likely to be the case when demand parameters are more important. An equity-efficiency trade-off is instead unlikely when considering equity across severity types if more severe patients tend to benefit more than low-severity ones from increases in quality.

Our study provides a theoretical framework to guide future empirical work. Future empirical studies should focus not only on testing the effect of competition on quality, but also its equity implications. This can be done, as discussed in Section 5, by developing measures of dispersions in quality and health outcomes within a given hospital catchment or market area, and then by relating these to changes in patient choice and in prices through consolidated econometric strategies.

References

- [1] Barigozzi, F., Burani, N., 2016. Competition and screening with motivated health professionals, *Journal of Health Economics*, 50, 358-371.
- [2] Beitia, A., 2003. Hospital quality choice and market structure in a regulated duopoly. *Journal of Health Economics*, 22, 1011–1036.
- [3] Bloom, N., Propper, C., Seiler, S., Van Reenen, J., 2015. The impact of competition on management quality: evidence from public hospitals. *The Review of Economic Studies*, 82, 457-489.
- [4] Brekke, K.R., Nuscheler, R., Straume, O.R., 2006. Quality and location choices under price regulation. *Journal of Economics & Management Strategy*, 15, 207–227.
- [5] Brekke, K.R., Nuscheler, R., Straume, O.R., 2007. Gatekeeping in health care. *Journal of Health Economics*, 26, 149–170.
- [6] Brekke, K.R., Siciliani, L., Straume, O.R., 2011. Hospital competition and quality with regulated prices. *Scandinavian Journal of Economics*, 113, 444–469.
- [7] Brekke, K.R., Siciliani, L., Straume, O.R., 2012. Quality competition with profit constraints. *Journal of Economic Behavior & Organization*, 84, 642–659.
- [8] Brekke, K.R., Gravelle, H., Siciliani, L., Straume, O.R., 2014. Patient Choice, Mobility and Competition Among Health Care Providers, Chapter 1, 1-26, in R. Levaggi and M. Montefiori

- (eds.), *Health Care Provision and Patient Mobility, Developments in Health Economics and Public Policy*, 12, Springer-Verlag.
- [9] Cookson, R., Laudicella, M., Li Donni, P., 2013. Does hospital competition harm equity? Evidence from the English National Health Service, *Journal of Health Economics*, 32, 410-422.
- [10] Cooper, Z., Gibbons, S., Jones, S., McGuire, A., 2011. Does hospital competition save lives? evidence from the English NHS patient choice reforms. *The Economic Journal*, 121(554), F228-F260.
- [11] Dafny, L., 2005. How do hospitals respond to price changes? *American Economic Review*, 95,1525-1547.
- [12] Dalton, D., 2014. Examining new options and opportunities for providers of NHS care, *The Dalton Review*, Department of Health and Social Care.
- [13] Calem, P.S., Rizzo, J.A., 1995. Competition and specialization in the hospital industry: an application of Hotelling's location model. *Southern Economic Journal*, 61, 1182-1198.
- [14] EXPH (Expert Panel on Effective Ways of Investing in Health), 2015. Investigating Policy Options regarding Competition among Providers of Health Care Services in EU Member States. European Union, http://ec.europa.eu/health/expert_panel/index_en.htm.
- [15] Farrar, S., et al., 2009. Has payment by results affected the way that English hospitals provide care? Difference-in-differences analysis, *British Medical Journal*, 339, b3047.
- [16] Gaynor, M., 2006. What do we know about competition and quality in health care markets? *Foundations and Trends in Microeconomics*, Vol. 2, Issue 6.
- [17] Gaynor, M., Moreno-Serra, R., Propper, C., 2013. Death by market power: Reform, competition, and patient outcomes in the National Health Service. *American Economic Journal: Economic Policy*, 5, 134-166.
- [18] Gaynor, M., Town, R., 2011. Competition in health care markets, *Handbook of Health Economics*, Chapter 9, 499-637.
- [19] Gowrisankaran, G., Town, R.J., 1999. Estimating the quality of care in hospitals using instrumental variables. *Journal of Health Economics*, 18, 747-767.

- [20] Gravelle, H., 1999. Capitation contracts: access and quality. *Journal of Health Economics*, 18, 3, 315–340.
- [21] Karlsson, M., 2007. Quality incentives for GPs in a regulated market. *Journal of Health Economics*, 26, 699–720.
- [22] Kessler, D., McClellan, M., 2000. Is hospital competition socially wasteful? *Quarterly Journal of Economics*, 115, 577–615.
- [23] Kessler, D., Geppert, J. 2005. The effects of competition on variation in the quality and cost of medical care. *Journal of Economics and Management Strategy*, 14, 575-589.
- [24] Ma, C.A., Burgess, J.F., 1993. Quality competition, welfare, and regulation. *Journal of Economics*, 58, 153–173.
- [25] OECD, 2012. *Competition in Hospital Services*, OECD Policy Roundtables, Directorate for Financial and Enterprise Affairs Competition Committee.
- [26] Sari, N., 2002. Do competition and managed care improve quality? *Health Economics*, 11, 571–584.
- [27] Sculpher, M., Gafni, A., 2001. Recognizing diversity in public preferences: The use of preference sub-groups in cost-effectiveness analysis. *Health Economics*, 10, 317-324.
- [28] Shen, Y.-S., 2003. The effect of financial pressure on the quality of care in hospitals. *Journal of Health Economics*, 22, 243-269.
- [29] Siciliani, L., Chalkley, M., Gravelle, H., 2017. Competition policy in five European countries. What can be learned for health policy in England, *The Health Foundation Working Paper* 1, 1-28.
- [30] Tay, A., 2003. Assessing competition in hospital care markets: the importance of accounting for quality differentiation. *RAND Journal of Economics*, 34, 786–814.
- [31] Wagstaff, A., Van Doorslaer, E., 2000. Equity in healthcare finance and delivery. *Handbook of Health Economics*, Vol. 1, 1803-1862.
- [32] Wolinsky, A., 1997. Regulation of duopoly: Managed competition vs regulated monopolies. *Journal of Economics & Management Strategy*, 6, 821–847.

Appendix

Second-order and stability conditions

The second-order conditions of the hospitals' profit-maximising problem are given by

$$\frac{\partial^2 \pi_i}{\partial q_i^2} = \sum_k \left((p - c_i^k) \frac{\partial^2 D_i^k}{\partial q_i^2} - 2 \frac{\partial c_i^k}{\partial q_i} \frac{\partial D_i^k}{\partial q_i} - \frac{\partial^2 c_i^k}{\partial q_i^2} D_i^k \right) - \frac{\partial^2 C}{\partial q_i^2} < 0, \quad (\text{A1})$$

where

$$\frac{\partial^2 D_i^h}{\partial q_i^2} = \frac{\lambda}{2t} \frac{\partial^2 B^h}{\partial q_i^2}; \quad \frac{\partial^2 D_i^l}{\partial q_i^2} = \frac{1 - \lambda}{2t} \frac{\partial^2 B^l}{\partial q_i^2}. \quad (\text{A2})$$

The Nash equilibrium is stable if

$$H := \frac{\partial^2 \pi_i}{\partial q_i^2} \frac{\partial^2 \pi_j}{\partial q_j^2} - \frac{\partial^2 \pi_i}{\partial q_j \partial q_i} \frac{\partial^2 \pi_j}{\partial q_i \partial q_j} > 0. \quad (\text{A3})$$

Comparative statics

The effects of a marginal price change on equilibrium qualities are given by

$$\begin{bmatrix} \frac{\partial^2 \pi_1}{\partial q_1^2} & \frac{\partial^2 \pi_1}{\partial q_2 \partial q_1} \\ \frac{\partial^2 \pi_2}{\partial q_1 \partial q_2} & \frac{\partial^2 \pi_2}{\partial q_2^2} \end{bmatrix} \begin{bmatrix} dq_1 \\ dq_2 \end{bmatrix} + \begin{bmatrix} \frac{\partial D_1}{\partial q_1} \\ \frac{\partial D_2}{\partial q_2} \end{bmatrix} dp = 0, \quad (\text{A4})$$

which implies

$$\frac{\partial q_1^*}{\partial p} = \frac{1}{H} \left[\frac{\partial D_2}{\partial q_2} \frac{\partial^2 \pi_1}{\partial q_2 \partial q_1} - \frac{\partial D_1}{\partial q_1} \frac{\partial^2 \pi_2}{\partial q_2^2} \right] > 0, \quad (\text{A5})$$

and

$$\frac{\partial q_2^*}{\partial p} = \frac{1}{H} \left[\frac{\partial D_1}{\partial q_1} \frac{\partial^2 \pi_2}{\partial q_1 \partial q_2} - \frac{\partial D_2}{\partial q_2} \frac{\partial^2 \pi_1}{\partial q_1^2} \right] > 0 \quad (\text{A6})$$

where

$$\frac{\partial^2 \pi_1}{\partial q_2 \partial q_1} = \frac{1}{2t} \left(\lambda \frac{\partial c_1^h}{\partial q_1} \frac{\partial B^h}{\partial q_2} + (1 - \lambda) \frac{\partial c_1^l}{\partial q_1} \frac{\partial B^l}{\partial q_2} \right) > 0 \quad (\text{A7})$$

and

$$\frac{\partial^2 \pi_2}{\partial q_1 \partial q_2} = \frac{1}{2t} \left(\lambda \frac{\partial c_2^h}{\partial q_2} \frac{\partial B^h}{\partial q_1} + (1 - \lambda) \frac{\partial c_2^l}{\partial q_2} \frac{\partial B^l}{\partial q_1} \right) > 0. \quad (\text{A8})$$

The effects of a marginal change in *transportation costs* on equilibrium qualities are given by

$$\begin{bmatrix} \frac{\partial^2 \pi_1}{\partial q_1^2} & \frac{\partial^2 \pi_1}{\partial q_2 \partial q_1} \\ \frac{\partial^2 \pi_2}{\partial q_1 \partial q_2} & \frac{\partial^2 \pi_2}{\partial q_2^2} \end{bmatrix} \begin{bmatrix} dq_1 \\ dq_2 \end{bmatrix} + \begin{bmatrix} \frac{\partial^2 \pi_1}{\partial r \partial q_1} \\ \frac{\partial^2 \pi_2}{\partial r \partial q_2} \end{bmatrix} dr = 0, \quad (\text{A9})$$

which implies

$$\frac{\partial q_1^*}{\partial r} = \frac{1}{H} \left[\frac{\partial^2 \pi_2}{\partial r \partial q_2} \frac{\partial^2 \pi_1}{\partial q_2 \partial q_1} - \frac{\partial^2 \pi_1}{\partial r \partial q_1} \frac{\partial^2 \pi_2}{\partial q_2^2} \right] \quad (\text{A10})$$

and

$$\frac{\partial q_2^*}{\partial r} = \frac{1}{H} \left[\frac{\partial^2 \pi_1}{\partial r \partial q_1} \frac{\partial^2 \pi_2}{\partial q_1 \partial q_2} - \frac{\partial^2 \pi_2}{\partial r \partial q_2} \frac{\partial^2 \pi_1}{\partial q_1^2} \right] \quad (\text{A11})$$

where

$$\begin{aligned} \frac{\partial^2 \pi_1}{\partial r \partial q_1} &= \frac{\lambda}{2t^2} \left[(p - c_1^h) \frac{\partial B^h}{\partial q_1} - (B^h(q_1) - B^h(q_2)) \frac{\partial c_1^h}{\partial q_1} \right] \\ &\quad + \frac{1-\lambda}{2t^2} \left[(p - c_1^l) \frac{\partial B^l}{\partial q_1} - (B^l(q_1) - B^l(q_2)) \frac{\partial c_1^l}{\partial q_1} \right] \end{aligned} \quad (\text{A12})$$

and

$$\begin{aligned} \frac{\partial^2 \pi_2}{\partial r \partial q_2} &= \frac{\lambda}{2t^2} \left[(p - c_2^h) \frac{\partial B^h}{\partial q_2} + (B^h(q_1) - B^h(q_2)) \frac{\partial c_2^h}{\partial q_2} \right] \\ &\quad + \frac{1-\lambda}{2t^2} \left[(p - c_2^l) \frac{\partial B^l}{\partial q_2} + (B^l(q_1) - B^l(q_2)) \frac{\partial c_2^l}{\partial q_2} \right]. \end{aligned} \quad (\text{A13})$$

By applying the first-order conditions, (6), we can simplify and rewrite (A12)-(A13) as

$$\frac{\partial^2 \pi_1}{\partial r \partial q_1} = \frac{1}{2t} \left(\lambda \frac{\partial c_1^h}{\partial q_1} + (1-\lambda) \frac{\partial c_1^l}{\partial q_1} + 2 \frac{\partial C}{\partial q_1} \right) > 0 \quad (\text{A14})$$

and

$$\frac{\partial^2 \pi_2}{\partial r \partial q_2} = \frac{1}{2t} \left(\lambda \frac{\partial c_2^h}{\partial q_2} + (1-\lambda) \frac{\partial c_2^l}{\partial q_2} + 2 \frac{\partial C}{\partial q_2} \right) > 0, \quad (\text{A15})$$

which implies that $\partial q_1^*/\partial r > 0$ and $\partial q_2^*/\partial r > 0$.

Proofs

Proof of Proposition 1. Notice first that assumptions A2 and A3 imply $\bar{c}_1 = \bar{c}_2$ and $\partial^2 C / \partial q_1^2 = \partial^2 C / \partial q_2^2$. (i) Since both ψ_2 and ψ_1 go to zero as $\partial c_i^k / \partial q_i$ goes to zero, while $\bar{p}_1 \geq \bar{p}_2$ (by assumption A1) and $\bar{B}_1 > \bar{B}_2$, for any value of $\partial c_i^k / \partial q_i$, as long as $B^k(\cdot)$ is strictly concave, the statement in the first part of the Proposition is true by monotonicity. (ii) Suppose that $\partial^2 B^k / \partial q_i^2 \rightarrow 0$, which

implies $\bar{B}_1 \rightarrow \bar{B}_2$ and $\bar{p}_1 \rightarrow \bar{p}_2$. It follows from (16) that $\partial\Delta/\partial p > 0$ if $\psi_2 > \psi_1$, which is true for $\partial c_1^k/\partial q_1 < \partial c_2^k/\partial q_2$. By continuity, this result also holds for a sufficiently low degree of concavity of $B^k(\cdot)$. *Q.E.D.*

Proof of Proposition 2. Notice first that assumptions A2 and A3 imply $\bar{c}_1 = \bar{c}_2$ and $\partial^2 C/\partial q_1^2 = \partial^2 C/\partial q_2^2$. Suppose that $\partial c_i^k/\partial q_i \rightarrow 0$, which implies $\psi_1 \rightarrow 0$ and $\psi_2 \rightarrow 0$. Furthermore, suppose $\partial^2 B^k/\partial q_i^2 \rightarrow 0$, which implies $\bar{B}_1 \rightarrow \bar{B}_2$ and $\bar{p}_1 \rightarrow \bar{p}_2$. It follows from (21) that $\partial\Delta/\partial r > 0$ if $\xi_1 > \xi_2$, which, from (19)-(20), is true if $\partial C/\partial q_1 > \partial C/\partial q_2$. Because of the convexity of C , this condition holds for all $q_1^* > q_2^*$. By continuity, this result also holds for sufficiently low values of $\partial c_i^k/\partial q_i$ and for a sufficiently low degree of concavity of $B^k(\cdot)$. *Q.E.D.*

Equilibrium qualities in Tables 3 and 4

The numerical examples in Table 3 are based on a health benefit function given by

$$B^k(q_i) = \alpha_k + \beta_k q_i - \frac{\gamma}{2} q_i^2, \quad k = l, h, \quad (\text{A16})$$

along with the cost functions given in Example 1. The resulting Nash equilibrium is given by

$$q_1^* = \frac{\eta(p - c_1)}{\gamma(p - c_1) + 2kt}, \quad (\text{A17})$$

$$q_2^* = \frac{\eta(p - c_2)}{\gamma(p - c_2) + 2kt}, \quad (\text{A18})$$

where $\eta := \lambda\beta_h + (1 - \lambda)\beta_l$.

The numerical examples in Table 4 are based on a health benefit function given by (A16) with $\gamma = 0$, along with the cost functions given in Example 2. The resulting Nash equilibrium is given by

$$q_1^* = \frac{\eta(p(\eta(c_1 + 2c_2) + 2kt) - 3tc_1c_2) - 2kt^2c_1}{\eta(4kt(c_1 + c_2) + 3\eta c_1c_2) + 4k^2t^2}, \quad (\text{A19})$$

$$q_2^* = \frac{\eta(p(\eta(2c_1 + c_2) + 2kt) - 3tc_1c_2) - 2kt^2c_2}{\eta(4kt(c_1 + c_2) + 3\eta c_1c_2) + 4k^2t^2}. \quad (\text{A20})$$

Most Recent Working Paper

NIPE WP 01/2018	Siciliani, Luigi e Odd Rune Straume , “Competition and Equity in Health Care Market”, 2018
NIPE WP 13/2017	Aguiar-Contraria, Luís, Maria Joana Soares e Rita Sousa , “California’s Carbon Market and Energy Prices: A Wavelet Analysis”, 2017
NIPE WP 12/2017	Mustapha Olalekan Ojo, Luís Aguiar-Contraria e Maria Joana Soares , "A time-frequency analysis of the Canadian macroeconomy and the yield curve", 2017.
NIPE WP 11/2017	Sousa, Rita , Adérito Santana e Inês Mourão, “Low-Emission Energy Outlook in a Small Island Developing States – The case of Sao Tome And Principe”, 2017
NIPE WP 10/2017	Vareiro, Laurentina, J. Cadima Ribeiro e Paula Remoaldo, “Destination attributes and tourist’s satisfaction in a cultural destination”, 2017
NIPE WP 09/2017	Amado, Cristina , Annastiina Silvennoinen e Timo Teräsvirta, "Modelling and forecasting WIG20 daily returns", 2017
NIPE WP 08/2017	Almeida, André, Hugo Figueiredo, João Cerejeira, Miguel Portela, Carla Sá e Pedro Teixeira , "Returns to Postgraduate Education in Portugal: Holding on to a Higher Ground?", 2017
NIPE WP 07/2017	Magalhães, P. e Luís Aguiar-Contraria , "Procedural Fairness and Economic Voting", 2017
NIPE WP 06/2017	Veiga, Francisco José e Linda Gonçalves Veiga , "Term limits and voter turnout", 2017
NIPE WP 05/2017	Brekke, Kurt R., Tor Helge Holmas, Karin Monstad e Odd Rune Straume , "Competition and physician behaviour: Does the competitive environment affect the propensity to issue sickness certificates?", 2017
NIPE WP 04/2017	Sochirca, Elena e Francisco José Veiga , "Measuring political rivalry and estimating its effect on economic growth", 2017
NIPE WP 03/2017	Esteves, Rosa-Branca e Sofia Cerqueira, "Behavior-Based Pricing under Imperfectly Informed Consumers", 2017
NIPE WP 02/2017	Esteves, Rosa-Branca e Joana Resende, "Personalized Pricing with Targeted Advertising: Who are the Winners?", 2017
NIPE WP 01/2017	Bohn, Frank e Francisco José Veiga , “Political Opportunism and Countercyclical Fiscal Policy in Election-year Recessions”, 2017
NIPE WP 13/2016	Bernardino, Susana, J. Freitas Santos e Sandra Fagundes, “ Uma abordagem qualitativa ao papel do capital social no desenvolvimento de iniciativas sociais em Portugal”, 2016
NIPE WP 12/2016	Castro, Vitor , “Functional components of public expenditure, fiscal consolidations and the economic activity”, 2016
NIPE WP 11/2016	Castro, Vitor , “On the behaviour of the functional components of government expenditures during fiscal consolidations”, 2016
NIPE WP 10/2016	Amorim, Andreia e Rosa-Branca Esteves , "Retention Strategies in a Switching Cost Model", 2016
NIPE WP 09/2016	Natália P. Monteiro e Odd Rune Straume , “Are cooperatives more productive than investor-owned firms? Cross-industry evidence from Portugal”, 2016
NIPE WP 08/2016	Aidt, Toke S., Vitor Castro e Rodrigo Martins, “Shades of red and blue: Political ideology and sustainable development”, 2016
NIPE WP 07/2016	Veiga, Linda Gonçalves e Francisco Veiga , "Term limits at the local government level: evidence from a natural experiment", 2016
NIPE WP 06/2016	Castro, Vítor e Rodrigo Martins, “The Electoral Dynamics of Human Development”, 2016
NIPE WP 05/2016	Bastos, Paulo, Natália P. Monteiro e Odd Rune Straume , “Foreign acquisition and internal organization”, 2016
NIPE WP 04/2016	Esteves, Rosa-Branca , Qihong Liu e Jie Shuai, “Behavior-Based Price Discrimination in a Multi-Dimensional Preferences Market”, 2016
NIPE WP 03/2016	Bleaney, Michael e Manuela Francisco , “Real Exchange Rate Volatility: Is Sub-Saharan Africa Different?”, 2016
NIPE WP 02/2016	Castro, Vítor e Rodrigo Martins, “Politically driven cycles in fiscal policy: In depth analysis of the functional components of government expenditures”, 2016