Interuniversity Research Centre on Local and Regional Finance

CIFREL Working Papers

Contracting unverifiable quality in healthcare: The importance of political stability for relational contracts

Marco Buso, Berardino Cesi, Silvia Coretti,
Gilberto Turati

Working Paper n. 01/2024

CIFREL is an Interuniversity research centre that conducts applied research on local and regional governments and more generally on public economics.
The current members of the Centre are: the Department of Economics and Finance of the Università Cattolica del Sacro Cuore, the Department of Economics and Finance of the University of Brescia, the Department of Economics and Finance of the University of Ferrara, the Department of Economics and Business (DISEI) and the Department of Law and Political, Economic and Social Sciences (DIGSPES) of the University of Piemonte Orientale, the Department of Economics and Statistics Cognetti de Martiis and the Department of Economics, Social Studies, Applied Mathematics and Statistics of the University of Torino.

Contacts:

CIFREL
Università Cattolica del Sacro Cuore
Via Necchi 5
20123 Milano
Telephone: 0039.02.7234.2976
e-mail: dip.economiaefinanza@unicatt.it
web: https://centridiricerca.unicatt.it/cifrel_index.html

Contracting unverifiable quality in healthcare: the importance of political stability for relational contracts*

M. Buso † B. Cesi‡ S. Coretti § G. Turati ¶

March 15, 2024

Abstract

We consider an infinitely repeated game between a public purchaser of a health service and a semi-altruistic hospital when some quality dimensions of the service are non-contractible. We discuss how a *Pay-for-Performance Relational Contract* (P4P-RC) can induce the hospital to deliver positive unverifiable quality. We find that the optimal conditions for both price and quantities of the P4P-RC converge to the first-best the higher the stability of the interaction between the purchaser and the hospital. Using the length of tenure of regional politicians in Italy as a proxy for a stable interaction, we empirically test the relationship between proxies of healthcare service quality and political stability from 1996 to 2020. We find evidence supporting the view that unverifiable quality increases in the political stability of the regional governments.

Keywords: relational contracts; political stability; healthcare quality

JEL codes: H57, L41, C73

*We wish to thank Domenico Lisi, Vanessa Cirulli, Stefania Cosci, and the participants at the SIEP 2019 and SIE 2019 Conference for comments on very early drafts of this manuscript. Usual disclaimers apply.

†Università di Padova, Department of Economics and Management "Marco Fanno"

‡Università La Sapienza, Department of Political Science

§Università La Sapienza, Department of Economics and Law

¶Università Cattolica del Sacro Cuore, Department of Economics and Finance

1 Introduction

Defining, contracting, and evaluating the quality of healthcare services is a major concern worldwide. According to the paradigm originally proposed by Donabedian (1988), "quality" in healthcare is a multidimensional concept including (i) the *context* in which care is delivered, (ii) the *process* of care, from diagnosis to treatment, and (iii) the *outcome* of care. For instance, considering hospital care services, the *structural quality* of services is defined by both hotel services (e.g., size, comfort, quiet, and cleanliness of rooms) and medical staff (e.g., technical abilities and skills of doctors); the *procedural quality* of services refers to the provision of appropriate treatments to patients; the *outcome quality* of services refers to both patients' satisfaction and their health status (like mortality rates and readmissions), which are connected to both structural and procedural quality.

Given the multifaceted nature of quality, contracting for quality poses serious challenges for buyers (either patients directly, or public and private insurers). It is crucial to distinguish between verifiable service characteristics and those that cannot be verified. Verifiable characteristics (such as the quality of hotel services or the skills of doctors) can be either regulated by public authorities (as ex-ante conditions to access the market) or written in contractual clauses, which can then be enforced by the power of an independent court (Eggleston 2005). Conversely, characteristics concerning procedural or outcome quality, like the appropriateness of treatments or patients' health status, are often harder to review, assess, and verify even by an independent board of physicians. Writing *complete* contracts on these dimensions of quality is challenging, implying difficulties in inducing providers to deliver a satisfactory level of quality.[1]

If contracts are incomplete because at least some dimensions of quality cannot be verified, decisions by the medical staff can be biased by economic incentives, resulting in different treatment patterns for the same health problem (e.g., Chandra et al. (2011)). Relevant examples include differences across providers in the preferred treatment for prostate cancer (Wilt et al. 2008); geographical variation in tonsillectomy rates (Wennberg & Gittelsohn 1973); the use of percutaneous coronary interventions (PCI) for stable coronary disease, lacking proven substantial clinical benefit (Boden et al. 2007, Weintraub et al. 2008). Contract incompleteness makes enforcing the quality provision in all its dimensions by standard contractual clauses almost impossible (or exceedingly costly).

How to handle non-contractible quality dimensions in healthcare then? The theoretical literature typically suggests two solutions: (i) the selection of altruistic (public and nonprofit) providers (e.g., Hansmann (1988)), characterized by an objective function increasing in all quality dimensions; (ii) the use of contractual clauses based on (almost) verifiable proxies of outcome quality (e.g., readmission rates, mortality, ...) to define Pay-for-Performance (P4P) payment schemes (see, e.g., Benbassat & Taragin (2000), Fischer et al. (2014), Talsma et al. (2014) Peluso et al. (2019*a*)). The reality of many healthcare systems worldwide suggests that the first solution is the most common: many hospitals around the world are public or private not-for-profit, especially in contexts where they can exploit market power, albeit P4P schemes have also started to emerge in recent years.

[1] Evidence of unverifiable quality in healthcare contracts is confirmed by several papers. See, e.g., (De Luca et al. 2021, Kaarboe & Siciliani 2011, Beitia 2003, Eggleston 2005, Dumont et al. 2008, Newhouse 2002, Smith & York 2004).

Both solutions show important limitations: on the one hand, public and private nonprofit hospitals are characterized by inefficiencies (likely driven by the lack of residual claimants) that influence the overall outcome; on the other hand, all measurable indicators that can be used in P4P schemes (also for public and nonprofit hospitals) are valid only for some specific treatments and they all need to be adjusted for patients' characteristics (a procedure requiring detailed information in the hands of providers).[2]

In this paper, we discuss *relational contracts* as an alternative solution for handling the problem of non-verifiable dimensions of quality in healthcare. Following Levin (2003), we define a relational contract as an agreement whose enforcement does not come from the power of an external court, but from the value providers and purchasers assign to their future interaction.[3] We first set up an infinitely repeated game in which one purchaser (a public insurer) and one provider (a semi-altruistic hospital) agree on a Pay-for-Performance-Relational Contract (P4P-RC) to provide a health service with both verifiable and unverifiable quality characteristics. We show that a P4P-RC allows the enforcement of an optimal level of unverifiable quality *when the interaction between the public purchaser and the hospital is stable over time.*

Then, we look for supporting empirical evidence for this result, considering the Italian NHS as our testing ground. In Italy, public insurers contract with both public and private hospitals (either for-profit or not-for-profit) but contracts are largely incomplete in terms of quality of care. In addition, the NHS is managed at the regional level and this introduces regional variability in political stability and the share of private producers contracting with public insurers. Both sources of variability can be exploited to study whether and how the stability of the relationship between the public purchaser and the hospitals affects the quality of care. To this end, we build a novel data set including measures of various dimensions of quality, as well as the tenure and personal characteristics of both regional governors and regional ministries of health. Governors and ministries are the key political figures in the contractual relationship between the public insurer and the hospitals. Our results show that longer tenure is associated with reduced inappropriateness and enhanced outcome quality for some of the proxies used to measure quality, supporting the view that political stability facilitates the achievement of improved quality with P4P-RC.

The contribution of our paper is twofold: first, we contribute to the theoretical literature studying P4P schemes to enforce unverifiable quality even in the presence of a non-altruistic provider. Eggleston (2005) and Kaarboe & Siciliani (2011) are two of the most recent works studying the incentive scheme able to induce providers to deliver both verifiable and unverifiable quality.[4] In particular, after discussing examples of unverifiable quality in the

[2]An example of an experimental P4P scheme introduced in Italy is discussed in Peluso et al. (2019*a*). The authors find a positive effect on the hospitals' performance limited to outcomes that can be more influenced by managers (like the number of readmissions, transfers, and returns to the surgery room) but no significant changes for the number of voluntary discharges and mortality.

[3]Relational contracts have been pioneered by Bull (1987) and MacLeod & Malcomson (1989) and applied in several fields: labour market (Baker et al. 2002, Wennberg & Gittelsohn 1973, Calzolari et al. 2017, Rayo 2007, Taylor & Plambeck 2007, Andrews & Barron 2016), regulation (Cesi et al. 2012), procurement (Albano et al. 2023, 2017, Calzolari & Spagnolo 2009, Doni et al. 2006), environmental regulation (Cesi & D'Amato 2023) and experimental economics (Fehr et al. 2007, Bigoni et al. 2014). See also Malcomson (2016) for an extensive review.

[4]Beitia (2003) studies the optimal market structure, oligopoly or monopoly, in a context of unverifiable

health system in the UK and the US, Eggleston (2005) introduces a P4P scheme as a device to enforce non-contractible quality. She shows that while a P4P scheme may increase the verifiable quality dimension (to the benefit of patients), it may concurrently decrease the non-verifiable quality level (to the detriment of patients), resulting in an overall welfare effect that is ambiguous. Kaarboe & Siciliani (2011) integrate and generalize the model by Eggleston (2005) studying a P4P scheme in a static scenario by setting up a sequential game between the purchaser and an altruistic provider. In this game, the purchaser first sets the payments, and then the provider determines quality (both verifiable and unverifiable). They show that the desirability of a P4P depends on (i) the complementarity/substitutability between verifiable and unverifiable quality and (ii) the provider's degree of altruism.

We add to this literature by studying the characteristics of an optimal *relational* contract able to enforce quality in a dynamic context where quality is unverifiable and the provider is semi-altruistic. Removing altruism is important, as the assumption of altruism in this literature has been confirmed to have a crucial effect on the providers' performance.[5] Although the assumption is not granted for all the providers of health services, in a theoretical framework altruism serves as the motivating factor for providers to deliver unverifiable quality. In a static scenario, this mitigates the moral hazard problem arising from the lack of external enforcement of unverifiable quality, with clear cross implications for the verifiable component.

Second, we provide an empirical investigation into the impact of political stability and extended tenure on various dimensions of healthcare quality. Since structural quality measures are required to enter the NHS market, our focus is on proxies for procedural quality (like the share of C-sections out of total deliveries and the share of Percutaneous Coronary Interventions delivered on time) and outcome quality (like patients' satisfaction mapped through official surveys).

The remainder of the paper is structured as follows: Section 2 formalizes the theoretical model, Section 3 provides the empirical analysis, and Section 4 concludes.

2 The Model

We consider an infinitely repeated game with two active players, the public purchaser P of health services and a hospital H that delivers the service. At all t, with $t = 0, ..., \infty$, the purchaser awards the delivery of the service by proposing a P4P scheme to the provider requiring qualities $(q_1, q_2) \in [0, +\infty\}$ at prices (p_1, p_2) and a transfer T. Both qualities are observable, however, q_1 is verifiable and q_2 is unverifiable (hence, not contractible, and not enforceable by a court of law).

The static game of our model is in line with the most recent literature (see, e.g., Kaarboe & Siciliani (2011)). The public purchaser receives a benefit of $B(q_1, q_2)$ from delivering the

quality by letting the regulator use a two-part tariff in which the hospital is paid a variable part depending on the number of patients choosing the hospital.

[5] Olivella & Siciliani (2017) find that altruism affects quality and whether the quality is observable matters in terms of the provider's incentive. Siciliani (2009) studies how prices affect quantity when public providers vary in altruism. Makris & Siciliani (2013) explain how the level of altruism affects quantity (relative to the first best) under adverse selection.

service to insured patients, that is increasing in quality, $B_{q_i}(q_i, q_j) > 0$ and $B_{q_i q_i}(q_i, q_j) \leq 0$, with $i = 1, 2$ ($j \neq i$). We also assume that qualities can be either substitutes (complements) for patients' benefits: $B_{q_i q_j}(.) < 0$ ($B_{q_i q_j}(.) > 0$). Under substitability (complementarity), more q_j reduces (increases) the marginal benefits of q_i.

Purchaser payoff W at time t is given by the benefit B net of payment to the hospital:

$$W(q_1, q_2) = B(q_1, q_2) - T - p_1 q_1 - p_2 q_2,$$

In line with the literature, we assume the hospital to be semi-altruistic, meaning that it also cares about the benefit of patients besides the financial incentives provided by the full payment net of costs. The degree of altruism is measured by α: the higher α, the higher the interest for patients. The hospital total payoff π is given by:

$$\pi(q_1, q_2) = \alpha B(q_1, q_2) + T + p_1 q_1 + p_2 q_2 - \phi(q_1, q_2),$$

when awarded the contract for the service, 0 otherwise. The hospital cost for delivering the service is $\phi(q_i, q_j)$, satisfying $\phi_{q_i}(q_i, q_j) > 0$ and $\phi_{q_i q_i}(q_i, q_j) > 0$, with $i = 1, 2$ ($j \neq i$). We also assume that the two quality dimensions can be either substitutes (complements) in the hospital cost function: $\phi_{q_i q_j}(.) > 0$ ($\phi_{q_i q_j}(.) < 0$). Under substitability (complementarity), more q_j increases (decreases) the marginal costs of q_i. The public purchaser has complete information: it knows the cost of delivering the quality of the service and its benefit is common knowledge.

The dynamic game we consider is an infinite repetition of the following sequential stage game:

Stage 1 (*The payment definition stage*): the purchaser sets prices p_1 and p_2 for each unit of quality and a fixed transfer T;

Stage 2 (*The execution stage*): once the payment is defined, the hospital chooses qualities and delivers the service. Qualities are observed and all payoffs are collected.

In what follows, we first find and compare the first-best and the second-best solution of the stage game. We then move to analyze the relational contract in a repeated game.[6]

2.1 First-best complete contract

The first-best solution is derived in a setting where the purchaser maximizes welfare W assuming both qualities are observable and verifiable, then contractible. Since qualities are both verifiable, it is possible to set a specific price for each dimension of quality. The following Definition describes the first-best solution:

[6] Considering the unverifiability of quality, the placement of payment at different stages has no impact on the equilibrium. Moreover, the recurring nature of the game ensures that payment timing, whether at the current game's end or the subsequent subgame's beginning, does not alter the subgame perfect equilibrium characteristics.

Definition 1 The first best solution denoted by $(q_1^{FB}, q_2^{FB}, p_1^{FB}, p_2^{FB})$ is characterized by the following FOCs:

$$p_1^{FB} = B_{q_1}\left(q_1^{FB}\left(p_1^{FB}\right), q_2^{FB}\left(p_2^{FB}\right)\right)$$
$$p_2^{FB} = B_{q_2}\left(q_1^{FB}\left(p_1^{FB}\right), q_2^{FB}\left(p_2^{FB}\right)\right)$$
$$(1+\alpha) B_{q_1}\left(q_1^{FB}\left(p_1^{FB}\right), q_2^{FB}\left(p_2^{FB}\right)\right) = \phi_{q_1}\left(q_1^{FB}\left(p_1^{FB}\right), q_2^{FB}\left(p_2^{FB}\right)\right)$$
$$(1+\alpha) B_{q_2}\left(\phi\left(q_1^{FB}\left(p_1^{FB}\right), q_2^{FB}\left(p_2^{FB}\right)\right)\right) = \phi_{q_2}\left(q_1^{FB}\left(p_1^{FB}\right), q_2^{FB}\left(p_2^{FB}\right)\right)$$

Proof. See Appendix. ■

The intuition is simple: for each quality dimension, the first-best prices p^{FB} are equal to the marginal benefits, while the first-best qualities are obtained by equalizing marginal benefits to marginal costs. Qualities and payments are all contractible and the first-best P4P contract corresponds to a standard complete contract.

2.2 Second-best incomplete contract in the static game

We now turn to the case in which q_2 is not verifiable. As q_2 is then not contractible, following the argument in Kaarboe & Siciliani (2011), we assume that the contract will fix a payment $T - p^{SB} q_1$, where the variable part of the payment depends only on the verifiable quality q_1. By solving for the equilibrium of the static game, we have the following result:

Proposition 1 *The sub-game equilibrium $(q_1^{SB}, q_2^{SB}, p^{SB})$ of the static game is:*

$$\alpha B_{q_2}(q_1, q_2) = \phi_{q_2}(q_1, q_2)$$

$$p^{SB} + \alpha B_{q_1}(q_1, q_2) = \phi_{q_1}(q_1, q_2)$$

$$p^{SB} = B_{q_1}\left(q_1^{SB}(p), q_2^{SB}(p)\right) + \frac{\frac{dq_2}{dp}}{\frac{dq_1}{dp}} B_{q_2}\left(q_1^{SB}(p), q_2^{SB}(p)\right)$$

$$T^{SB} = \phi\left(q_1^{SB}\left(p^{SB}\right), q_2^{SB}(p)\right) - p^{SB} q_1^{SB} - \alpha B\left(q_1^{SB}\left(p^{SB}\right), q_2^{SB}\left(p^{SB}\right)\right)$$

with hospital and purchaser's payoff given by $\pi^{SB}\left(q_1^{SB}, q_2^{SB}, p^{SB}, T^{SB}\right)$ *and* $W\left(q_1^{SB}, q_2^{SB}\right)$, *and* $\frac{\frac{dq_2}{dp}}{\frac{dq_1}{dp}} = \frac{\alpha B_{q_1 q_2} - \phi_{q_1 q_2}}{-\alpha B_{q_2 q_2} + \phi_{q_2 q_2}}$

Proof. See Appendix. ■

The static equilibrium replicates the result in Kaarboe & Siciliani (2011). In the static game, with $\alpha > 0$, the hospital delivers the verifiable quality q_1 by comparing marginal costs and marginal benefits, which depends now on the price $p^S B$ and α, whereas in the condition on unverifiable quality q_2 appears only the parameter α. By comparing the first- with the second-best quality dimensions, we note that the levels of both qualities are distorted, but for different reasons. On the one hand, the distortion of q_1 depends on the value $\frac{\frac{dq_2}{dp}}{\frac{dq_1}{dp}} B_{q_2}$ in the optimal condition on price. Since the sign of $\frac{dq_1}{dp}$ is strictly positive, the distortion depends on the ambiguous sign of $\frac{dq_2}{dp}$. On the other hand, the level of q_2 is distorted because there are no specific rewards for the hospital. An increase in p^{SB} decreases (increases) q_2 when qualities are substitutes (complements) in benefit and providers costs: $B_{q_1 q_2}(.) < 0$ and $\phi_{q_1 q_2}(.) > 0$. The ambiguity in the sign of the overall effect $(\alpha B_{q_1 q_2} - \phi_{q_1 q_2})$ strictly depends on the degree of altruism when qualities are complements in benefits but substitutes in costs. Clearly, for a hospital caring only about financial rewards (i.e., $\alpha = 0$), the hospital finds it optimal to deliver only the verifiable quality and set $q_2 = 0$. Notice that the condition for the price of the verifiable quality is downward distorted with respect to the first best. In the next section, we show that a relational contract is able to mitigate the distortions in the static setting.

2.3 P4P Relational contract

We now consider a repeated game composed of an infinite repetition of the previous stage game, in which both the public purchaser and the hospital are characterized by the same discount factor $\delta \in [0, 1)$. Standard analysis of repeated games shows that the discount factor reflects not only the players' intertemporal preferences but also other circumstances, such as the frequency and the probability of continuing the game in the following periods between the same players.

The P4P-RC contract defines the level of qualities, transfer, and prices (q_1^*, q_2^*, T^*, p_1^*, p_2^*). Although q_2^* is not verifiable, differently from the static contract, in a P4P-RC it is "relationally contractible" as suggested Levin (2003).[7]

As in Levin (2003), we assume that both the public purchaser and the hospital use trigger strategies s_P and s_H which are defined as follows:

Purchaser (s_P): Set T^*, p_1^*, p_2^* at time t if up to time $t-1$, the provider has delivered q_1^* and q_2^*; otherwise set p^{SB} and T^{SB} as in the Nash equilibrium of the static game forever.

Hospital (s_H): Set q_1^* and q_2^* at time t if up to the first period of time t the purchaser has set p_1^*, p_2^* and T^*; otherwise set q_1^{SB} and q_2^{SB} as in the Nash equilibrium of the static game forever.

The values p^{SB}, T^{SB}, q_1^{SB}, and q_2^{SB} denote the punishment path the two players will play after deviation from the cooperative path. Note that this punishment equilibrium is the Nash equilibrium of the static game in Proposition 1.

[7] In the agency problem in Levin (2003) the principal may decide to reward the agent with a bonus according to some unverifiable performance but the bonus cannot be claimed by the firm. In particular, should the principal decide not to award the bonus the agent is only free not to perform as the principal requires, nothing more. By the same argument, the principal cannot legally claim any performance if not freely chosen by the agent.

This infinitely repeated game is characterized by multiple equilibria. In line with Levin (2003), we let the P4P-RC be defined by the sub-game (perfect) equilibrium values q_1^*, q_2^*, T^*, p_1^*, p_2^* obtained as the solution to the following purchaser's maximization problem:

$$\max_{p_1^*,p_2^*,q_1^*,q_2^*} V = \frac{1}{1-\delta} W\left(q_1^*, q_2^*\right)$$

where V is the intertemporal utility of the purchaser along the "cooperative" path when both players stick to the strategy s_P and s_H, subject to the following incentive compatibility constraint (IC) for the hospital:

$$\frac{1}{1-\delta}\pi^C\left(q_1^*, q_2^*, p_1^*, p_2^*\right) \geq \pi^D\left(q_1^*, \widehat{q}_2\left(.\right), p_1^*, p_2^*\right) + \left(\frac{\delta}{1-\delta}\right)\pi^{SB}\left(q_1^{SB}, q_2^{SB}, p^{SB}, T^{SB}\right) \qquad (1)$$

where: π^C and π^D are the hospital payoffs in the case of "cooperation" and "deviation", respectively; $\widehat{q}_2$ denotes an optimal deviation from the cooperative path (because of the verifiability of q_1, the hospital can only cheat on q_2), where

$$\widehat{q}_2\left(p_1^*, p_2^*, q_1^*\right) = \arg\max_{q_2} \pi\left(q_1^*, q_2, p_1^*, p_2^*\right)$$

that gives

$$\alpha B_{q_2}\left(q_1^*, q_2\right) = \phi_{q_2}\left(q_1^*, q_2\right) \qquad (2)$$

The left-hand side of the incentive compatibility constraint in Eq. (1) is the discounted provider's profit along the cooperative path. Note that, since the unverifiable quality is relationally contractible, it enters the hospital profit. The right-hand side represents the payoff in the case of deviation from the cooperative path. Once the payment is set, a cheating hospital can deviate from the optimal unverifiable quality by delivering $\widehat{q}_2$ at the *execution stage*. This current "deviation" profit is represented in the first part of the right-hand side of (1). The choice of $\widehat{q}_2$ does not reduce the payment but it affects $B\left(.\right)$ and $\phi\left(.\right)$ in the second stage.[8] It is possible to show that a higher degree of altruism reduces the incentive to under-deliver unverifiable quality, as specified in the following Corollary:

Corollary 2 *When $q_2^* > 0$ and $\widehat{q}_2 > 0$, then $\widehat{q}_2 < q_2^*$ if $\alpha < \delta\frac{\phi_{q_2}\left(q_1^*,\widehat{q}_2\right)}{\phi_{q_2}\left(q_1^*,q_2^*\right)-\phi_{q_2}\left(q_1^*,\widehat{q}_2\right)}$.*

When $\alpha = 0$, the best deviation for the hospital is setting $\widehat{q}_2 = 0$ because unverifiable quality is valuable only through future cooperation with the public purchaser.

The incentive to deviate from the cooperative path works through the difference between the deviation profit $\pi^D(q_1^*, \widehat{q}_2, T^*, p_1^*, p_2^*)$ and the cooperative profit $\pi^C(q_1^*, q_2^*, T^*, p_1^*, p_2^*)$, which are defined as follows:

$$\pi^D\left(q_1^*, \widehat{q}_2\right) = \alpha B\left(q_1^*, \widehat{q}_2\right) + T + p_1^* q_1^* + p_2^* q_2^* - \phi\left(q_1^*, \widehat{q}_2\right)$$

$$\pi^C\left(q_1^*, q_2^*\right) = \alpha B\left(q_1^*, q_2^*\right) + T + p_1^* q_1^* + p_2^* q_2^* - \phi\left(q_1^*, q_2^*\right)$$

[8] The difference $\alpha B\left(q_1, q_2\right) - \phi\left(q_1, q_2\right)$ is concave in q_2 and it is possible to see that the deviation quality is positive and it depends on the degree of altruism.

Note that the deviation is profitable when $\pi^D(q_1^*, \widehat{q}_2) > \pi^C(q_1^*, q_2^*)$, which requires

$$\Lambda = \alpha\left[B(q_1^*, \widehat{q}_2) - B(q_1^*, q_2^*)\right] - \left[\phi(q_1^*, \widehat{q}_2) - \phi(q_1^*, q_2^*)\right] > 0 \tag{3}$$

When (1) holds, the hospital should stick to the cooperative strategy than deviating and being punished thereafter. In line with the standard approach of repeated games, constraint (1) can be rewritten as:

$$\delta \geq \frac{\pi^D(q_1^*, \widehat{q}_2(.)) - \pi^C(q_1^*, q_2^*)}{\pi^D(q_1^*, \widehat{q}_2(.)) - \pi^{SB}(q_1^{SB}, q_2^{SB}, p^{SB}, T^{SB})} \tag{4}$$

As standard in relational contracting, Eq. (4) gives the lowest discount factor such that the two players stick to their cooperative strategies that enforce the P4P-RC as a sub-game perfect equilibrium of the game. A similar IC constraint should be added to ensure that also the public purchaser does not deviate from the cooperative strategy offering T^*, p_1^*, p_2^*. However, due to the sequential actions, the IC constraint for the public purchaser always holds because if the purchaser deviates at the first stage it will be punished by the hospital in the same period, without inducing any current gain from the deviation.[9]

The following proposition defines the optimal P4P-RC:

Proposition 3 *Let:*

$$\overline{\delta} = \frac{\Lambda}{\left[\pi^D(q_1^*, \widehat{q}_2) - \pi^{SB}(q_1^{SB}, q_2^{SB})\right]} \tag{5}$$

When $\delta \geq \overline{\delta}$, the strategies s_P and s_H define a self enforcing P4P-RC entailing (q_1^*,q_2^*,T^*) and (p_1^*, p_2^*) such that:

$$B_{q_1}(q_1^*, q_2^*) = \frac{1-\delta}{\alpha+\delta}\left(\alpha B_{q_1}(q_1^*, \widehat{q}_2) - \phi_{q_1}(q_1^*, \widehat{q}_2)\right) + \frac{1}{\alpha+\delta}\phi_{q_1}(q_1^*, q_2^*) \tag{6}$$

$$B_{q_2}(q_1^*, q_2^*) = \frac{1}{\alpha+\delta}\phi_{q_2}(q_1^*, q_2^*) \tag{7}$$

$$p_2^* = B_{q_2}(q_1^*, q_2^*) \tag{8}$$

$$p_1^* = B_{q_1}(q_1^*, q_2^*) \tag{9}$$

with T^* satisfying the binding IC.

Proof. See the Appendix. ■

Proposition 3 shows our main result: the higher the weight given to the future interaction (the discount factor δ), the higher the hospital's incentive to stick to the required qualities (hence, the smaller the distortions from the first best outcome). A low discount factor (a less patient hospital) induces an *upward* distortion in the condition for the unverifiable quality q_2. Results for the verifiable quality q_1 are less clear; when $\alpha = 0$, we can have either (i)

[9] If instead the hospital and the purchaser played their actions simultaneously in each period, then the P4P-RC should need to satisfy both IC constraints, one for each player.

a *downward* distortion when verifiable and unverifiable qualities are substitutes in costs, or (ii) an *upward* distortion with complementarity in costs.

Note that, although the optimal price satisfies the first-best condition, this does not imply that the optimal price is set at the specific first-best level. In fact, the optimal qualities affect the first-best conditions to fine-tune the provider incentives to cooperate and, at the same time, minimize the distortions. Thus, we cannot implicitly consider the *downward* or *upward* distortions in the quality conditions as cases of *under* or *over* provision of quality with respect to the first-best, respectively.

To define the willingness of the purchaser to substitute away verifiable and unverifiable quality, we can rewrite (6)-(7) as follows:

$$\frac{B_{q_1}\left(q_1^*, q_2^*\right)}{B_{q_2}\left(q_1^*, q_2^*\right)} = \frac{\phi_{q_1}\left(q_1^*, q_2^*\right) + (1-\delta)\left[\alpha B_{q_1}\left(q_1, \widehat{q}_2\right) - \phi_{q_1}\left(q_1, \widehat{q}_2\right)\right]}{\phi_{q_2}\left(q_1^*, q_2^*\right)} \tag{10}$$

Eq. (10) depends on the hospital incentive to deviate from the unverifiable quality, weighted by the time preferences $(1-\delta)$. When the hospital cares a lot about the future contracts with the purchaser (i.e., it is almost "infinitely" patient, $\delta \to 1$), we obtain the first best ratio. However, the lower δ, the higher the weight given by the purchaser to the provider's incentive to deviate when choosing unverifiable quality. A low discount factor induces the hospital not to care enough about future interaction with the purchaser (in terms of both cooperative profit and punishment). Hence, the public purchaser needs to "fine-tune" the required levels of qualities specified in the relational contract to tackle this incentive to deviate. In particular, a low discount factor induces the purchaser to require less verifiable quality to mitigate the incentive to reduce unverifiable quality.

2.4 Testable predictions

Our repeated game between a public purchaser and a hospital delivering healthcare services is meant to replicate the interactions emerging in the reality of many healthcare systems around the world. Contracts are mostly incomplete since the purchaser is unable to specify the procedural and outcome quality of the services. Since contracts are incomplete, in addition to contracting with public hospitals (for which α is large and positive), relational contracts based on the value of the ongoing relationship emerge as a standard solution in the dynamic game between the purchaser and the non-altruistic providers (for which α is zero or close to zero).

In this framework, the length of the relationship between public purchasers and hospitals is influenced by political elections, which can terminate the experience of managers acting as public purchasers appointed by elected politicians. Political stability is then crucial to developing a long-term relationship between the purchaser and the hospitals, in particular private non-altruistic hospitals. Variations in political stability across different purchasers will then allow us to investigate the effect of the stability of interaction between purchasers and hospitals on the optimal provision of both unverifiable and verifiable qualities.

For what concerns the level of *unverifiable* quality, Corollary 2 states that a less altruistic provider under-provides unverifiable quality when deviating from the P4P-RC. Since the incentive to deviate is decreasing in δ, we expect *political stability to improve the provision* *of unverifiable quality in the presence of non-altruistic hospitals.*

For what concerns the level of *verifiable* quality, the model does not provide a clear prediction on the effect of political stability, which depends on several dimensions such as the degree of complementarity or substitutability between verifiable and unverifiable qualities. Consider, for instance, substitutability in costs. Looking at Eq. 10, which captures the willingness of the purchaser to substitute away verifiable and unverifiable quality, we can observe that when the provider is non-altruistic ($\alpha = 0$), the numerator of (10) is lower than the numerator in the first best (when $-(1-\delta)\,\phi_{q1}\left(q_1, \widehat{q}_2\right) < 0$), meaning that, relative to the first best, the purchaser is less willing to give up unverifiable quality for more verifiable quality *ceteris paribus*.[10] This incentive depends also on δ: when δ is close to one (in the case of "infinitely" patient players), the first best is restored; when $\delta \to 0$, the downward distortion increases. Therefore, political stability should increase the provision of verifiable quality toward the first best if qualities are substitutes in costs and $\alpha = 0$, while in the case of altruistic hospitals, this effect can be mitigated and even reversed. Hence, we expect *the effect of political stability on the level of verifiable quality to be unclear*, even though it can be positive in the case of non-altruistic providers.

3 Case study: the Italian NHS

We look at the Italian NHS to find supporting evidence for our model. The Italian healthcare system is a public insurance scheme, tax-funded by the Central government and managed at the regional level (Turati (2013)). Both the main principles of the system and the core basic package of services (to be evenly provided across the national territory) are defined and mostly funded by the Central government. Within this framework, regions are largely autonomous in defining the organization of the provision of care, deciding for instance the structure of the public hospital network, and if and how to contract with private providers.

Since the early Nineties, the Italian NHS adopted a quasi-market model that has been adapted to local characteristics by each regional government. According to this model, public purchasers —in Italy, Local Health Authorities (LHAs), i.e., public bodies operating within each region and financed by regional governments— buy services from both public and private hospitals competing for patients, who receive services free of charge. One source of differentiation is the role each regional government decides to leave to private providers in supplying services for the public insurer. Private providers have to be accredited by regional governments, and they are granted public funds for services delivered on behalf of the NHS. The process of "accreditation" is aimed at verifying *structural* quality characteristics of hospitals.

Every year, LHAs contract with public and private hospitals, allocating their budget for hospital activity among independent public hospitals, public hospitals directly managed by the LHA, and accredited private hospitals. In particular, with independent public hospitals and private accredited hospitals, the purchaser contracts the number and type of admissions, as well as the restrictions (overall ceiling and tariff caps and cuts) in case of excess

[10] The reason is that requiring a higher level of verifiable quality (on which the provider cannot cheat) affects the incentive to cheat on the level of unverifiable quality. The higher the required level of verifiable quality, the higher the incentive to cheat and save costs ($\phi_{q_1}\left(q_1, \widehat{q}_2\right)$).

production. Admissions are paid under a prospective payment system based on Diagnosis Related Groups (DRGs) (Fabbri & Robone 2010, Brenna & Spandonaro 2015). The Central government defines the reference DRG tariffs to be meant as the maximum amount of money that the National Health Service is willing to pay for that particular service. However, regions can modify national tariffs for several reasons, such as enforcing the use of a specific technology and encouraging or hindering the provision of a certain service by decreasing the reimbursement rate relative to its best alternative.

3.1 Data and descriptive evidence

Our empirical analysis is based on aggregate regional-level data.[11] We build a new database merging relevant data from several sources. The main database for regional data (health outcomes, supply, and patients' satisfaction) is the *Health for All* OECD database. Data on more detailed health outcomes come from the *Programma Nazionale Esiti* (National Outcomes Program), a database managed by the Agency for Regional Healthcare Services (Agenas) and the Ministry of Health since 2008. Data on DRG tariffs come from the reports published periodically by the Italian association of firms producing medical devices (*Confindustria Dispositivi Medici*, former Assobiomedica). We also collected information on political variables from the Ministry of Interior database. The longest period covered by our data extends through the years from 1996 to 2020; the length of the period covered in the analysis depends exclusively on data availability.

The main variables we are interested in are proxies for the quality of hospital services and the length of the relationship accounting for the discount factor δ. Precise and well-established empirical measures of verifiable and (even more) unverifiable quality of healthcare do not exist in both the management practice and in the literature.[12] Building on Donabedian (1988), we consider different measures of quality to capture both *procedural* and *outcome* quality.[13] In particular, we select five proxies of quality. Unfortunately, there are no clear rules of thumb in the literature to identify which measures of quality are verifiable and which are not. In general, variables reflecting the structural quality of hospital buildings and human resources should be considered verifiable (e.g., Eggleston (2005)), and in fact, they are used in the accreditation process of private providers. Variables mapping procedural quality should be verifiable, but allowing clinical discretion makes this dimension of quality almost unverifiable (e.g., Hurwitz (1995)). Finally, outcome quality is also difficult to verify, both

[11] Unfortunately, data relative to proxies of quality at the hospital level are still largely unavailable to researchers. Notice that, since the governance of local health care systems is delegated to the two autonomous provinces of Trento and Bolzano, we consider the two provincial governments separately in the following analysis instead of the region Trentino-Alto Adige. This raises the number of units to 21.

[12] This is not to say that one cannot find examples in the literature of papers attempting to measure quality. For instance, looking at the differences between public and private hospitals in Italy, Moscone et al. (2020) use 30-day mortality and emergency readmission as measures of clinical quality for very specific treatments like AMI, stroke, and hip fracture. Both measures refer to *outcome* quality but it is hard to find a consensus on whether these two proxies truly reflect the quality of services, especially from the side of public health scholars. See, e.g., Fischer et al. (2011) and Press et al. (2013).

[13] As already remarked, the specification of *structural* quality parameters is embedded in the accreditation process.

in terms of health and subjective evaluation of services.[14]

Two proxies for procedural quality reflect service appropriateness. One variable measures the share of patients with Acute Myocardial Infarction (AMI) receiving Percutaneous Transluminal Coronary Angioplasty (PTCA) within two days (*ptca*). The variable indicates whether or not medical staff comply with clinical guidelines suggesting the importance of a swift intervention to guarantee an optimal outcome.[15] A second, more traditional, proxy for appropriateness measures the share of C-section rates on the overall number of deliveries (*C-section*). As for PTCA, clinical guidelines suggest the use of C-sections in the presence of specific characteristics and needs of the mother and the baby (e.g., in terms of presentation). However, a large literature suggests an excess of C-sections related to several non-clinical factors, like the higher financial incentives relative to more traditional vaginal deliveries (e.g., Francese et al. (2014)). The share of C-sections is very heterogeneous across regions in Italy (reaching one out of two deliveries in some Southern regions), suggesting that inappropriate use of C-sections is widespread in certain regional contexts where institutional quality is low (e.g., De Luca et al. (2021)).

As for outcome quality, we use three variables measuring patients' self-reported evaluations of different dimensions of hospital services. Data are collected by the Italian National Institute of Statistics (Istat) via an official survey, asking participants to indicate their level of satisfaction using a Likert scale. We consider the share of patients "satisfied" or "very satisfied" with the quality of hygienic services, medical assistance, and nursing assistance in hospitals (*hygienic*, *sat. nursing*, *sat. medical*, respectively). The first variable reflects a characteristic of the setting in which services are provided, while the latter two reflect the quality of human resources working in regional hospitals.

We look at regional elections to build a proxy for the length of the relationship between public purchasers and the hospitals. While contracts are awarded by LHAs, all public managers are appointed by the Regional Council chaired by the regional Governor, who is the key political figure in regional politics and is directly elected by citizens. We then look at the length of the Governor's tenure as a proxy for δ in the theoretical model. We define the variable *tenure* as the number of years in office since the elections. The use of this variable, which is a novelty in the literature on relational contracts, is appropriate when discussing contracts awarded by elected bodies. Indeed, the more governors remain in charge, the more likely (and therefore more valued by hospitals) will be future interactions. Since there is a figure in the Regional Council specifically dedicated to healthcare, the regional Ministry of Health (MoH from now on), we also define the variable *tenure* measuring the length of MoH's tenure as an additional proxy for δ.

In studying the relationship between quality and the length of tenure, we control for (i) regional socioeconomic characteristics (the number of resident foreigners, the female unemployment rate, the GDP, and the occupation rate), (ii) the characteristics of the local healthcare system (the level of public health spending, the share of private accredited hospital beds over total beds for acute care beds, the adoption by the regional government of a set of DRG tariffs and a set of tariffs differentiated according to the type of hospital,

[14] Unsurprisingly, P4P schemes are based on objective outcomes such as mortality rates. See, e.g., discussion in Peluso et al. (2019*b*).

[15] Information on this variable is available only for the period from 2008 to 2015.

the presence of an external commissioner to substitute the regional governor in the case of Recovery Plans[16]), (iii) the personal characteristics of regional politicians (whether they are medical doctors, whether they belong to the same party/coalition governing at the national level, and whether they belong to a center-left party). We also consider additional covariates when analyzing specific dependent variables. In particular, we also control for the mean age of women at delivery, the spontaneous abortion rate, and the six-day neonatal mortality to control for the women's and newborns' health status in the models for *C-section* rates; the average number of patients admitted to each hospital for AMI per region per year in specifications that study determinants of *ptca*.

[16]Recovery Plans are imposed by the central government to regional governments recording large deficits. The harder version of these plans requires the appointment of an external commissioner. See the discussion in, e.g., Bordignon et al. (2020).

Table 1: Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Dependent variables: proxies for quality q_1 and q_2					
C-section rate (*C-section*)	524	33.51	9.44	14.12	62.41
Patients receiving PTCA in 2 days (*ptca*)	189	31.23	11.93	0	62.86
Satisfaction with hygienic services (*hygienic*)	422	31.41	14.70	5.24	73.88
Satisfaction with medical assistance (*sat. medical*)	422	38.66	13.72	11	76.54
Satisfaction with nursing assistance (*sat. nursing*)	422	38.41	14.71	10.45	78.79
Control variables					
Proxies for δ					
Governor's tenure (*tenure*)	520	3.68	3.70	0	21
MoH's tenure (*tenure*)	525	2.43	2.50	0	13
Personal features of incumbents					
Political alignment governor (*alignment*)	490	.52	.5	0	1
Political alignment MoH (*alignment*)	525	.36	.48	0	1
Governor medical doctor (*meddoc*)	525	.08	.27	0	1
Moh medical doctor (*meddoc*)	525	.21	.41	0	1
Governor belonging to centre left party (*centleft*)	525	.61	.49	0	1
Moh belonging to centre left party (*centleft*)	525	.43	.49	0	1
Characteristics of regional health care system					
Adoption of DRG tariffs (*tariffario*)	520	.965	.183	0	1
Differentiated tariffs by hospitals (*difftar*)	520	.558	.497	0	1
Supervised recovery plan (*srp*)	525	.095	.294	0	1
Per capita public health spending (*pubspend*)	517	1647.9	372.7	779.3	2495.0
Share of private accredited hospital beds - acute care (*private*)	525	14.49	11.05	0	45.87
Regional socio-economic characteristics					
Number of foreigners over population (%) (*foreigners*)	504	5.01	3.42	.31	12.08
Female unemployment (*unemplf*)	517	12.72	7.74	2.25	40.48
Occupation rate (%) (*occup*)	517	44.154	7.023	30.35	58.87
Per Capita Gross Domestic Product (*gdp*)	520	25,210.17	7,529.19	10,526.9	48,551.35
Additional control variables for C-section					
Mothers' age at delivery (*agedelivery*)	522	31.34	.95	28.5	33.15
Spont. abortion rate per 1,000 live births (*abortionrate*)	486	121.77	21.40	38.2	197.61
Six day neonatal mort. per 10,000 live births (*neonmort*)	466	8.921	4.869	0	29.95
Additional control variables for AMI					
Average number of patients admitted for AMI (*AMIadm*)	187	215.205	66.888	61.5	394.276

Descriptive statistics for all the variables used in the analysis are reported in Table 1. All measures of quality display a large range of variability. C-sections represent 33% of deliveries on average, a share higher than conventional benchmarks (15%) estimated by the WHO to define appropriate C-sections (e.g., Francese et al. (2014)), with a range from 14 to 62%. The average share of patients satisfied or very satisfied with hospital services is 31.4% for the conditions of hygienic services, 38.6% for medical assistance, and 38.4% for nursing assistance. These numbers suggest that users, in general, are not very satisfied with the quality of hospital services. However, again, the range of variability is impressive: in some regions, only around 10% (or even less) of patients are generally satisfied, while this share goes well above 70% in some others. Finally, considering patients admitted for AMI, only 31.23% are treated according to guidelines, receiving PTCA within 2 days. The range of variability goes from zero to about 63%, suggesting a large heterogeneity in terms of procedural quality.

As for our main independent variables, *tenure* for regional governors shows an average of 3.6 years, while the average for the regional Ministry of Health (MoH) is 2.4 years. These values reflect the typical five-year mandate for MoH (with *tenure* taking values from zero to 4), while governors are more likely to remain a second term in office. Figure 1 provides a visual description of the two variables across regions. The figure emphasizes the large variability in the data. First, about 3/4 of governors are characterized by a tenure longer than one term, reaching two terms in most cases. This is not the case for the MoH: half of them experience no more than one term. Second, there are some regions in which the governor and/or the MoH experience an unusually long duration of their tenure. For instance, in Emilia-Romagna, both the governor and the MoH experienced three terms in office; in Lombardy and the small province of Bolzano, the governor experienced more than three terms. Interestingly, in Umbria, the MoH has experience in office longer than the regional governor. We also control for personal characteristics of the governor and the MoH: some of them are medical doctors (7% and 20% of cases, respectively); 52% (36%) of governors (MoH) are politically aligned with the Central government, and 61% (43%) belong to a center-left coalition.

As for the other covariates, Figure 2 provides information on the size of the regional market for healthcare services and the characteristics of the market for hospital services. The size of the market managed by the public purchasers is represented using the per-capita public health spending (the average is 1,647 euros). Except for Bolzano, where spending is historically higher (partly for the larger degree of autonomy), spending is quite homogeneous across regions, with a variability that is largely dictated by the growth over time of spending. However, the structure of the market, defined by the share of private accredited hospital beds over total beds for acute care, looks very different across the regions. First, the share of private hospital beds is limited for many regions below 20% (the average is 14%); hence, in most cases, aggregate quality is largely determined by public hospitals, characterized by a large α . Second, in some regions like Lombardy and Tuscany, the share of private hospital beds is stable over time (around 20% and 10%, respectively), reflecting clear choices of the regional governments. Third, changes over time have been substantial in Southern regions like Calabria and Campania, two regions where the share of private hospital beds is the highest compared to other regions. Interestingly, Calabria is one of the regions where the tenure of both the governor and the MoH is the lowest, not exceeding one term in office. For the other variables describing the characteristics of the healthcare system, nearly all regions have implemented the new prospective payment system based on DRG tariffs between 1997 and 2000. However, only half of them allow for differentiated tariffs based on the type of hospital. In a few cases, some regions are under a supervised recovery plan, imposing choices toward the reduction of fiscal deficit and the improvement of outcomes (e.g., Bordignon et al., 2020).

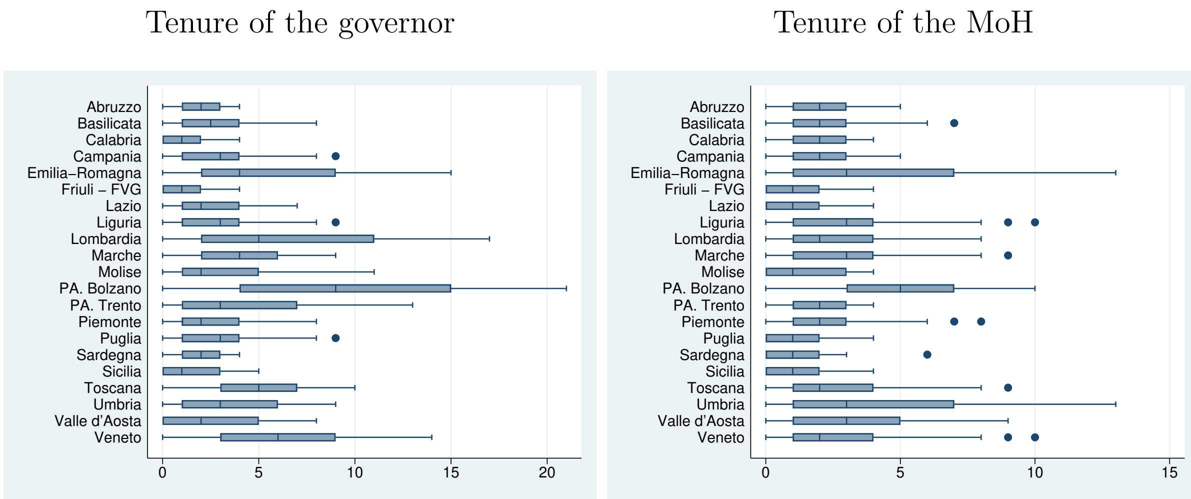

Figure 1: Circles indicate outside value, the upper adjacent line the maximum value, the upper hinge the 75th percentile, the line in the box the median, the lower hinge the 25th percentile, the lower adjacent value the minimum value.

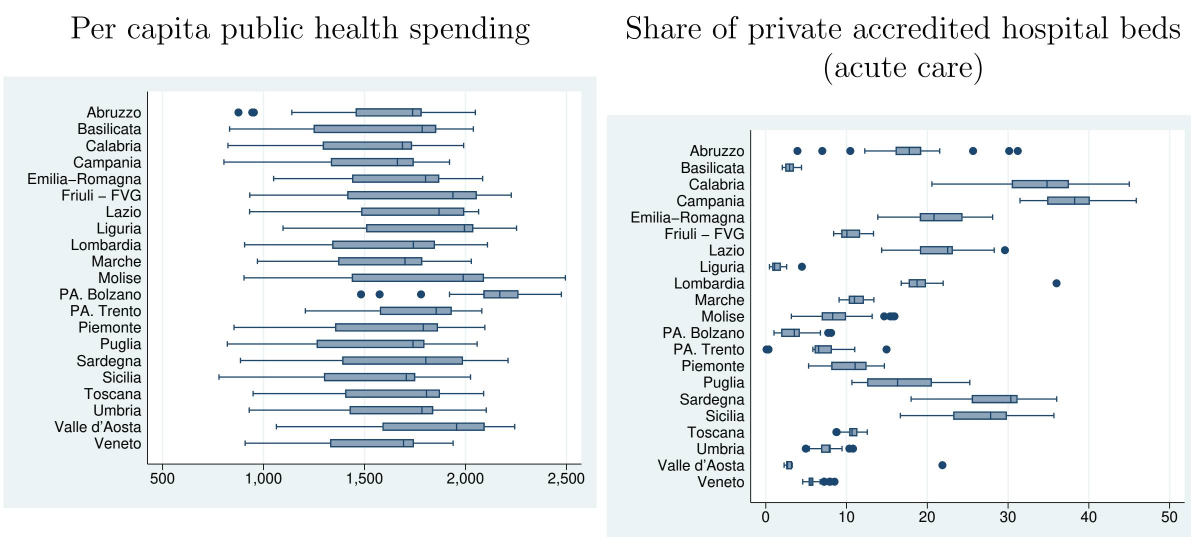

Figure 2: Circles indicate outside value, the upper adjacent line the maximum value, the upper hinge the 75th percentile, the line in the box the median, the lower hinge the 25th percentile, the lower adjacent value the minimum value.

3.2 Empirical strategy

To identify the effect of the length of tenure on the quality of services, we begin considering a very simple model:

$$q^{k}_{\ i,t} = c + \beta_1 tenure_{i,t} + \gamma X_{i,t} + \alpha_i \sum r_i + \alpha_t \sum y_t + \epsilon_{i,t}; \qquad k = (1, ..., 5) \tag{11}$$

where ${q^k}_{i,t}$ stands for one of the five k dimensions of quality in region i in year t. The main variable we are interested in is $tenure_{i,t}$, which is the time, measured in years, spent by the same regional governor (or the regional MoH) in office. Identification of β_1 is allowed by the large variability observed in Figure 1. The vector $X_{i,t}$, is a set of control variables, with specific controls added for specific dimensions of quality. r and y are region- and year-fixed effects, respectively. ϵ is the error term.

Eq. 11 relies on an "aggregate" measure of quality ${q^k}_{i,t}$ which represents a regional-yearly average proxy of the quality supplied by each hospital (either public or private, for-profit or not-for-profit) providing services in each region. However, the theoretical model suggests the importance of the length of tenure especially for semi-altruistic providers. We then augment Eq. 11 with an additional variable obtained interacting the market share of private accredited hospital beds over total beds (*private*) with (*tenure*):

$$\begin{aligned} q_{i,t} = c + \beta_1 tenure_{i,t} + \beta_2 tenure_{i,t} \times private_{i.t} + \\ + \gamma X_{i,t} + \alpha_i \sum r_i + \alpha_t \sum y_t + \epsilon_{i,t}; \qquad k = (1, ..., 5) \end{aligned} \tag{12}$$

All the models have been estimated using OLS. Given the small number of regions (as a possible dimension for clustering), we opted for robust standard errors in all the specifications.

3.3 Results

Table 2 and Table 3 report the estimates of Eq. 11 and Eq. 12 separately for the regional governor and the MoH.[17] In both tables, each column considers one of the five proxies for quality.

[17] The complete tables with coefficients for all the controls are in Appendix B.

Table 2: The link between Governor's tenure and quality of healthcare services

VARIABLES	C-section	hygienic	sat. medical	sat. nursing	ptca
Panel A: no interaction (Eq. 11)					
tenure	-0.1146***	0.0222	0.1840*	0.0459	-0.0968
	(0.034802)	(0.124963)	(0.109083)	(0.112743)	(0.105419)
Panel B: interaction w/private (Eq. 12)					
tenure	-0.1543**	-0.1688	0.0732	-0.1183	0.0447
	(0.068376)	(0.197575)	(0.171260)	(0.187441)	(0.155359)
$tenure \times private$	0.0033	0.0156	0.0091	0.0134	-0.0119
	(0.004365)	(0.011065)	(0.012178)	(0.012606)	(0.009994)
Observations	411	374	374	374	170
R-squared	0.96	0.80	0.78	0.81	0.88
Women's/newborns' health	YES	NO	NO	NO	NO
Additional controls AMI	NO	NO	NO	NO	YES
Regional socio-economic char.	YES	YES	YES	YES	YES
Characteristics reg. healthcare system	YES	YES	YES	YES	YES
Personal features of the incumbent	YES	YES	YES	YES	YES
Region FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Notes: OLS. Robust standard errors in parentheses. Levels of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Proxies for variables are: C-section rates (C-section), share of people satisfied or very satisfied with hygienic services (hygienic), with medical assistance (sat. medical), and with nursing assistance (sat. nursing), share of patients receiving PTCA in two days (ptca).

Table 3: The link between MoH's tenure and quality of healthcare services

VARIABLES	C-section	hygienic	sat. medical	sat. nursing	ptca
Panel A: no interaction (Eq. 11)					
tenure	-0.1405***	-0.2012	-0.2626	-0.2725	0.3331*
	(0.045337)	(0.197006)	(0.182204)	(0.174422)	(0.194611)
Panel B: interaction w/private (Eq. 12)					
tenure	-0.1246	-0.4193	-0.1959	-0.1752	0.6066**
	(0.080824)	(0.292319)	(0.281563)	(0.262266)	(0.301382)
$tenure \times private$	-0.0015	0.0201	-0.0061	-0.0090	-0.0249
	(0.006449)	(0.017499)	(0.018260)	(0.016984)	(0.016013)
Observations	439	394	394	394	186
R-squared	0.95	0.81	0.80	0.83	0.83
Women's/newborns' health	YES	NO	NO	NO	NO
Additional controls AMI	NO	NO	NO	NO	YES
Regional socio-economic char.	YES	YES	YES	YES	YES
Characteristics of reg. healthcare system	YES	YES	YES	YES	YES
Personal features of the incumbent	YES	YES	YES	YES	YES
Region FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Notes: OLS. Robust standard errors in parentheses. Levels of significance: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Proxies for variables are: C-section rates (C-section), share of people satisfied or very satisfied with hygienic services (hygienic), with medical assistance (sat. medical), and with nursing assistance (sat. nursing), share of patients receiving PTCA in two days (ptca).

Looking at Table 2, Panel A, the coefficient for the governor's tenure takes the expected sign, suggesting a positive association with quality, for almost all the proxies, but it is statistically significant only in two out of five cases. The coefficient for C-sections is negative as expected and significant at the 1% level: a repeated and durable interaction between the public purchaser and the hospitals is associated with a reduction of a potentially inappropriate service and thus is associated with an increase in the level of quality. All the coefficients for self-reported satisfaction with the quality of services are positive, but only the coefficient relative to medical assistance is also statistically significant at the 10% level: a longer tenure increases users' satisfaction with medical assistance, and so does the level of perceived quality. The coefficient for *ptca* is instead negative, contrary to what we expected, but not statistically significant. Results in Table 2, Panel B, partially confirm these findings. Only the coefficient for C-sections remains negative and statistically significant at the 5% level. The coefficient for self-reported satisfaction with medical assistance remains positive but turns statistically non-significant. The interaction term $tenure \times private$ is never statistically significant.

We obtain more nuanced results when considering the tenure of the MoH in Table 3. The association with the C-section rates persists, but only in the model without interaction (Panel A): the coefficient is still negative and statistically significant at the 1% level. Contrary to what we expected, coefficients for the three proxies measuring self-reported satisfaction are all negative but statistically non-significant in both specifications. More importantly, now the coefficient for *ptca* is also positive as expected and statistically significant in both specifications. When including the interaction term in Panel B, the magnitude doubles.

Results for C-sections are the only coherent between the governor and the MoH: a longer tenure is associated with a decline in the provision of potentially inappropriate services. The share of C-sections is the most common indicator of inappropriateness: it is very easy to communicate and understand also by politicians and by the general public, and it is monitored yearly by the Ministry of Health, making it a good candidate for building a reputation by hospitals. This is the reason emphasized by our theoretical model: hospitals that expect to interact with the same public purchaser in the next years are keener on behaving properly and building a positive reputation over time. This interpretation adds to those already available in the literature finding the same result, shedding new light on available findings: for instance, Francese et al. (2014) comment that experience in office (*tenure* in our case) can facilitate a regional monitoring activity on hospitals' behavior or the bargaining process with hospital managers in the case of decisions affecting the hospital network. Similarly, De Luca et al. (2021) find that low institutional quality of regional governments (meaning weak rule of law, lack of good governance, and corruption) reinforces hospital incentives to provide less quality.

Results for *ptca* stand up instead as a feature of the MoH. The longer their tenure, the higher the share of patients treated according to guidelines, and the higher the quality. One possible explanation for this result relies on the more "technical" nature of this indicator, as opposed to the share of C-sections, which is more common and used in political debates. The share of MoH who are medical doctors is three times larger than the share of governors who are medical doctors. Hence, reputation building with doctors might require concentrating on more technical indicators.

Finally, the absence of any statistically significant effect on self-reported evaluation of quality is likely to confirm that these measures are hardly observed and discussed at the regional level. Reputation building requires an implicit agreement between the purchaser and the hospitals on the variables to be used to measure quality, and these three variables – which are provided by the National Institute of Statistics (ISTAT) through a yearly survey – have never been used in political debates despite being a good proxy of perceived quality by users of health services.

4 Conclusions

In this paper, we explore the effect of repeated interactions between a public purchaser and providers of health services on the provision of unverifiable quality. Our findings demonstrate that the establishment of a relational contract, grounded in reputation building, enables the public purchaser to approach the first-best outcome the longer the relationship with hospitals.

By using Italian regional data, we define several proxies for quality and we empirically explore how the incentive power of relational contracts depends on political stability. Considering the institutional features of the Italian NHS – a tax-funded public scheme where contracts are written by public purchasers appointed by regional governments – we measure political stability with the tenure of both regional governors and regional Ministries of Health. We estimate the relationship between tenure duration and quality proxies. OLS estimates suggest a positive association between aggregate quality outcomes at the regional level and tenure, especially for variables that are commonly considered good indicators of inappropriateness, such as the share of C-sections. Interestingly, we find that more technical indicators, like the one based on Percutaneous Transluminal Coronary Angioplasty (PTCA), are affected by the length of tenure of the MoH, who, in many instances, is a medical doctor as opposed to the regional governor.

Our results offer insights from both positive and normative perspectives. According to a positive perspective, they help explain why – in the presence of largely incomplete contracts between public purchasers (in Italy, the Local Health Authorities) and public and private hospitals – we obtain relatively good outcomes and performances in several regions. Political stability can be the endogenous mechanism sustaining reputation building: hospitals deliver good outcomes to establish their reputation, concurrently helping local politicians to sustain their reputation with voters, further reinforcing the reputation-building mechanism with hospitals.

According to a normative perspective, they suggest considering relational contracts as a tool to enforce good quality in the provision of healthcare services. The two mechanisms on which we are currently relying are the ownership of hospitals (favoring altruistic providers, like public and, at least to some extent, private not-for-profit hospitals) and P4P incentive payment schemes. One might consider incorporating reputation-building as an additional mechanism to enforce quality. This calls for a definition of a commonly agreed set of indicators to be strictly monitored by regional regulators when evaluating hospital performance. These indicators should be selected with the awareness of the potential influence of hospital managers on them when making their choices. Needless to say, if reputation can be built by focusing on selected indicators, managers will favor these measures against those not mapped by regulators. Hence, the selection of indicators should be carefully considered to avoid strategic considerations.

References

Albano, G. L., Cesi, B. & Iozzi, A. (2017), 'Public procurement with unverifiable quality: The case for discriminatory competitive procedures', *Journal of Public Economics* **145**, 14–26.

Albano, G. L., Cesi, B. & Iozzi, A. (2023), 'Teaching an old dog a new trick: Reserve price and unverifiable quality in repeated procurement', *Journal of Economics & Management Strategy* **32**(2), 377–399.

Andrews, I. & Barron, D. (2016), 'The allocation of future business: Dynamic relational contracts with multiple agents', *American Economic Review* **106**(9), 2742–59.

Baker, G., Gibbons, R. & Murphy, K. J. (2002), 'Relational contracts and the theory of the firm', *The Quarterly Journal of Economics* **117**(1), 39–84.

Beitia, A. (2003), 'Hospital quality choice and market structure in a regulated duopoly', *Journal of Health Economics* **22**(6), 1011–1036.

Benbassat, J. & Taragin, M. (2000), 'Hospital readmissions as a measure of quality of health care: advantages and limitations', *Arch Intern Med.* **160**(8), 1074–81.

Bigoni, M., Spagnolo, G. & Valbonesi, P. (2014), 'Sticks and carrots in procurement: an experimental exploration', *The BE Journal of Economic Analysis & Policy* **14**(3), 893–936.

Boden, W. E., O'Rourke, R. A., Teo, K. K., Hartigan, P. M., Maron, D. J., Kostuk, W. J., Knudtson, M., Dada, M., Casperson, P., Harris, C. L., Chaitman, B. R., Shaw, L., Gosselin, G., Nawaz, S., Title, L. M., Gau, G., Blaustein, A. S., Booth, D. C., Bates, E. R., Spertus, J. A., Berman, D. S., Mancini, G. J. & Weintraub, W. S. (2007), 'Optimal medical therapy with or without pci for stable coronary disease', *New England Journal of Medicine* **356**(15), 1503–1516.

Bordignon, M., Coretti, S., Piacenza, M. & Turati, G. (2020), 'Hardening subnational budget constraints via administrative subordination: The italian experience of recovery plans in regional health services', *Health Economics* **29**(11), 1378–1399.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1002/hec.4138*

Brenna, E. & Spandonaro, F. (2015), 'Regional incentives and patient cross-border mobility: evidence from the italian experience', *International journal of health policy and management* **4**(6), 363–372.

Bull, C. (1987), 'The existence of self-enforcing implicit contracts', *The Quarterly journal of economics* **102**(1), 147–159.

Calzolari, G., Felli, L., Koenen, J., Spagnolo, G. & Stahl, K. O. (2017), 'Relational contracts, competition and innovation: Theory and evidence from german car manufacturers'.

Calzolari, G. & Spagnolo, G. (2009), 'Relational procurement', *Unpublished paper* .

Cesi, B. & D'Amato, A. (2023), 'Relational voluntary environmental agreements with unverifiable emissions', *The B.E. Journal of Economic Analysis Policy* **23**(4), 1045–1071.
URL: *https://doi.org/10.1515/bejeap-2022-0464*

Cesi, B., Iozzi, A. & Valentini, E. (2012), 'Regulating unverifiable quality by fixed-price contracts', *The BE Journal of Economic Analysis & Policy* **12**(1).

Chandra, A., Cutler, D. & Song, Z. (2011), *Who Ordered That? The Economics of Treatment Choices in Medical Care*, Vol. 2 of *Handbook of Health Economics*, Elsevier, chapter 0, pp. 397–432.

De Luca, G., Lisi, D., Martorana, M. & Siciliani, L. (2021), 'Does higher institutional quality improve the appropriateness of healthcare provision?', *Journal of Public Economics* **194**, 104356.
URL: *https://www.sciencedirect.com/science/article/pii/S0047272720302206*

Donabedian, A. (1988), 'Quality of care. how can it be assessed?', *JAMA* **260**(12), 1743–1748.

Doni, N. et al. (2006), 'Optimal information releasing in multidimensional public procurement', *Rivista di Politica Economica* **96**(1/2), 235.

Dumont, E., Fortin, B., Jacquemet, N. & Shearer, B. (2008), 'Physicians' multitasking and incentives: Empirical evidence from a natural experiment', *Journal of Health Economics* **27**(6), 1436 – 1450.

Eggleston, K. (2005), 'Multitasking and mixed systems for provider payment', *Journal of Health Economics* **24**(1), 211 – 223.

Fabbri, D. & Robone, S. (2010), 'The geography of hospital admission in a national health service with patient choice', *Health Economics* **19**(9), 1029–1047.

Fehr, E., Klein, A. & Schmidt, K. M. (2007), 'Fairness and contract design', *Econometrica* **75**(1), 121–154.

Fischer, C., Anema, H. A. & Klazinga, N. S. (2011), 'The validity of indicators for assessing quality of care: a review of the European literature on hospital readmission rate', *European Journal of Public Health* **22**(4), 484–491.
URL: *https://doi.org/10.1093/eurpub/ckr165*

Fischer, C., Lingsma, H. F., van de Mheen, P. J. M., Kringos, D. S., Klazinga, N. S. & Steyerberg, E. W. (2014), 'Is the Readmission Rate a Valid Quality Indicator? A Review of the Evidence', *PLOS ONE* **9**(11), 1–9.

Francese, M., Piacenza, M., Romanelli, M. & Turati, G. (2014), 'Understanding inappropriateness in health spending: The role of regional policies and institutions in caesarean deliveries', *Regional Science and Urban Economics* **49**, 262–277.

Hansmann, H. (1988), 'Ownership of the Firm', *Journal of Law, Economics, and Organization* **4**(2), 267–304.

Hurwitz, B. (1995), 'Clinical guidelines and the law: advice, guidance or regulation?', *Journal of Evaluation in Clinical Practice* **1**(1), 49–60.
URL: *https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1365-2753.1995.tb00007.x*

Kaarboe, O. & Siciliani, L. (2011), 'Multi-tasking, quality and pay for performance', *Health Economics* **20**(2), 225–238.

Levin, J. (2003), 'Relational Incentive Contracts', *American Economic Review* **93**(3), 835–857.

MacLeod, W. B. & Malcomson, J. M. (1989), 'Implicit contracts, incentive compatibility, and involuntary unemployment', *Econometrica: Journal of the Econometric Society* pp. 447–480.

Makris, M. & Siciliani, L. (2013), 'Optimal incentive schemes for altruistic providers', *Journal of Public Economic Theory* **15**(5), 675–699.

Malcomson, J. M. (2016), 'Relational incentive contracts with persistent private information', *Econometrica* **84**(1), 317–346.

Moscone, F., Siciliani, L., Tosetti, E. & Vittadini, G. (2020), 'Do public and private hospitals differ in quality? evidence from italy', *Regional Science and Urban Economics* **83**, 103523.
URL: *https://www.sciencedirect.com/science/article/pii/S0166046218304800*

Newhouse, J. P. (2002), 'Why is there a quality chasm?', *Health Affairs* **21**(4), 13–25.

Olivella, P. & Siciliani, L. (2017), 'Reputational concerns with altruistic providers', *Journal of health economics* **55**, 1–13.

Peluso, A., Berta, P. & Vinciotti, V. (2019*a*), 'Do pay-for-performance incentives lead to a better health outcome?', *Empirical Economics* **56**(6), 2167—2184.

Peluso, A., Berta, P. & Vinciotti, V. (2019*b*), 'Do pay-for-performance incentives lead to a better health outcome?', *Empirical Economics* **56**(6), 2167–2184.

Press, M. J., Scanlon, D. P., Ryan, A. M., Zhu, J., Navathe, A. S., Mittler, J. N. & Volpp, K. G. (2013), 'Limits of readmission rates in measuring hospital quality suggest the need for added metrics', *Health Affairs* **32**(6), 1083–1091.
URL: *https://doi.org/10.1377/hlthaff.2012.0518*

Rayo, L. (2007), 'Relational incentives and moral hazard in teams', *The Review of Economic Studies* **74**(3), 937–963.

Siciliani, L. (2009), 'Paying for performance and motivation crowding out', *Economics letters* **103**(2), 68–71.

Smith, P. C. & York, N. (2004), 'Quality incentives: The case of u.k. general practitioners', *Health Affairs* **23**(3), 112–118.

Talsma, A. K., Lingsma, H. F., Steyerberg, E. W., Wijnhoven, B. P. L. & Van Lanschot, J. J. B. (2014), 'The 30-day versus in-hospital and 90-day mortality after esophagectomy as indicators for quality of care', *Annals of surgery* **260**(2), 267—273.
URL: *https://doi.org/10.1097/SLA.0000000000000482*

Taylor, T. A. & Plambeck, E. L. (2007), 'Supply chain relationships and contracts: The impact of repeated interaction on capacity investment and procurement', *Management science* **53**(10), 1577–1593.

Turati, G. (2013), *The Italian Servizio Sanitario Nazionale: A Renewing Tale of Lost Promises*, Palgrave Macmillan UK, London, pp. 47–66.
URL: *https://doi.org/10.1057/9781137291875*$_3$

Weintraub, W. S., Spertus, J. A., Kolm, P., Maron, D. J., Zhang, Z., Jurkovitz, C., Zhang, W., Hartigan, P. M., Lewis, C., Veledar, E., Bowen, J., Dunbar, S. B., Deaton, C., Kaufman, S., O'Rourke, R. A., Goeree, R., Barnett, P. G., Teo, K. K. & Boden, W. E. (2008), 'Effect of pci on quality of life in patients with stable coronary disease', *New England Journal of Medicine* **359**(7), 677–687.

Wennberg, J. & Gittelsohn, A. (1973), 'Small area variations in health care delivery', **182**(4117), 1102–1108.

Wilt, T. J., MacDonald, R., Rutks, I., Shamliyan, T. A., Taylor, B. C. & Kane, R. L. (2008), 'Systematic review: comparative effectiveness and harms of treatments for clinically localized prostate cancer', *Annals of internal medicine* **148**(6), 435–448.

A: Theoretical Appendix

Proof of Definition 1

Proof. Proceeding by backward induction, we first derive the optimal qualities set at the second stage by the provider that maximizes its profit:

$$p_1 + \alpha B_{q_1}(q_1, q_2) = \phi_{q_1}(q_1, q_2) \tag{A.1}$$

$$p_2 + \alpha B_{q_2}(q_1, q_2) = \phi_{q_2}(q_1, q_2) \tag{A.2}$$

Thereafter, we solve the maximization of W given the optimal levels of quality set by the provider (A.1) and (A.2):

$$W_{p_1} = (1+\alpha) B_{q_1(.)} \frac{dq_1}{dp_1} - \phi_{q_1}(q_1, q_2) \frac{dq_1}{dp_1} = 0$$

$$W_{p_2} = (1+\alpha) B_{q_2}(.) \frac{dq_2}{dp_2} - \phi_{q_2}(q_1, q_2) \frac{dq_2}{dp_2} = 0$$

By using (A.1) and (A.2) we have:

$$W_{p_1} = \frac{dq_1}{dp_1}\left(B_{q_1}(.) - p_1\right) = 0$$

$$W_{p_2} = \frac{dq_2}{dp_2}\left(B_{q_2}(.) - p_2\right) = 0$$

Therefore, the first-best prices are:

$$p_1^{FB} = B_{q_1}\left(q_1^{FB}, q_2^{FB}\right)$$

$$p_2^{FB} = B_{q_2}\left(q_1^{FB}, q_2^{FB}\right)$$

When substituting the first-best prices in (A.1) and (A.2), we have the first-best qualities. ■

Proof of Proposition 1

Proof. The proof is straightforward. The model is solved by backwards induction, starting with the provider's choice of quality levels. At the last stage the provider finds a benefit from the delivery of unverifiable quality only through altruism. In particular, the solution of the static game is:

Stage 2: For a given p, the provider optimally sets q_2^{SB} and q_1^{SB} to maximize:

$$\pi(q_1, q_2) = \alpha B(q_1, q_2) + T + p^{SB} q_1 - \phi(q_1, q_2)$$

Thus, the optimal qualities $q_1(p)$ and $q_2(p)$ are defined below:

$$\alpha B_{q_2}(q_1, q_2) = \phi_{q_2}(q_1, q_2) \tag{A.3}$$

$$p + \alpha B_{q_1}(q_1, q_2) = \phi_{q_1}(q_1, q_2) \tag{A.4}$$

Stage 1: Purchaser's maximization:

$$\max_{p} (1+\alpha) B(q_1, q_2) - T - pq_1$$

$$\begin{aligned} & \text{s.t.} \\ U &\geq 0 \\ & \text{or} \\ T + pq_1 &\geq \phi(q_1, q_2) \end{aligned}$$

Since constraint on quality is binding, we have:

$$\max_{p} (1+\alpha) B(q_1(p), q_2(p)) - \phi(q_1(p), q_2(p))$$

under the constraints from the provider's maximization problem (A.3) and (A.4):

$$\alpha B_{q_2}(q_1, q_2) - \phi_{q_2}(q_1, q_2) = 0; q_2 > 0 \tag{A.5}$$

$$p + \alpha B_{q_1}(q_1, q_2) - \phi_{q_1}(q_1, q_2) = 0; q_1 > 0 \tag{A.6}$$

By maximizing we obtain:

$$\begin{aligned} &\frac{dq_1}{dp}\left((1+\alpha) B_{q_1}\left(q_1^{SB}(p), q_2^{SB}(p)\right) - \phi_{q_1}\left(q_1^{SB}(p), q_2^{SB}(p)\right)\right) + \\ &+ \frac{dq_2}{dp}\left((1+\alpha) B_{q_2}\left(q_1^{SB}(p), q_2^{SB}(p)\right) - \phi_{q_2}\left(q_1^{SB}(p), q_2^{SB}(p)\right)\right) = 0 \end{aligned}$$

Then, using (A.5) and (A.6), we have:

$$p^{SB} = B_{q_1}\left(q_1^{SB}(p), q_2^{SB}(p)\right) + \frac{\frac{dq_2}{dp}}{\frac{dq_1}{dp}} B_{q_2}\left(q_1^{SB}(p), q_2^{SB}(p)\right)$$

The optimal quality is given by:

$$(1+\alpha) B_{q_1}\left(q_1^{SB}(p), q_2^{SB}(p)\right) = \phi_{q_1}\left(q_1^{SB}(p), q_2^{SB}(p)\right)$$

$$(1+\alpha) B_{q_2}\left(q_1^{SB}(p), q_2^{SB}(p)\right) = \phi_{q_2}\left(q_1^{SB}(p), q_2^{SB}(p)\right)$$

The solution of T is given by the binding constraint. ■

Proof of Proposition 3

Proof. Consider the following maximization problem[18]:

$$L = \frac{1}{1-\delta}\left(B\left(q_1,q_2\right)-T-p_1q_1-p_2q_2\right)+\lambda\left(\delta-\frac{\alpha\left(B\left(q_1,\widehat{q}_2\right)-B\left(q_1,q_2\right)\right)+\phi\left(q_1^*,q_2^*\right)-\phi\left(q_1^*,\widehat{q}_2\right)}{\alpha B\left(q_1^*,\widehat{q}_2\right)+T+p_1q_1^*+p_2q_2^*-\phi\left(q_1^*,\widehat{q}_2\right)}\right)$$

From the maximization we obtain:

$$\frac{dL}{dq_1} = \frac{1}{1-\delta}\left(B_{q_1}\left(q_1,q_2\right)-p_1\right)-\frac{\lambda}{\left(\alpha B\left(q_1^*,\widehat{q}_2\right)+T+pq_1^*+pq_2^*-\phi\left(q_1^*,\widehat{q}_2\right)\right)^2} \tag{A.7}$$

$$\Big[\left(\alpha\left(B_{q_1}\left(q_1,\widehat{q}_2\right)-B_{q_1}\left(q_1,q_2\right)\right)+\phi_{q_1}\left(q_1,q_2\right)-\phi_{q_1}\left(q_1,\widehat{q}_2\right)\right)\left(\alpha B\left(q_1^*,\widehat{q}_2\right)+T+p_1q_1^*+p_2q_2^*-\phi\left(q_1^*,\widehat{q}_2\right)\right)$$

$$-\left(\alpha B_{q_1}\left(q_1,\widehat{q}_2\right)-\phi_{q_1}\left(q_1,\widehat{q}_2\right)+p_1\right)\left(\alpha\left(B\left(q_1,\widehat{q}_2\right)-B\left(q_1,q_2\right)\right)+\phi\left(q_1^*,q_2^*\right)-\phi\left(q_1^*,\widehat{q}_2\right)\right)\Big] \leq 0,$$

$$q_1 \geq 0, \qquad \frac{dL}{dq_1}q_1 = 0$$

$$\frac{dL}{dq_2} = \frac{1}{1-\delta}\left(B_{q_2}\left(q_1,q_2\right)-p_2\right)-\frac{\lambda}{\left(\alpha B\left(q_1^*,\widehat{q}_2\right)+T+p_1q_1^*+p_2q_2^*-\phi\left(q_1^*,\widehat{q}_2\right)\right)^2} \tag{A.8}$$

$$\Big[\left(-\alpha B_{q_2}\left(q_1,q_2\right)+\phi_{q_2}\left(q_1,q_2\right)\right)\left(\alpha B\left(q_1^*,\widehat{q}_2\right)+T+p_1q_1^*+p_2q_2^*-\phi\left(q_1^*,\widehat{q}_2\right)\right)+$$

$$-p_2\left(\alpha\left(B\left(q_1,\widehat{q}_2\right)-B\left(q_1,q_2\right)\right)+\phi\left(q_1^*,q_2^*\right)-\phi\left(q_1^*,\widehat{q}_2\right)\right)\Big] \leq 0,$$

$$q_2 \geq 0, \qquad \frac{dL}{dq_2}q_2 = 0$$

$$\frac{dL}{dp_1} = -\frac{1}{1-\delta}q_1+\lambda\frac{\alpha\left(B\left(q_1,\widehat{q}_2\right)-B\left(q_1,q_2\right)\right)+\phi\left(q_1^*,q_2^*\right)-\phi\left(q_1^*,\widehat{q}_2\right)}{\left(\alpha B\left(q_1^*,\widehat{q}_2\right)+T+p_1q_1^*+p_2q_2^*-\phi\left(q_1^*,\widehat{q}_2\right)\right)^2}q_1^* \leq 0, \tag{A.9}$$

$$p_1 \geq 0, \qquad \frac{dL}{dp_1}p_1 = 0$$

$$\frac{dL}{dp_2} = -\frac{1}{1-\delta}q_2+\lambda\frac{\alpha\left(B\left(q_1,\widehat{q}_2\right)-B\left(q_1,q_2\right)\right)+\phi\left(q_1^*,q_2^*\right)-\phi\left(q_1^*,\widehat{q}_2\right)}{\left(\alpha B\left(q_1^*,\widehat{q}_2\right)+T+p_1q_1^*+p_2q_2^*-\phi\left(q_1^*,\widehat{q}_2\right)\right)^2}q_2 \leq 0, \tag{A.10}$$

$$p_2 \geq 0, \qquad \frac{dL}{dp_2}p_2 = 0$$

[18] we rule out the index $*$ to ease the proof.

$$\frac{dL}{d\lambda} = \delta - \frac{\alpha\left(B\left(q_1, \widehat{q}_2\right) - B\left(q_1, q_2\right)\right) + \phi\left(q_1^*, q_2^*\right) - \phi\left(q_1^*, \widehat{q}_2\right)}{\alpha B\left(q_1^*, \widehat{q}_2\right) + T + p_1 q_1^* + p_2 q_2^* - \phi\left(q_1^*, \widehat{q}_2\right)} \geq 0, \tag{A.11}$$

$$\lambda \geq 0, \qquad \frac{dL}{d\lambda} p_2 = 0$$

Note that by the envelop theorem and (2) we have $\alpha B_{\widehat{q}_2}\left(q_1, \widehat{q}_2\right) - \phi_{\widehat{q}_2}\left(q_1, \widehat{q}_2\right) = 0$. The case $q_1 = 0$, which gives $\frac{dL}{dq_1} < 0$, clearly contradicts our assumption on the utility. The same holds for q_2. From Eqs.(A.7), (A.8) and (A.9), it follows that $\lambda > 0$. If this is not the case, we will have that $\frac{dL}{dp_2} = \frac{dL}{dp_1} = \frac{dL}{dq_1} = \frac{dL}{dq_2} = 0$, which clearly contradicts our hypothesis that the first best is out of reach. Therefore from $\frac{dL}{d\lambda} = 0$ we obtain Eq. (5). Considering the interior solution, by dividing (A.8) and (A.7) we obtain:

$$\begin{aligned}
\frac{B_{q_2}\left(q_1, q_2\right) - p_2}{B_{q_1}\left(q_1, q_2\right) - p_1} = \Big[&\left(-\alpha B_{q_2}\left(q_1, q_2\right) + \phi_{q_2}\left(q_1, q_2\right)\right)\left(\alpha B\left(q_1^*, \widehat{q}_2\right) + T + p_1 q_1^* + p_2 q_2^* - \phi\left(q_1^*, \widehat{q}_2\right)\right) + \\
&- p_2\left(\alpha\left(B\left(q_1, \widehat{q}_2\right) - B\left(q_1, q_2\right)\right) + \phi\left(q_1^*, q_2^*\right) - \phi\left(q_1^*, \widehat{q}_2\right)\right) \Big] \\
&/ \Big[\left(\alpha\left(B_{q_1}\left(q_1, \widehat{q}_2\right) - B_{q_1}\left(q_1, q_2\right)\right) + \phi_{q_1}\left(q_1, q_2\right) - \phi_{q_1}\left(q_1, \widehat{q}_2\right)\right)\left(\alpha B\left(q_1^*, \widehat{q}_2\right) + T + p_1 q_1^* + p_2 q_2^* - \phi\left(q_1^*, \widehat{q}_2\right)\right) - \\
&\left(\alpha B_{q_1}\left(q_1, \widehat{q}_2\right) + p_1 - \phi_{q_1}\left(q_1, \widehat{q}_2\right)\right)\left(\alpha\left(B\left(q_1, \widehat{q}_2\right) - B\left(q_1, q_2\right)\right) + \phi\left(q_1^*, q_2^*\right) - \phi\left(q_1^*, \widehat{q}_2\right)\right) \Big]
\end{aligned} \tag{A.12}$$

by dividing (A.10) and (A.8) and using (5) we obtain the optimal condition for q_2

$$B_{q_2}\left(q_1, q_2\right) = \frac{1}{\alpha + \delta} \phi_{q_2}\left(q_1, q_2\right) \tag{A.13}$$

Similarly, by dividing (A.9) and (A.7) and using (5) we obtain the optimal condition for q_1:

$$B_{q_1}\left(q_1, q_2\right) = \frac{1-\delta}{\delta}\left(\alpha B_{q_1}\left(q_1, \widehat{q}_2\right) - \phi_{q_1}\left(q_1, \widehat{q}_2\right)\right) - \frac{1}{\delta}\left(\alpha B_{q_1}\left(q_1, q_2\right) - \phi_{q_1}\left(q_1, q_2\right)\right) \tag{A.14}$$

From Eqs. (A.9) and (5) we obtain:

$$\lambda = \frac{1}{1-\delta} \frac{\alpha\left(B\left(q_1, \widehat{q}_2\right) - B\left(q_1, q_2\right)\right) + \phi\left(q_1^*, q_2^*\right) - \phi\left(q_1^*, \widehat{q}_2\right)}{\delta^2} \tag{A.15}$$

By using (A.15) and (A.13) into (A.8) as well as (A.15) and (A.14) into (A.7), we have the optimal prices as in the Proposition.

■

Proof of Corollary 2

Proof. Consider $\widehat{q}_2 < q_2^*$. Since q_1^* does not change, assumption of concavity of B in q_2 implies

$$B_{q_2}\left(q_1^*, \widehat{q}_2\right) > B_{q_2}\left(q_1^*, q_2^*\right) \tag{A.16}$$

(7) and (2) give $\frac{1}{\alpha}\phi_{q_2}\left(q_1^*,\widehat{q}_2\right) > \frac{1}{\alpha+\delta}\phi_{q_2}\left(q_1^*,q_2^*\right)$; convexity of ϕ in q_2 implies $\phi_{q_2}\left(q_1^*,q_2^*\right) > \phi_{q_2}\left(q_1^*,\widehat{q}_2\right)$, therefore (A.16) holds if $\alpha < \delta\frac{\phi_{q_2}(q_1^*,\widehat{q}_2)}{\phi_{q_2}(q_1^*,q_2^*)-\phi_{q_2}(q_1^*,\widehat{q}_2)}$. Consider $\widehat{q}_2 > q_2^*$. Concavity of B in q_2 now implies

$$B_{q_2}\left(q_1^*,\widehat{q}_2\right) < B_{q_2}\left(q_1^*,q_2^*\right) \tag{A.17}$$

that, because of (7) and (2), implies $\frac{1}{\alpha+\delta}\phi_{q_2}\left(q_1^*,q_2^*\right) > \frac{1}{\alpha}\phi_{q_2}\left(q_1^*,\widehat{q}_2\right)$. Convexity of ϕ in q_2 induces $\phi_{q_2}\left(q_1^*,q_2^*\right) < \phi_{q_2}\left(q_1^*,\widehat{q}_2\right)$, implying that (A.17) cannot hold.

The case $\widehat{q}_2 = q_2^*$ implies $\phi_{q_2}\left(q_1^*,q_2^*\right) = \phi_{q_2}\left(q_1^*,\widehat{q}_2\right)$ and $B_{q_2}\left(q_1^*,\widehat{q}_2\right) = B_{q_2}\left(q_1^*,q_2^*\right)$ but, under $\delta > 0$ we have the contradiction because $\frac{1}{\alpha+\delta}\phi_{q_2}\left(q_1^*,q_2^*\right) = \frac{1}{\alpha}\phi_{q_2}\left(q_1^*,\widehat{q}_2\right)$ cannot hold. The result clearly arises when both $\widehat{q}_2$ and q_2^* are interior solutions of their maximization problems. Should we focus on corner solution therefore the case $\widehat{q}_2 > q_2^*$ might be possible. ■

B: Empirical Appendix

Table B.1: The link between Governor's tenure and quality of healthcare services
Panel A: no interaction (Eq. 11)

VARIABLES	C-section	hygienic	sat. medical	sat. nursing	ptca
tenure	-0.1146***	0.0222	0.1840*	0.0459	-0.0968
	(0.034802)	(0.124963)	(0.109083)	(0.112743)	(0.105419)
foreigners	-0.6709***	1.6183**	0.7735	1.3539**	1.1873
	(0.184065)	(0.658638)	(0.654015)	(0.632484)	(1.415183)
unemplf	-0.0660	-0.0515	0.0199	-0.0377	-0.2730
	(0.064883)	(0.173477)	(0.181670)	(0.172812)	(0.337167)
gdp	-0.0003**	-0.0011**	-0.0008**	-0.0007**	-0.0003
	(0.000137)	(0.000476)	(0.000384)	(0.000374)	(0.000467)
occup	0.2280	0.8796	0.7684	0.3721	-0.7721
	(0.169222)	(0.581596)	(0.575705)	(0.565301)	(1.072976)
tariffario	-1.3830*	1.3202	-2.1053	0.6478	
	(0.708157)	(2.689801)	(2.433790)	(2.687727)	
difftar	1.2487**	-2.1789	-2.6144*	-2.9171**	-1.9238
	(0.491468)	(1.613817)	(1.514386)	(1.479792)	(2.172434)
srp	1.1968**	-2.4896	-1.6228	-0.5243	4.3417*
	(0.552322)	(2.010984)	(1.846497)	(1.883061)	(2.459279)
pubspend	0.0048**	0.0084	0.0053	0.0120*	0.0190
	(0.001945)	(0.007562)	(0.007019)	(0.006672)	(0.012613)
private	-0.0342	0.0451	0.1327	0.1795*	0.3455**
	(0.042418)	(0.110449)	(0.106663)	(0.105767)	(0.158370)
samecoalitiongov	0.4949*	-0.8058	-0.2798	-0.9412	-2.0180**
	(0.271379)	(0.875477)	(0.845355)	(0.851386)	(0.954626)
meddoc	2.2699**	1.1794	-1.9740	0.0803	-6.8446**
	(0.909799)	(2.454295)	(2.236007)	(2.235474)	(3.442553)
centleft	0.4799	0.0132	-1.6322	0.1920	1.8278
	(0.341420)	(1.096150)	(1.143328)	(1.148064)	(1.350188)
agedelivery	-1.6233**				
	(0.821425)				
abortionrate	0.0186**				
	(0.009091)				
neonmort	0.0421				
	(0.039322)				
AMIadm					0.0389***
					(0.011562)
Constant	62.0669**	1.6574	16.7007	24.0776	14.4517
	(28.544484)	(25.519956)	(25.105778)	(24.648038)	(57.678279)
Observations	411	374	374	374	170
R-squared	0.96	0.80	0.78	0.81	0.88
Region FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Notes. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Dependent variables are: C-section rates (C-section), satisfaction with hygienic services (hygienic), satisfaction with medical assistance (sat. medical), satisfaction with nursing assistance (sat. nursing), share of patients receiving PTCA in two days (ptca).

Table B.2: The link between Governor's tenure and quality of healthcare services
Panel B: interaction w/private (Eq. 12)

VARIABLES	C-section	hygienic	sat. medical	sat. nursing	ptca
tenure	-0.1543**	-0.1688	0.0732	-0.1183	0.0447
	(0.068376)	(0.197575)	(0.171260)	(0.187441)	(0.155359)
tenure_private	0.0033	0.0156	0.0091	0.0134	-0.0119
	(0.004365)	(0.011065)	(0.012178)	(0.012606)	(0.009994)
private	-0.0359	0.0344	0.1266	0.1703	0.3470**
	(0.042470)	(0.110565)	(0.107785)	(0.106471)	(0.159795)
foreigners	-0.6780***	1.5889**	0.7565	1.3286**	1.3289
	(0.184657)	(0.655984)	(0.653609)	(0.632447)	(1.451982)
unemplf	-0.0578	-0.0151	0.0410	-0.0064	-0.2764
	(0.066273)	(0.174133)	(0.180680)	(0.171142)	(0.343041)
gdp	-0.0003**	-0.0011**	-0.0008*	-0.0007*	-0.0003
	(0.000137)	(0.000478)	(0.000387)	(0.000380)	(0.000468)
occup	0.2344	0.9122	0.7873	0.4001	-0.8614
	(0.169297)	(0.579703)	(0.572183)	(0.557270)	(1.105114)
tariffario	-1.3519*	1.3627	-2.0807	0.6844	
	(0.704262)	(2.721170)	(2.446418)	(2.691219)	
difftar	1.2287**	-2.2298	-2.6439*	-2.9608**	-2.3372
	(0.496136)	(1.615341)	(1.519098)	(1.486225)	(2.222171)
srp	1.2553**	-2.2240	-1.4687	-0.2960	3.6266
	(0.572257)	(2.029248)	(1.852168)	(1.887136)	(2.569737)
pubspend	0.0050**	0.0093	0.0059	0.0129*	0.0180
	(0.001949)	(0.007691)	(0.007179)	(0.006862)	(0.012449)
alignment	0.5017*	-0.7342	-0.2383	-0.8797	-2.1980**
	(0.272441)	(0.869348)	(0.845865)	(0.847343)	(0.975381)
meddoc	2.2817**	1.1205	-2.0081	0.0297	-6.7037*
	(0.921079)	(2.485561)	(2.250380)	(2.249608)	(3.582751)
centleft	0.4637	-0.1016	-1.6987	0.0933	1.8429
	(0.339035)	(1.099529)	(1.154640)	(1.155866)	(1.331603)
agedelivery	-1.6288**				
	(0.824638)				
abortionrate	0.0181**				
	(0.009067)				
neonmort	0.0420				
	(0.039215)				
AMIadm					0.0391***
					(0.011528)
Constant	61.8760**	-0.9442	15.1920	21.8411	20.2423
	(28.581399)	(25.571955)	(24.949146)	(24.217855)	(58.568666)
Observations	411	374	374	374	170
R-squared	0.96	0.80	0.78	0.81	0.88
Region FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Notes. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Dependent variables are: C-section rates (C-section), satisfaction with hygienic services (hygienic), satisfaction with medical assistance (sat. medical), satisfaction with nursing assistance (sat. nursing), share of patients receiving PTCA in two days (ptca).

Table B.3: The link between MoH's tenure and quality of healthcare services
Panel A: no interaction (Eq. 11)

VARIABLES	C-section	hygienic	sat. medical	sat. nursing	ptca
tenure	-0.1405***	-0.2012	-0.2626	-0.2725	0.3331*
	(0.045337)	(0.197006)	(0.182204)	(0.174422)	(0.194611)
foreigners	-0.9504***	1.4263**	0.3103	0.9189	0.7356
	(0.190693)	(0.639645)	(0.647290)	(0.609162)	(1.560338)
unemplf	-0.0320	-0.0551	0.1135	0.0424	-0.3157
	(0.064856)	(0.166957)	(0.179772)	(0.166389)	(0.396139)
gdp	-0.0001	-0.0011**	-0.0007*	-0.0006*	0.0000
	(0.000125)	(0.000449)	(0.000365)	(0.000340)	(0.000440)
occup	0.2006	1.1510**	1.1096**	0.5385	-1.4818
	(0.169497)	(0.555316)	(0.564088)	(0.520458)	(0.967962)
tariffario	-1.4644**	1.5467	-0.9598	1.4347	
	(0.714330)	(2.765411)	(2.556769)	(2.711875)	
difftar	1.6184***	-0.6880	-0.5929	-1.3964	1.7494
	(0.440908)	(1.534038)	(1.559746)	(1.413775)	(1.932526)
srp	0.9678*	-2.4846	-1.4482	-0.5366	0.8416
	(0.543363)	(1.913935)	(1.760122)	(1.753331)	(2.600967)
pubspend	0.0039*	0.0041	-0.0006	0.0077	0.0034
	(0.001972)	(0.007487)	(0.006721)	(0.006380)	(0.013938)
private	-0.0547	-0.0039	0.1042	0.1827*	0.4542***
	(0.043027)	(0.113522)	(0.107347)	(0.105361)	(0.160378)
alignment	0.5577**	-1.1561	-1.0378	-1.1791	-1.0541
	(0.253729)	(0.861046)	(0.815163)	(0.790942)	(1.018638)
meddoc	0.8185***	1.3223	1.5293	-0.2255	-3.1802
	(0.310209)	(1.016500)	(0.985502)	(0.983296)	(1.938968)
centleft	1.1751***	2.0309	1.4664	1.6313	1.3213
	(0.367229)	(1.264245)	(1.202283)	(1.155641)	(1.270026)
agedelivery	-2.1358***				
	(0.762123)				
abortionrate	0.0113				
	(0.008505)				
neonmort	0.0155				
	(0.037563)				
AMIadm					0.0426***
					(0.010951)
Constant	76.3617***	-5.0727	4.0313	16.8669	64.3176
	(26.898557)	(23.800833)	(24.812417)	(22.716150)	(58.302515)
Observations	439	394	394	394	186
R-squared	0.95	0.81	0.80	0.83	0.83
Region FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Notes. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.
Dependent variables are: C-section rates (C-section), satisfaction with hygienic services (hygienic), satisfaction with medical assistance (sat. medical), satisfaction with nursing assistance (sat. nursing), share of patients receiving PTCA in two days (ptca).

Table B.4: The link between MoH's tenure and quality of healthcare services
Panel B: interaction w/private (Eq. 12)

VARIABLES	C-section	hygienic	sat. medical	sat. nursing	ptca
tenure	-0.1246	-0.4193	-0.1959	-0.1752	0.6066**
	(0.080824)	(0.292319)	(0.281563)	(0.262266)	(0.301382)
tenure_private	-0.0015	0.0201	-0.0061	-0.0090	-0.0249
	(0.006449)	(0.017499)	(0.018260)	(0.016984)	(0.016013)
private	-0.0524	-0.0357	0.1140	0.1969*	0.4837***
	(0.043540)	(0.117436)	(0.112422)	(0.109585)	(0.164080)
foreigners	-0.9553***	1.4981**	0.2884	0.8869	0.7720
	(0.191897)	(0.643195)	(0.656674)	(0.615627)	(1.567117)
unemplf	-0.0317	-0.0556	0.1136	0.0426	-0.3296
	(0.065163)	(0.165982)	(0.180666)	(0.167359)	(0.397199)
gdp	-0.0001	-0.0012**	-0.0007*	-0.0006*	0.0000
	(0.000125)	(0.000453)	(0.000370)	(0.000341)	(0.000453)
occup	0.1995	1.1991**	1.0948*	0.5171	-1.5442
	(0.169572)	(0.555688)	(0.564891)	(0.520245)	(0.976086)
tariffario	-1.4730**	1.7360	-1.0178	1.3502	
	(0.709356)	(2.789781)	(2.569131)	(2.711029)	
difftar	1.6292***	-0.7300	-0.5801	-1.3777	1.4641
	(0.444012)	(1.535433)	(1.562649)	(1.415842)	(1.939541)
srp	0.9705*	-2.4729	-1.4518	-0.5418	0.4882
	(0.542569)	(1.915389)	(1.763275)	(1.758230)	(2.643192)
pubspend	0.0038*	0.0048	-0.0008	0.0074	0.0021
	(0.001996)	(0.007624)	(0.006832)	(0.006460)	(0.014338)
alignment	0.5562**	-1.1125	-1.0512	-1.1986	-1.2643
	(0.254071)	(0.859667)	(0.818535)	(0.789739)	(1.060730)
meddoc	0.8111***	1.4266	1.4974	-0.2720	-3.1981*
	(0.309267)	(1.015255)	(0.987743)	(0.972763)	(1.931067)
centleft	1.1691***	2.1304*	1.4359	1.5869	1.3984
	(0.363392)	(1.262561)	(1.193670)	(1.149451)	(1.243106)
agedelivery	-2.1265***				
	(0.769087)				
abortionrate	0.0115				
	(0.008465)				
neonmort	0.0159				
	(0.037819)				
AMIadm					0.0437***
					(0.010822)
Constant	76.1060***	-7.4277	4.7526	17.9175	69.3939
	(27.095301)	(23.872972)	(24.799134)	(22.719118)	(59.425964)
Observations	439	394	394	394	186
R-squared	0.95	0.81	0.80	0.83	0.84
Region FE	YES	YES	YES	YES	YES
Year FE	YES	YES	YES	YES	YES

Notes. Robust standard errors in parentheses. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$. Dependent variables are: C-section rates (C-section), satisfaction with hygienic services (hygienic), satisfaction with medical assistance (sat. medical), satisfaction with nursing assistance (sat. nursing), share of patients receiving PTCA in two days (ptca).