Provider Altruism in Incentives Contracts: Medicare's Quality Race

Galina Besstremyannaya[1], **Sergei Golovan**[2]

[1] National Research University Higher School of Economics,
11, Pokrovsky blv., Moscow, 101000, Russian Federation.
E-mail: gbesstremyannaya@hse.ru

[2] New Economic School,
3, Nobelya st., Moscow, 121205, Russian Federation.
E-mail: sgolovan@nes.ru

The paper analyzes the impact of provider altruism and motivation on the outcomes of pay-for-performance reimbursement in healthcare, where a fixed price contract on quantity is supplemented with a relative performance contract on quaility. We develop a theoretical model which forecasts the crowding out of most altruistic providers. Using the example of Medicare's nationwide natural experiment with a relative performance contract on quality (the data for 3000 acute care hospitals in 2004–2017, with the incentives contract implemented since 2013), we conduct an empirical test of the model predictions. We assume that altruism is heterogeneous across hospitals and the values of altruism in each hospital are higher for quality measures which are strongly associated with the patient's benefit. The analysis employs dynamic panel data estimations to account for «habit-formation» and we exclude pre-reform and post-reform «regression-to-the-mean» effects by modeling the time-dependent long-term mean as a function of hospital characteris-

The paper was prepared in the framework of the HSE University's Basic Research Program. We thank Alexis Belianin, Jay Bhattacharya, Katharina Blankart, Ruben Enikolopov, Heike Hennig-Schmidt, Sergei Izmalkov, Irina Ivashkovskaya, Martin Karlsson, Sergey Kokovin, Haruko Noguchi, Anna Obizhalova, Hideo Owan, Vladimir Smirnov, Anastasia Stepanova, Anton Suvorov, Yuta Toyama, Alexei Zakharov, participants of the European Economic Association 2015, 2021 and 2022 Annual Congresses, and participants of seminars at New Economic School (2015), HSE/NES Political Economy (2016), CINCH Health Economics Research Center at University of Duisburg-Essen (2019), HSE Corporate Finance Center (2019–2020), Waseda Institute of Political Economy Empirical Microeconomics (2021), HSE International Laboratory for Experimental and Behavioral Economics (2022) for helpful comments. The research stems from the working paper written while Besstremyannaya was a visiting researcher at CINCH Health Economics Research Center at University of Duisburg-Essen (Essen, Germany) in December 2018 – February 2019.

Galina Besstremyannaya – Senior researcher, International Laboratory for Macroeconomic Analysis, Associate Professor, Faculty of Economic Sciences.
Sergei Golovan – Dotsent.

The article was received: 09.01.2022/The article is accepted for publication: 08.09.2022.

tics. We focus on highest-quality hospitals and discover a deterioration of quality measures, which may be linked to the patient's benefit (communication of patients with medical personnel and ability to receive help promptly). It may be interpreted as an illustration of the fact that relative performance incentive contracts may be associated with crowding altruistic providers out of the healthcare market.

Key words: contract theory; incentives; altruism; dynamic panels.

JEL Classification: C22, C23, D21, D22, I18.

DOI: 10.17323/1813-8691-2022-26-3-375-403

For citation: Besstremyannaya G., Golovan S. Provider Altruism in Incentives Contracts: Medicare's Quality Race. *HSE Economic Journal.* 2022; 26(3): 375–403.

1. Introduction

Public contracting with firms under asymmetric information about their technology provides a classic example of an agency problem, where the government as a principal can achieve social optimum in terms of product's quantity and agent's efforts through nonlinear prices [Laffont, Tirole 1993]. However, firms face a number of competing objectives, and this context of multi-tasking may result in trade-offs between quality, quantity and efforts [Hölmstrom, Milgrom, 1991]. A solution has been found in incentives contracts on quality, stemming from pay-for-performance in managerial economics. However, both theoretical literature and natural experiments point to deteriorating performance of the front-runners in such contracts [Bénabou, Tirole, 2006; Murdock, 2002; Prendergast, 1999]. Yet, altruistic agents would be interested in a social value of their performance per se. There is limited theoretical literature on altruism in public good games and piece-rate incentives contracts [Makris, Siciliani, 2014; Buurman, Dur, 2012; Siciliani, 2009], but little is known about the influence of altruism on the outcomes of relative performance schemes.

This paper analyzes the impact of provider altruism on the effects of pay-for-performance remuneration in healthcare, where a fixed price contract on quantity is supplemented with a relative performance contract on quality. Our theoretical model predicts that altruism may lead to quality decrease among subgroups of the high-performing providers. In an empirical application to the U.S. Medicare's nationwide natural experiment with value-based purchasing for inpatient acute care since 2013, we show that quality dimensions, which are linked to patient's benefit, demonstrate deterioration among top-performing hospitals. In other words, altruism may become a reason for crowding out on the healthcare market.

The novelty of our empirical approach is severalfold. We use dynamic panel data estimations to account for «habit-formation» and the analysis excludes pre-reform and post-reform «regression-to-the-mean» effect by modeling the time-dependent long-term mean as a function of hospital characteristics. The analysis employs longitudinal datasets on each quality measure of all acute-care Medicare's hospitals before and after the reform (fiscal years 2004–2017). The data are supplemented with patient case-mix, ownership, share of Medicare population and various hospital control variables, coming from: Medicare's Impact Files, Final Rules, Provider of Service Data, and Provider Utilization and Payment Data.

The remainder of the paper is structured as follows. Section 2 reviews the literature on incentives contracts and altruism in healthcare. Section 3 explains incentives regulation for inpatient care, as implemented in the U.S. within Medicare's value-based purchasing. Section 4 provides a theoretical model and outlines its predictions. Section 5 describes the data. The results of the estimation are given in Section 6. Section 7 is concerned with the discussion of the impact of altruism on the behavior of providers. Appendix A presents the derivation of the comparative statics of total quality, while Appendix B shows results of the empirical analysis with dynamic panel data models.

2. Incentives Contracts in Healthcare

The market for healthcare is centered on consumers who incur the risk of getting sick and receive a benefit B from the healthcare system in case of illness [Ellis, McGuire, 1986]. Healthcare providers are hospitals, clinics, physician groups or individual practices. Providers are interested in their profits but also show at least some degree of ethical behavior (altruism) and motivation (pride). Specifically, provider altruism towards patients implies concern about patient's benefit, i.e. patient's improvement of health and well-being [Hennig-Schmidt et al., 2011; McGuire, 2000].

The regulation on the healthcare markets requires the identification of products and determination of a reasonable cost for each product. This is accomplished with the help of a restricted number of medically justified groups (diagnosis-related groups, DRGs), with a statistically stable distribution of resource consumption within each group [Thompson, Averill, Fetter, 1979]. DRGs are constructed according to the following criteria: 1) coding of patient's diagnosis (e.g. according to international classification of diseases, ICD), 2) the cost of treating a patient with such diagnosis should be predictable. Examples of DRG are pneumonia, heart failure, cesarean section[1]. This classification is a core part of a prospective payment system (PPS) – a method of reimbursement that provides fixed payments for a patient with a given DRG. The approach exploits the yardstick competition model by Shleifer (1985), which establishes a fixed price contract for each firm dependent on the costs of similar firms and independent of the firm's own price. The contract has been expanded to a more efficient cost-sharing principle, which allows incorporating outlier cases: a fixed price and a fraction of actually incurred costs [Laffont, Tirole, 1993].

The contract provides for optimal output but does not offer a remedy to quality and quantity trade-offs. A solution to the problem of potential quality decrease under yardstick competition may be discovered in incentives regulation. Its origins may be found in the early 1980s when various performance targets were employed for enhancing the quality of natural monopolies and telecommunications [Joskow, Schmalensee, 1986]. Regarding incentives contracts in healthcare, a piece-rate pay-for-performance (also called «payment by results») started with quality measures for family practices in the U.K. These targets were established in 2004 and covered patient experience of care, management of chronic diseases and practice organization [Campbell et al., 2009]. There is limited theoretical and experimental literature on altruism in

[1] One DRG commonly includes several ICD codes and the total number of DRGs used in different countries varies from several hundred (761 Medicare Severity Diagnosis Related Groups as of 2019) and a few thousand (2,873 Diagnosis Procedure Combinations in Japan as of 2020 [Hayashida et al., 2021]).

such piece-rate incentives contracts: altruism is expected to cause differential quality dynamics: quality of providers with low and high altruism increases, while providers with medium altruism are subject to motivation crowding out [Siciliani, 2009; Green, 2014].

However, little is known about the effect of altruism in relative performance contracts, which have become widespread in various countries. For instance, these contracts have been applied for acute-inpatient care in Japan since 2003 (length-of-stay relative performance), in the U.S. as a pilot program in 2006–2008, and as a nationwide program since 2013 (benchmarking of inpatient quality measures).

3. Medicare's Incentives Contract

Value-based purchasing, started in fiscal year 2013, applies to discharges within the inpatient prospective payment system for acute-care Medicare's hospitals (with an exception of two states: Puerto Rico and Maryland). The reform decreases Medicare's DRG-based payment to each hospital by a factor α and redistributes the accumulated fund across all participating hospitals. The adjustment coefficient is calculated as:

$$1+\left(slope\frac{\mathrm{tps}_i}{100}-1\right)\cdot\alpha,$$

where i is the index for hospital, tps_i is hospital's *total performance score* ($0 \le \mathrm{tps}_i \le 100$) and $slope$ is the slope of a linear exchange function, which makes the incentives scheme budget-neutral for the regulator (the values of slope vary from 2 to 3 in various years). Hospitals with adjustment coefficient below 1 suffer a los efficient above 1 implies that hospitals are rewarded under the reform. The factor α is uniform s, while adjustment co across hospitals: $\alpha = 0.01$ in 2013 and is annually increased by 0.0025 in 2014–2017.

The total performance score is a weighted average of the scores for several domains:

1. Clinical process of care (12–13 measures).

Each measure is the percentage of patient cases for which the corresponding clinical requirement is satisfied (i.e. certain type of therapy provided within a given time interval). Examples of measures are: percutaneous coronary intervention received within 90 minutes of hospital arrival of patient with acute myocardial infarction; blood cultures performed in the emergency department prior to initial antibiotic received in hospital by a pneumonia patient; initial antibiotic selection for community acquired pneumonia in immunocompetent patient. (See full list of measures and their definitions in Table 1).

2. Patient experience of care (8 measures).

Each measure is the percentage of discharged patients who gave the most positive («top-box») response to the corresponding question (e.g. communication with doctors, nurses, medical staff, assessment of cleanliness and quietness of hospital environment).

3. Outcome of care (3–5 measures, domain added in 2014).

Measures of outcome of care are 30-day mortality rates (hazard rates) for patients with each of the three conditions: AMI, heart failure and pneumonia.

4. Efficiency (1 measure – Medicare spending per beneficiary, domain added in 2014).

The domain score is the sum of the scores for each measure. The score for each measure *m* is computed as the maximum of points for hospital's improvement and achievement. Achievements points a_i^m ($0 \le a_i^m \le 10$) are assigned with a stepwise function, which positions a hospital within the empirical distribution of a quality measure:

$$a_i^m = \begin{cases} 10, & \text{if } y_i^m \ge m_b, \\ \text{Round}\left[\dfrac{9\left(y_i^m - m_a\right)}{m_b - m_a} + 0.5\right], & \text{if } m_a \le y_i^m < m_b, \\ 0, & \text{if } y_i^m < m_a, \end{cases}$$

where y_i^m is the value of measure *m* for hospital i in the current period, m_b is benchmark, m_a is achievement threshold for measure *m*. The benchmark and achievement threshold are respectively set as the mean of the top decile (or 95th percentile) and the median in the empirical distribution of y^m, according to the survey in the baseline period.

The pricing schedule includes improvement points, yet they are irrelevant for the front-runner hospitals, i.e. hospitals with at least 9 achievement points out of 10.

4. Theoretical Model

4.1. Overview

Under the principal-agent approach on the healthcare market, a principal (a government or a social planner) contracts agents (providers) on behalf of consumers (patients). In this paper we follow the original approach by Ellis and McGuire (1986) who extend the concept of altruism from an individual (i.e. physician or nurse) to an organization: an agent is a «provider of service» and agent's actions include «hiring labor». The approach is employed in theoretical analyses of organizational altruism in recent health economics literature [Markis, Siciliani, 2013; Brekke et al., 2011; 2012; 2021], where a hospital is regarded as an agent with its parameter of altruism.

In view of the interrelation between individual and organizational altruism in hospitals, it becomes necessary to assume that hospital managers are also altruistic. Indeed, according to Rose-Ackerman (1986) altruistic managers contribute to the altruistic character of an organization[2]. In this regard, this paper relies on the changes in the management of Medicare's hospitals, which bridge the gap between the incentives of hospitals and physicians [Centers for Medicare and Medicaid Services, 2007].

In our model, providers have a type-specific altruism θ with pdf $f(\theta)$. The parameter θ would only be homogeneous in the industries with strict social norms [Makris, Siciliani, 2013], so we follow the approach which allows providers to differ with respect to their altruism [Liu,

[2] Overall, an argument about the transformation of individual altruism into organizational altruism may be found in [Rose-Ackerman, 1996]. She uses examples of healthcare institutions, universities who attract consumers through the provision of high quality of services so this quality per se becomes important to the organization («a teaching hospital may concentrate on difficult, interesting cases; a university might screen for high intelligence and strong high school records», p. 720).

Ma, 2013; Makris, Siciliani, 2013]. Providers have an additively separable utility function with respect to their profit and patient's benefit[3]. Provider's program in the presence of altruism is amended by reservation utility or limited liability constraint: $U_0 \geq 0$ [Makris, Siciliani, 2013]. Indeed, unlimited altruism may lead to bankruptcy[4].

We model a relative performance contract by introducing a linear pricing rule, which rewards/punishes providers for quality q above/below the standard q_0. The use of linear contracts based on relative performance is motivated by numerous real-life examples which include the US Medicare program and its variants in other countries: the UK, France, Italy, New Zealand and Korea [Peluso et al., 2019; Kristensen et al., 2016; Bisiaux, Chi, 2014; Bousquet et al., 2014; Sutton et al., 2012; Buetow, 2008].

For simplicity the model in this paper centers upon a fixed price contract. However, it can be generalized to a cost-sharing contract by replacing a fixed price a by a two-part tariff $a+bC(q)$, where $C(q)$ is the cost function and $0 \leq b \leq 1$.

4.2. Altruism and motivation

The approach is based on the Siciliani (2009) formulation of the Bénabou and Tirole (2006) intrinsic motivation: an agent with quality above a benchmark $\tilde{q}$ derives a fixed utility of good reputation under an incentives contract and has a linear disutility, proportional to monetary reward for quality. The major differences in our study are as follows. While Siciliani (2009) analyzes piece-rate pay for absolute performance, this paper models a linear contract on relative performance. Additionally, Siciliani (2009) focuses on the change in piece-rate pay for quality within the pay-for-performance contract, while our objective is comparative statics under a changeover from a prospective pay to pay-for-performance.

Provider's profit is defined as

$$\pi = a + \gamma(q - q_0) - C(q),$$

where $\gamma \in [0, +\infty)$ is incentives contract[5], a is fixed pay, $q_0 \geq 0$ is the absolute standard: agents with quality q above/below the absolute standard are subject to a reward/penalty under incentives contract, $C(q)$ is the cost function with $C_q > 0$, $C_{qq} > 0$, $C_{qqq} \geq 0$ [6] and $C(0) = C_q(0) = 0$. Note that in case of the Medicare contract, $q_0 = 100/slope$.

[3] The separability assumption is shared by the theoretical literature and comes from Blomqvist (1997).

[4] A limited liability form of the minimum profit constraint [Makris, Siciliani, 2013] equates $\pi_0 = 0$. It should be noted that the analysis may be extended to a case where $\pi_0 = 0$ is a certain number (either positive or negative), not very large in the absolute value [Choné, Ma, 2011).

[5] The continuous variable γ enables analyzing an agent's behavior under anticipation of the reform (i.e. $0 < \gamma < 1$).

[6] Regarding the third derivative, we follow the approach of Laffont and Tirole (1993, ch. 4, p. 215). This is a reasonable extension of the empirical regularity that quality improvements are costlier for high-

Provider's objectives are own profit π and altruistic concern about patient's benefit B, so $U = \pi + \theta B(q)$. Following Ellis and McGuire (1986), we assume that $B_q > 0$, $B_{qq} \leq 0$. As cost function and benefit function reflect opposite phenomena in terms of their curvature, we extend the approach of Laffont and Tirole (1993) about $C_{qqq} \geq 0$ and suppose that $B_{qqq} \leq 0$. The assumption about the aforementioned signs of the third derivatives is shown in [Olivella, Siciliani, 2017, p. 5] to be a sufficient condition for a good altruistic provider to decrease quality under an incentive contract. Similarly, in our model, the assumption is required to prove that altruism decelerates quality increase under the incentive contract with altruistic providers.

The important feature of our model is the participation constraint $\pi \geq 0$: a hospital chooses quality level $q \in Q$, so that the hospital's profits are non-negative. Given that the profit function is strictly concave, the set Q which satisfies the participation constraint, is either empty or constitutes a segment $Q = \left[\underline{q}(\gamma), \overline{q}(\gamma)\right]$.

Note that the inequality $\pi \geq 0$ simplifies to $a - C(q) \geq 0$ under $\gamma = 0$, which means that the set Q is non-empty. So, using the continuity argument, we conclude that the solution of hospital's maximization problem exists for $\gamma \in \left[0, \overline{\gamma}\right]$ (where $\overline{\gamma}$ can be infinite).

The interior solution for the problem of $\max_q U$ in absence of motivation is given by

$$\max_q a + \gamma(q - q_0) - C(q) + \theta B(q)$$
$$s.t. \pi \geq 0,$$

where $q^*(\theta, \gamma)$ comes from the FOC:

$$\gamma + \theta B_q = C_q .$$

The zero-profit equation may have two solutions: $\underline{q}(\gamma)$ and $\overline{q}(\gamma)$. Comparing the FOC with the condition for the profit maximization ($\gamma = C_q$) and using the fact that $B(q)$ is an increasing concave function, we conclude that $q^*(\theta, \gamma)$ is greater than the profit-maximizing value for $\theta > 0$. So the solution $\underline{q}(\gamma)$ is not chosen.

Consider the zero-profit equation:

$$a + \gamma(q - q_0) - C(q) = 0.$$

quality firms than for low-quality firms (i.e. of the fact that the second derivative is non-negative). Specifically, Laffont and Tirole (1993) use the notations $\psi(s)$ for a disutility ψ of providing quality s and write that $\psi'''(s) \geq 0$. A similar supposition that the third derivative of disutility of cost-reducing efforts is positive is required for a deterministic solution of the regulator's program [Laffont, Tirole, 1993, p. 107, 135].

Owing to the implicit function theorem

$$\frac{\partial \overline{q}(\gamma)}{\partial \gamma} = \frac{\overline{q}(\gamma) - q_0}{C_q - \gamma}.$$

$\overline{q}(\gamma)$ lies to the right of the value of q that maximizes the profits π, therefore the denominator in the equation for the implication of the implicit function theorem is positive. Subsequently, $\overline{q}(\gamma)$ decreases in γ if the value of the absolute standard q_0 is large. This is the case with the Medicare reform: the available data show that the value of $q_0 = 100/slope$ is above the empirical mean, and above the empirical median of TPS in most of post-reform years. So the number of hospitals where the aggregate quality is below q_0 is above 50 percent and may reach even 70–75 percent.

The upper limit $\overline{q}(\gamma)$ can be binding for high values of θ. The supposition goes in line with Proposition 1 on p. 648 of Brekke et al. (2012), where optimal quality increases under regulation for high values of θ, and decreases otherwise.

As regards our model, two points support the conjecture about the binding character of $\overline{q}(\gamma)$. Firstly, the higher is θ, the smaller is the profit. So the constraint is likely to be binding for high θ Secondly, as was mentioned earlier in footnote 4 about the limited liability constraint, the framework may be extended to the model where a hospital has a positive profit. But if the profit falls in the next period due to the change in the slope of the constraint, the constraint will be biding as the hospital regrets having a larger profit in the previous period. These two points communicate that the binding character of $\overline{q}(\gamma)$ is possible anywhere but is more likely for higher θ.

Define $\check{\theta}(\gamma)$ as the value for which $q^*(\theta,\gamma) = \overline{q}(\gamma)$. Following the logics about the binding character of $\overline{q}(\gamma)$, we obtain that $\check{\theta}(\gamma) < \overline{\theta}$. The fact implies that $\overline{q}(\gamma)$ is attainable by some providers with high θ (i.e. their altruism is less than the upper bound $\overline{\theta}$). This is the key supposition of the model, and it leads to the conclusion about quality deterioration of these providers under the incentive contract. The justification for the supposition may be found in the references to the evidence on the healthcare and education markets, and the numerical examples related to Proof of Proposition 5 of Brekke et al. (2012)[7]. The supposition also goes in line with Olivella and Sicialini (2017), where the highest quality is offered by healthcare providers.

The second order condition is satisfied[8] and implies that $\theta B_{qq} - C_{qq} < 0$. Using the implicit function theorem to differentiate q^* in γ and θ, we obtain comparative statics:

[7] Where the quality can even be overprovided.

[8] The SOC holds automatically if $\theta \geq 0$, since $C_{qq} > 0$ and $B_{qq} \leq 0$.

$$\frac{\partial q^*}{\partial \gamma} = -\frac{1}{\theta B_{qq} - C_{qq}} \geq 0,$$

$$\frac{\partial q^*}{\partial \theta} = -\frac{B_q}{\theta B_{qq} - C_{qq}} \geq 0,$$

$$\frac{\partial^2 q^*}{\partial \gamma \partial \theta} = \frac{B_{qq} + \frac{\partial q^*}{\partial \theta}\left(\theta B_{qqq} - C_{qqq}\right)}{\left(\theta B_{qq} - C_{qq}\right)^2} = \frac{\theta B_{qq}^2 - B_{qq}C_{qq} - \theta B_q B_{qqq} + B_q C_{qqq}}{\left(\theta B_{qq} - C_{qq}\right)^3} \leq 0.$$

The comparative statics of quality by altruistic providers may be summarized as the set of results below.

***Result* 1:** The partial derivative of the optimal quality in γ is non-negative because of the second order condition. So, given an interior solution, there is a quality increase under incentives contract γ.

***Result* 2:** Similarly, since the partial derivative of the optimal quality is non-negative in θ, the optimum quality is higher for higher altruism.

***Result* 3:** Altruism decelerates quality increase as the mixed partial derivative of the optimal quality in γ and θ is non-positive.

Now we follow Siciliani (2009) to add the discrete motivation component to the linearly separable utility function of an agent under the assumptions: 1) the reputation is associated with quality above a certain benchmark value $\tilde{q} \geq 0$, 2) the optimal value of quality $q^*\left(\overline{\theta},\gamma\right) \geq \tilde{q}$, so the group with $q^*\left(\theta,\gamma\right) \geq \tilde{q}$ is nonempty, 3) there exists $\hat{\theta}\left(\gamma\right)$, so that $U\left(\hat{\theta}, q^*\left(\hat{\theta},\gamma\right)\right) - U\left(\hat{\theta},\tilde{q}\right) = \left(V - wp\right)\gamma$, where $q^*\left(\tilde{\theta},\gamma\right) = \tilde{q}$.

The bound for high or low reputation is fixed exogenously [Olivella, Sicilian, 2017; Siciliani, 2009]. This fact as well as the discrete motivation component may be viewed as a simplification of the Bénabou and Tirole (2006) model. But such simplification may be considered reasonable for the healthcare sector. Indeed, consumers do not bear sufficient medical knowledge to make judgements about the quality of healthcare. Therefore, they have to rely on exogenous signals of good or bad reputation[9].

In our application, the reputation is observed under non-zero values of the reform parameter γ, so the full utility function becomes:

$$U = \pi + \theta B\left(q\right) + 1\left(q > \tilde{q}\right)\left(V - wp\right)\gamma,$$

[9] Siciliani (2009, p. 69) notes the exogenous bounds for reputation and the discrete motivation component: «…when patients comment about the skills of the doctors, they make statements of the type: "that doctor is good", "she is very good", or "she is a bad doctor". Therefore judgements of the type "good", "very good" or "bad" are in itself discrete, rather than continuous. A continuous specification of the reputational effect would imply extremely accurate statements from the patients.»

where $0 \le \underline{\theta} \le \theta \le \overline{\theta}$ is the parameter of altruism with pdf $f(\theta)$, V is a constant, reflecting utility of high reputation, w is marginal disutility of monetary reward p which come from the Bénabou and Tirole (2006) model.

Compute total amount of quality:

$$Q = \int_{\underline{\theta}}^{\hat{\theta}(\gamma)} q^*(\theta,\gamma) f(\theta) d\theta + \int_{\hat{\theta}(\gamma)}^{\tilde{\theta}(\gamma)} \tilde{q} f(\theta) d\theta + \int_{\tilde{\theta}(\gamma)}^{\check{\theta}(\gamma)} q^*(\theta,\gamma) f(\theta) d\theta + \int_{\check{\theta}(\gamma)}^{\overline{\theta}} \overline{q}(\gamma) f(\theta) d\theta.$$

Here we use the fact that $q^* < \tilde{q}$ is chosen by providers with $\theta(\gamma) \in \left[\underline{\theta}, \hat{\theta}(\gamma)\right)$, $q^* = \tilde{q}$ is the quality by providers with $\theta(\gamma) \in \left[\hat{\theta}(\gamma), \tilde{\theta}(\gamma)\right]$, while providers with $\theta(\gamma) \in \left(\tilde{\theta}(\gamma), \check{\theta}(\gamma)\right)$ have $q^* < \tilde{q}$. (The argument is similar to Siciliani (2009).) The novel part of our model is the fact that the remaining providers have $\theta(\gamma) > \check{\theta}(\gamma)$ and choose $\overline{q}(\gamma)$ owing to the participation constraint. (They would like to choose $q^*\left(\theta(\gamma),\gamma\right) > \overline{q}(\gamma)$ but this leads to negative profits).

Figure 1 provides a graphical illustration of the optimal amount of quality chosen by providers with different values of $\theta(\gamma)$. The lowest value of $q^*(\theta)$ is associated with the lowest value of altruism $\underline{\theta}$. There are four parts of the curve for the optimal quality on Figure 1:

1) q^* increases in altruism for providers with low altruism, i.e. when $\theta(\gamma) \in \left[\underline{\theta}, \hat{\theta}(\gamma)\right]$,

2) q^* is flat in altruism for providers with median altruism, i.e. when $\theta(\gamma) \in \left[\hat{\theta}(\gamma), \tilde{\theta}(\gamma)\right]$,

3) q^* increases in altruism for providers with high altruism, i.e. when $\theta(\gamma) \in \left(\tilde{\theta}(\gamma), \check{\theta}(\gamma)\right)$,

4) q^* is flat in altruism for providers with the highest altruism, i.e. when $\theta(\gamma) \in \left[\check{\theta}(\gamma), \overline{\theta}\right]$.

It may be noted that our results differ from the findings on a piece-rate incentive contract in Siciliani (2009). Specifically, we have a flat value of quality for providers with the highest altruism at the fourth part of the curve. Yet, this part of the curve is absent in Siciliani (2009), where quality goes up on the whole segment $\theta(\gamma) \in \left[\tilde{\theta}(\gamma), \overline{\theta}\right]$.

Finally, we proceed to the comparative statics of the total amount of quality Q by motivated providers with different values of altruism. As is shown in Appendix A, the change in total quality owing to the incentive contract equals:

$$\frac{dQ}{d\gamma}=\underbrace{\int_{\underline{\theta}}^{\hat{\theta}(\gamma)}\frac{\partial q^{*}(\theta,\gamma)}{\partial\gamma}f(\theta)d\theta}_{+}+\underbrace{\int_{\tilde{\theta}(\gamma)}^{\check{\theta}(\gamma)}\frac{\partial q^{*}(\theta,\gamma)}{\partial\gamma}f(\theta)d\theta}_{+}+\underbrace{\int_{\check{\theta}(\gamma)}^{\overline{\theta}}\frac{\partial\overline{q}(\gamma)}{\partial\gamma}f(\theta)d\theta}_{-}+$$

$$+\underbrace{\left(\tilde{q}-q^{*}\left(\hat{\theta}(\gamma),\gamma\right)\right)}_{+}\underbrace{\frac{f\left(\hat{\theta}(\gamma)\right)}{B\left(q^{*}\left(\hat{\theta}(\gamma),\gamma\right)\right)-B(\tilde{q})}}_{+}\left(\underbrace{\left(\tilde{q}-q^{*}\left(\hat{\theta}(\gamma),\gamma\right)\right)}_{+}+\underbrace{V-wp}_{-}\right).$$

The main conclusion drawn from the comparative statics of total quality is given in Result 4, which highlights heterogeneity in the provider response to the incentives contract.

***Result* 4:** Providers with $\theta\in\left[\underline{\theta},\hat{\theta}(\gamma)\right]$ and $\theta\in\left[\tilde{\theta}(\gamma),\check{\theta}(\gamma)\right]$ increase their quality and providers with $\theta\in\left(\check{\theta}(\gamma),\overline{\theta}\right]$ decrease their quality.

Other implications of the comparative statics are listed below and are shown in Figure 2. Firstly, there is a change in the number of providers on the interval $\left(\hat{\theta}(\gamma),\tilde{\theta}(\gamma)\right)$, as its limit points depend on γ. For instance, $\hat{\theta}(0)<\hat{\theta}(1)$, so some of the former providers with $\tilde{q}$ now provide $q^{*}<\tilde{q}$. At the same time, $\tilde{\theta}(0)>\tilde{\theta}(1)$ and some providers with $\tilde{q}$ now provide $q^{*}<\tilde{q}$.

Secondly, there is a certain convergence of quality across providers: quality goes up at the lowest tail of the distribution of altruism where $\theta\in\left[\underline{\theta},\hat{\theta}(\gamma')\right]$ and at a part of the highest tail with $\theta\in\left[\tilde{\theta}(\gamma'),\check{\theta}(\gamma')\right]$. Quality may decrease for medium types of altruism where $\theta\in\left[\hat{\theta}(\gamma'),\hat{\theta}(\gamma)\right]$. Finally, owing to participation constraint, quality goes down for the highest types: $\theta\in\left[\tilde{\tilde{\theta}}(\gamma'),\overline{\theta}\right]$.

The total aggregate effect of incentive across all providers is not necessarily positive. For instance, if $\left[\left(\tilde{q}-q^{*}\left(\hat{\theta}(\gamma),\gamma\right)\right)+(V-wp)\right]$ is negative and rather large in absolute terms, the third and the fourth summand in the equation for $\frac{dQ}{d\gamma}$ may be larger than the first two positive summands. In this case the total effect of the incentives contract is negative.

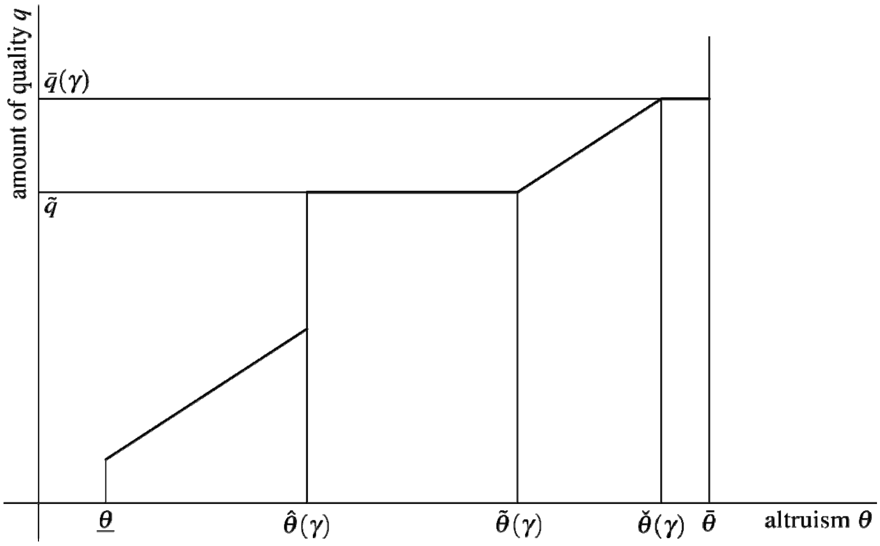

Fig. 1. Optimal amount of quality by providers with different values of the parameter of altruism

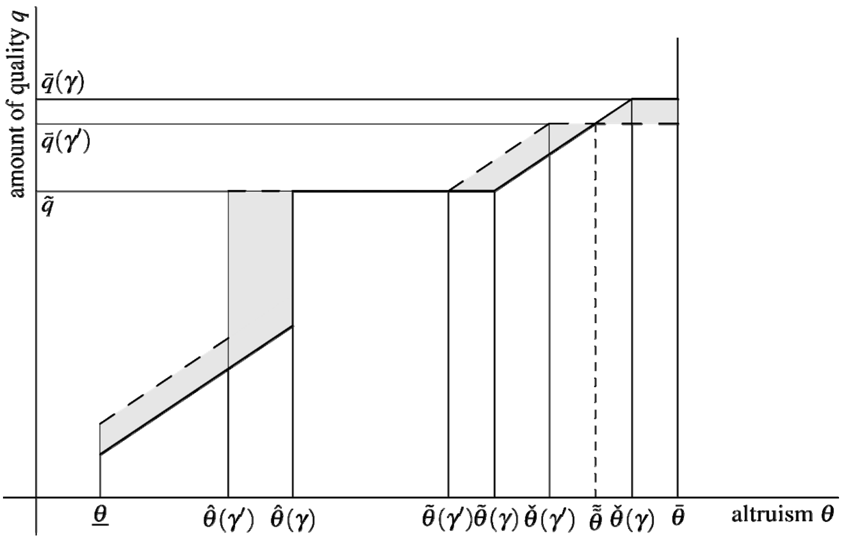

Fig. 2. Change in the optimal amount of quality under the changeover from policy γ (solid line) to γ' (dotted line)

Note that this paper assumes that the optimal quality of each hospital is non-verifiable by the principal and hence is non-contractible. Instead, the principal establishes q_0 as the absolute standard of quality. The principal then incentivizes hospitals to reach this standard by the payment which is inversely proportional to the distance from the standard.

4.3. Empirical approach

4.3.1. Dynamic panel data model

The analysis is applied to the data on quality at the US Medicare hospitals. We employ dynamic models to describe the evolution of quality. Specifically, we conjecture that a hospital strongly adheres to its practice patterns, so the value of each quality measure depends on its values in the previous periods. The data is likely to exhibit mean reversion as the observations of the quality measures are susceptible to random errors [Oxholm et al., 2018]. To separate the effect of value-based purchasing reform and the potential mean reversion, we distinguish the pre-treatment and post-treatment long-term means in the autocorrelation models.

Specifically, we extend the autocorrelation specification of Hamilton (1994), assuming there are two distinct long-term means in the pre-reform and post-reform periods, and allowing each long-term mean to be a function of hospital variables $\mathbf{x}$. Denote $\tilde{\mathbf{x}}_{it} = \left[\mathbf{x}_{it} 1\left(t \geq t_0\right), \mathbf{x}_{it} 1\left(t < t_0\right)\right]$, where t_0 is the year of incentivizing the given quality measure y through a contract and $\mathbf{x}$ is a vector of hospital control variables, which does not include a constant.

We study the dynamics of each quality measure separately and focus on the treatment effect of the reform for the groups of hospitals. In particular, the top decile and percentiles 95–100 approximate the groups of front-runner agents.

The analysis is based on the second-order[10] autocorrelation model:

$$
\begin{aligned}
y_{it} - \tilde{\mathbf{x}}_t \boldsymbol{\beta} - \mu = {} & \alpha_1 \left(y_{i,t-1} - \tilde{\mathbf{x}}_{i,t-1} \boldsymbol{\beta} - \mu\right) + \alpha_2 \left(y_{i,t-1} - \tilde{\mathbf{x}}_{i,t-1} \boldsymbol{\beta} - \mu\right) r_{it} + \\
& + \alpha_3 \left(y_{i,t-2} - \tilde{\mathbf{x}}_{i,t-2} \boldsymbol{\beta} - \mu\right) + \alpha_4 \left(y_{i,t-2} - \tilde{\mathbf{x}}_{i,t-2} \boldsymbol{\beta} - \mu\right) r_{it} + \tilde{\mathbf{x}}_{it} \gamma + \\
& + \tilde{\mathbf{x}}_{i,t-1} r_{it} \delta + \theta r_{it} + \tilde{\mathbf{x}}_{i,t-2} r_{it} \lambda + \tilde{\mathbf{x}}_{i,t-2} \kappa + \nu_i + \varepsilon_{it}.
\end{aligned} \tag{1}
$$

The dependent variable, y_{it}, is quality measure, r_{it} is the reform dummy which equals unity if hospital i participates in value-based purchasing in year t, ν_i are hospital fixed effects, ε_{it} are i.i.d. with zero mean. The terms in the third line are included for identification. Interaction terms $\left(y_{i,t-s} - \tilde{\mathbf{x}}_{i,t-s} \boldsymbol{\beta} - \mu\right) r_{it}$, $s = 1,2$ capture the effect of the reform conditional on the pre-reform value of the dependent variable.

It should be noted that we separate mean reversion owing to the incentives contract per se (through variable r) and owing to the non-financial impact of incentives contract, such as peer-effects and public reporting (through time-variable t_0). When our sensitivity analysis relaxes the assumptions about two types of mean reversion, we clearly find regression-to-the-mean owing to t_0, or both t_0 and r.

For convenience, we collect terms in (1) and rewrite:

[10] According to the results of the Arellano-Bond (1991) test, the first order lag did not allow excluding serial correlation.

$$y_{it} = c_0 + c_1 y_{i,t-1} + c_2 y_{i,t-1} r_{it} + c_3 r_{it} + \tilde{\mathbf{x}}_{it}\mathbf{c}_4 + \tilde{\mathbf{x}}_{i,t-1}\mathbf{c}_5 + \tilde{\mathbf{x}}_{i,t-1}\mathbf{c}_6 r_{it} + \\ + c_7 y_{i,t-2} + c_8 y_{i,t-2} r_t + \tilde{\mathbf{x}}_{i,t-2}\mathbf{c}_9 + \tilde{\mathbf{x}}_{i,t-2}\mathbf{c}_{10} r_{it} + \nu_i + \varepsilon_{it}. \quad (2)$$

Equation (2) is estimated using Arellano and Bover (1995) and Blundell and Bond (1998) estimator, with robust variance-covariance matrix [Windmeijer, 2005]. The reform and its interaction terms with $\tilde{\mathbf{x}}$ and $y_{i,t-s}$ are treated as predetermined variables, which means that lagged levels and lagged differences of y_{it}, r_{it}, $\tilde{\mathbf{x}}_{i,t-s} r_{it}$ and $y_{i,t-s} r_{it}$ ($s = 1,2$) are used as instruments for the difference equation.

Equating coefficients in (1) and (2) we obtain: $\boldsymbol{\beta} = -\mathbf{c}_5 / c_1$ and $\mu = c_0 / (1 - c_1 - c_7)$.

The effect of the incentives contract is analyzed across different percentile groups of hospitals, sorted according to their quality (we use deciles and percentiles 95–100).

The average treatment effect of the reform (at group means) is estimated by setting $r = 1$ in (2):

$$\tau = c_3 + c_2 \bar{y}_{t-1} + c_8 \bar{y}_{t-2} + \bar{\tilde{\mathbf{x}}}_{t-1}\mathbf{c}_6 + \bar{\tilde{\mathbf{x}}}_{t-2}\mathbf{c}_{10}.$$

4.3.2. Hypotheses

We assume that altruism is heterogeneous across hospitals and the values of altruism in each hospital are higher for quality measures which are strongly associated with a patient's benefit.

Following the predictions of the theoretical model about the decrease of quality at top-quality altruistic and motived providers (owing to participation constraint), we conjecture that the average treatment effect in the top percentile groups would be positive for measures unrelated to patient's benefit (and hence, not linked with altruistic behavior of providers) while it would be negative for quality measures that are strongly related to patient's benefit.

The data and our interviews with physicians (held at San Mateo and Santa Clara counties, California in 2015–2016) show that quality measures of the clinical process of care domain are not linked to a patient's health outcomes. Therefore, hospitals would not have altruistic concerns for performance on these measures.

On the other hand, patient experience of care measures would reveal more altruism. Specifically, it is plausible to assume that measures associated with altruistic behavior are: communication between patients and medical personnel, and patient's ability to receive help promptly.

Hypothesis I: $\tau \geq 0$ for the measures of clinical process of care at top-quality hospitals.

Hypothesis II: $\tau < 0$ at top-quality hospitals for the measures of patient experience of care linked to patient's benefit.

5. Data

5.1. Hospital quality measures

The data for quality measures and the reform participation come from the *Hospital Compare data archive* by the Centers for Medicare & Medicaid Services with data for value-based purchasing in 2013–2017. Our analysis focuses on Medicare's acute-care hospitals, as the incentives contract applies exclusively to this subgroup. The dichotomous variable for reform participation equals unity in fiscal years from 2013 onwards if a hospital is listed as a value-based purchasing hospital in the corresponding year (Table 1). We regard the reform participation as a predetermined variable. It is associated with hospital's desire to invest in the resources for data collection and data validation, along with the overall intent to focus on quality improvement [Centers for Medicare and Medicaid Services, 2007].

The unit for the time period in our analysis is a fiscal year. Medicare's acute care hospitals are incentivized to report the measures of clinical process of care since 2004 and patient experience of care since July 2007[11]. Therefore, our analysis uses the data for the clinical process of care between 2004 and 2015 (measures are not collected in later years) and for patient experience of care between 2007 and 2017.

Measures in outcome of care and in efficiency domains are not studied in this paper: they were incentivized only in 2014, so we lack post-reform data to fit second order autocorrelation model with interaction terms.

5.2. Hospital and patient control variables

Hospital characteristics are taken from the hospital files by *Hospital Compare*, which contain variables on hospital location and ownership. The number of hospital beds, share of Medicare's discharges, resident-to-bed ratio, ownership and the dichotomous variables for urban location come from Medicare's *Impact Files*. *Provider of Service Data* are exploited for the variables on the history of Medicare affiliation, number of Medicare certified beds, numbers of doctors, nurses and residents.

As regards patient control variables, we take the casemix variable from the *Impact Files*, which supplement the *Final Rules* on Medicare's payments to each hospital for a given fiscal year. The variable reflects the relative weight of each DRG in financial terms and it enables to control for the composition of patient cases in view of an objective link between the severity of illness and required hospital's resources. The disproportionate share index, coming from the *Impact Files*, accounts for the share of low-income patients. The lists of hospital and patient control variables are motivated by results of the Centers for Medicare and Medicaid Services (2007) interviews with hospitals and hospital associations, as well as the findings in the literature on the pilot programs [Damberg et al., 2014].

---

[11] Enactment of the Deficit Reduction Act of 2005, HCAHPS Fact sheet, 2012; section 501(b) of the Medicare Prescription Drug, Improvement and Modernization Act of 2003, Hospital Quality Initiative Overview, 2008.

Table 1.

Descriptive statistics for Medicare's acute-care hospitals in 2004–2017

Variable	Definition	Obs	Mean	St.Dev	Min	Max
Patient experience	**of care measures (2008–2017)**					
Comp-1-ap	Nurses always communicated well	34880	76.574	6.279	16	100
Comp-2-ap	Doctors always communicated well	34880	80.081	5.295	23	100
Comp-3-ap	Patients always received help as soon as they wanted	34876	64.013	8.896	8	100
Comp-4-ap	Pain was always well controlled	31714	69.266	5.747	0	100
Comp-5-ap	Staff always gave explanation about medicines	34854	61.478	6.855	2	100
Comp-1-yp	Yes, staff did give patients discharge information	34868	83.664	5.362	27	100
Clear-hsp-ap	Room was always clean	34880	70.656	7.444	7	100
Quiet-hsp-ap	Hospital always quiet at night	32652	58.873	10.568	0	100
Hsp-rating-910	Patients who gave hospital a rating of 9 or 10 (high)	34877	68.017	9.453	10	100
Clinical process	**of care measures (2004–2015)**					
AMI-8a	Primary percutaneous coronary intervention received within 90 minutes of hospital arrival (acute myocardial infection, AMI)	8163	66.186	24.594	0	100
HF-1	Discharge instructions (heart failure)	34285	77.075	25.148	0	100
PN-3b	Blood cultures performed in the emergency department prior to initial antibiotic received in hospital (pneumonia)	18857	88.219	9.955	0	100
PN-6	Initial antibiotic selection for community acquired pneumonia in immunocompetent patient (pneumonia)	35019	88.203	11.780	0	100
SCIP-Card2	Surgery patients on beta-blocker therapy prior to arrival who received a beta-blocker during the perioperative period	30993	90.459	15.108	0	100
SCIP-Inf1	Prophylactic antibiotic received within 1 hour prior to surgical incision	25739	95.718	8.900	0	100

Continues

Variable	Definition	Obs	Mean	St.Dev	Min	Max
SCIP-Inf2	Prophylactic antibiotic selection for surgical patients	30882	87.838	16.424	0	100
SCIP-Inf3	Prophylactic antibiotics discontinued within 24 hours after surgery end time	7471	91.376	14.712	0	100
SCIP-Inf4	Cardiac surgery patients with controlled 6 A.M. postoperative blood glucose	25854	90.338	14.885	0	100
SCIP-VTE2	Surgery patients who received appropriate venous thromboembolism prophylaxis within 24 hours prior to surgery to 24 hours after surgery	18408	93.313	12.082	0	100
Reform	**dummy (2004–2017)**					
r	=1 in 2013 onwards if was a value-based purchasing hospital in the corresponding fiscal year	48523	0.419	0.493	0	1
Hospital	**characteristics (2004–2017)**					
public	=1 if managed by federal, state or local government, or hospital district or authority	48523	0.142	0.349	0	1
emergency	=1 if emergency hospital	48523	0.932	0.252	0	1
urban	=1 if urban hospital	48523	0.735	0.441	0	1
resbed	=1 resident-to-bed ratio	46205	0.060	0.154	0	1.996
beds	Number of beds	40769	184.814	171.444	1	1928
medicare share	Share of Medicare cases	44956	0.455	0.156	0.001	1
Patient	**characteristics (2004–2017)**					
casemix	Transfer-adjusted casemix index	46205	1.416	0.321	0.356	4.363
Dsh	Disproportionate share index, reflecting the prevalence of low-income patients	46205	0.271	0.174	0	1.492

6. Results

6.1. Identification

The results are presented for six out of nine patient experience of care measures and for six out of nine clinical process of care measures because the Arellano and Bond (1991) test rejects the hypothesis about the absence of order two serial correlation in the first differenced errors for other three patient experience of care measures and for other three clinical process of care measures[12]. Owing to unavailability of long time-series for post-reform data, we cannot estimate higher order lags and limit our analysis to models with the above-mentioned 12 measures. The stationary conditions for the AR(2) process, namely the requirements of $c_1 + c_7 < 1$, $c_7 - c_1 < 1$ and $|c_7| < 1$, are satisfied for each of these 12 models. The coefficients for lagged dependent variables and other explanatory variables are presented in Tables B1 and B2 in Appendix B.

6.2. Effect of altruism

Concerning patient experience of care, value-based purchasing has a negative effect for three measures. The reform decreases quality for front-runner hospitals: percentiles 95–100 of measures on nurse communication (Comp-1-ap), the ability for patients to receive help quickly (Comp-3-ap) and discharge instructions (Comp-6-yp). As regards the first two measures, their values also go down at hospitals above the 80th percentile (Table 2). Note that we are unable to identify the effect of the incentives contract on a similar measure: communication with doctors (Comp-2-ap), since the dynamic panel data analysis could not be applied to it.

Quality measures associated with clean room (Clear-hsp-ap) and overall rating of hospital (Hsp-rating-910) do not decrease across hospitals in top percentiles (Figure 2 and Table 2).

None of clinical process of care measures deteriorates across hospitals in top percentiles and any other percentile groups. The measures on prophylactic antibiotics, selection for surgical patients (SCIP-Inf2) and on surgery patients on beta-blocker therapy prior to arrival who received a beta-blocker during the perioperative period (SCIP-Card2) increase in all percentiles owing to the reform, and the absolute value of the effect is higher in the lowest percentiles 0–10 and the highest percentiles 90–100 and 95–100 (Figure 3 and Table 3).

To sum up, under assumption that communication of medical personnel with patients (nurse communication and discharge instructions) and patient ability to receive help promptly are quality measures that are strongly associated with provider altruism, we may not reject *Hypothesis* I or *Hypothesis* II and conjecture that social preferences are revealed in the behavior of hospitals with the highest quality.

[12] Clean-hsp-ap and Quiet-hsp-ap albeit measured separately, are viewed as one measure «Cleanliness and quietness of hospital environment» in The Final Rule.

6.3. Internal validity and limitations

The analysis exploits longitudinal data, which may lack observations in certain years. Overall, the panels are unbalanced but 85–93% of hospitals would have observations in each year. As robustness check, we conducted analysis with balanced panels and discovered similar distribution of the dependent variables and negligible difference in the values for the coefficients for the explanatory variables.

The major limitation of our analyses is the lack of patient-level data. Nonetheless, the quality measures exploited in the empirical part of the paper are risk-adjusted according to major patient control variables, for instance age, education and co-morbidities [Medicare.gov, 2017; HCAHPS, 2013]. Yet, our estimates do not fully control for individual socio-demographic characteristics, which may influence treatment patterns.

It should be noted that altruism towards the social value of a product in our approach captures only certain aspects of the «other-regarding behavior» and may be interpreted as heterogeneous personal norms associated with an action which is desirable by a principal [Fischer, Huddart, 2008]. However, social norms that introduce behavioral and physiological effects to pure economic motives of incentives contracts are much broader. Examples include intrinsic, ethical, cooperative, reciprocal and other social issues [Arce, 2013].

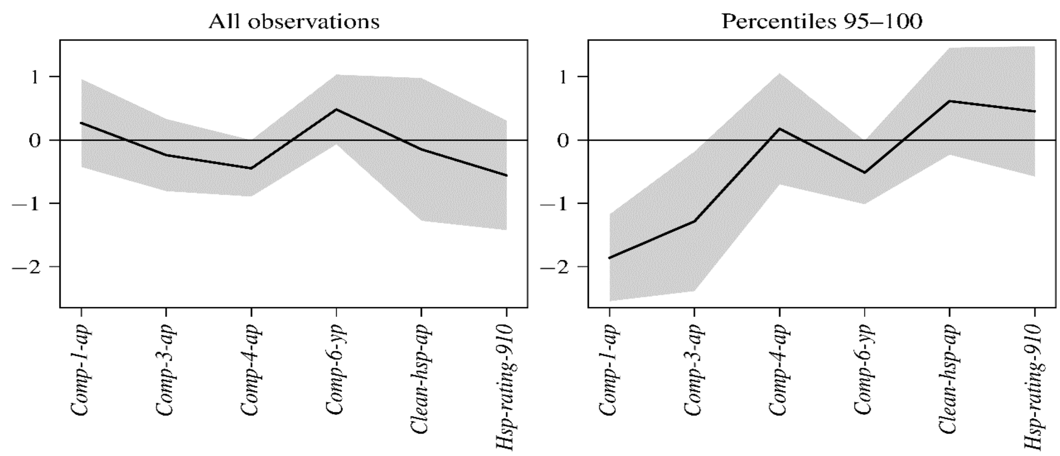

Fig. 3. The average treatment effect of the reform for patient experience of care measures with the 95% confidence intervals

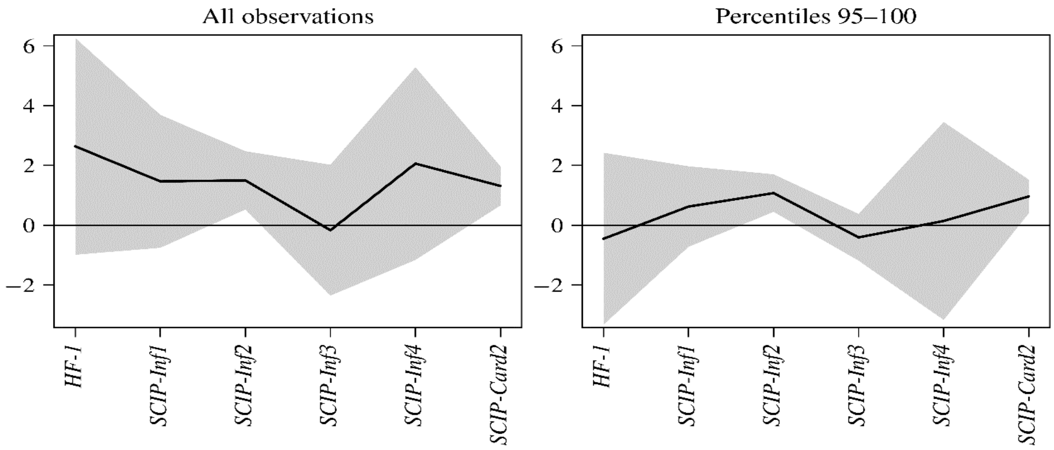

Fig. 4. The average treatment effect of the reform for clinical process of care measures with the 95% confidence intervals

Table 2.

Average treatment effect of incentives contract in percentile groups for each patient experience of care measure in dynamic panel data model

	Comp-1-ap	Comp-3-ap	Comp-4-ap	Comp-6-yp	Clean-hs-ap	Hsp-rating-910
τ_{0-10}	2.064***	0.510	−0.839	1.506***	−0.718	−1.307**
	(0.558)	(0.644)	(0.526)	(0.413)	(0.940)	(0.596)
τ_{10-20}	1.199***	0.205	−0.654*	1.008***	−0.458	−0.957*
	(0.468)	(0.505)	(0.362)	(0.361)	(0.842)	(0.503)
τ_{20-30}	0.854*	0.061	−0.582*	0.775***	−0.366	−0.799
	(0.453)	(0.430)	(0.308)	(0.329)	(0.769)	(0.484)
τ_{30-40}	0.568	−0.065	−0.547**	0.636**	−0.255	−0.693
	(0.429)	(0.373)	(0.279)	(0.320)	(0.721)	(0.478)
τ_{40-50}	0.365	−0.175	−0.491*	0.499	−0.176	−0.602
	(0.403)	(0.332)	(0.254)	(0.320)	(0.679)	(0.469)
τ_{50-60}	0.155	−0.271	−0.430*	0.373	−0.110	−0.506
	(0.378)	(0.286)	(0.221)	(0.301)	(0.596)	(0.494)
τ_{60-70}	−0.052	−0.371	−0.432**	0.253	−0.055	−0.418
	(0.364)	(0.262)	(0.219)	(0.299)	(0.524)	(0.490)
τ_{70-80}	−0.248	−0.492**	−0.355*	0.126	0.035	−0.303
	(0.343)	(0.233)	(0.199)	(0.279)	(0.450)	(0.511)
τ_{80-90}	−0.561*	−0.668***	−0.255	−0.018	0.165	−0.170
	(0.299)	(0.240)	(0.207)	(0.267)	(0.382)	(0.506)

Continues

	Comp-1-ap	Comp-3-ap	Comp-4-ap	Comp-6-yp	Clean-hs-ap	Hsp-rating-910
τ_{90-100}	–1.431*** (0.297)	–1.070*** (0.431)	0.032 (0.342)	–0.364 (0.252)	0.462 (0.380)	0.216 (0.475)
τ_{95-100}	–1.861*** (0.349)	–1.283** (0.562)	0.178 (0.448)	–0.513** (0.255)	0.614 (0.431)	0.452 (0.525)
All observations	0.269 (0.354)	–0.239 (0.290)	–0.448** (0.227)	0.482* (0.281)	–0.148 (0.575)	–0.558 (0.441)

Notes: Robust standard errors in parentheses. *, ** and *** denote significance at 0.1, 0.05 and 0.01 level, respectively. The deciles of hospitals approximate different levels of quality with respect to each measure. τ_{0-10}, τ_{10-20} etc. denote the effects in percentiles 0–10, 10–20 etc. with respect to each measure. Mean of top decile is regarded by the Medicare's pricing schedule as the best performance, so we approximate hospital group with the highest quality as percentiles 95–100.

Table 3.

Average treatment effect of incentives contract in percentile groups for each clinical process of care measure in dynamic panel data model

	HF-1	SCIP-Inf1	SCIP-Inf2	SCIP-Inf3	SCIP-Inf4	SCIP-Card2
τ_{0-10}	8.016*** (2.721)	4.868 (3.519)	5.260** (2.339)	0.172 (2.981)	7.849 (4.876)	2.864** (1.384)
τ_{10-20}	5.037** (2.240)	2.560 (1.873)	2.174*** (0.816)	–0.116 (1.853)	3.339 (2.766)	1.782*** (0.614)
τ_{20-30}	3.952* (2.208)	1.892 (1.419)	1.535*** (0.507)	–0.095 (1.524)	2.559 (2.532)	1.465*** (0.418)
τ_{30-40}	3.081 (2.122)	1.547 (1.169)	1.230*** (0.359)	–0.170 (1.270)	1.726 (2.320)	1.273*** (0.356)
τ_{40-50}	2.445 (2.063)	1.252 (0.979)	1.049*** (0.275)	–0.243 (1.148)	1.325 (2.087)	1.141*** (0.345)
τ_{50-60}	1.932 (2.045)	1.007 (0.829)	0.883*** (0.207)	–0.254 (0.958)	0.938 (2.097)	1.074*** (0.347)
τ_{60-70}	1.320 (1.933)	0.812 (0.679)	0.757*** (0.174)	–0.291 (0.848)	0.440 (2.108)	0.996*** (0.368)
τ_{70-80}	0.851 (1.881)	0.625 (0.580)	0.663*** (0.149)	–0.348 (0.717)	0.198 (2.061)	0.885*** (0.367)
τ_{80-90}	0.243 (1.708)	0.493 (0.518)	0.689*** (0.158)	–0.420 (0.581)	–0.105 (2.083)	0.891*** (0.279)

Continues

	HF-1	SCIP-Inf1	SCIP-Inf2	SCIP-Inf3	SCIP-Inf4	SCIP-Card2
τ_{90-100}	–0.309 (1.539)	0.515 (0.600)	1.009*** (0.288)	–0.462 (0.431)	0.128 (1.651)	0.941*** (0.284)
τ_{95-100}	–0.453 (1.466)	0.624 (0.686)	1.074*** (0.320)	–0.406 (0.398)	0.142 (1.688)	0.962*** (0.286)
All observations	2.638 (1.850)	1.466 (1.134)	1.499*** (0.496)	–0.167 (1.116)	2.061 (1.643)	1.312*** (0.332)

Notes: Robust standard errors in parentheses. *, ** and *** denote significance at 0.1, 0.05 and 0.01 level, respectively. The deciles of hospitals approximate different levels of quality with respect to each measure. τ_{0-10}, τ_{10-20} etc. denote the effects in percentiles 0–10, 10–20 etc. with respect to each measure. Mean of top decile is regarded by the Medicare's pricing schedule as the best performance, so we approximate hospital group with the highest quality as percentiles 95–100.

7. Discussion and Conclusion

Incentives contracts cause unintended effects for most capable agents: their performance deteriorates owing to intrinsic behavior, conformism and slacking efforts. Altruistic agents, however, would be interested in the social value of their performance per se. Concerning empirical work in healthcare, we are not aware of any literature which would explicitly quantify heterogeneity in agents' altruism with real data. The common approach would be a revelation of a mean relative weight of the altruistic component in the utility function of providers, which may be accomplished through field experiments or analysis of prescription records [Galizzi et al., 2015].

Overall, there is limited theoretical literature on social preferences in public good games and piece-rate incentives contracts, but little is known about the effect of altruism on the outcomes of relative performance reimbursement. Yet, incentives contracts based on relative performance become increasingly widespread in public industries, where the number of agents is large and the distribution of their outcomes is precisely unknown.

The novelty of the present paper is the analysis of the impact of provider altruism and motivation on the outcomes of pay-for-performance reimbursement in healthcare, where a fixed price contract on quantity is supplemented with a relative performance contract on quality. The theoretical part of the paper defines altruism as the term $\theta B(q)$ in the utility function of the healthcare provider. The parameter of altruism θ is assumed to be provider-specific. We proposed a theoretical model, forecasting the adverse effects for providers with the highest quality, where motivation and altruism towards the social value of performance cause quality deterioration.

The predictions of the model are tested in the empirical part of the paper, which uses Medicare's hospital level administrative panel data on a recent changeover to an incentive contract for 3000 acute care hospitals in 2013–2017. The analysis exploits the dynamic panel data approach and accounts for potential mean reversion before and after the reform. The empirical part paper defines altruism as incorporation of the values of certain measures of patient experience of care into the utility function of hospitals. At the same time, the values of the measures of the clinical process of care are assumed not to enter the $\theta B(q)$ component of the hospital's utility function.

As regards highest-quality hospitals, we discover deterioration of quality measures which may be linked to patient's benefit (communication of patients with medical personnel and patient ability to receive help promptly). These empirical results are consistent with the theoretical model. They may be interpreted as an illustration of the fact relative performance incentive contracts may be associated with motivation crowding out owing to the existence of altruism on the healthcare market.

Moreover, this paper has demonstrated differential behavioral response to incentives regulations in healthcare by high-quality and low-quality providers in the US Medicare. Therefore, the knowledge about altruism may be exploited by a social planner for extracting rents and providing subsidies based on an altruistic type of healthcare provider.

Appendix A.

Comparative statics of optimal quality by motivated and altruistic providers

Compute total amount of quality:

$$Q=\int_{\underline{\theta}}^{\hat{\theta}(\gamma)} q^*(\theta,\gamma) f(\theta)d\theta+\int_{\hat{\theta}(\gamma)}^{\tilde{\theta}(\gamma)} \tilde{q} f(\theta)d\theta+\int_{\tilde{\theta}(\gamma)}^{\check{\theta}(\gamma)} q^*(\theta,\gamma) f(\theta)d\theta+\int_{\check{\theta}(\gamma)}^{\bar{\theta}} \bar{q}(\gamma) f(\theta)d\theta .$$

Differentiating Q in γ and using Leibnitz integral rule yields:

$$\frac{dQ}{d\gamma}=\int_{\underline{\theta}}^{\hat{\theta}(\gamma)} \frac{\partial q^*(\theta,\gamma)}{\partial\gamma} f(\theta)d\theta+q^*\left(\hat{\theta}(\gamma),\gamma\right) f\left(\hat{\theta}(\gamma)\right)\frac{\partial\hat{\theta}(\gamma)}{\partial\gamma}+\tilde{q} f\left(\tilde{\theta}(\gamma)\right)\frac{\partial\tilde{\theta}(\gamma)}{\partial\gamma}-$$

$$-\tilde{q} f\left(\hat{\theta}(\gamma)\right)\frac{\partial\hat{\theta}(\gamma)}{\partial\gamma}+\int_{\tilde{\theta}(\gamma)}^{\check{\theta}(\gamma)} \frac{\partial q^*(\theta,\gamma)}{\partial\gamma} f(\theta)d\theta-q^*\left(\tilde{\theta}(\gamma),\gamma\right) f\left(\tilde{\theta}(\gamma)\right)\frac{\partial\tilde{\theta}(\gamma)}{\partial\gamma}+$$

$$+q^*\left(\check{\theta}(\gamma),\gamma\right) f\left(\check{\theta}(\gamma)\right)\frac{\partial\check{\theta}(\gamma)}{\partial\gamma}+\int_{\check{\theta}(\gamma)}^{\bar{\theta}} \frac{\partial\bar{q}(\gamma)}{\partial\gamma} f(\theta)d\theta-\bar{q}(\gamma) f\left(\check{\theta}(\gamma)\right)\frac{\partial\check{\theta}(\gamma)}{\partial\gamma}$$

where each line corresponds to each summand in equation for the total quality Q.

Terms $q^*\left(\tilde{\theta}(\gamma),\gamma\right)) f\left(\tilde{\theta}(\gamma)\right)\frac{\partial\tilde{\theta}(\gamma)}{\partial\gamma}$ and $\tilde{q}f\left(\tilde{\theta}(\gamma)\right)\frac{\partial\tilde{\theta}(\gamma)}{\partial\gamma}$ cancel out by definition of $\tilde{\theta}(\gamma)$. Similarly, terms $q^*\left(\check{\theta}(\gamma),\gamma\right)f\left(\check{\theta}(\gamma)\right)\frac{\partial\check{\theta}(\gamma)}{\partial\gamma}$ and $\overline{q}f\left(\check{\theta}(\gamma)\right)\frac{\partial\check{\theta}(\gamma)}{\partial\gamma}$ cancel out by definition of $\check{\theta}(\gamma)$. So

$$\frac{dQ}{d\gamma}=\int_{\underline{\theta}}^{\hat{\theta}(\gamma)}\frac{\partial q^*(\theta,\gamma)}{\partial\gamma}f(\theta)d\theta+\int_{\tilde{\theta}(\gamma)}^{\check{\theta}(\gamma)}\frac{\partial q^*(\theta,\gamma)}{\partial\gamma}f(\theta)d\theta+\int_{\check{\theta}(\gamma)}^{\overline{\theta}}\frac{\partial\overline{q}(\gamma)}{\partial\gamma}f(\theta)d\theta+$$
$$+\left[q^*\left(\hat{\theta}(\gamma),\gamma\right)-\tilde{q}\right]f\left(\hat{\theta}(\gamma)\right)\frac{\partial\hat{\theta}(\gamma)}{\partial\gamma}.$$

Applying the implicit function theorem to

$$U\left(\hat{\theta},q^*\left(\hat{\theta},\gamma\right)\right)-U\left(\hat{\theta},\tilde{q}\right)-(V-wp)\gamma=0$$

yields

$$\frac{\partial\hat{\theta}}{\partial\gamma}=\frac{\left(\tilde{q}-q^*\left(\hat{\theta}(\gamma),\gamma\right)\right)+(V-wp)}{B\left(q^*\left(\hat{\theta}(\gamma),\gamma\right)\right)-B(\tilde{q})}.$$

Therefore, the change in quality simplifies to

$$\frac{dQ}{d\gamma}=\underbrace{\int_{\underline{\theta}}^{\hat{\theta}(\gamma)}\frac{\partial q^*(\theta,\gamma)}{\partial\gamma}f(\theta)d\theta}_{+}+\underbrace{\int_{\tilde{\theta}(\gamma)}^{\check{\theta}(\gamma)}\frac{\partial q^*(\theta,\gamma)}{\partial\gamma}f(\theta)d\theta}_{+}+\underbrace{\int_{\check{\theta}(\gamma)}^{\overline{\theta}}\frac{\partial\overline{q}(\gamma)}{\partial\gamma}f(\theta)d\theta}_{-}+$$
$$+\underbrace{\left(\tilde{q}-q^*\left(\hat{\theta}(\gamma),\gamma\right)\right)}_{+}\frac{f\left(\hat{\theta}(\gamma)\right)}{\underbrace{B\left(q^*\left(\hat{\theta}(\gamma),\gamma\right)\right)-B(\tilde{q})}_{+}}\left(\underbrace{\left(\tilde{q}-q^*\left(\hat{\theta}(\gamma),\gamma\right)\right)}_{+}+\underbrace{V-wp}_{-}\right).$$

Appendix B.

Estimation with dynamic panel data models

Table B1.

Coefficients for explanatory variables in dynamic panel data model with excluded mean reversion (Patient experience of care measures)

	Comp-1-ap	Comp-3-ap	Comp-4-ap	Comp-6-yp	Clean-hs-ap	Hsp-rating-910
$L(y)$	0.779*** (0.024)	0.684*** (0.022)	0.496*** (0.028)	0.642*** (0.025)	0.608*** (0.025)	0.636*** (0.021)
$L^2(y)$	0.184*** (0.014)	0.216*** (0.016)	0.135*** (0.019)	0.236*** (0.020)	0.111*** (0.015)	0.151*** (0.015)
r	13.444*** (2.731)	2.975 (2.496)	–4.560 (3.857)	12.310*** (3.113)	–6.750** (2.936)	–4.258** (1.891)
$L(\text{beds1})\cdot r$	0.002 (0.002)	0.001*** (0.000)	0.000* (0.000)	0.000 (0.001)	0.002 (0.003)	0.003 (0.002)
$L(\text{beds2})\cdot r$	0.001 (0.002)	–	–	–0.000 (0.001)	0.002 (0.003)	0.001 (0.002)
$L^2(\text{beds1})\cdot r$	0.001*** (0.000)	0.000 (0.002)	–0.002 (0.002)	0.001*** (0.000)	0.002*** (0.000)	–0.004** (0.002)
$L^2(\text{beds2})\cdot r$	–	–0.001 (0.002)	–0.003* (0.002)	–	–	–0.004** (0.002)
$L(\text{public1})\cdot r$	–0.162 (0.468)	1.806** (0.900)	1.010 (0.809)	–	–	–
$L(\text{public2})\cdot r$	–	–	–	0.346 (0.542)	0.184 (0.782)	–1.019 (0.694)
$L^2(\text{public1})\cdot r$	–	0.489 (0.301)	–	0.287* (0.167)	–	–
$L^2(\text{public2})\cdot r$	–0.472** (0.193)	–	–0.044 (0.245)	–	–0.305 (0.258)	–0.320 (0.275)
$L(y)\cdot r$	–0.278*** (0.037)	–0.077** (0.033)	–0.054 (0.043)	–0.119*** (0.041)	0.029 (0.036)	0.002 (0.032)
$L^2(y)\cdot r$	0.102*** (0.028)	0.022 (0.027)	0.115*** (0.032)	–0.024 (0.030)	0.058** (0.024)	0.053** (0.026)
emergency	0.073 (0.304)	0.288 (0.467)	0.284 (0.326)	–0.144 (0.358)	0.800** (0.330)	0.847** (0.393)
urban	–0.700** (0.339)	–1.269*** (0.417)	–1.472*** (0.411)	–0.867*** (0.284)	–1.168** (0.461)	–0.253 (0.441)

Continues

	Comp-1-ap	Comp-3-ap	Comp-4-ap	Comp-6-yp	Clean-hs-ap	Hsp-rating-910
medicare share	−0.054 (0.869)	−2.197* (1.197)	−2.403** (0.946)	−1.697** (0.794)	−0.309 (1.134)	−2.573** (1.135)
casemix	0.118 (0.529)	2.038*** (0.735)	1.602** (0.690)	1.138** (0.563)	1.892*** (0.710)	2.342*** (0.908)
resbed	1.828 (1.337)	−4.134 (2.573)	−1.053 (2.132)	0.233 (1.333)	0.629 (1.871)	1.204 (1.334)
dsh	−0.585 (0.879)	2.806 (1.795)	0.167 (1.049)	−0.720 (1.055)	−0.412 (1.304)	−2.662* (1.374)
beds1	−0.002 (0.002)	−0.004** (0.002)	−0.003** (0.002)	−0.003** (0.001)	−0.003* (0.002)	−0.006*** (0.002)
beds2	0.000 (0.001)	−0.003 (0.002)	−0.001 (0.001)	0.000 (0.001)	0.000 (0.002)	−0.000 (0.002)
$L(\text{beds1})$	0.004** (0.002)	0.006** (0.002)	0.002 (0.002)	0.004** (0.002)	0.005* (0.003)	0.007*** (0.002)
$L(\text{beds2})$	–	0.001 (0.003)	0.002 (0.002)	–	–	–
$L^2(\text{beds1})$	−0.002 (0.002)	−0.001 (0.002)	0.000 (0.002)	−0.001 (0.002)	−0.005** (0.002)	−0.003 (0.002)
$L^2(\text{beds2})$	−0.003* (0.002)	–	–	−0.002 (0.001)	−0.006*** (0.002)	–
public1	−0.217 (0.169)	0.102 (0.258)	−0.095 (0.231)	0.035 (0.166)	−0.421* (0.221)	−0.291 (0.241)
public2	1.019** (0.516)	0.701 (0.688)	−0.142 (0.694)	0.001 (0.632)	0.144 (0.701)	1.340* (0.771)
$L(\text{public1})$	−0.447** (0.201)	−0.596** (0.294)	−0.384 (0.240)	−0.313 (0.209)	−0.235 (0.251)	−0.290 (0.291)
$L(\text{public2})$	−1.237** (0.550)	−2.638*** (0.934)	−1.015 (0.895)	−0.241 (0.613)	−0.152 (0.711)	−1.378* (0.770)
$L^2(\text{public1})$	−0.197 (0.206)	−0.607** (0.306)	−0.533* (0.285)	−0.607** (0.239)	−0.495* (0.301)	0.088 (0.291)
$L^2(\text{public2})$	0.267 (0.453)	1.567* (0.890)	1.003 (0.778)	−0.444 (0.557)	0.002 (0.755)	1.285* (0.662)
constant	4.265* (2.248)	5.820** (2.550)	25.965*** (3.050)	11.545*** (2.250)	18.826*** (2.716)	13.870*** (2.096)

Continues

	Comp-1-ap	Comp-3-ap	Comp-4-ap	Comp-6-yp	Clean-hs-ap	Hsp-rating-910
Observations	21066	21062	21056	21056	21066	21064
Hospitals	3340	3340	3339	3338	3340	3339
Arellano–Bond test statistic	–1.238	–1.635	–1.993	–1.615	–1.948	0.685

Note: Robust standard errors in parentheses. *, ** and *** denote significance at 0.1, 0.05 and 0.01 level, respectively. Arellano-Bond test statistics for absence of order 2 serial correlation in the first-differenced errors.

* *
*

References

Arce D.G. (2013) Principals' Preferences for Agents with Social Preferences. *Journal of Economic Behavior & Organization,* 90, pp. 154–163.

Arellano M., Bond S. (1991) Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies,* 58, pp. 277–297.

Arellano M., Bover O. (1995) Another Look at the Instrumental Variable Estimation of Error-Components Models. *Journal of Econometrics,* 68, 1, pp. 29–51.

Bénabou R., Tirole J. (2006) Incentives and Prosocial Behavior. *American Economic Review,* 96, 5, pp. 1652–1678.

Besstremyannaya G. (2016) Differential Effects of Declining Rates in a Per Diem Payment System. *Health Economics,* 25, 12, pp. 1599–1618.

Bisiaux R., Chi Y.-L. (2014) Republic of Korea: Value Incentive Programme. *Paying for Performance in Health Care* (eds. C. Cashin, Y.-L. Chi, P.C. Smith, M. Borowitz, S. Thomson). World Health Organization, pp. 251–264.

Blomqvist Å. (1997) Optimal Non-Linear Health Insurance. *Journal of Health Economics,* 16, 3, pp. 303–21.

Blundell R., Bond S. (1998) Initial Conditions and Moment Restrictions in Dynamic Panel Data Models. *Journal of Econometrics,* 87, 1, pp. 115–143.

Bousquet F., Bisiaux R., Chi Y.-L. (2014) France: Payment for Public Health Services. *Paying for Performance in Health Care* (eds. C. Cashin, Y.-L. Chi, P.C. Smith, M. Borowitz, S. Thomson). World Health Organization, pp. 141–156.

Brekke K.R., Canta C., Siciliani L., Straume O.R. (2021) Hospital Competition in a National Health Service: Evidence from a Patient Choice Reform. *Journal of Health Economics*, 79.

Brekke K.R., Siciliani L., Straume O.R. (2011) Hospital Competition and Quality with Regulated Prices. *Scandinavian Journal of Economics*, 113, 2, pp. 444–469.

Brekke K.R., Siciliani L., Straume O.R. (2012) Quality Competition with Profit Constraints. *Journal of Economic Behavior & Organization*, 84, 2, pp. 642–659.

Buetow S. (2008) Pay-for-performance in New Zealand Primary Health Care. *Journal of Health Organization and Management,* 22, 1, pp. 36–47.

Buurman M., Dur R. (2012) Incentives and the Sorting of Altruistic Agents into Street-Level Bureaucracies. *Scandinavian Journal of Economics,* 114, 4, pp. 1318–1345.

Campbell S.M., Reeves D., Kontopantelis E., Sibbald B., Roland M. (2009) Effects of Pay for Performance on the Quality of Primary Care in England. *New England Journal of Medicine,* 361, 4, pp. 368–378.

Centers for Medicare and Medicaid Services (2007) *Report to Congress: Plan to Implement a Medicare Hospital Value-Based Purchasing Program*. U.S. Department of Health; Human Services. Available at: https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/AcuteInpatientPPS/downloads/hospitalVBPplanRTCFinalsubmitted2007.pdf

Choné P., Ma C.A. (2011) Optimal Health Care Contract Under Physician Agency. *Annales d'Économie et de Statistique,* 101/102, pp. 229–256.

Christianson J.B., Leatherman S., Sutherland K. (2008) Lessons from Evaluations of Purchaser Pay-for-performance Programs. *Medical Care Research and Review,* 65, 6 Suppl, pp. 5S–35S.

Damberg C.L., Sorbero M.E., Lovejoy S.L., Martsolf G.R., Raaen L., Mandel D. (2014) *Measuring Success in Health Care Value-Based Purchasing Programs: Findings from an Environmental Scan, Literature Review, and Expert Panel Discussions*. Research Report. Santa Monica, CA: RAND Corporation. Available at: http://www.rand.org/pubs/research_reports/RR306.html

Eijkenaar F., Emmert M., Scheppach M., Schöffski O. (2013) Effects of Pay for Performance in Health Care: A Systematic Review of Systematic Reviews. *Health Policy,* 110, 2, pp. 115–130.

Ellis R.P., McGuire T.G. (1986) Provider Behavior Under Prospective Reimbursement: Cost Sharing and Supply. *Journal of Health Economics,* 5, 2, pp. 129–151.

Eriksson T., Villeval M.C. (2008) Performance-Pay, Sorting and Social Motivation. *Journal of Economic Behavior & Organization,* 68, 2, pp. 412–421.

Fischer P., Huddart S. (2008) Optimal Contracting with Endogenous Social Norms. *American Economic Review,* 98, 4, pp. 1459–1475.

Galizzi M.M., Tammi T., Godager G., Linnosmaa I., Wiesen D. (2015) *Provider Altruism in Health Economics*. Discussion Paper. Tampere, Finland: National Institute for Health; Welfare.

Green E.P. (2014) Payment Systems in the Healthcare Industry: An Experimental Study of Physician Incentives. *Journal of Economic Behavior & Organization,* 106, pp. 367–378.

Hamilton J.D. (1994) *Time Series Analysis*. Princeton: Princeton University Press.

Hayashida K., Murakami G., Matsuda S., Fushimi K. (2021) History and Profile of Diagnosis Procedure Combination (DPC): Development of a Real Data Collection System for Acute Inpatient Care in Japan. *Journal of Epidemiology*, 31, 1, pp. 1–11.

HCAHPS online (2013) *HCAHPS Fact Sheet (HCAHPS Hospital Survey) August 2013*. Available at: http://www.hcahpsonline.org/files/August_2013_HCAHPS_Fact_Sheet3.pdf

Hennig-Schmidt H., Selten R., Wiesen D. (2011) How Payment Systems Affect Physicians' Provision Behaviour – an Experimental Investigation. *Journal of Health Economics,* 30, 4, pp. 637–646.

Hölmstrom B., Milgrom P. (1991) Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics, & Organization,* 7, pp. 24–52.

Joskow P.L., Schmalensee R. (1986) Incentive Regulation for Electric Utilities. *Yale Journal on Regulation,* 4, pp. 1–49.

Kairies-Schwarz N. (2014) *Altruism Heterogeneity and Quality Competition Among Healthcare Providers*. Ruhr Economic Papers. Available at: http://www.cinch-essen.de/files/REP_14_507.pdf

Kristensen S.R., Siciliani L., Sutton M. (2016) Optimal Price-Setting in Pay for Performance Schemes in Health Care. *Journal of Economic Behavior & Organization*, 123, pp. 57–77.

Laffont J.-J., Tirole J. (1993) *A Theory of Incentives in Procurement and Regulation*. MIT Press.

Liu T., Ma C.A. (2013) Health Insurance, Treatment Plan, and Delegation to Altruistic Physician. *Journal of Economic Behavior & Organization,* 85, pp. 79–96.

Makris M., Siciliani L. (2013) Optimal Incentive Schemes for Altruistic Providers. *Journal of Public Economic Theory,* 15, 5, pp. 675–699.

Makris M., Siciliani L. (2014) *Should Incentive Schemes Be High-Powered or Low-Powered in the Presence of Motivated Agents?* Available at: http://www.bris.ac.uk/media-library/sites/cmpo/migrated/documents/luigisiciliani.pdf

McGuire T.G. (2000) Physician Agency. *Handbook of Health Economics* (eds. A.J. Culyer, J.P. Newhouse), 1, part A, pp. 461–536.

Medicare.gov. (2017) *Timely & Effective Care Measures*. Available at: https://www.medicare.gov/hospitalcompare/Data/Measures.html

Murdock K. (2002) Intrinsic Motivation and Optimal Incentive Contracts. *RAND Journal of Economics,* 33, 4, pp. 650–671.

Nawata K., Kawabuchi K. (2013) Evaluation of the DPC-Based Inclusive Payment System in Japan for Cataract Operations by a New Model. *Mathematics and Computers in Simulation,* 93, pp. 76–85.

Olivella P., Siciliani L. (2017) Reputational Concerns with Altruistic Providers. *Journal of Health Economics*, 55, pp. 1–13.

Oxholm A.S., Kristensen S.R., Sutton M. (2018) Uncertainty about the Effort–Performance Relationship in Threshold-based Payment Schemes. *Journal of Health Economics*, 62, pp. 69–83.

Peluso A., Berta P., Vinciotti V. (2019) Do Pay-for-Performance Incentives Lead to a Better Health Outcome? *Empirical Economics*, 56, 6, pp. 2167–2184.

Prendergast C. (1999) The Provision of Incentives in Firms. *Journal of Economic Literature,* 37, pp. 7–63.

Rose-Ackerman S. (1986) Altruistic Nonprofit Firms in Competitive Markets: The Case of Day-care Centers in the United States. *Journal of Consumer Policy*, 9, 3, pp. 291–310.

Rose-Ackerman S. (1996) Altruism, Nonprofits, and Economic Theory. *Journal of Economic Literature*, 34, 2, pp. 701–728.

Shleifer A. (1985) A Theory of Yardstick Competition. *RAND Journal of Economics,* 16, pp. 319–327.

Siciliani L. (2009) Paying for Performance and Motivation Crowding Out. *Economics Letters,* 103, 2, pp. 68–71.

Sutton M., Nikolova S., Boaden R., Lester H., McDonald R., Roland M. (2012) Reduced Mortality with Hospital Pay for Performance in England. *New England Journal of Medicine,* 367, 19, pp. 1821–1828.

Thompson J.D., Averill R.F., Fetter R.B. (1979) Planning, Budgeting, and Controlling–One Look at the Future: Case-Mix Cost Accounting. *Health Services Research,* 14, 2, pp. 111–125.

Windmeijer F. (2005) A Finite Sample Correction for the Variance of Linear Efficient Two-Step GMM Estimators. *Journal of Econometrics,* 126, pp. 25–51.