Journal of Economics & Management Strategy

WILEY

Journal of Economics & Management Strategy

ORIGINAL ARTICLE OPEN ACCESS

Optimal Healthcare Contracts: Theory and Empirical Evidence

Paolo Berta[1,2] | Gianni De Fraja[3,4] | Stefano Verzillo[5]

[1]Department of Statistics, University of Milan-Bicocca, Milano, Italy | [2]CRISP - Interuniversity Research Centre on Public Services, University of Milan-Bicocca, Milano, Italy | [3]Nottingham School of Economics, Nottingham, UK | [4]C.E.P.R., London, UK | [5]European Commission, Joint Research Centre (JRC), Ispra, Varese, Italy

Correspondence: Stefano Verzillo (stefano.verzillo@ec.europa.eu)

Received: 12 July 2024 | **Revised:** 5 March 2026 | **Accepted:** 25 April 2026

Keywords: cream skimming | diagnosis related groups | healthcare contracts | hospitals | lombardy | patients choice

ABSTRACT

In this paper, we investigate the contracts offered by a large healthcare purchaser to health service providers. Contracts are based on the DRG principle that all hospitalizations in a diagnosis group are reimbursed at the same rate. This principle is relaxed in practice, as in several cases, the amount reimbursed within each DRG exhibits considerable variability. We build a theoretical model which explains this variability as the attempt of the health authority to ensure appropriate matching between hospitals and patients. We test the model using a very large and detailed administrative data set for the largest region in Italy. In line with our theoretical results, we show that the state funded purchaser offers providers a system of incentives such that, as required by optimality, that providers which treat more patients receive a higher average reimbursement per treatment, suggesting therefore that they treat on average more difficult patients and are compensated for doing so.

JEL Classification: I11, I18, D82, H42

1 | Introduction

The basis upon which many providers of health care across the world are paid for the services they provide is the system of Diagnosis Related Groups (DRG). With this system, reimbursements to health providers for the treatments they perform are based on an exhaustive list which determines the notional cost of each one of several hundred different health treatments or diagnoses. This menu of tariffs is fixed prior to patients' admission and is updated yearly in line with technological advances, and the purchaser's priorities.[1]

First introduced for Medicare payments in 1983 (Fetter 1991), the DRG mechanism is progressively replacing the retrospective reimbursement system, whereby providers are simply refunded all the costs they have incurred in treating the patients they admit. In contrast to the latter, a DRG mechanism reduces healthcare costs (Ellis and McGuire 1990), since it gives healthcare providers little incentive to perform costly unnecessary treatments. It does, on the other hand, give them an incentive to turn down patients whose diagnoses suggest they will require more complex and expensive care. Known as "cream skimming," this practice is considered a drawback of a DRG system for its potential to make it impossible for the most serious patients to find a provider willing to treat them under the terms of the mechanism.[2]

In this paper, we study the details of the DRG system which governs the purchase of health services by the publicly funded health purchaser of Lombardy, the wealthiest and most populous region in Italy. The authority offers, on a take-it-or leave-it basis, the same menu of reimbursements to all the 150 or so hospitals of the region, some public, some private non-profit, some for-profit, which compete to provide health care.

A key complication in the design of a DRG mechanism to be offered on equal terms to a very large number of hospitals

https://doi.org/10.1111/jems.70035

serving a large population is the obvious observation that not all hospitalizations in a given diagnosis group are equally costly: idiosyncratic differences among patients require some to receive more expensive treatments to achieve a given health outcome. Given the large number of providers, it is also inevitable that they differ in their ability to treat a given diagnosis. Thus, a key role of the DRG mechanism is to ensure an efficient allocation of patients to hospitals. In short, paradoxically, the reimbursement menu should be designed to *manage cream skimming* so as to ensure a socially efficient hospital-patient matching across the region. We show in Section 3 that, to achieve this efficient matching, the tariff for each DRG must be set by the purchaser at a finely tuned level: high enough to ensure that the hospitals best suited to supply the DRG are willing to provide it, but low enough to discourage less suitable hospitals to admit patients who need the treatment.

This fine tuning is done in the Lombardy region's DRG system in two distinct ways. First, 236 of the 528 separate DRGs form 118 pairs: one without complications and one with complications, the latter commanding a higher reimbursement, and applying when a precisely specified set of conditions is present *and* demonstrably affects the cost of treatment.[3] Second, since this coarse split is insufficient to capture fully the many nuanced differences among hospitalizations which remain, the Lombardy tariff book adjusts the “base” reimbursement for each DRG according to a specific set of additional characteristics of each hospitalization and each hospital.

However, this process cannot be pushed too far. Apart from the fact that listing and pricing each possible minute combination of observed characteristics would be exceedingly cumbersome, the theoretical model in Section 3 points to an insurmountable conceptual difficulty. While some of the patients' characteristics that determine cost differences between treatments within the same DRG can be measured accurately, others are observed by the provider but not by the purchaser, or, if they are observed by both parties, they are too vague or too subjective and cannot be precisely and objectively measured. Therefore, any algorithm to be used to compute the reimbursements due to the hospitals can take into account the former, such as the patient's BMI or specific pre-existing conditions, but not the latter, like “poor general health” or hard-to-quantify mental health problems.[4] Intuition strongly suggests a positive correlation between these two sets of characteristics: in terms of the above examples, patients with an excessive BMI are more likely to have poor general health. It is of course impossible to verify in practice any such correlation, since by definition only one of the set of characteristics can be measured at all.

The main theoretical contribution of Section 3 is that, given this positive correlation between the contractible and the non-contractible characteristics of a hospitalization, the socially optimal allocation of patients to hospitals requires that there should necessarily be a positive correlation between average reimbursement and the complexity of the treatment. This is because the health authority must choose adjustments to the reimbursements for each given DRG such that providers which are better able to treat more complex hospitalizations receive a higher *average* reimbursement for that DRG. These higher reimbursements act as compensation and reward to provide incentives for the hospitals which are more capable of treating more complex patients to indeed admit a more complex, and hence more costly, patient mix. Proposition 2 shows that the additional reimbursements for the *observable* adverse characteristics of the hospitalization are designed to exceed the additional cost due to these characteristics. This is to compensate for the hospital's additional costs due to the *unobservable* characteristics of the patient. A specific example illustrates this idea. The contract regulating the reimbursements by the Lombardy to its providers specifies that hospitals can claim €115 if they need to carry out a hemodialysis. The idea behind Proposition 2 is that this additional sum *exceeds* the cost of the hemodialysis in order to compensate the hospital for the expected additional cost due to higher likelihood of complex non-contractible combinations of costly characteristics of the hospitalization which patients needing a hemodialysis are likely to have. A hospital less able to provide this DRG might well be able to cover the cost of the contractible hemodialysis, but may find the additional non-contractible costs excessive and so it will try to turn down patients who require these costs.

We take our model to the administrative data set of the universe of the treatments carried out in 2013 in the contractual environment set up by the Lombardy healthcare purchaser. We first confirm the widespread use of adjustments to the basic menu of prices with the observation, in Section 2, that the pattern of reimbursements exhibits considerable systematic variation within each DRG. In Section 5, we indeed find that, consistent with the conditions for socially optimal matching between patients and hospitals obtained theoretically in Section 3, hospitals which treat more patients for a given DRG are paid more, on average, for performing this DRG. This effect is stronger, Table 4, Column (5), when the partition of the set of all hospitalization is made coarser by grouping together the pairs of DRG with and without complications. It remains significant when the full panoply of ICD-9-CM fixed effects is introduced. Both these results are consistent with the mechanism at the basis of our theoretical model. Also consistent with this mechanism are the results of two placebo regressions, where we show that the average reimbursement for a given DRG is *not* positively correlated with the number of patients treated either for daily hospitalizations, or for emergency treatments, where patients' choice and cream skimming are, respectively, less important and less feasible. The signs of the coefficients of the controls point to a positive correlation between the contractible characteristics of a hospitalization and the non-contractible ones. This is further confirmation for both our theoretical analysis and for our interpretation of the results as an indication of the possible optimality of the contract offered by the Lombardy region to its providers.

We underline that we are merely establishing a *positive correlation* between *reimbursement* received for a given DRG and the *number of patients treated* for that DRG, that is its capacity to attract patients, which we label *attractiveness*. This positive correlation may be due to one or both of two possibilities: (i) at least some patients who need a certain treatment have at least some information about the treatment they may receive at their local hospital, and, the more serious their condition, and hence the more expensive their treatment, the harder they will try to avoid hospitals they believe to be less capable of offering a treatment appropriate to their more serious diagnosis, and (ii) hospitals that are relatively less able to treat more complex cases of a certain DRG will be keener to try to turn down the most complex patients for that DRG, namely, those for whom the reimbursement is more likely to be insufficient to cover their expected higher cost of treatment. Note that (i) implies that patients respond to hospital perceived level of service, as they do in other contexts (Luft et al. 1990, or Gutacker et al. 2016), and (ii) implies that hospitals respond to the incentives provided by the competitive features of their contracts, as shown for example by Einav et al. (2018) and Eliason et al. (2018). For practical purposes, it does not matter which of the two mechanisms has the stronger effect. Social efficiency depends on the matching pattern, not on what determines it: the aim of the design of the reimbursement menu is to ensure that hospitals' and patients' preferences are aligned: hospitals' with specialism in some DRG must prefer to attract the more serious patients who need that DRG the most, and they, in turn, precisely because their diagnosis is more serious, are keen to be treated by these more specialized hospitals. At the same time, the health authority must avoid leaving excessive room for gaming, in the form of upcoding[5] and unbundling.

Providing evidence of correlation between reimbursement and number of patients treated at the DRG-hospital level, the first indispensable step toward understanding how contracts are determined, is not an easy process. It requires teasing out from the data the choices of patients and hospitals in a plausible and robust manner. In Section 6 we suggest a route to identify other plausible influences on the rules for reimbursement chosen by the region and provide two plausible examples.

The paper is organized as follows. In Section 2, we describe in detail the systematic variation exhibited by the reimbursements in the Lombardy DRG system. In Section 3, we extend the model in De Fraja (2000) to the case of multiple DRGs and multiple budget-constrained hospitals. The empirical strategy is described in Section 4. Section 5 presents our results, which confirm the theoretical hypothesis of a positive correlation between average reimbursement and the fraction of the potential patients treated for a given DRG. In Section 6, we study how to identify possible links between exogenous factors and the reimbursement rules. Section 7 concludes. An online appendix contains additional theoretical material and results.

2 | The Data: Deviations From the Tariff

Beginning thirty years ago, a series of reforms devolved to the Italian regions the responsibility for the financing and the provision of healthcare. Lombardy, the largest region in Italy,[6] home to one sixth of Italy's population and accounting for over one fifth of her GDP, and one of the wealthiest and best educated in Europe, was among the first fully to embrace these reforms. In 1997, it created a quasi-market for healthcare built on the pillars of (i) separation between the monopsonistic purchaser, Lombardy's regional government, and the providers, the region's accredited hospitals, (ii) competition among providers, be they private or public, (iii) a transparent mechanism of reimbursements, the same for all providers, based on the international DRG system, and (iv) patients' freedom to choose the hospital where they are treated (Brenna 2011; Fabbri and Robone 2010).

In this paper, we evaluate the performance of this quasi-market in 2013, when the system had been in place for over a decade. To this aim, we use an administrative data set which covers the universe of Lombardy hospitalizations in the year. For each hospitalization, the data report the hospital where it took place, the dates of admission and discharge, whether or not it was an emergency admission, the principal diagnosis and up to five co-diagnoses, and the main procedure provided to the patient and up to 5 secondary procedures. Both diagnoses and procedures are coded according to the Ninth Revision International Classification of Diseases, Clinical Modification (ICD-9-CM 1996). It also reports whether a special-care unit was used, and the wards where the patient was cared for. This information determines the DRG code assigned to the hospitalization, which a complex algorithm combines with the information on diagnoses and procedures to compute the payment to the provider as reimbursement for the hospitalization. The data include some demographic information on the patient, including their age, gender, their marital status, their place of birth, the referring physician and their municipality and postcode of residence, which allows us to define the patient's location as the intersection of the two.[7] The data set, due to its administrative nature, contains no information on outcomes, other than whether re-admission or death took place within 30 days. Table 1 reports summary statistics on the number of hospitalizations and hospitals in the region; Table 2 reports those for the variables we used in the empirical analysis.

The tariff for each DRG is a schedule characterized by four parameters. (i) The "daily" reimbursement, which applies when the hospitalization requires less than a 24-h stay in hospital. (ii) The standard tariff, that is the flat reimbursement, independent of the number of nights the patient spends in hospital, for hospitalizations that require up to a threshold number of overnight stays. (iii) This DRG specific threshold is the third parameter. (iv) The additional per day reimbursement rate, for hospitalizations that need to exceed the threshold. All these four parameters[8] of the schedule vary by DRG, and their descriptive statistics are collected in Table 3. More details are in the Appendix.

In this paper, we aim to identify factors associated with deviations from the contractual tariff, for those reimbursed under the standard part of the tariff, (ii) in the above paragraph. To ensure comparability across DRGs with widely different tariffs, we calculate the *percentage deviation from the contractual tariff*: formally, if $\text{reimbursement}_{ith}$ is the actual reimbursement to hospital h for hospitalization i allocated to DRG t, and if tariff_t is the tariff for DRG t, then we define the "deviation" variable for hospitalization i under DRG t in hospital h as

$$R_{ith} = \frac{\text{reimbursement}_{ith} - \text{tariff}_t}{\text{tariff}_t} \times 100. \tag{1}$$

In principle, the DRG system of a fixed reimbursement per hospitalization would ensure that the tariff matches exactly the cost, making $R_{ith} = 1$ for every hospitalization i. The data,

TABLE 1 | Summary statistics for DRGs, DRG- hospital pairs, and hospitals.

	Min	Max	Avg	St. Dev
DRG				
Number of hospitalizations	2	47,774	1,276.3	3,002.7
Number of hospitals providing it	1	133	68.2	33.3
Min deviation from tariff: R_{ith} in Equation (1)	−100.0	20.6	−54.8	30.3
Mean deviation from the tariff	−53.7	51.4	−5.9	10.4
Max deviation from tariff: R_{ith} in Equation (1)	−14.2	282.2	6.1	20.9
DRG-Hospital				
Number of hospitalizations	2	2,443	18.7	56.9
Min deviation from tariff: R_{ith} in Equation (1)	−100.0	142.4	−23.5	31.1
Mean deviation from the tariff	−95.2	210.0	−6.9	15.4
Max deviation from tariff: R_{ith} in Equation (1)	−95.2	282.2	−0.7	14.0
Hospital				
Number of hospitalizations	23	27,193	4,441.9	4,612.8
Number of DRG per hospital	1	487	237.4	133.7

Note: Summary statistics for the regression sample, which includes 529 DRGs, the 36,080 DRG-hospital pairs, and 151 hospitals. The row "Min (Mean, Max) deviation from tariff" measure the deviation from the DRG tariff of the hospitalization with the least (the average, the most) reimbursement for each DRG, across all hospitals: the min is the percentage deviation of the lowest such reimbursements (−100% in the case) the max the highest (+26% here). Similarly for the other rows.

TABLE 2 | Summary statistics.

	Daily		Standard		Extra Threshold	
Variable	**mean**	**sd**	**mean**	**sd**	**mean**	**sd**
Independent variables						
Reimbursement in €	1442.3	1395.3	3612.0	4113.2	7547.1	12909.8
Deviation from tariff: R_{ith} in Equation (1)	−28.19	37.34	−3.79	18.50	30.60	24.60
Hospital-DRG characteristics						
Attractiveness (see Equation (18) below)	0.382	0.243	0.485	0.228	0.503	0.230
Hospital-DRG FE (see Column (6) in Table 4)	0.458	0.074	0.448	0.072	0.435	0.068
Patient Observable Characteristics						
Age of patient	48.1	24.1	51.3	28.5	60.6	24.6
Female patient	0.466	0.499	0.532	0.499	0.554	0.497
At least one comorbidity	0.157	0.364	0.371	0.483	0.530	0.499
Death within 30 days	0.041	0.197	0.044	0.204	0.107	0.309
Length of hospital stay (normalized)	0.596	0.197	0.285	0.255	1.938	1.463
Emergency admission	0.040	0.196	0.106	0.308	0.181	0.385
Distance from location to hospital	20.8	16.3	18.9	15.4	17.5	14.9
Never married	0.297	0.457	0.280	0.449	0.184	0.387
Born in the municipality of residence	0.219	0.414	0.233	0.422	0.230	0.421
Transfer to extra wards	0.009	0.093	0.097	0.295	0.292	0.454
Other Controls						
Number of local hospitals (Competition)	2.9	2.8	2.7	2.7	2.6	2.5
Number of DRG in hospital (total)	168.3	75.1	326.6	116.5	172.2	92.2
Number of observations	131,499		698,549		51,669	

Note: Summary statistics of the dataset used in the regression. See main text for explanation of the definition and the construction of the variables, and for the observations excluded from the data set.

however, reveal that only a tiny fraction of all hospitalizations are reimbursed within 0.01% of the tariff. These deviations of the actual reimbursements from the DRG tariffs are of course not arbitrary, but instead a systematic relaxation of the fixed reimbursement principle of the DRG mechanism, explicitly stipulated in the contractual arrangements offered by the region to its healthcare providers (more details are in Appendix A. See also Busse et al. 2013, or, for lower income countries, Mathauer

TABLE 3 | The Lombardy DRG prospective payment system in 2013.

Variable	N	Mean	St. Dev.	Min	Max
(i) Base refund for day treatment (€)	529	3,133	7,068	25	76,240
(ii) Standard refund for treatment (€)	529	5,628	9,221	428	101,344
(iii) Threshold stay for standard refund (days)	529	22.98	18.87	2	138
(iv) Refund per day beyond the threshold (€)	529	238	171	45	1,149

Note: Summary statistics for the parameters of the prospective payment system adopted by the Lombardy region in 2013 to reimburse hospitalizations completed in the region's hospitals and classified as one of the 529 DRGs in our sample.

and Wittenbecher 2013). In some DRG systems, individual negotiation on the specifics of the hospitalizations performed by a provider may lead payer and provider to agree reimbursements which differ from the tariff (Cooper et al. 2019). Such individual negotiations do not occur in the Lombardy system, where all providers face the same menu of DRG tariffs, irrespective of their location, ownership, or size: the region is a monopsonistic purchaser and all hospitals are reimbursed according to the same "take-it-or-leave-it" tariff menu. Conceptually, this rigid and highly prescriptive system of permitted deviations for the DRG tariff is intended to replace the margin for *ad hoc* negotiations for particularly difficult cases which occurs in other DRG systems. Both are instances of small steps in the direction of the cost reimbursement system, the region Lombardy's likely better suited to the large number of providers and patients, and clearly best suited to reduce the potential influence of self-interested politicians.

In practice, the Lombardy system adjusts the fixed DRG reimbursement in two conceptually distinct ways. Some DRGs are split into two, with different codes, according to whether or not there are complications, and also according to some characteristics of the patient. For example, "Other Disorders of the Eye" are classified as DRG 46 for adults with complications, as DRG 47 for adult without complications and as DRG 48 for children. Their tariff vectors are (€189.19, €2401.4, 10days, € 253.25), (€189.19, €1483.48, 15days, € 154.15), and (€202.2, €1420.42, 13days, € 165.17). The process of splitting each DRG into different typologies with different tariffs cannot, however, be pushed too far, as it has the potential to generate a very large and unwieldy increase in the length of the menu. Thus, the region adopts a second, more flexible way to make reimbursements reflect the complexity of the treatment. The "headline" tariff for a given DRG is adjusted by adding or subtracting supplementary amounts depending on the values of some observable parameters. These capture both the likely complexity of the patient's condition, attributable to her pre-existing complications, comorbidities, discharge status, and the characteristics of the particular treatment provided, such as details of the diagnosis, use of special materials, the adopted procedures, hospitalization settings, and so on. In addition, the reimbursement is reduced for procedures delivered in clearly defined, simpler "low-intensity surgery" settings, or as a reflection of the overall services provided by the hospital.[9] Divergence can occur also as a consequence of subsequent adjustments, either because the outcome falls short of a target quality, for example, if a patient is re-admitted within 40 days, indicating a possibly premature discharge, or following checks on the admission and treatment records. These also take the role of deterring and sanctioning upcoding. Section C in the online appendix elaborates on the general contractual arrangements and provides some specific detailed examples.

3 | Theoretical Background

Aside from the vast expansion of the price list it would cause, the process to contractually adjust reimbursements in order to reflect the characteristics of the hospitalization is subject to an inescapable conceptual limit. While the presence of some specified complicating factors can be determined precisely, patients also differ in characteristics, general state of health, fitness, attitude, behaviors, and so on, which a medical check-up would reveal with a reasonable degree of precision to the trained eye, but which are impossible to measure or describe in detail *ex-ante*, and hence cannot be part of the contractual algorithm which determines the reimbursement for a given treatment.

To model in a compact formal manner, these two conceptually distinct sources of variation in the cost of treating a patient, we posit in this section that the cost of treating DRG t affecting a given patient is given by $c^t(u, k)$, where u captures the set of the patient's *unobservable* characteristics, and $k = 1, \ldots, K$ is an index for a combination of values of *observable* factors affecting the cost of treatment, such as age, sex, pre-existing conditions, presence of A&E departments, provision of specific diagnostic examinations, use of specialized materials, and so on. The latter, on the other hand, can be objectively quantified, and so the contractual reimbursement schedule can explicitly include a different reimbursement for each possible values of k. To do so requires that there should be a finite, if large, number of possible values of k. This determines a full vector of adjustments to the tariff for the various DRGs so that we can define

$$p_{tk}, k = 1, \ldots, K, t \in T, \qquad (2)$$

as the overall reimbursement for DRG t comprising the "base" reimbursement and the adjustment associated to the set of characteristics $k, k = 1, \ldots, K$.[10] On the other hand, the nuanced nature of the set of unobservable characteristics captured by u suggests that it should vary in a continuum: $u \in [\underline{u}, \bar{u}] \subset \mathbb{R}$.

A further potential source of variation in the cost of providing a given DRG arises from differences among hospitals. Some are better run than others, better able to attract train and retain outstanding personnel, and so on, and are therefore able to provide at a lower cost the necessary treatment to a patient whose diagnosis has a given complexity, or, conversely, to provide the required quality of treatment at a lower cost.[11] We model this heterogeneity in a simple but plausible and intuitive manner by multiplying the "base" cost function $c^t(u,k)$ for DRG t by a hospital-DRG idiosyncratic parameter γ_t^h. The empirical analysis below posits this parameter to be known only to the hospital and to (some of) the local patients, not to the purchaser (or the analyst). To sum up, the cost of providing DRG t in hospital h for a patient with non-contractible and contractible characteristics u and k is

$$\gamma_t^h c^t(u,k). \tag{3}$$

Denoting partial derivatives with a subscript, with a slight abuse of notation in the case of k, we order u and k so that $c_u^t(\cdot) > 0$ and $c_k^t(\cdot) \geqslant 0$: higher values of u or k indicate more expensive patients to treat. We assume $c_{uu}^t(\cdot) \geqslant 0$ and $c_{uk}^t(\cdot) \geqslant 0$: the two sets of a patient's characteristics have similar effects on cost of provision. We also introduce the technical assumption that $c_{uuk}^t(\cdot) \leqslant 0$: this suggests that the rate of increase in cost is "not too high".[12]

The characterization of hospital h's cost to provide DRG t in Equation (3) has two important features. The first is the assumption we make that u and k are positively correlated. This of course cannot be tested directly, as u is not observed, but it stands to reason that individuals with unfavorable observable characteristics, such as the need for special materials or techniques, are also more likely to have unfavorable non-observable characteristics, such as poor general health, poor fitness, and negative attitudes and behaviors. Moreover, as we point out in Section 5, some of our findings provide indirect evidence for the positive sign of the correlation.

We capture this formally by positing that the distribution of u in $[\underline{u}, \overline{u}]$ depends on k. Specifically, let $\Phi_k^{ht}(u)$ be the distribution of u for provider h's potential patients needing DRG t characterized by the observable k, with $\Phi_k^{ht}(\underline{u}) = 0, \Phi_k^{ht}(\overline{u}) = 1$ and density $\phi_k^{ht}(u) = \Phi_k^{ht\prime}(u)$. The hospital superscript h indicates that different hospitals serve in general populations with different distributions of the unobservable characteristics. The positive correlation between k and u is captured formally as follows.

Assumption 1. For every $k = 2, \ldots, K$, for every $h = 1, \ldots, H$, $\Phi_{k-1}^{ht}(u) > \Phi_k^{ht}(u)$ for $u \in (\underline{u}, \overline{u})$. That is $\Phi_k^{ht}(u)$ first order stochastically dominates $\Phi_{k-1}^{ht}(u)$.

In words, when the observable characteristics indicate higher cost of treatment, then it is more likely that the unobservable characteristics also push the cost up. This assumption is illustrated in Figure 1. The diagram depicts the density function of u for two groups of patients with different levels of the observable characteristics, $k-1$ and k. To illustrate the positive correlation between k and u, consider a certain level of "difficulty" for a given DRG, the level implied by the vertical line, drawn at the abscissa $\hat{u}$. As the picture shows, for the observable characteristics k, which are worse than $k-1$, there are more patients with a value of u worse, that is higher than the threshold. Assumption 1 requires that this is true for every possible choice of the threshold $\hat{u}$.

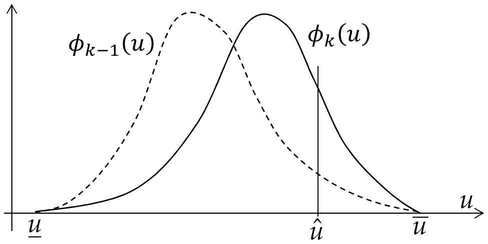

FIGURE 1 | First order stochastic dominance. *Note:* The dashed line is the density of patients with low cost observable characteristics (low k); the solid line for higher k patients. There are proportionally more patients with u below a given value among those in the low k group. [Color figure can be viewed at wileyonlinelibrary.com]

The second noteworthy feature we impose on Equation (3) is that of constant returns to scale up to a given level: provider h cannot treat more than $\varkappa_t^h$ patients for DRG t. Together with the linear role, we set for the efficiency parameter γ_t^h, this implies that hospital h's cost of providing DRG t to each of its first $\varkappa_t^h$ patients is independent of the number of patients hospital h is treating for DRG t, and is infinity beyond $\varkappa_t^h$. In practice, there is flexibility as patients can be admitted in different wards, and a more realistic constraint would take this possibility into account, and be defined at ward or even at hospital level.[13] As Equation (1) below shows, capacity constraints change the solution to a hospital's problem only in a limited way, and so we posit that provider h faces capacity constraints in every DRG. These are naturally expressed by the formal statement that for each DRG t, the total number of patients treated by provider h must not exceed the exogenously given capacity $\varkappa_t^h$:

$$\sum_{k=1}^{K} \Phi_k^{ht}\left(u_{kt}^h\right) \leqslant \varkappa_t^h, t = 1, \ldots, T. \tag{4}$$

The last institutional detail to consider is a consequence of the health authority's budget constraint, that is the upper bound to the Lombardy region overall health expenditure in each year. To remain within the budget, the region imposes on each hospital h a cap, denoted by P^h, to the total reimbursement that hospital h can receive in the year. This, which is not a standard feature in other DRG systems, in general implies that hospitals are unable to admit all their potential patients. Given this cap, each hospital, to the extent that they are able to, will choose the set of patients which allows them to minimize their costs: faced by the need to choose between two patients whose treatment would bring the same reimbursement, would clearly prefer to treat the one with the lower cost.[14] Thus, the hospitals' choice variable is the set of patients admitted for each of the DRGs. While in theory a hospital is required to treat every patient who seek admission, in practice the presence of a budget constraint forces them to try to turn some away. They may avoid being explicit, for example, by providing advice and information to GPs or even by advising patients about the suitability of their treatment facilities to the patient's specific diagnosis. Of course it may still happen that a hospital ends up treating a patient whom it would prefer to turn down, for example, to avoid bad publicity or a lawsuit. The cost of treating these "unavoidable" patients is equivalent to a reduction in the available budget, P^h, and so it requires no formal change in the model.

Proposition 1. *Let T be the set of DRGs provided by hospital h. Then hospital h treats patients characterized by parameters (u, k) if and only if $u \leq u_{kt}^h$, where u_{kt}^h is the solution to*

$$\frac{\gamma_t^h c_u^t(u,k)}{1-\lambda^h} = p_{tk} - \tau_t^h, k = 1, \ldots, K, t \in T, \tag{5}$$

and $\lambda^h \geq 0$ is a hospital specific Lagrange multiplier, whose value depends on the overall cap set by the provider, P^h, and $\tau_t^h \geq 0$ is a hospital-DRG specific Lagrange multiplier, whose value depends on the DRG specific capacity constraint, and is 0 whenever the capacity constraint is not binding. Moreover, if the capacity constraint is not binding:

$$\frac{du_{kt}^h}{d\gamma_t^h} = -\frac{c_u^t(\cdot)}{\gamma_t^h c_{uu}^t(\cdot)} < 0, k = 1, \ldots, K, t \in T, \tag{6}$$

$$\frac{du_{kt}^h}{dp_{tk}} = \frac{1-\lambda^h}{\gamma_t^h c_{uu}^t(\cdot)} > 0, k = 1, \ldots, K, t \in T, \tag{7}$$

$$\frac{du_{kt}^h}{d\lambda^h} = -\frac{c_u^t(\cdot)}{c_{uu}^t(\cdot)(1-\lambda^h)} < 0, k = 1, \ldots, K, t \in T. \tag{8}$$

Proof of Proposition 1. To lighten notation, we normalize the number of potential patients to 1 and the cost parameter correspondingly. Begin by noting that, if provider h decides to treat for DRG tn_{kt}^h patients with observable characteristics k, then it will treat those with unobservable characteristics below u_{kt}^h, where u_{kt}^h satisfies

$$n_{kt}^h = \int_{\underline{u}}^{u_{kt}^h} \phi_k^{ht}(u)\mathrm{d}u = \Phi_k^{ht}\left(u_{kt}^h\right),$$

This is because, given that provider h observes the type of each patient and knows the overall distribution of the characteristics of the patients who require the treatment during the year, it can choose to treat only those with the most favorable characteristics, viz. those with $u \leqslant u_{kt}^h$.

Thus, the problem of provider h can simply be stated as:

$$\begin{aligned} &\max_{\left\{\left\{u_{kt}^h\right\}_{k=1}^K\right\}_{t=1}^T} \sum_{t=1}^T \sum_{k=1}^K \left(p_{tk}\Phi_k^{ht}\left(u_{kt}^h\right) - \int_{\underline{u}}^{u_{kt}^h} \gamma_t^h c^t(u,k)\phi_k^{ht}(u)\mathrm{d}u\right) \\ &\text{s.t.:} \sum_{t=1}^T \sum_{k=1}^K p_{tk}\Phi_k^{ht}\left(u_{kt}^h\right) \leqslant P^h, \text{ and} (4). \end{aligned} \tag{9}$$

The solution to problem (A2) is a standard constrained maximization, whose Lagrangean is given by:

$$\begin{aligned} \mathcal{L} = &\sum_{t=1}^T \sum_{k=1}^K \left(p_{tk}\Phi_k^{ht}\left(u_{kt}^h\right) - \int_{\underline{u}}^{u_{kt}^h} \gamma_t^h c^t(u,k)\phi_k^{ht}(u)\mathrm{d}u\right) \\ &- \lambda^h\left(\sum_{t=1}^T \sum_{k=1}^K p_{tk}\Phi_k^{ht}\left(u_{kt}^h\right) - P^h\right) \\ &- \sum_{t=1}^T \tau_t^h\left(\sum_{k=1}^K \Phi_k^{ht}\left(u_{kt}^h\right) - \kappa_t^h\right), \end{aligned} \tag{10}$$

with the complementarity slackness constraints, $\tau_t^h\left(\sum_{k=1}^K \Phi_k^{ht}(u_{kt}^h) - \kappa_t^h\right) = 0$ for $t = 1, \ldots, T$. Differentiation of Equation (10) gives:

$$\begin{aligned} \frac{\partial\mathcal{L}}{\partial u_{kt}^h} = &\, p_{tk}\phi_k^{ht}\left(u_{kt}^h\right) - \gamma_t^h c^t\left(u_{kt}^h, k\right)\phi_k^{ht}\left(u_{kt}^h\right) - \lambda^h p_{tk}\phi_k^{ht}\left(u_{kt}^h\right) \\ &- \tau_t^h\phi_k^{ht}\left(u_{kt}^h\right) = 0 \end{aligned}$$

In the case of a DRG t with a non-binding capacity constraint, $\tau_t^h = 0$ and Equation (5) follows immediately, after some straightforward rearranging. When instead the capacity constraint is binding, the solution is such that

$$\sum_{k=1}^K \Phi_k^{ht}\left(u_{kt}^h\right) - \kappa_t^h = 0,$$

which give rise to a sub-problem where the capacity is allocated to the various values of $k = 1, \ldots, K$, with a shadow tariff of $p_{tk} - \tau_t^h$,

$$\begin{aligned} &\max_{\left\{u_{kt}^h\right\}_{k=1}^K} \sum_{k=1}^K \left(\left(p_{tk} - \tau_t^h\right)\Phi_k^{ht}\left(u_{kt}^h\right) - \int_{\underline{u}}^{u_{kt}^h} \gamma_t^h c^t(u,k)\phi_k^{ht}(u)\mathrm{d}u\right) \\ &\text{s.t.:} \sum_{k=1}^K \Phi_k^{ht}\left(u_{kt}^h\right) = \kappa_t^h. \end{aligned}$$

This establishes the first part. Next, note the total differentiation of Equation (5) yields

$$\frac{c_u^t(\cdot)}{1-\lambda^h}\mathrm{d}\gamma_t^h + \frac{\gamma_t^h c_u^t(\cdot)}{(1-\lambda^h)^2}\mathrm{d}\lambda^h + \frac{\gamma_t^h c_{uu}^t(\cdot)}{1-\lambda^h}\mathrm{d}u_{kt}^h = \mathrm{dp}_{\mathrm{tk}}.$$

Rearranging the above yields Equation (6)–(8) and completes the proof. $\square$

The proofs of this and all the other formal results in this section is in the Appendix.[15] The comparative statics effects in (6)–(8) are intuitive: the optimal number of patients in a given DRG is reduced by an increase in the cost of treating that DRG, (6), by a tightening of a capacity constraint for that DRG (this follows from the positive relationship between τ_t^h and κ_t^h), and is increased by an increase in the reimbursement, (7), and by an increase in the total payment for the hospital, again from the relationship between λ^h and P^h, which is instead negative. Proposition 1 highlights the interdependence between DRGs: the number of patients admitted for DRG t depends of the overall distribution of u in all other DRGs $t' \neq t$. This shifts admission decisions up the hierarchy, from the DRG/ward level to the whole hospital, and makes the problem effectively a static one, the determination of the entire vector of "thresholds" of seriousness for all the DRGs it provides. This can be thought of as being done by the hospital management at the start of the year, based on the menu of tariffs set by the health authority: these thresholds are then implemented at DRG level. Random fluctuations in demand for the various DRGs compensate in expectation. Given that γ_t^h is hospital h's cost parameter for DRG t, Equation (6) in Proposition 1 says that a hospital which has a lower unit cost from treating a given DRG will at the optimum treat more patients and more difficult patients.

This interpretation of Proposition 1 informs our econometric approach: the unobserved cost parameter for each hospital–DRG is correlated both with the observable number of patients treated and with the average complexity of those patients, and through this with the average reimbursement to that hospital for that DRG. The rest of the analysis in this section formally establishes this spurious correlation between the number of patients and the average reimbursement which we then determine in our empirical analysis.

Corollary 1. *Let N_t^h be the total number of patients treated for DRG t in hospital h, and let $\overline{p}_t^h$ be the average reimbursement to hospital h for the treatments it performs under DRG t. We have*

$$N_t^h = \sum_{k=1}^{K} \Phi_k^{ht}\left(u_{kt}^h\right), \tag{11}$$

$$\overline{p}_t^h = \frac{\sum_{k=1}^{K} p_{tk}\, \Phi_k^{ht}\left(u_{kt}^h\right)}{\sum_{k=1}^{K} \Phi_k^{ht}\left(u_{kt}^h\right)}. \tag{12}$$

The health authority maximizes the *gain* in utility of its citizens,[16] subject to an exogenously given budget constraint $\mathcal{P} > 0$.[17] Formally, denoting by $w_t(u, k)$ the utility gain obtained by a patient with characteristics (u, k) for receiving the treatment appropriate to DRG t, and by $u_{kt}^h(P^h)$ the cut-off chosen by hospital h for patients with observable characteristics k who seek DRG t when faced by an overall budget constraint P^h and prices $\left\{\{p_{tk}\}_{k=1}^{K}\right\}_{t\in T}$, the grand problem of the health authority is given by

$$\begin{aligned} &\max_{\left\{\{p_{tk}\}_{k=1}^{K}\right\}_{t\in T},\{P^h\}_{h=1}^{H}} \sum_{h=1}^{H}\sum_{t\in T}\sum_{k=1}^{K}\int_{\underline{u}}^{u_{kt}^h(P^h)} w_t(u,k)\phi_k^{ht}(u) \\ &\quad \mathrm{d}u, \\ &\text{s.t} : \sum_{h=1}^{H}\sum_{t\in T}\sum_{k=1}^{K}\Phi_k^{ht}\left(u_{kt}^h(P^h)\right)p_{tk} \leqslant \mathcal{P}. \end{aligned} \tag{13}$$

In solving problem (13), the health authority will take into account that some patients may wish to choose their hospital, and that, as noted above, some hospitals are unable to turn down patients whom they would prefer not to treat.[18] The solution to the full health authority problem (13) is outside the scope of this paper. Instead, we *assume* that socially optimal reimbursements exist, and then focus our analysis on some of the features that the schedule of these reimbursements must satisfy, and explore the consequences of these features for the choice made by the hospitals.

We can now state the main result of this section (its proof is in the Appendix).

Proposition 2. *Let $K > 1$. (i) The number of patients N_t^h decreases with γ, and (ii) for suitably chosen reimbursements satisfying*

$$p_{tk} > p_{t,k-1}, k = 2, \dots, K, t \in T, \tag{14}$$

the average reimbursement $\overline{p}_t^h$ decreases with γ.

The requirement that the reimbursement vector satisfy Equation (14) is that, at the optimal solution, the purchaser chooses reimbursements for each DRG which increase in a very precise way with the severity of the observed characteristics k. While Proposition 2 is very general, it does not attempt to establish existence, let alone uniqueness. Existence cannot be taken for granted, given that the ex-post allocation of patients to hospitals is endogenously determined, as patients turned down by one hospital will seek admission to different hospitals, altering its distribution, and therefore so are the functions $\Phi_k^{ht}(u)$ which determine hospital h's equilibrium demand for DRG t. What the proposition establishes is that whenever a reimbursement vector exists that satisfies the first-order conditions, then it must also satisfy Equation (14). Moreover, Equation (14) is necessary but not sufficient: the mere observation of reimbursements satisfying Equation (14) does not imply optimality. Instead, reimbursements need to be finely tuned to the demand and supply conditions to ensure the appropriate threshold for each DRG in each hospital: reimbursements must be high enough to induce providers with a low value of the cost parameter γ_t^h to accept "expensive" patients, while at the same time low enough to dissuade higher γ_t^h providers from pretending to have a lower γ_t^h instead and admitting the "expensive" patients. Given the correlation between the idiosyncratic parameters k and u, the average reimbursement of a hospital with lower cost for this DRG will be higher than that one with higher cost. A simple example may illustrate why. Consider a single DRG. Suppose that some patients have multiple sclerosis (MS), which is observable and contractible, and be in good or bad general health, which is non-contractible. Let p^H be the reimbursement for treating a patient with MS, and let and p^L be that for a patient who does not have MS. Finally, let c_{YX}^H (c_{YX}^L) be the cost of treating a patient who has (has not) MS, respectively, incurred by a hospital of type X, where $X \in \{E, I\}$ indicates a low cost, "efficient" and a high cost, "inefficient" hospital, for treating a patient of general health Y, where Y is either G (good), or B (bad).

Suppose the socially optimal allocation of patients is for the high cost hospitals to treat only patients in good general health, and for the low cost hospitals to treat any patient. This allocation results if the reimbursements p^L and p^H are such that

$$p^L > c_{GI}^L, \tag{15}$$

$$c_{BI}^L > p^H > c_{BE}^H. \tag{16}$$

We can now verify that Proposition 2 holds. Let π^H (π^L) be the probability of a patient having good general health if she has (has not) MS, respectively, and let the proportions of patients with MS be α. Finally, suppose that if the patients were randomly allocated each type of hospital would have one half of the patients. Then, if high cost hospitals are able to refuse treatment to patients whose cost of treatment is higher than the reimbursement, which is the case if Equations (15) and (16) hold, the low cost hospitals treat $\frac{1-\alpha}{2}\pi^L + \frac{\alpha}{2}\pi^H +$

$(1-\alpha)(1-\pi^L)+\alpha(1-\pi^H)$ patients: their share of the patients with good general health plus all the patients with poor general health, who are turned away by the high cost hospitals. This is more than $\frac{1-\alpha}{2}\pi^L + \frac{\alpha}{2}\pi^H$, the number of patients admitted by the high cost hospitals, in line with part (i) of Proposition 2.

Regarding the average reimbursement, part (ii), the low cost hospitals' is higher if

$$\frac{\frac{1-\alpha}{2}\pi^L p^L + \frac{\alpha}{2}\pi^H p^H + (1-\alpha)(1-\pi^L)p^L + \alpha(1-\pi^H)p^H}{\frac{1-\alpha}{2}\pi^L + \frac{\alpha}{2}\pi^H + (1-\alpha)(1-\pi^L) + \alpha(1-\pi^H)} > \frac{\frac{1-\alpha}{2}\pi^L p^L + \frac{\alpha}{2}\pi^H p^H}{\frac{1-\alpha}{2}\pi^L + \frac{\alpha}{2}\pi^H},$$

which is true as long as $(\pi^L - \pi^H)(p^H - p^L) > 0$: this is the case if $\pi^L > \pi^H$, that is if having MS makes it less likely to be in good general health, which is the case if there is positive correlation between these characteristics.

A simple noteworthy consequence of Proposition 1 is that an increase in the budget of a hospital leads to a *lower* average reimbursement for all its DRGs.

Corollary 2. *At the solution of problem (A2), an increase in* P^h *determines a reduction in* $\overline{p}_t^h$*, for every* $k = 1, \ldots, K$.

4 | Econometric Strategy

4.1 | The Estimated Model

Our empirical strategy is based on the theoretical analysis of Section 3. We compare the matching of patients and hospitals that actually happened in 2013, with the one that would have occurred if patients simply went to their nearest hospital. This choice is viewed as the default in the literature on patients' choice and hospital competition (Kessler and McClellan 2000; LeGrand 2009; Brekke et al. 2011; Moscelli et al. 2016, 2021, among many others) when patients have no information on the service they would receive at the available hospitals and if hospitals were unable to select patients. In this default case, differences in the average reimbursement for a given DRG would be due only to differences in the characteristics of their local patients, which, given the relatively large and homogenous population, are likely to be random.

On the contrary, the theoretical model in Section 3 predicts that, if patients (or their GPs, as in Beckert 2018), have informed opinions on their local hospitals, they are more likely to eschew them if they think they will not be adequately treated, and if the region is choosing a socially optimal contract, we should observe correlation between the average reimbursement received by the number of patients treated by a hospital for a given DRG. To the extent that a more complex patient (i) is more likely to be keen to choose hospitals which can provide adequate treatment, because they have a steeper trade-off between quality of the care and convenience of location, and (ii) a hospital less capable of treating this patient adequately is more likely to try to turn her away, because they are less likely able to cover the cost of the hospitalization, then the Proposition 2 above shows that the allocation of patients to hospitals is not random: instead more complex hospitalizations for a given DRG are concentrated in some hospitals in a systematic way. Of course, we do not observe the complexity of a hospitalization, but only the measurable component of this complexity, and we assume both that complexity is positively correlated with this measurable component and that it can therefore be proxied by the reimbursement attached to its observable characteristics.

This implies that some hospitals attract *more patients and more complex patients* and receive therefore a higher average reimbursement from the health authority, provided that the reimbursement rules set by the health authority satisfies the optimality necessary condition (14), and that the observable and unobservable characteristics of a hospitalization are correlated (Assumption 1). This creates a spurious correlation between two *observable* variables, the hospitals' ability to attract patients needing treatment for that DRG and its willingness to treat them, *and* the adjustments to the reimbursement schedule determined by the idiosyncratic characteristics of a given hospitalization. The aim of this paper is to detect this correlation in the data, and thus to show that the price schedule imposed by the regional health authority is in line with the theoretical optimal contracts derived in Section 3. We do not aim to establish causality: a hospital may employ a consultant known by patients to be excellent who prefers to work there for family reasons, and therefore may attract more patients who need this consultant's DRGs. Conversely, patients in a certain area may be more affected by a pathology needing a certain DRG, causing the local hospital to invest more resources in this DRG. Our data, a 1 year cross-section snapshot, do not permit a robust causal identification such as one exploiting a credible instrumental variable suggested by a quasi-experiment due to a random or policy variation. Conversely, the very broad nature of our data, which covers all DRGs, is unsuitable to study crucial details of a specific treatment, along the lines, for example, of Chandra and Staiger's 2020 comparison of the effects of intensive and non-intensive care in heart attacks. Nevertheless, establishing correlation is the first step toward a causal analysis while more suitable data become available.

Toward this aim, we fit this simple cross-section OLS regression to the administrative data set described in Section 2:

$$R_{ith} = \rho A_{th} + \beta X_{ith} + F_{ith} + \varepsilon_{ith}. \tag{17}$$

In (17), the dependent variable, R_{ith} is the percentage deviation of the reimbursement paid to hospital h for hospitalization i from the tariff for DRG t, defined in (1) as $R_{ith} = \frac{\text{reimbursement}_{ith} - \text{tariff}_t}{\text{tariff}_t} \times 100$.

The focus of the paper is on ρ, the coefficient of A_{th}, the first variable on the right-hand side, defined in the next subsection as the "mutual" *attractiveness* of hospital h for the patients who need DRG t. The vector X_{ith} on the RHS are controls for relevant characteristics of the hospitalization and of the patient, and F_{ith} is a set of fixed effects.

4.2 | Attractiveness

As anticipated, attractiveness A_{th} is a combined measure of how keen the local patients who need DRG t are to be treated at hospital h and how willing hospital h is to admit them. As our measure of attractiveness, we compute the *share of the patients who need DRG* t *and are local to hospital* h, *who are treated by hospital* h.

We base our measure on local patients for two reasons. In the first place, they are more likely to have reliable information on the health care available at the local hospital. Second, and conceptually more importantly, the theoretical analysis of Section 3 implicitly takes as given the catchment area of the hospital. Any measure of attractiveness must depend on the size of the hospitals' pool of potential patients. After all, around four in ten patients are treated at the local hospital, and so a hospital located near many people who need a given DRG is bound to have a higher number of treatments, without necessarily being more attractive than a hospital in a sparsely populated area which attracts all of its few potential patients.

We proceed as follows. For given hospital h and DRG t, we compute N_{th}^{p}, the total number of hospitalizations for DRG t in Lombardy which were such that the relative location of the patient and hospital h made hospital h local for the patient. Of these, we ask how many were actually treated at hospital h: let N_{th}^{c} the number of those who were. Then hospital h's attractiveness for DRG t is simply the ratio between these:

$$A_{th} = \frac{N_{th}^{c}}{N_{th}^{p}}. \tag{18}$$

While the definition of local hospital is straightforward in theory, it may lead to measurement error if not dealt with carefully. The simplest and most intuitive definition of a local hospital is the one nearest to the patient's location. This, however, may not always be appropriate. In the first place, we measure distance from the centroid of her location, defined above as the intersection of municipality and postcode, as we are unable to pinpoint exactly a patient's home address, and so in some cases, our imputation does not pick a patient's nearest hospital. Second, for some patients, there may be several hospitals located at similar distances. When this happens, it may be inaccurate to classify only one of these hospitals as local: it is easy to imagine situations where a large hospital is conveniently built on a town's ringroad, but is further in distance from the town's residents than a smaller hospital located in the town center. Appendix F, online, provides some specific examples. For this reason, we extend the definition of local hospital for the patient in hospitalization i to any hospital located within 150% of the distance between the patient's residence and her nearest hospital.[19] This implies, of course, that many patients will have multiple local hospitals. This is accounted for in Equation (18), by counting a patient who has n local hospitals as $1/n$-th for each of her local hospitals.[20] Because of its central role in the paper, this variable is the main focus of the robustness analysis we conduct in Section 5.2. All the different measures we report there (and all those we have consider but do not report) give qualitatively similar results.

4.3 | Controls

The vector X_{ith} contains the other covariates we control for. At the hospitalization level, we include the distance of the patient's location from the hospital, their gender and age, squared to account for non-linearities due to young and elderly patients being potentially more complex, whether the patient was born in their municipality of residence, which may indicate different ability or opportunity to move to preferable locations, or different access to information, a categorical indicator for their marital status, as it may affect health (see Kalmijn 2017, for a recent survey) or ability to travel. Moscelli et al. (2021) use the extent of patient choice as a measure of competition: thus, we control for the number of hospitals which are defined as "local" for the patient in hospitalization i. We also include the full set of the patient's comorbidities,[21] and a dummy indicating whether the patient was transferred between wards in the course of the hospitalization. We include the length of their hospital stay. To account for the very wide differences between the standard stay for different procedures shown in Table 3, we normalize the length as the ratio between the number of days from admission to discharge, after the first and the standard treatment threshold for that DRG.[22]

Just above 9% of the hospitalizations are for patients who reside in a different Italian region, or abroad.[23] For this reason, we include among the controls the percentage of the hospitalizations for hospital h and DRG t which are carried out on patients whose residence is outside the region. We find that it does not affect the average reimbursement and hence neither the patient mix.[24]

Table 2 above reports summary statistics for these controls, and other values of interest. Summary statistics for the comorbidities are in the online appendix (Table A3).

We include in the specification in Equation (17) several sets of fixed effects to control for other possible sources of correlation between reimbursement and attractiveness. Even though we have normalized both the price and the length of stay, there may well remain specific characteristics of the DRG, and so we include a DRG fixed effect. We also include a ward-hospital fixed effect, which obviously includes the hospital fixed effect, to account for unobserved heterogeneity across hospitals and wards in the same hospital. The location fixed effects account for different distribution of relevant characteristics of the patients across Lombardy. These are also captured by the GP fixed effect, which captures both the anecdotal evidence that in the larger municipalities with more than one physician, each tends to serve patients of similar location and social class, and GPs' possible different propensities to refer patients. Together, these two variables afford considerable granularity, given that we have 8,877 different physicians and 1,544 municipalities.[25] Finally, we include fixed effects for the month when admission took place: apart from holiday periods, and the seasonality of the admission patterns, it may also account for potentially different admission policies by hospitals as the end of the financial year approaches and the tightness of the budget constraint P^{h} described after Assumption 1 in Section 3 becomes better known to the hospital.

We end this section by describing our sample. From the universe of about 1.5 million admissions and discharges, we drop all the hospitalizations excluded from the DRG reimbursement mechanism, such as those for palliative care or rehabilitations, and those classified as "day-hospital". We also exclude all the hospitaliszations where the patient's residence is outside Lombardy, or not recorded, and, to avoid births and neo-natal interventions, those where the patient's recorded age is 0. This reduces the sample to $990,777$ observations. In our main regression, we also exclude the $110,066$ emergency admissions, as it seems plausible to assume that neither the hospital not the patient have much say in the admission,[26] and keep only those that are reimbursed on the standard portion of the tariff: more than 1 day, but below the threshold. This leaves $666,831$ hospitalizations. We use some of the excluded observations in robustness and placebo tests.

5 | Results

5.1 | Attractiveness Matters

Our main results are reported in Table 4. We begin, in the first column, with a skeleton regression: the reimbursement given by the Lombardy region for hospitalization i is regressed only on the attractiveness of the hospital where the hospitalization took place, the distance of the patient's residence to the hospital, and DRG fixed effects. Results change, Column (2), when fixed effects for the hospital-ward combinations are added. This includes hospital fixed effects, and at 1329 different values, is a finer partition than the 152 hospitals. Distance matters, but attractiveness does not. The role of these variables is reversed, raw distance disappearing, and attractiveness playing a strong role, when we include the covariates which may be correlated with the observable characteristics k: in column (3), we add the patients and DRG covariates, including the comorbidities dummies. In column (4), our main regression, we add the demographic fixed effects, determined as the combination of the fixed effect for the referring physician and the municipality of residence as explained at the end of Section 4. This is the specification we keep in the rest of the table and in the robustness tests reported in Table 5 in Section 5.2.

The positive sign of the attractiveness coefficient A is our main result. It confirms that hospitals which are "better" at attracting and admitting local patients are reimbursed more on average. In the spirit of theoretical analysis in Section 3, we can state that the menu of tariffs used to alter the fixed DRG reimbursement does respect the optimality requirement we derived in Proposition 2. This requirement is that observable conditions which make a patient's treatment more complex increase the reimbursement for the treatment of that patient in a finely tuned fashion, specifically one that induces the hospitals which are better able to provide more complex treatments to admit these more expensive patients, while at the same time deterring other hospitals to admit them.

To get a handle on the size of the effect of attractiveness, consider the average patient treated for the average DRG. Suppose this DRG becomes one standard deviation more attractive at the Poppleton University Hospital (PUH). Then the patient mix of this DRG at PUH will change. The average reimbursement for the treatment received by PUH for this DRG would increase by the product of the estimated coefficient (1.8109, from Table 4, Column 4) times the standard deviation of the attractiveness (0.228, from the fourth column in Table 2). This product is 0.4125. This is, an increase in average reimbursement equivalent to the average patient admitted for this DRG at PUH being "replaced" with one whose hospitalizsation requires a reimbursement 8.025% more expensive.[27] This is a large effect, underlining the considerable role played by the contractual adjustments to the principle of the DRG system that all treatments for a given DRG are reimbursed at the same rate.

In the next two columns, we attempt to tease out the role of the refinements to the DRG system introduced by the Lombardy region. In Column (5), we shorten the menu by bunching into one the two separate DRGs which form the 118 DRG pairs of a diagnosis without and with complications. We expect this to amplify the effect of attractiveness, as patients with complicating factors would be more likely both to be more informed on hospitals' relative quality, and to be more willing to seek admission in the "better" ones. This is what we find. In view of the central role of the concept of attractiveness, it is essential to convince the readers that it is robust to even radically different conceptual approaches to its construction. We report numerous robustness tests in Table 5, and we begin in Column (6) in Table 4, where the regression is the same as Column (4), with the measure of attractiveness constructed as follows. We begin by estimating a multinomial logit (MNL) choice model of patient hospital selection, with the patient's choice set including all the region's hospitals provide the DRG she needs. The patient's utility depends on travel distance as well as a full set of hospital-specific dummy variables. The estimated, DRG-specific, hospital fixed effects therefore capture the patients' average willingness to incur the hospital-specific travel cost required to be treated at each given hospital, given their home location. We are thus justified in interpreting the estimated fixed effects as a revealed-preference measure of hospital attractiveness and we use them as alternative measures of attractiveness. The correlation table in footnote 32 indicates that this measure is not highly correlated to the measures of (18) used in the main regression. We normalized it to make it have the same range as the latter. The estimates in Column (6) show that results are robust even to this alternative measure of attractiveness. The higher coefficient is likely due to the difference in distribution.

In the last regression of the table, reported in Column (7) we test whether our results are robust to a substantial extension the set of covariates. This column includes the full set of ICD-9-CM fixed effects: of the approximately 13,000 different codes in the ICD-9-CM system, the data in our main regression sample include 6,014 for the main diagnosis, 2517 for the main procedures, and 1123 identifying diagnosis which complicate the hospitalization. Since many of the adjustments to the base DRG are due to specific diagnoses and defined procedures reported via these codes, we would expect that the explanatory power of our measure of attractiveness should be reduced, which it does. This remains significant despite this extremely precise refinement of the set of control variables given by the inclusion of these fixed effects. We take this as a strong indication of the importance of the non-contractible characteristics, captured by

TABLE 4 | Deviation of reimbursements from Tariff and Hospital attractiveness.

Variable	(1) No controls	(2) DRG & Hosp.	(3) Indiv. Controls	(4) Main	(5) DRG No Compl.	(6) MNL- FE	(7) ICD-9
Attractiveness	2.4186***	0.8866	1.8007***	1.8093***	2.3479***	8.2682***	0.7358**
	0.407	0.572	0.564	0.565	0.636	2.537	0.29
Distance to hosp	−0.3918***	−0.1272***	−0.052	−0.0497	−0.0346	−0.0817**	−0.0239
	0.04	0.027	0.033	0.033	0.039	0.032	0.029
Competition		0.0161*	0.0105	0.0085	0.0204	0.0089	−0.0024
		0.009	0.011	0.011	0.013	0.011	0.009
Female			0.2980***	0.3193***	0.3938***	0.3281***	0.2467***
			0.06	0.059	0.066	0.059	0.053
Age			0.0280***	0.0328***	0.0293**	0.0337***	0.0392***
			0.01	0.01	0.011	0.010	0.009
Age^2			−0.1046	−0.1374*	−0.1164	−0.1450*	−0.1726**
			0.081	0.082	0.091	0.081	0.074
Born in municipality			−0.0439	−0.035	−0.0652	−0.0352	−0.013
			0.056	0.055	0.064	0.055	0.052
Never married			−0.1625**	−0.1246*	−0.1261	−0.1280*	−0.0796
			0.069	0.069	0.08	0.069	0.064
Length of stay			0.2165***	0.2161***	0.2531***	0.2173***	0.2073***
			0.007	0.007	0.008	0.006	0.006
Transfers			0.8892***	0.8885***	1.1969***	0.8541***	0.5719***
			0.187	0.187	0.234	0.186	0.16
% Out-of-region		−1.8291	−0.9513	−0.8437	−1.2844	−1.0645	−1.2328
		1.242	1.237	1.234	1.53	1.168	0.759
Fixed Effects Included							
DRG	✓	✓	✓	✓	✓	✓	✓
Hospital-ward		✓	✓	✓	✓	✓	✓
Admission Month			✓	✓	✓	✓	✓
Municipality			✓	✓	✓	✓	✓
GP				✓	✓	✓	✓
ICD-9							✓
Observations	666,841	655,864	655,864	655,606	655,606	663,137	653,562
R-squared	0.212	0.262	0.330	0.340	0.328	0.340	0.432

Note: Attractiveness is measured as the proportion of local patients who are treated in the hospital, where local means with 1.5 of the distance of the nearest hospital to the patient's location. In Column (1), Equation (17) is estimated controlling for the distance to the hospital, and DRG fixed effects. Column (2) adds competition and the hospital-DRG percentage of patients that reside outside Lombardy as controls as well as hospital fixed effects. Column (3) adds patients' characteristics such as gender, age, whether the patient reside in the same municipality they were born, whether they were ever married, the length of the patient's stay in hospital, normalized by the maximum length allowed by the contract, whether they were transferred to a different ward, the comorbidity dummies, municipality of residence dummies and the month of the year. Column (4), our preferred specification, adds the GP fixed effects. Column (5) shortens the menu by bunching into one the two separate DRGs pairs of without and with complications. Column (6) uses the hospital fixed effects as a measure of the hospitals' attractiveness and Column (7) includes the full set of ICD-9 fixed effects. In all specifications errors are clustered at the DRG–hospital pair.
***$p < 0.01$; **$p < 0.05$; *$p < 0.1$.

u in the theoretical model of Section 3, which lends support to our assumption of a positive correlation between observable and unobservable characteristics of the hospitalization, k and u.

The coefficients for most of the other covariates are plausible, and consistent across column (4)–(7). The distance between the patient and a hospital explains the reimbursement only when it stands in for omitted variables. It is interesting to note that omitting attractiveness from column (4) does not change the coefficient for distance: this is −0.0636*, with a standard error of 0.033. It is probably premature to interpret this as a prevalence of hospitals' over patients' choice, and further research is required to determine a causal direction in the role of our attractiveness variable in the allocation of patients. Older and female patients appear to determine a higher reimbursement, but living close to the place of birth does not. Patients who are not married are reimbursed less,[28] although omitting this categorical variable simply increases the coefficient for age,

TABLE 5 | Robustness checks and heterogeneity analysis.

	(1)	(2)	(3)	(4)	(5)	(6)
	Travel Time	Nearest	Near/Town	1.4 of near	10^{th} percentile	MNL
Attractiveness	2.0684*	1.0323**	0.8755**	1.7204*	4.4723*	1.1254*
	0.551	0.445	0.434	0.560	1.191	0.231
Observations	657,762	600,238	609,797	651,028	662,068	663,137
R-squared	0.339	0.341	0.341	0.339	0.339	0.34
	(7)	**(8)**	**(9)**	**(10)**	**(11)**	**(12)**
	ln(reimburs.)	25% of 50^{th}	50% of 25^{th}	5% of 90^{th}	Teaching	Non-teaching
Attractiveness	0.0620*	1.8070*	1.7747*	1.7119*	−0.4798	5.8600*
	0.011	0.584	0.580	0.556	0.958	1.351
Observations	662,068	650,805	651,504	651,349	165,556	495,636
R-squared	0.873	0.339	0.339	0.339	0.376	0.345
	(13)	**(14)**	**(15)**	**(16)**	**(17)**	**(18)**
	Public	For Profit	Non-profit	Daily	Over thresh	Emergency
Attractiveness	3.9555*	5.3551*	1.3027	−0.6481**	−0.1571	0.5735
	1.472	1.542	2.415	0.287	0.306	0.479
Observations	490,249	110,091	59,704	133,572	50,585	85,809
R-squared	0.337	0.416	0.451	0.969	0.902	0.366

Note: In all columns, Robustness tests, all specifications are identical in structure to Column (4) in Table 4 with only the coefficient for Attractiveness reported; the full results are in the online Appendix. In Column (1) distance is measured in time of travel, while in Columns (2) and (3) different definitions of "local hospital" are used: only the nearest hospital is considered local (2) or if it is the nearest, or in the patient's town (3). In Column (4) a hospital is considered local if its location is within 1.4 times the distance to the nearest. In Column (5) attractiveness is measured with the MNL model described in the text, and Column (6) considers local all hospitals which are in the top decile of the distance ranking. Column (7) replaces the outcome variable in (16), the percentage deviation from the tariff with the log of the price in euros. In Columns (8)-(10), we exclude hospital–DRG pairs with very few hospitalizations. Heterogeneity is explored in Columns (11) and (12) which split the sample in teaching and non-teaching hospitals, and in Columns (13)-(15), where it is divided according to hospital ownership, state-owned, private for profit and not-for-profit. Estimation of placebo regressions is in the remaining columns: the samples are the hospitalizations requiring less than one full day stay in the hospital (Column (13)), those that required longer than the DRG specific threshold (Column (14)), and the emergency admissions (Column (15)). In all specifications errors are clustered at the DRG–hospital pair.

*$p < 0.01$; **$p < 0.05$; ***$p < 0.1$.

without otherwise affecting the rest of the regression: this would suggest that "Never married" simply captures some of the coefficient for age. Reimbursement increases with the length of the hospital stay: recall that this is due only to the greater complexity of hospitalizations requiring a longer stay, as the sample for this regression is the hospitalizations reimbursed according to the standard portion of the tariff, which is contractually independent of the length of the hospital stay.

The positive and statistically significant coefficients of the length of stay and of the dummy indicating transfers between wards indicate that these are positively correlated with the unobservable component of the complexity of the treatment, lending further support to our assumption of a positive correlation between k and u, which form the basis for Proposition 2. Neither our proxy for competition nor the percentage of hospitalizations carried out on patients residents outside Lombardy have any significant effect on reimbursements. The estimated coefficients for the comorbidities are in Table A3.

5.2 | Robustness

Table 5 reports the value of the coefficient for attractiveness for the main regression, Column (4) in Table 4, in a number of other specifications or sub-samples: the full sets of coefficients are in Tables A4–A6 in the online appendix. The table is conceptually divided into three parts. The first, Columns (1)–(10), are robustness tests; the second part, Columns (11)–(15), investigates heterogeneity, and the last three columns are placebo tests. The table begins by considering different definitions of local hospitals. In Column (1), the distance between patients' residence and hospitals is measured by traveling time rather than distance: the coefficient is very similar, if anything slightly stronger.[29] In view of the potential for the cost of travel to correlate negatively with the seriousness of a patient condition, it is encouraging that it does not matter whether travel cost is computed as distance or as time. In the second column, every patients has only one local hospital, the nearest, rather than all those within 150% of the distance of the nearest. In Column (3), we add an administrative layer to the physical distance. Many hospitals are concentrated in cities, where the local transport network may be a more appropriate indication of the convenience of reaching hospitals from one's residence: we therefore define a hospital as local either if it is the nearest, or if it is in the patient's location. In Column (4), we consider local for a given DRG a hospital located within 140% of the nearest hospital providing that DRG. This reduces the number of local hospitals each patient has. In Column (5), we account for the possibility that rather than those closer than a given distance, patients may consider local a given number of hospitals. That is, we ranked all hospitals in the region that offer DRG t by distance to the patient's residence, and defined local to a patient all hospitals which are in the first decile of this ranking, rounded up: if there are N hospitals offering DRG t, patients consider local all the hospitals no farther than the hospital ranked $\lceil \frac{10}{N} \times 100 \rceil$.[30]

In the first five columns of Table 5, as well as in main specification, a hospital is considered either local or not local. All local hospitals are then deemed to be chosen with equal probability by each of its local patients. That is, if a patient has n local hospital, she chooses each with probability $\frac{1}{n}$. In Column (6), we estimate the model with a different measure of attractiveness, also based on a multi-nominal logit model, but different from the one we adopted in Column (6) of Table 4. Here, we posit a gradual decline of the probability of a certain patient going to a hospital as the distance from her residence varies. Formally, for each DRG, we estimate a separate multi-nominal logit model, with the chosen hospital as the dependent variable and distance from the hospital as the sole explanatory variable.[31] We then use the estimated coefficients to compute predicted choice probabilities for each patient, and rank the hospitals in their choice set from highest to lowest probability. Using these to compute the expected number of patients would, by construction, yield the actual number of patients admitted for each hospital–DRG pair. So we assume that if a hospital predicted probability is very different from the one just above it in the ranking, then this hospital is not considered: it, and all those below it in the ranking, are selected with zero probability. We put the threshold at double: so if hospital h has probability 0.18 of being chosen by patient i for DRG t, and hospital h', the next in patient i's estimated ranking has a probability lower than 0.09, then hospital h', and those below it in the ranking are chosen with zero probability. Probabilities for the remaining hospitals are then scaled up to add up to 1. From these adjusted probabilities, we compute the expected and the actual number of local patients who are treated in a given hospital, that is, N^c_{th} and N^p_{th} in (18). We finally use these to construct the measure of attractiveness used in Column (6).

In all these cases, the results are qualitatively unchanged, both for the measure of attractiveness, Table 5, and for the controls, Table A4.[32] The robustness continues in Column (7) of the table, where the reimbursement is measured in euros, logged, rather than deviation from the tariff given in (17), so the LHS variable is $\log(\text{reimbursement}_{ith})$. The results change little except that the R^2 is much higher: the reason is that much of the variability in the reimbursement is due to the DRG which determines it. This is a subtle point that underlines the importance of measuring the reimbursement variable appropriately and is worth a simple example to explain. Imagine that all hospitalizations are reimbursed exactly according to the tariff, with just a minuscule homoschedastic random error in each reimbursement. When the LHS variable is $\log(\text{reimbursement}_{ith})$, the R^2 would clearly be very close to 1, since essentially all the variation in reimbursements would be due to differences in the DRG tariff, and so fully accounted for by the DRG fixed effects. By the same token, when the dependent variable is the percentage deviation from the tariff all the variation is due to the random error, and the R^2 would be effectively 0. As it happens, R^2 when the DRG fixed effects are the only explanatory variables is 0.85 when the LHS variable is $\log(\text{reimbursement}_{ith})$, and 0.21 when it is the percentage deviation R_{ith}. This systematic variation suggests that DRGs do differ in their propensity to require deviations from the tariff and is confirmed in the separate regressions we run for the 230 most treated DRGs. Figure 2, shows the coefficients obtained from these regressions, and online Appendix B contains more details on this figure and further investigations of the differences among DRGs, including evidence that the DRGs driving the overall effect tend to be those delivered more frequently.

Our final set of robustness tests is in Columns (8)–(10). The data suggest that some DRGs are only occasionally performed in some hospitals. This may happen because some patients are discharged from a hospital for a DRG that the hospital does not normally perform.[33] These "outlier" observations may end up having a disproportionate importance in the data. The reason is that they affect the measurement error not just of that observation, but they may distort the computation of the attractiveness of many other hospitals offering that DRG.[34] This is so because attractiveness is built from the concept of nearest local hospital providing the DRG: the inclusion of an erroneous hospital providing the DRG, may "deprive" some other hospitals of their status as the nearest hospital or remove them from the 150% of the nearest hospital radius, and thus affect their being local of these hospital and hence the calculation of their attractiveness. Appendix F illustrates a specific case of this possibility. To confirm that these small DRG–hospital pairs are not driving the results, in Columns (16)-(18) in Table 5 we therefore carry out robustness tests by excluding hospitals that perform "few" DRGs: few is clearly a relative concept, as there is huge variability in the number of hospitalizations in the region in the various DRGs, as shown in Table 1. Thus in Column (16)–(18), we first rank hospitals in each DRG by the number of hospitalizations carried out in the year. Then we exclude from the regression sample all the "small" hospital–DRG pairs. Column (16)–(18) differ in the definition of "small": in Column (16) we exclude hospitals which carried out less than one quarter of the number of hospitalizations carried out by the median hospital in the set of hospitals which performed that DRG. In Column (17), those that carried out less than half the number of hospitalizations carried out by the median hospital for that DRG, and in Column (18) the hospital excluded are those that carried fewer than 5% of the number of hospitalizations carried out by the hospital at the 90-th percentile for that DRG. The effect of excluding these observations is very small on the overall data set, only about 1% of the hospitalizations are excluded, but given that the hospital performing them perform very few of those hospitalizations, the effect on the number of hospital–DRG pairs which are excluded from the regression is vastly disproportionate, around 10% of the total (see Table A6 in the online appendix). It is indicative of the robustness of our conclusions that the estimated coefficients are changed only marginally by these exclusions, as Columns (8)–(10) in Table 5 show.[35]

In the second part of Table 5, we split the sample with the aim to identify any heterogeneity among hospitals. One might expect teaching hospitals to follow different admission policies, in order to fulfill their medical training and research obligations. Indeed, when we estimate separately the sample of hospitalizations carried in hospitals which are formally affiliated with a

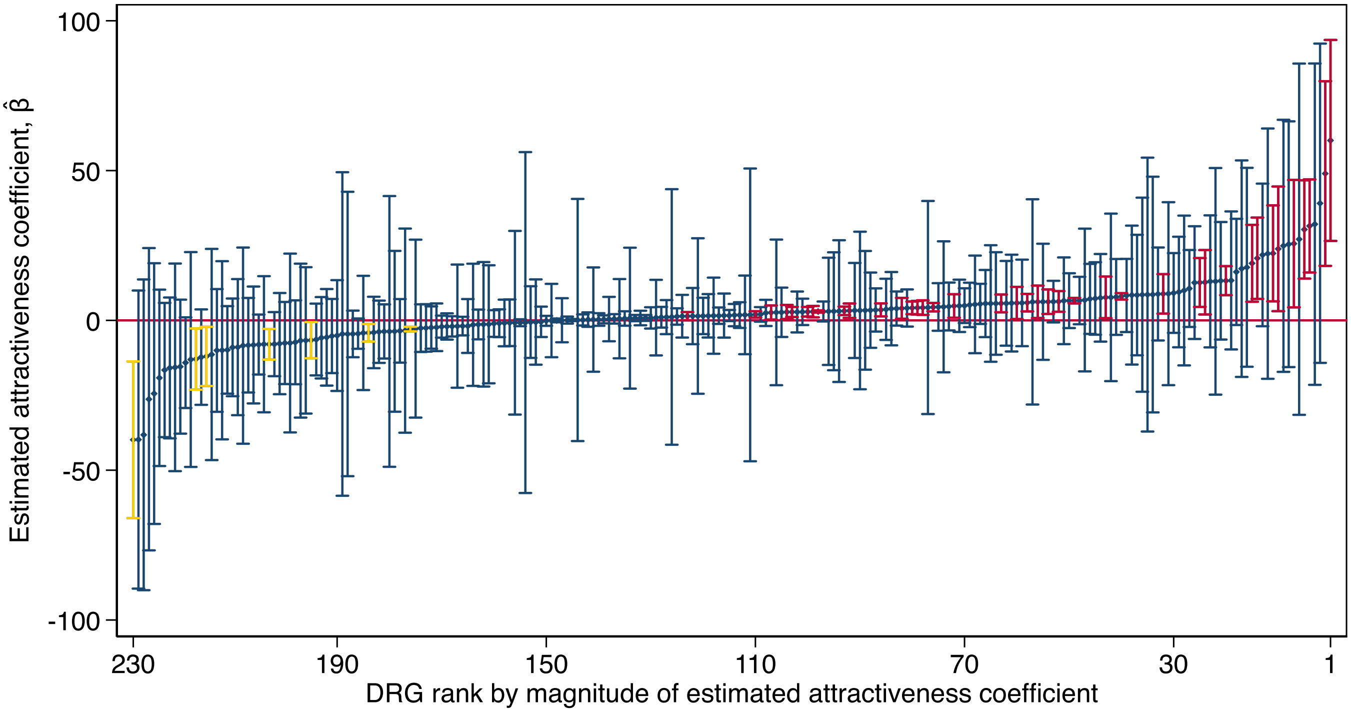

FIGURE 2 | Estimated $\hat{\beta}$ of reimbursement on attractiveness by DRG. *Note:* Estimated coefficients of attractiveness obtained by splitting the sample regressions by DRG. The horizontal axis is ordered according to the magnitude of the estimated attractiveness coefficients. We have colored in red the statistically significant and positive coefficients, in yellow those statistically significant and negative, and in blue those statistically not different from zero, at the standard 5% level. [Color figure can be viewed at wileyonlinelibrary.com]

university and those in non-teaching hospitals, we find that the attractiveness coefficient is qualitatively different in the two subsamples, see Columns (11) and (12); Table A5 shows that the coefficients for the other explanatory variables are similar.

In the lowest third, Table 5 explores the role of the hospitals' objective function, summarized in their ownership. Columns (13)–(15) report the results when the sample is split into the three types of hospitals, state-owned, private for profit, and not-for-profit. In other environments, these types of institutions have been shown to pursue different admission or discharge practices (e.g., Bayindir 2012; Eliason et al. 2018). This appears to be the case in our sample as well: public and private for profit hospitals differ, in that the attractiveness of the hospital is more strongly correlated with the reimbursement in public hospitals than in for-profit ones, and even more with non-profit. Understanding the motivations of public and private hospitals is beyond the scope of this paper, and so we limit ourselves to suggesting that this may be due to the relative concentration of for-profit providers on DRGs with certain characteristics. To the extent that these characteristics are associated with less "in-built" potential for deviation from the standard tariff, then we would expect a less precise estimation in private hospitals. This is indeed what we find in the following simple regression:

$$\begin{aligned}\text{ratio}_t = {} & \underset{0.049}{-0.107^{**}} + \underset{0.036}{0.034}\ln \text{tariff}_t - \underset{0.003}{0.0092^{***}} sd_t \\ & + \underset{0.034}{0.034} N_t - \underset{0.000}{0.001^{***}} \ln N_t^h \\ & + \underset{0.000}{0.0026^{***}} \bar{a}_t - \underset{0.017}{0.0236} FEM_t - \underset{0.033}{0.0135} TR_t \\ & + \varepsilon_t .\end{aligned} \tag{19}$$

In Equation (19), the numbers below the coefficients are the standard errors, and the stars identify the statistical significance with the usual convention; ratio_t is the proportion of hospitalizations under DRG t carried out in private hospitals, and the last variable on the RHS is the total reimbursement for all the hospitalizations under DRG t, $TR_t = \sum_i \sum_h \text{reimbursement}_{iht}$. The remaining variables, tariff_t, sd_t, N_t, N_t^h, $\bar{a}_t$, and FEM_t are defined as the standard tariff, the standard deviation of the distance of the reimbursements from the tariff, the number of hospitalizations (logged), the number of hospitals offering the DRG, the average age of the patients, and the proportion of female patients, all for DRG t.[36] Further research and new data may be needed to satisfactorily investigate the extent and the possible causes of qualitative differences in behavior in private and public hospitals.

Table 5 is completed by three placebo regressions, all with the same specification as the main regression, Column (4) in Table 4. In Column (13) we consider the sample of the hospitalizations requiring a short permanence in hospital (those in the first column in Table 2), likely to be simpler. The negative sign is hard to interpret, but may be due to the fact that to compute attractiveness we excluded these hospitalizations, on the ground that patients might be less inclined to put much thought in the choice of hospital when their treatment is likely to be very straightforward. The next column studies hospitalizations exceeding the threshold (whose summary statistics are reported in the last column in Table 2). Again, the attractiveness coefficient for these is not precisely estimated: this is likely due to the small sample, and to large differences among them. Next, in column (15) the sample is made of all the hospitalizations recorded as emergency admissions. In this case, the mechanisms identified in the theoretical model should be absent: hospitals have very little scope to turn away patients in an emergency situation, and conversely, patients' ability to choose one hospital in preference to another is very limited and less paramount in these cases. And indeed, in this case, this is what we find: the preference shown by patients for their local hospitals does not affect the average patient mix, proxied by the average reimbursement.

6 | Other Possible Determinants of Reimbursements

We end the paper by considering other possible routes affecting the nature of the contracts offered by the regional authority to its hospitals. We consider two examples, the influence of political pressure and the role of local house prices, though the methodological approach can be applied to other plausible candidates. With respect to the former, one can imagine an environment where local politicians try to influence the amount of funds that hospitals in their constituency receives from the regional center. To do so, they might utilize their more precise information on the health needs of the local population to alter the minute details of the reimbursement rules, namely, the extra-payments associated with the various *k* in Section 3. For example, imagine that (i) in a given province the local politician knew that those of his constituents who need the removal of kidney stones with ultrasonic energy bursts (DRG 210) are more likely than the regional average to suffer from renal failure, (ii) that he therefore might try to convince the regional politicians and officers designing the contracts to give a higher adjustments to the tariffs for DRG 210 to hospitalizations where renal failure is a complicating factor, and, (iii) if the provincial politician is in the same party as the politicians in charge at regional level, the latter are more likely to accede to the former's request. Then, if all these hold, one would expect a higher fixed effect for DRG 210; and also, possibly, other unobserved factors which could be captured by the hospital fixed effect. In sum, for hospitals in provinces controlled by the same party as the regional government, we would expect higher hospital fixed effects, and higher DRG fixed effects for the DRGs offered by the hospital. A more direct, different, route, and not necessarily alternative to this, is that a politician's ability to induce the regional health authority to treat upcoding leniently when is suspected in their local hospital may be higher if the politician belongs to same party as the regional government. This would also increase the average reimbursement to hospitals in the politician's constituency.

In Figure 3, we depict the fixed effects obtained from the main regression, Column (4) in Table 4. Each point represents a hospital. Its abscissa is simply the hospital's fixed effects. The ordinate is instead the weighted average of the DRG fixed effects of the DRGs available at the hospital, with weights the share of the hospital's total reimbursements attributable to each DRG. We distinguish public hospitals (circles) and private ones (triangles), whereas the size of the symbols describes the quartile the hospital is in, when the region's hospitals are ranked by total reimbursements. Finally, the color of the symbol illustrates the political colors of the council governing the province where hospital *h* is located: red for center-left, green for the regional Northern League, and blue for center-right.

It is hard to detect any pattern driven by color, size, or shape of the symbols. Our interpretation of this scatter is that provincial

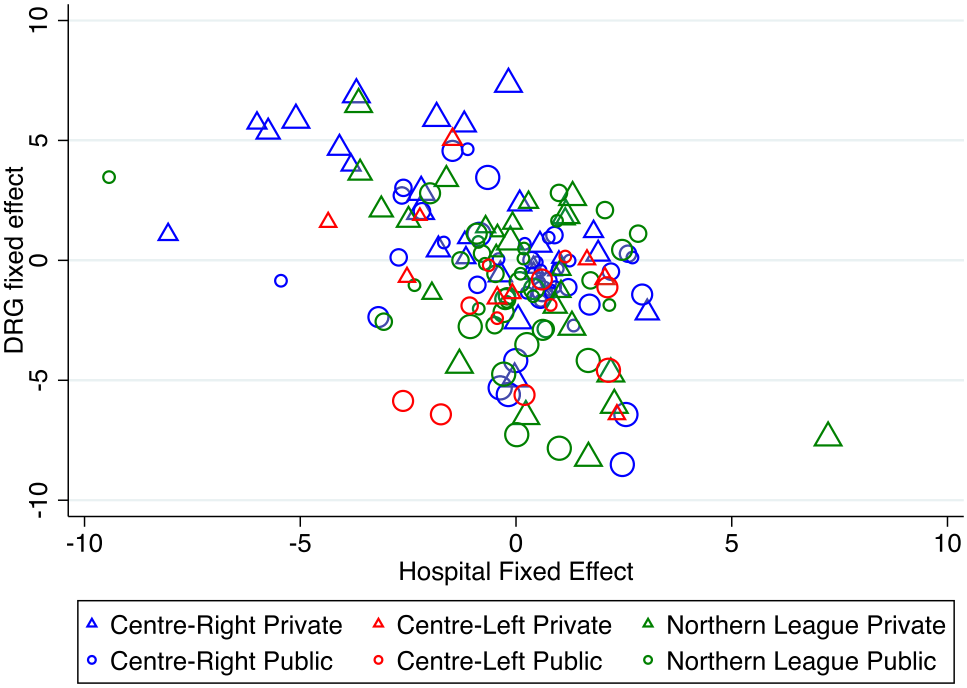

FIGURE 3 | Fixed effect and political color. *Note:* Hospital fixed effect (horizontal axis) and average DRG fixed effect for the DRGs provided by the hospital (vertical axis) for hospital located in areas with different parties enjoying political control. The size of the symbol indicates the size of the hospital. [Color figure can be viewed at wileyonlinelibrary.com]

politicians do not appear to influence the pricing decisions of the administration of the health authority. This could be due to the fact that the administration resists their effort, or that they do not try, perhaps because they do not have the information required to make effective suggestions, or because they perceive that their effort would neither be worth any benefit gained for their constituents, nor affect their voting pattern. This weak role of the fixed effects is confirmed by two very simple regressions we run. In each we estimate, the fixed effects with dummies indicating whether the political wing in power in the province where hospital h is located is the center-right coalition, $CR_h = 1$, or the Northern League $NL_h = 1$. with the center-left coalition set as the baseline case. We include as controls the quartile size dummies. These are the relevant coefficients of the estimated equations:

$$F_h^H = -\underset{.595}{.017} - \underset{.566}{.326}CR_h + \underset{.574}{.089}\mathrm{NL}_h + \varepsilon_h^H, \tag{20}$$

$$F_h^{\mathrm{DRG}} = -\underset{.824}{1.596} + \underset{.784}{1.974^{**}}\mathrm{CR}_h + \underset{.795}{.796}\mathrm{NL}_h + \varepsilon_h^{DRG}. \tag{21}$$

At the time the contract was finalized, the second half of 2012, the region as a whole was run by a coalition of the center-right and the Northern League, which had obtained a strong mandate in 2010, having already run the region for the previous 15 years. If the line of potential influence sketched above were effective, then those hospitals located in provinces controlled by the center-right and the Northern League should, other things equal, benefit from higher reimbursements, so that Figure 3 would show green and blue symbols toward the north-east of the diagram, and the red ones toward the south west. Similarly, we would expect the two regressions to have positive coefficients for CR_h and NL_h. While there is a hint in Equation (21) of a slightly higher reimbursement for DRGs that were more prevalent in hospitals in provinces controlled by the center-right, whereas the reimbursement to hospitals appears unaffected by these considerations, as none of the "party" coefficients are significant in Equation (20).

As anticipated, we close this section with a second possible route via which non-medical factors may affect the pattern of reimbursement offered by the regional health purchaser to its hospitals. Real estate price may affect hospital costs, for example in areas such as personnel recruitment, both medical, doctors and nurses, and ancillary personnel, from cleaners to receptionists and other workers. As above, we exploit the estimated fixed effects coefficients.

Figure 4 plots the hospital's fixed effect against the average housing price, measured in euros per square meter, in the municipality where the hospital is located. We run two separate regressions, one for the whole sample, one excluding the hospitals in Milan, which are easily identified as those with the same ordinate, as clearly they share the same housing costs. Note that the efficiency of the transport system within the urban area of Milan justifies using the same average real estate price for all the hospitals in the city. The results are reported in the caption, and show that the coefficient for average house price is not statistically significant in either regression. We feel justified in our interpretation that local real estate costs do not appear to

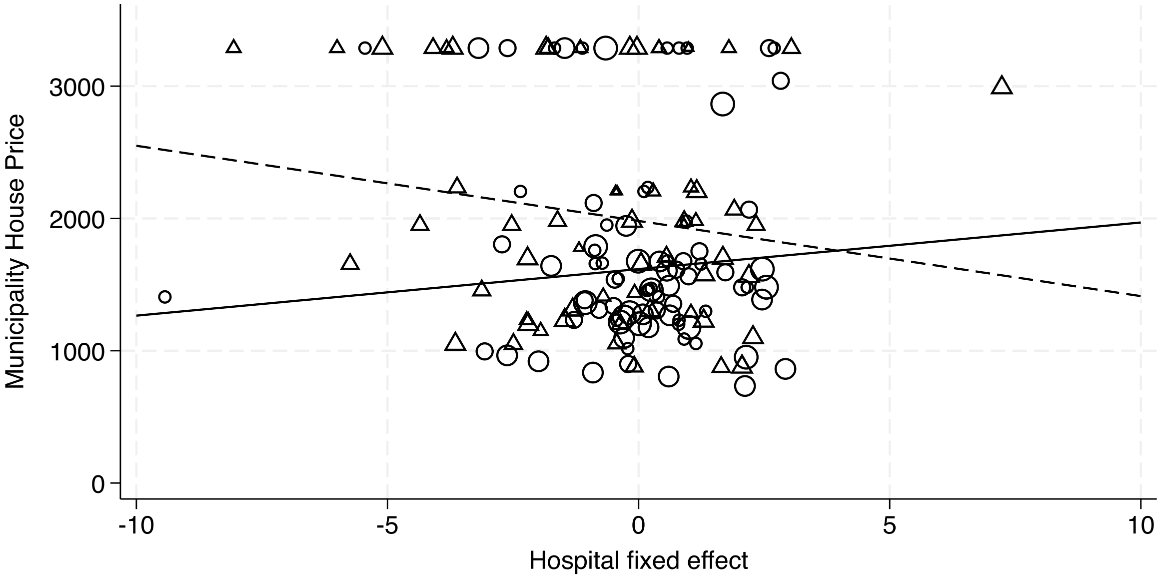

FIGURE 4 | Does house price matter? *Note:* The horizontal axis measures hospital fixed effect, and the vertical the municipality average housing prices, in euros per square meter. As in Figure 3, hospital ownership is indicated by the symbol: triangles for private hospitals, circles for public ones, and the symbol size indicates the size of the hospital. Housing price data are from the *Osservatorio del Mercato Immobiliare* of the Italian tax authority. The diagram shows two best fit lines: the dashed line, for the entire sample, is $F_h^H = -\underset{132.897}{1981.143^{***}} - \underset{30.626}{56.832} \times HP_h + \varepsilon_h^H$, and the solid line, where hospitals in Milan are excluded: $F_h^H = -\underset{82.966}{1617.301^{***}} + \underset{35.202}{21.721} \times HP_h + \varepsilon_h^H$. In both equations, F_h^H is a hospital fixed effect, as in Figure 3, and HP_h the average house price in the municipality where the hospital is located. Both include dummies for quartile of the size ranking. $^{***}p < 0.01$, and unstarred coefficients imply $p > 0.1$. [Color figure can be viewed at wileyonlinelibrary.com]

influence systematically the reimbursement decisions made by regional health authorities. To further investigate this, we run simple regressions of the hospital fixed effects on local housing prices, including dummies for size quartiles. As shown in the estimated equations below, the coefficients of interest are not statistically different from zero.

7 | Concluding Remarks

Purchasers of health services in many countries use DRG systems to reimburse the providers they contract with. The general principle of this system is that each hospitalization within a DRG is reimbursed equally. The efficiency properties of any system adopted in practice depend crucially on the minute details of the way in which this principle is implemented. This paper proposes a theoretical underpinning of the manner in which the DRG idea is applied by one large purchaser, a very populous Italian region, and subsequently tests the theory on the universe of the contracts offered by this purchaser, over a 1 year period of healthcare provision. We find suggestive evidence that the contracts offered by the health authority to its providers determined allocations of patients to hospitals such that “more attractive” hospitals, those which attract more patients, also admitted “more expensive” patients. This is in line with the pattern which would result when the contracts offered by the purchaser to the providers satisfy some simple optimality principles. In Section 6, we test other, not necessarily competing, plausible explanations for the emergence of a pattern in the region's reimbursements to its hospitals. We find that neither attempts by the politicians who control the health authority to favor their own parties, nor house prices help understand this pattern.

The theoretical analysis and the way in which it is used to interpret the available individual data on hospitalizations are clearly transferable to other environments which adopts a DRG system to reimburse providers, adjust the reimbursements dictated by the higher cost of provision for specific situations and for patients with relatively well defined characteristics and comorbidities, and where all providers face exactly the same menu of prices.

Acknowledgments

Early versions of this paper were presented at various conference and seminars, beginning with the European Health Economics Workshop EHEW in Oslo. We are particularly grateful for suggestions to Jesse Matheson, József Sákovics, Luigi Siciliani, Sebastian Fleitas, and Simona Grassi. We are also grateful to the anonymous co-editor and two referees of this journal for extensive comments which improved the paper substantially. This paper uses confidential data from the Healthcare Department of the Lombardy Region, Palazzo Lombardia, piazza città di Lombardia 1, Milano, Italy, whom we thank for their help. The authors (paolo berta, paolo.berta@unimib.it) are willing to assist researchers who wish to contact the office to access the data. All replicating codes and software are available at this link. The information and views set out in this paper are those of the authors and do not reflect the official opinion of the European Union. Neither the European Union institutions and bodies nor any person acting on their behalf may be held responsible for the use which may be made of the information contained therein.

Data Availability Statement

The data that support the findings of this study were made available to the University of Milan-Bicocca, Dept. of Statistics from the Lombardy Health Authority. Restrictions apply to the availability of these data due to the confidentiality of personal level data.

Endnotes

[1] Comprehensive discussions of the details of the implementation of DRG based payments in the US and in several European countries are Mihailovic et al. (2016), Cots et al. (2011), and Busse et al. (2011).

[2] Early theoretical contributions are Ellis and McGuire (1986), Newhouse (1996), and Sappington and Lewis (1999). Empirical analyses are numerous: Friesner and Rosenman (2009), Berta et al. (2010), Cheng et al. (2015) for three recent examples, in the US, Italy, and Australia. The potential for cream skimming is present whenever the features of a supplier's customers affect the supplier's cost of providing the good or service. Beside healthcare, other examples are financial markets (Ellis and McGuire 1986), regulation of utilities (Laffont and Tirole 1990), and the provision of education, whereby private schools compete with state schools by trying to attract the best students (Epple and Romano 2008 for a theoretical analysis, and Figlio and Stone 2001 and Altonji et al. 2015 for empirical ones).

[3] This is to reduce upcoding, the practice to use the presence of complicating factors, rather than their actual effect on the cost of treatment as a reason to choose a more expensive DRG for a given treatment. Vittadini et al. (2012) shows that the algorithm used by the Lombardy region attenuated this problem.

[4] In the conceptual framework provided by the incomplete contracts literature (Hart 1988 and Tirole 1999), these dimensions of patients' variability are “non-contractible”, that is, even though they are observed by both parties, they cannot be described in terms sufficiently unambiguous for a third party arbitrator easily to interpret and enforce them in the event of a dispute.

[5] In the year we study, the Lombardy region had an inspection policy requiring precise checks on the clinical handwritten notes of 10% of each provider's hospitalizations to verify the correspondence with the input in the discharge documentation upon which the reimbursement is based. The hospitalizations to be checked were not chosen by the provider. This percentage is well above the national average of 2%. We have however no information on either the number of hospitalization which were determined to have been upcoded, or the consequences for the hospitals deemed to have deliberately attempted to increase their revenues using this practice.

[6] Its population of 10 million makes it similar in size to Ohio, Portugal, or Sweden, with density as high as New Jersey's and the Netherlands. Milan, the region's administrative, financial, and cultural capital, accounts for approximately 15% of the region's population, and, correspondingly, does not monopolize healthcare provision, which is indeed available through a network of providers that covers the region's territory in the densely populated area south of the Alps. The map in Figure A4 in the Appendix, which depicts the location and size of hospitals in the region, illustrates this.

[7] There are 1544 municipalities, and 679 postcodes, both partitioning the region. Their intersection determines 1594 different locations; the average number of patients in a location is 3707, and the median is 1625.

[8] The parameters for each of the 538 DRGs are revised on a yearly basis, reflecting information on hospitals' costs, perceived distortions, or the purchaser's desire to influence the hospitals' overall supply.

[9] For example, the contract states that reimbursements to providers without an A&E (Accident and Emergency) department are reduced

by 3% relative to the tariff, those with a “high intensity” A&E department increased by 5%.

[10] In general the number of possible special cases K varies from DRG to DRG. Thus for many DRG–special case pairs (t, k), p_{tk} is constant.

[11] This corresponds to Gaynor et al.'s (2023) unobserved variability in the physicians' marginal costs of provision. A second dimension of heterogeneity between providers they model is their altruism. Unlike us, they have no unobserved heterogeneity in the patients, which we capture here with the parameter u.

[12] This restriction on the third derivative and that on the parameter γ_t^h entering multiplicatively the cost function (3) could be relaxed at the cost of greater algebraic complexity in the proof of Proposition 2.

[13] This set up is such that cost is constant up to κ_t^h and then it jumps to infinity. It would seem more plausible to posit a smooth function $\tilde{\gamma}_t^h(x)$, such that cost increases progressively as the capacity constraint is approached. An example of such a function is $\tilde{\gamma}_t^h(x) = \gamma_t^h + \frac{\exp(-\bar{\gamma}_t^h / x^2)}{\kappa_t^h - x}$ where γ_t^h is the cost of the first patient in DRG t, and κ_t^h is the maximum capacity, and $\bar{\gamma}_t^h$ is a positive parameter. As $\bar{\gamma}_t^h$ tends to $+\infty$, the cost function tends to the (reversed) L-shaped one we have posited. This functional form would complicate the notation but would not change the nature of the results of this section.

[14] This is obvious for-profit hospitals, but the literature has pointed out that it is plausible to postulate profit maximization even for not-for-profit hospitals, whether public or private, which are legally unable to distribute profits (Sloan 2000), as they may aim to generate a surplus on the treatments provided to devote to research or to caring for deserving patients (Dranove and White 1994). In our sample, 42 hospitals (almost 30% of the total) are for-profit, and they carry out 18% of the hospitalizations.

[15] In the current financial climate, it seems plausible to assume that hospitals are budget constrained, and so $\lambda^h > 0$. Should a hospital be “overfunded”, then Corollary 3 in the Appendix shows that λ^h would be 0 for such a hospital, and the budget constraint be slack at the optimum.

[16] It seems natural to assume the health authority to maximize a social welfare function satisfying anonymity, that is one that focuses only on the pattern of well-being, and not the identities of the people who end up at particular well-being levels (Adler 2016, p 267).

[17] This reflects the situation in Italy in the period we consider where regions did not have tax levying power. Thus the Lagrange multiplier of problem (13) is determined endogenously. Alternatively, if the health authority raised its own taxes, one would posit an exogenously given shadow cost of public funds.

[18] That is, problem (13) is the “subproblem” of a more general set-up where $\Phi_k^{ht}(u)$ and P^h are endogenously determined at the Nash equilibrium values of the game played by patients and hospitals. As we are only interested in characterizing the reimbursement vector $\{p_{tk}\}$, this is not limiting.

[19] That is if this patient's nearest hospital is 40km away, then all hospitals no further than 60km are considered local. All distances are measured along car routes. The correlation between geodesic and driving travel distances between pairs of centroids is 0.944, if road distance is calculated in minutes, and 0.983 if in kilometers. Thus results would be essentially identical had we used geodesic instead.

[20] It is as if we run a large Monte Carlo simulation randomly labeling patients with multiple local hospital to only one of them in each repetition and then averaging appropriately to obtain (18). Note also that some hospitals may have no local patients for a given DRG: to see this, consider a linear town on the segment $[0, 1]$ with hospital A in 0, hospital B in 0.2, and all patients to the right of 0.2. If some patients are admitted to hospital A, then hospital A provides this DRG, but has no local patients, and (18) is not defined. Hospitalizations in these hospitals–DRG pairs are dropped from the regression, but they constitute less than 0.1% of the total.

[21] These measures are obtained from information on the clinical diagnoses reported in the data by applying the Elixhauser algorithm (Elixhauser et al. 1998), which produces 30 dummies, each for a specific comorbidity, such as “obesity,” “cardiac arrhythmias,” “hypothyroidism,” and so on.

[22] This threshold is 60 days for a craniotomy on a young patient, and so a hospitalization for a craniotomy which requires 39 nights in hospital would be classified as having a length of stay of $(39-2)/(60-2) = 0.638$.

[23] By contrast the number of Lombardy residents who are treated in another Italian region is only 3.7% of the region's hospitalizations (Ministero della Salute 2014, Table 5.22 p234).

[24] Note that, if there are capacity constraint, this may create a spillover from patients whose treatment is reimbursed by a different region, similar in spirit to that from different insurers studied in Einav et al. (2020), whereby providers may reduce the number of treatments of regional patients in one DRG, to relax the budget constraint for other more profitable DRGs. The lack of significance of the coefficient for this variable encourages to think that this effect is at best limited.

[25] Approximately 10% of the hospitalizations do not report the information about the referring physician; for these, we impute the physician using information on the patient's municipality and postcode.

[26] Emergency hospitalizations are highly concentrated in some DRGs, with 12 of them having over 90% of them, and around 2/3 having less than 1% of their admissions classified as emergencies.

[27] To see, this, from Equation (1), using superscripts 0 and 1 to label the values pre- and post-increase in attractiveness, we have we have $\frac{R^0}{100} = \frac{\text{reimbursement}^0 - \text{tariff}}{\text{tariff}}$, from which tariff$= \frac{\text{reimbursement}^0}{1 + R^0 / 100}$. From the third column in Table 2, we see R_0 to be -3.79, and $R_1 = R_0 + 0.4125$. Given $\frac{R^1}{100} = \frac{\text{reimbursement}^1 - \text{tariff}}{\text{tariff}}$, we substitute to obtain $\text{reimbursement}^1 = \frac{R_1 + 100}{R_0 + 100} \times \text{reimbursement}^0 = 3627.5$. The PUH's revenues increase by $(3627.5 - 3612) = 15.5$€, for each of its 18.7 hospitalizations, which is equivalent to swapping the average hospitalization with one which is reimbursed at $15.5 \times 18.7 = 289.85$€ more, and $\frac{289.85}{3612} \cong \frac{8.025}{100}$.

[28] In all regressions the “Never married” category is the only one that is ever significantly different from the “Married” one, and so we do not report the others, such as divorced or widowed.

[29] Hardly surprising, given that the correlation between the values of A measured in the two methods is 0.936. We consider time a less clean measure of travel cost, both because the mode of transport may affect the time needed to travel to the hospital in a manner not orthogonal to the patient's complexity, and because today's travel times are likely to differ considerably more than the road distances from their 2013 values.

[30] Table A4 (Column 5a) shows that when the fifth percentile is considered instead, the coefficient obtained are practically identical to the main regression's (Column (4) in Table 4).

[31] DRGs with fewer than 100 hospitalizations are excluded.

[32] We use in this section eight overall measures of attractiveness, and the estimated coefficients using them are substantially similar. It may therefore be appropriate to report their correlation.

	A	A^{Time}	A^{Near}	A^{Town}	$A^{1.4}$	A^{MNL}	$A^{10\text{ th}}$	$A^{5\text{ th}}$	$A^{\text{MNL-FE}}$
A	1								
A^{Time}	0.952	1							
A^{Near}	0.898	0.857	1						
A^{Town}	0.869	0.822	0.992	1					
$A^{1.4}$	0.978	0.935	0.901	0.872	1				
A^{MNL}	0.560	0.575	0.494	0.491	0.527	1			
$A^{10\text{ th}}$	0.672	0.686	0.609	0.603	0.653	0.494	1		
$A^{5\text{ th}}$	0.986	0.943	0.914	0.884	0.991	0.554	0.659	1	
$A^{\text{MNL-FE}}$	0.305	0.307	0.279	0.281	0.299	0.274	0.268	0.303	1

Note: Correlations between the measure of attractiveness A, defined in Equation (18), and used in the main regression, Column (4) in Table 4, and in Table 5, and analogous measures built on different definitions of "local" hospital: A^{Time} is the measure obtained when distance is measured in travel time, A^{Near} is the when only the nearest hospital is considered local, A^{Town} when local hospitals are the nearest and any in the place of residence of the patient, $A^{1.4}$ when all those within 140% of the distance of the nearest are local, A^{MNL} the probability of a hospital being the local is the result of the MNL regression described in the text and used in Column (6) of Table 5, $A^{10\text{ th}}$ and $A^{5\text{ th}}$ are the measures obtained when any hospital is local if it is within the fifth and tenth percentile of the ranking by distance. The regressions corresponding to the four variables in the above table are and Columns (2), (6), and (3) in Tables 5 and A4. Finally, $A^{\text{MNL-FE}}$ is the hospital fixed effect obtained from the multi-nominal logit described in Section 5.1. For all coefficients, $p \leqslant 0.0005$.

[33] This may happen when a patient is admitted to a hospital for a treatment associated to a certain diagnosis (usually associated with a specific DRG at discharge), but she requires additional treatments during the hospitalization which may become the main reason of the hospitalization addressing her to a different DRG.

[34] While different, this distortion is analogous to that affecting the measures of competition based on the number of hospitals offering a given service within a given radius, which may be influenced by characteristics of the hospitals, such as their quality (e.g., Propper et al. 2004, p. 1254).

[35] We do not report other robustness tests we have carried out, as they do not alter the results: capturing demographic and socio-economic background with different combinations of location and referring physician fixed effect does, or excluding those who have died either during the hospitalization or within thirty days of discharge: Table 2 shows the latter to constitute around 4% of the sample.

[36] See Supporting Information S1: Equation (A8) for details. The sample is the regression sample, that is we exclude emergency hospitalizations, those on patients under 1 year of age, and those that exceed the threshold and the daily ones. Varying the controls included and using the coefficient of variation of the absolute value of the reimbursements does not alter the qualitative gist of the results reported in Equation (19). By the same token, excluding the non-profit hospitals leads to results similar, but slightly stronger than those reported in Equation (19).

[37] Suppose not: there exist hospitals h_0 and h_1 such that $\lambda^{h_0} > \lambda^{h_1}$. Then, from Equation (5) we have $\gamma_t^{h_0} c(u_{kt}^{h_0}, k) = p_{tk}(1 + \lambda^{h_0}) > p_{tk}(1 + \lambda^{h_0})$, by increasing slightly P^{h_1} and reducing slightly P^{h_0}, the health authority would remain within the budget $\mathcal{P}$, and determine (an overall reallocation of patients including) a shift of patients with observable characteristics k who need DRG t from hospital h_0 to hospital h_1. Since the latter has higher marginal cost in a sufficiently small neighborhood of $u_{kt}^{h_0}$ than the marginal cost of hospital h_1 in a sufficiently small neighborhood of $u_{kt}^{h_0}$, the health authority would be able to treat the same patients at a lower cost.

References

Adler, M. D. 2016. "Benefit–Cost Analysis and Distributional Weights: An Overview." *Review of Environmental Economics and Policy* 10, no. 2: 264–285.

Altonji, J. G., C. -I. Huang, and C. R. Taber. 2015. "Estimating the Cream Skimming Effect of School Choice." *Journal of Political Economy* 123, no. 2: 266–324.

Bayindir, E. E. 2012. "Hospital Ownership Type and Treatment Choices." *Journal of Health Economics* 31, no. 2: 359–370.

Beckert, W. 2018. "Choice in the Presence of Experts: The Role of General Practitioners in Patients' Hospital Choice." *Journal of Health Economics* 60: 98–117.

Berta, P., G. Callea, G. Martini, and G. Vittadini. 2010. "The Effects of Upcoding, Cream Skimming and Readmissions on the Italian Hospitals Efficiency: A Population-Based Investigation." *Economic Modelling* 27, no. 4: 812–821.

Brekke, K. R., L. Siciliani, and O. R. Straume. 2011. "Hospital Competition and Quality With Regulated Prices." *Scandinavian Journal of Economics* 113, no. 2: 444–469.

Brenna, E. 2011. "Quasi-Market and Cost-Containment in Beveridge Systems: The Lombardy Model of Italy." *Health Policy* 103, no. 2: 209–218.

Busse, R., A. Geissler, A. Aaviksoo, et al. 2013. "Diagnosis Related Groups in Europe: Moving Towards Transparency, Efficiency, and Quality in Hospitals?" *British Medical Journal* 346: f3197.

Busse, R., A. Geissler, W. Quentin, and M. Wiley. 2011. *Diagnosis-Related Groups in Europe: Moving Towards Transparency, Efficiency and Quality in Hospitals*. Open University Press.

Chandra, A., and D. O. Staiger. 2020. "Identifying Sources of Inefficiency in Healthcare." *Quarterly Journal of Economics* 135, no. 2: 785–843.

Cheng, T. C., J. P. Haisken-DeNew, and J. Yong. 2015. "Cream Skimming and Hospital Transfers in a Mixed Public-Private System." *Social Science & Medicine* 132: 156–164.

Cooper, Z., S. V. Craig, M. Gaynor, and J. Van Reenen. 2019. "The Price Ain't Right? Hospital Prices and Health Spending on the Privately Insured." *Quarterly Journal of Economics* 134, no. 1: 51–107.

Cots, F., P. Chiarello, X. Salvador, and W. Quentin. 2011. "DRG-Based Hospital Payment: Intended and Unintended Consequences." In *Diagnosis-Related Groups in Europe: Moving Towards Transparency, Efficiency and Quality in Hospitals*, 75–92. https://d1wqtxts1xzle7.cloudfront.net/42490748/Diagnosis-Related_Groups_in_Europe_-_Mov20160209-6632-b8i70e-libre.pdf?1455045380=&response-content-

disposition=inline%3B+filename%3DDiagnosis_related_groups_in_Europe_movin.pdf&Expires=1779100970&Signature=fufOFQSvk2ECPhveJMYDKK~tQFNx8Qj72C5hLlWhMB-sIQR5NjrBqkDRgelyqIPri4prEvqL2fdOWPLSHDUNpxCCu3V8RTRG-nal2mfpW7Q1MpvZzUbCbWaaZpi59BsiHDq3J~uAsabbsvSFL-qyhdwPj~5pqIaURfmtLKjB7EgD~9tdTFNSctmHkmRF1yMjsJI9j3x8-zy1lGZ2106-pv-XzgBajo-qiDBUUV1z9~-arwWvcoZ30-wthPR4R3-H5cvsWe8jpj0wuXoWD5-Am-veQRAqXSU~t4W21p1r~m5s7LvnA3zL3nTQGbJKH4l936BJSEQVJKx8RXc3py1hLa~O8w__&Key-Pair-Id=APKAJLOHF5GGSLRBV4ZA#page=106.

De Fraja, G. 2000. “Contracts for Health Care and Asymmetric Information.” *Journal of Health Economics* 19, no. 5: 663–677.

De Fraja, G. 2016. “Optimal Public Funding for Research: A Theoretical Analysis.” *RAND Journal of Economics* 47, no. 3: 498–528. https://doi.org/10.1111/1756-2171.12135.

Dranove, D., and W. D. White. 1994. “Recent Theory and Evidence on Competition in Hospital Markets.” *Journal of Economics and Management Strategy* 3: 169–209.

Einav, L., A. Finkelstein, Y. Ji, and N. Mahoney. 2020. “Randomized Trial Shows Healthcare Payment Reform Has Equal-Sized Spillover Effects on Patients Not Targeted by Reform.” *Proceedings of the National Academy of Sciences* 117, no. 32: 18939–18947.

Einav, L., A. Finkelstein, and N. Mahoney. 2018. “Provider Incentives and Healthcare Costs: Evidence From Long-Term Care Hospitals.” *Econometrica* 86, no. 6: 2161–2219.

Eliason, P. J., P. L. E. Grieco, R. C. McDevitt, and J. W. Roberts. 2018. “Strategic Patient Discharge: The Case of Long-Term Care Hospitals.” *American Economic Review* 108, no. 11: 3232–3265.

Elixhauser, A., C. Steiner, D. R. Harris, and R. M. Coffey. 1998. “Comorbidity Measures for Use With Administrative Data.” *Medical Care* 36, no. 1: 8–27.

Ellis, R. P., and T. G. McGuire. 1986. “Provider Behavior Under Prospective Reimbursement.” *Journal of Health Economics* 5, no. 2: 129–151.

Ellis, R. P., and T. G. McGuire. 1990. “Optimal Payment Systems for Health Services.” *Journal of Health Economics* 9, no. 4: 375–396.

Epple, D., and R. Romano. 2008. “Educational Vouchers and Cream Skimming.” *International Economic Review* 49, no. 4: 1395–1435.

Fabbri, D., and S. Robone. 2010. “The Geography of Hospital Admission in a National Health Service With Patient Choice.” *Health Economics* 19, no. 9: 1029–1047.

Fetter, R. B. 1991. “Diagnosis Related Groups: Understanding Hospital Performance.” *Interfaces* 21, no. 1: 6–26.

Figlio, D. N., and J. A. Stone. 2001. “Can Public Policy Affect Private School Cream Skimming?” *Journal of Urban Economics* 49, no. 2: 240–266.

Friesner, D. L., and R. Rosenman. 2009. “Do Hospitals Practice Cream Skimming?” *Health Services Management Research* 22, no. 1: 39–49.

Gaynor, M., N. Mehta, and S. Richards-Shubik. 2023. “Optimal Contracting With Altruistic Agents: Medicare Payments for Dialysis Drugs.” *American Economic Review* 113, no. 6: 1530–1571.

Gutacker, N., L. Siciliani, G. Moscelli, and H. Gravelle. 2016. “Choice of Hospital: Which Type of Quality Matters?” *Journal of Health Economics* 50: 230–246.

Hart, O. D. 1988. “Incomplete Contracts and the Theory of the Firm.” *Journal of Law, Economics, & Organization* 4, no. 1: 119–139.

ICD-9-CM. 1996. *International Classification of Diseases, 9th Revision, Clinical Modification*. Medicode.

Kalmijn, M. 2017. “The Ambiguous Link Between Marriage and Health: A Dynamic Reanalysis of Loss and Gain Effects.” *Social Forces* 95, no. 4: 1607–1636.

Kessler, D. P., and M. B. McClellan. 2000. “Is Hospital Competition Socially Wasteful?” *Quarterly Journal of Economics* 115, no. 2: 577–615.

Laffont, J. -J., and J. Tirole. 1990. “Adverse Selection and Renegotiation in Procurement.” *Review of Economic Studies* 57, no. 4: 597–625.

LeGrand, J. 2009. “Choice and Competition in Publicly Funded Health Care.” *Health Economics, Policy and Law* 4, no. 4: 479–488.

Luft, H. S., D. W. Garnick, D. H. Mark, et al. 1990. “Does Quality Influence Choice of Hospital?” *JAMA* 263, no. 21: 2899–2906.

Mathauer, I., and F. Wittenbecher. 2013. “Hospital Payment Systems Based on Diagnosis-Related Groups: Experiences in Low-And Middle-Income Countries.” *Bulletin of the World Health Organization* 91: 746–756A.

Mihailovic, N., S. Kocic, and M. Jakovljevic. 2016. “Review of Diagnosis-Related Group-Based Financing of Hospital Care.” *Health Services Research and Managerial Epidemiology* 3: 2333392816647892.

Ministero della Salute. 2014. “Rapporto annuale sull'attività di ricovero ospedaliero.” Report, Dir. Gen. della Programmazione sanitaria.

Moscelli, G., H. Gravelle, and L. Siciliani. 2021. “Hospital Competition and Quality for Non-Emergency Patients in the English Nhs.” *RAND Journal of Economics* 52, no. 2: 382–414.

Moscelli, G., L. Siciliani, N. Gutacker, and H. Gravelle. 2016. “Location, Quality and Choice of Hospital: Evidence From England 2002–2013.” *Regional Science and Urban Economics* 60: 112–124.

Newhouse, J. P. 1996. “Reimbursing Health Plans and Health Providers: Efficiency in Production Versus Selection.” *Journal of Economic Literature* 34, no. 3: 1236–1263.

Propper, C., S. Burgess, and K. Green. 2004. “Does Competition Between Hospitals Improve the Quality of Care?: Hospital Death Rates and the NHs Internal Market.” *Journal of Public Economics* 88, no. 7–8: 1247–1272.

Sappington, D. E. M., and T. R. Lewis. 1999. “Using Subjective Risk Adjusting to Prevent Patient Dumping in the Health Care Industry.” *Journal of Economics & Management Strategy* 8, no. 3: 351–382.

Sloan, F. A. 2000. “Not-for-Profit Ownership and Hospital Behavior.” In Handbook of Health Economics, edited by A. J. Culyer and J. P. Newhouse, volume 1 B, 1141–1174. Elsevier.

Tirole, J. 1999. “Incomplete Contracts: Where Do We Stand?” *Econometrica* 67, no. 4: 741–781.

Vittadini, G., P. Berta, G. Martini, and G. Callea. 2012. “The Effect of a Law Limiting Upcoding on Hospital Admissions: Evidence From Italy.” *Empirical Economics* 42, no. 2: 563–582.

Supporting Information

Additional supporting information can be found online in the Supporting Information section.
Supporting File: jems70035-sup-0001-supp-info.pdf.

Appendix A

Appendix: Proofs of the Results in Section 3

Proof of Proposition 1. To lighten notation, we normalize the number of potential patients to 1 and the cost parameter correspondingly. Begin by noting that, if provider h decides to treat for DRG t n_{kt}^h patients with observable characteristics k, then it will treat those with unobservable characteristics below u_{kt}^h, where u_{kt}^h satisfies

$$n_{kt}^h = \int_{\underline{u}}^{u_{kt}^h} \phi_k^{ht}(u)\mathrm{d}u = \Phi_k^{ht}\left(u_{kt}^h\right), \tag{A1}$$

This is because, given that provider h observes the type of each patient and knows the overall distribution of the characteristics of the patients who require the treatment during the year, it can choose to treat only those with the most favorable characteristics, *viz.* those with $u \leqslant u_{kt}^{h}$.

Thus the problem of provider h can simply be stated as:

$$\max_{\left\{\left\{u_{kt}^{h}\right\}_{k=1}^{K}\right\}_{t=1}^{T}} \sum_{t=1}^{T}\sum_{k=1}^{K}\left(p_{tk}\Phi_{k}^{ht}\left(u_{kt}^{h}\right) - \int_{\underline{u}}^{u_{kt}^{h}} \gamma_{t}^{h} c^{t}(u,k)\phi_{k}^{ht}(u)\mathrm{d}u\right)$$
$$\text{s.t.}: \sum_{t=1}^{T}\sum_{k=1}^{K} p_{tk}\Phi_{k}^{ht}\left(u_{kt}^{h}\right) \leqslant P^{h}. \tag{A2}$$

The proposition follows in a straightforward manner from the solution to problem (A2). Total differentiation of Equation (5) yields Equation (6), after some straightforward rearranging. ☐

An immediate consequence of the fact that the constraint in Equation (A2) is an inequality constraint is the following.

Corollary 3. *There exists* $P^{M} > 0$ *such that if* $P^{h} > P^{M}$, *then* $\lambda^{h} = 0$, *and so* $\gamma_{t}^{h} c_{u^{h}kt}^{\,t}(u,k) = p_{tk}$ *for every* $k = 1, \ldots, K$ *and* $t \in T$.

Proof of Corollary 3. This follows immediately from the slack complementarity condition in (A2): $\left(\sum_{t=1}^{T}\sum_{k=1}^{K} p_{tk}\Phi_{k}^{ht}(u_{kt}^{h}) - P^{h}\right)\lambda^{h} = 0$. ☐

Proof of Corollary 1. Simply differentiate (A1) in the proof of Proposition 1, and compute the average reimbursement by dividing the total reimbursement by N_{t}^{h}. ☐

Proof of Proposition 2. To begin, it is fairly simple to show that at any solution, $\lambda^{h} = \lambda^{h'}$ for every $h, h' = 1, \ldots, H$, that is, all the hospitals' Lagrange multipliers, λ^{h} in Equation (5), are identical.[37] From this, taking into account Equation (5) and the assumption that $c_{uu}^{t}(\,\cdot\,) \geqslant 0$, there follows in turn that $u_{kt}^{h_0} < u_{kt}^{h_1}$ if and only if $\gamma_{u}^{h_0} < \gamma_{u}^{h_1}$: if a patient is admitted at a given hospital, then she would also be admitted at a hospital with a lower cost parameter. Because the observable characteristics k are labeled to be increasing in the cost of treatment, it follows naturally that the optimal reimbursements chosen by the health authority satisfy, for DRG-characteristics with a strictly positive number of hospitalizations in equilibrium: this is Equation (14). To continue, note that the solution to Equation (5) depends on γ_{t}^{h}, which we make explicit in what follows by writing $u_{kt}^{h} = u_{kt}^{h}(\gamma)$. (i), the first part of the statement of the proposition, is established immediately by differentiation:

$$\frac{\partial N_{t}^{h}}{\partial \gamma} = \sum_{k=1}^{K} \phi_{k}^{ht}\left(u_{kt}^{h}(\gamma)\right)\frac{\mathrm{d}u_{kt}^{h}(\gamma)}{\mathrm{d}\gamma} = -\sum_{k=1}^{K} \phi_{k}^{ht}\left(u_{kt}^{h}(\gamma)\right)\frac{c_{u}(\,\cdot\,)}{\gamma_{t}^{h} c_{uu}(\,\cdot\,)} < 0.$$

Consider (ii) next. We begin by noting that differentiation of $\frac{c_{u}(\cdot)}{\gamma_{t}^{h} c_{uu}(\cdot)}$ yields

$$\frac{\partial}{\partial k}\left(\frac{\mathrm{d}u_{kt}^{h}(\gamma)}{\mathrm{d}\gamma}\right) = \frac{-1}{\gamma_{t}^{h} c_{uu}(\,\cdot\,)}\left(c_{uk}(\,\cdot\,) - \frac{c_{u}(\,\cdot\,)c_{uuk}(\,\cdot\,)}{c_{uu}(\,\cdot\,)}\right). \tag{A3}$$

This, given the technical assumption on $c_{uuk}(\,\cdot\,)$, is negative.

We next proceed by induction on K, the number of possible observable types. We begin by denoting by $\overline{p}_{t}^{hk}$ the average reimbursement for DRG t received by hospital h when the number of possible combinations of factors affecting the reimbursement is k. We compact notation in the proof by omitting the subscript t and the superscript h. We begin the induction argument by showing that the assert holds for $K = 2$. The algebraic section of the online appendix shows that developing Equation (12) for $K = 2$, one obtains

$$\frac{\partial \overline{p}^{2}}{\partial \gamma} = \frac{(p_2 - p_1)\phi_2(\,\cdot\,)\phi_1(\,\cdot\,)}{(\Phi_1(\,\cdot\,) + \Phi_2(\,\cdot\,))^2}\left(\frac{\partial u_2}{\partial \gamma}\frac{\Phi_1(\,\cdot\,)}{\phi_1(\,\cdot\,)} - \frac{\partial u_1}{\partial \gamma}\frac{\Phi_2(\,\cdot\,)}{\phi_2(\,\cdot\,)}\right).$$

Now note that we have (i) $p_2 > p_1$ by Equation (14), (ii) $\frac{\partial u_2}{\partial \gamma} < \frac{\partial u_1}{\partial \gamma} < 0$, from Equation (A3), and (iii) Lemma 1 in De Fraja (2016) shows that if Φ_2 first order stochastically dominates Φ_1, then $\frac{\Phi_1(\cdot)}{\phi_1(\cdot)} > \frac{\Phi_2(\cdot)}{\phi_2(\cdot)}$, so that the last term in the above is negative. Therefore, $\frac{\partial \overline{p}^{2}}{\partial \gamma} < 0$, which establishes the first step of the induction process.

Next, assuming to have shown that the assert holds for $K - 1$, we show that it holds also for K. Write the average reimbursement (12) as

$$\overline{p}^{K} = \frac{\sum_{k=1}^{K-1} p_k\Phi_k(\,\cdot\,) + p_K\Phi_K(\,\cdot\,) + \overline{p}^{K-1}\Phi_K(\,\cdot\,) - \overline{p}^{K-1}\Phi_K(\,\cdot\,)}{\sum_{k=1}^{K}\Phi_k(\,\cdot\,)}. \tag{A4}$$

Next, use the fact that

$$\overline{p}^{K-1} = \frac{\sum_{k=1}^{K-1} p_k\Phi_k(\,\cdot\,)}{\sum_{k=1}^{K-1}\Phi_k(\,\cdot\,)},$$

and so

$$\sum_{k=1}^{K-1} p_k\Phi_k(\,\cdot\,) = \overline{p}^{K-1}\sum_{k=1}^{K-1}\Phi_k(\,\cdot\,),$$

to write Equation (A4) as:

$$\overline{p}^{K} = \overline{p}^{K-1} + \frac{(p_K - \overline{p}_{K-1})\Phi_K(\,\cdot\,)}{\sum_{k=1}^{K}\Phi_k(\,\cdot\,)}. \tag{A5}$$

Now differentiate (we refer again to Supporting Information S1):

$$\frac{\partial \overline{p}^{K}}{\partial \gamma} = \frac{\partial \overline{p}_{K-1}}{\partial \gamma}\left(1 - \frac{\Phi_K(\,\cdot\,)}{\sum_{k=1}^{K}\Phi_k(\,\cdot\,)}\right)$$
$$+ \frac{p_K - \overline{p}^{K-1}}{\left(\sum_{k=1}^{K}\Phi_k(\,\cdot\,)\right)^2}\sum_{k=1}^{K}\left(\frac{\partial u_K}{\partial \gamma}\frac{\Phi_k(\,\cdot\,)}{\phi_k(\,\cdot\,)} - \frac{\partial u_k}{\partial \gamma}\frac{\Phi_K(\,\cdot\,)}{\phi_K(\,\cdot\,)}\right) \tag{A6}$$
$$\phi_k(\,\cdot\,)\phi_K(\,\cdot\,).$$

Recall that we have $\frac{\partial \overline{p}^{K-1}}{\partial \gamma} < 0$ by the induction hypothesis, and $p_K > \overline{p}^{K-1}$ by Equation (14), the assumption that reimbursements are increasing with k; thus $\overline{p}^{K-1}$ is a weighted average of reimbursements all lower than p_K. Next consider the terms in the sum: $\phi_k(\,\cdot\,)\phi_K(\,\cdot\,)$ are all positive. The same result on first order stochastic dominance used in the first step of the induction process shows that the terms in the brackets are all negative except when $k = K$, in which case it is 0. This shows that $\frac{\partial \overline{p}^{K}}{\partial \gamma} < 0$ and completes the proof. ☐

Proof of Corollary 2. This holds because the Lagrange multiplier λ_h in Equation (5) is lower and so the cut-off value of u_{kt}^{h} decreases, and the marginal hospitalizations are refunded at a lower rate, thus lowering the average reimbursement. ☐