Reimbursement Schemes for Hospitals: The Impact of Case and Firm Characteristics*

Eberhard Feess
Victoria University of Wellington

Helge Müller
University of Marburg

Ansgar Wohlschlegel†
University of Portsmouth

Abstract

We contribute to the debate on high-powered versus low-powered incentives in regulation by studying their heterogeneous impacts on different subpopulations, using data from the introduction of a high-powered prospective payment system (PPS) for hospital reimbursement in Germany. While no overall effect on quality or cost saving is found, our results support hypotheses drawn from an incentive and selection perspective: PPS reduces the length of stay of older relative to younger patients, of more severe relative to less severe cases, and in smaller relative to larger hospitals. Hospitals which adopted PPS earlier provide higher quality under PPS as proxied by the case-specific readmission rate. Our study also contributes to the health economic literature on hospital reimbursement as our data permits us to identify the treatment effect via different timings of adoption of PPS and to use a more accurate quality measure by following patients even when readmitted to other hospitals.

JEL Classification: I11, D22, I18

Keywords: Hospital reimbursement; high-powered incentives; readmission; length of stay.

*We are grateful for helpful comments by an anonymous referee, as well as Arnaud Chevalier, Thomas Gall, Fernanda Leite Lopez de Leon, Robin Lumsdaine, Mustafa Ozer, Hannah Schildberg-Hörisch, Guido Schwerdt, Christian Traxler, and participants at the Royal Economic Society and Verein für Socialpolitik conferences and seminars at Bamberg, Konstanz, Marburg, Massey, Paderborn, Portsmouth, Regensburg and WHU Vallendar.

†Corresponding author. Postal Address: Portsmouth Business School, Richmond Building, Portland Street, Portsmouth PO1 3DE, United Kingdom. Email: Ansgar.Wohlschlegel@port.ac.uk

1 Introduction

The impact of policy interventions often varies considerably across different subpopulations. For instance, treatment effects of education programs differ among individuals with different abilities (List and Rasul (2011)), and introducing high-powered incentive schemes in the electricity industry leads to a large variety of reactions by market players (Armstrong and Sappington (2006); Ter-Martirosyan and Kwoka (2010)). If firms are differently affected by changes in the incentive structure, then their reactions are likely to differ as well. Against this backdrop, we investigate how the impact of moving from a cost reimbursement scheme (fee-for-service; FFS) to a prospective payment system (PPS) for hospitals depends on case characteristics (such as the patient-specific complication risk and age) and firm characteristics (such as the size and the ownership of the hospital). Our original data include all about 194,000 in-hospital cases between 2001 and year-end 2008 reimbursed by a German health insurance company. Germany moved to PPS between 2003 and 2006, so that we have cases before and after the system change for each hospital. The dependent variables in our analysis are cost reductions proxied by the length of stay in hospital (LOS) and the quality of service measured by the one-year readmission rate to any hospital with the same main diagnosis.

We find no impact of the system change on the *average* LOS and readmission rate, but results differ significantly between patients and hospitals in ways suggested by the economic theory of incentives and selection: patients older than 65, who are likely to have lower opportunity costs, are discharged relatively earlier after the system change than younger patients. The same holds for patients with more severe secondary diagnoses. Interestingly, these changes in LOS had no impact on long-term readmission rates.[1] However, we find that the readmission rate decreased in hospitals that switched early to the new system; relative to those hospitals which delayed the adoption of the new system as long as possible. This supports the intuition that hospitals expecting to do well

[1]Comparable results are recently found by Batty and Ippolito (2017) who show that lower reimbursements for non-insured patients lead to a considerable reduction in LOS but not to a decline in quality measured by several proxies.

under the new system were eager to adopt it, whereas hospitals likely to struggle with PPS sought to delay the system change. Furthermore, there is some tentative evidence that private hospitals responded stronger to the change in the incentive structure than public hospitals.

Theoretically, the impact of substituting FFS by PPS on the hospitals' incentives is ambiguous (see the overview in Ellis and Miller (2009)): Under FFS, hospitals are reimbursed for all justifiable costs, so that there are low incentives for cost-cutting. Financial incentives for quality provision are also weak as readmissions lead to additional reimbursements. PPS sets strong incentives for cost savings as payments are to a large degree fixed (see section 2 for details)[2] but also incentives for quality provision as readmissions with the same main diagnosis within a period of 30 days trigger no additional reimbursements. Since thus both cost-cutting and quality incentives are larger under PPS, overall effects are unclear. As discussed in section 2, we can nevertheless draw hypotheses on the case and hospital specific impact of the system change.

The advantage of using data from a health insurance company instead of data from single hospitals only is that we can follow the patient's history also when she is readmitted to a different hospital. This is important as, under PPS, hospitals have incentives to trigger readmissions to other hospitals in order to avoid case consolidations. We have patient specific information about primary and secondary diagnoses, in-hospital complications, dates of admissions and discharge, surgeries and treatments, discharge categories, sex and age. In addition, we collected hospital-specific information such as the number of beds and the ownership structure from various official statistics. For all hospitals, we have observations under both systems but partially at different times as hospitals had some discretionary power on when to adopt the new system. This allows us to apply a difference-in-differences approach, where we consider cases under FFS as control group and cases under PPS as treatment group. Furthermore, we control for unobserved het-

[2]In Germany, this has triggered serious concerns about "bloody discharges" (See e.g. Deutscher Bundstag, Drucksache 16/6083), which were reinforced by the fact that the liability system for malpractice plays a lower role compared to the US (Kennedy and Grubb (2000) and Faure and Koziol (2001)).

erogeneity by estimating coefficients with hospital fixed effects and for technical progress and other changes over time with a time trend.

Most papers analyzing the impact of PPS restrict attention to LOS and find that the average LOS has declined (e.g. DesHarnais, Kobrinski, Chesney, Long, Ament, and Fleming (1987); Ellis and McGuire (1996); see Cutler and Zeckhauser (2000) for an overview). Most closely related to our paper are studies which also can apply a difference-in-differences approach as not all hospitals are affected by the system change at the same time. For the period in China considered by Yip and Eggleston (2001), six out of fourteen hospitals included in their data switched from FFS to PPS. They find that PPS reduced expenditures and LOS but do not extend to quality measures. Similar results are obtained by Zhang (2010), who also restricts attention to cost measures and LOS. In contrast to these papers, we find no significant overall impact on the system change on LOS,[3] but we do find differences among patients and hospitals.

Several different measures such as time per patient spend, the degree of satisfaction of patients as stated in questionnaires and mortality rates have been used in the literature to proxy the quality of treatment (see IGES (2013) for the most recent available descriptive statistics on these measures for Germany) but the most commonly used proxy is readmission rates. Many studies are confined to descriptive statistics (Louis, Yuen, Braga, Cicchetti, Rabinowitz, Laine, and Gonnella (1999)), only one hospital (Rich and Freedland (1988)) descriptive cross-country analyses (see Busse, Schreyögg, and Smith (2008) and Stargardt (2008)) or to a few diagnoses, or they lack a control group as all health care provides were forced to switch to PPS at the same time. Kahn, Draper, and Keeler (1992) were among the first to use the readmission rate as a quality measure but restrict attention to six medical conditions and neither account for time-trends nor for readmissions to other hospitals. Cutler (1995) shows that distinguishing between different time horizons is important as the mortality rate in the hospital itself or shortly after discharge

[3]The Federal Statistical Office of Germany, Statistisches Bundesamt (2017), reports that the time trend in the reduction of the LOS has weakened and that DRG-adjusted LOS has basically been stable in the last years.

is higher with PPS, but that there is no difference in the one-year-mortality rate. In the context of policy changes which have lead to adjustments in the remuneration for many diagnoses, Dafny (2005)) finds that hospitals respond strongly to incentives, albeit mainly for re-coding, while actual treatments are only slightly modified.

More recently, a few papers apply a difference-in-differences approach for readmissions. Busato and von Below (2010) utilize the fact that PPS was introduced in Swiss cantons at different times. They find higher readmission rates for a 90-day period in DRG areas but emphasize that they compare readmission rates for the whole population without controlling for case- or patient-specific characteristics. Hamada, Sekimoto, and Imanaka (2012) use data from Japan and report higher 30-day readmission rates under DRG. However, their study suffers from selection problems as hospitals had complete discretion on whether and when they switch to PPS. In their study on Taiwan, Cheng, Chen, and Tsai (2012) can avoid the selection problem as the assignment to treatment and control groups is exogenous. While their studiy is confined to patients listed in the cardiovascular surgical DRG category, we consider many different main diagnoses to identify average effects of the system change. Furthermore, we control for secondary diagnoses, hospital fixed effects and the timing when the new system was adopted.[4]

The remainder of this paper is structured as follows: Section 2 describes the reform in Germany and derives hypotheses on its impact on the length of stay and the readmission rate. We describe the data in Section 3. Section 4 explains our estimation strategy. 5 presents our main results. Robustness checks with respect to the treatment of secondary diagnoses and the period for the readmission rate are discussed in section 6. Section 7 concludes.

[4]Gaynor, Moreno-Serra, and Propper (2013) exploit a policy change in the UK that led to higher competition in the health care industry to apply a difference-in-differences approach and find evidence that higher competition has no cost effect, but increases quality.

2 The Reform and Incentives for Hospitals

Reimbursement Schemes Before and After the Reform. Before 2003, German hospitals were reimbursed according to a mixed system in which most payments were made on a fee-for-service (FFS) base. About 80% of hospitals' budgets originated from reimbursements for every day a patient stayed in hospital (Lüngen and Lauterbach (2003)). The exact amount of this per-day reimbursement depended on the costs for the patient's treatment and on whether the total occupancy of a hospital during a year was above or below a certain negotiated level. The prospective payment system (PPS) was enacted by the Health Reform Act ('*GRV-Gesundheitsreformgesetz*') in 2000, which required hospitals to switch to the new system by 1st January 2004; the earliest possible switching date being the 1st January 2003. Hospitals could ask to postpone the introduction of PPS under certain circumstances. The requirements for getting permission turned out to be quite mild until the 31st August 2004, but got increasingly demanding thereafter. By 1st January 2006, almost all hospitals in our dataset had introduced PPS.[5]

The German PPS is a version of the Australian 'AR-DRG' system (see Mihailovic, Kocic, and Jakovljevic (2016) for a recent overview on different DRG-systems). For each case, hospitals are reimbursed a fixed amount depending on the diagnosis related group (DRG) of the case and the severity of secondary diagnoses. Secondary diagnoses are transformed into the "Patient Clinical Complexity Level" (PCCL), which is an index that can take values from zero (no complications requiring additional expenditures) to four (highest level of complications)[6]. If the actual LOS is far above or below the average LOS of that DRG, a surcharge or deduction applies, respectively. If a patient is readmitted to the same hospital within 30 days with the same main diagnosis, or with a complication that occurred due to the treatment of the original diagnosis, both cases are consolidated. Consolidation means that hospital stays are reimbursed as a single case, i.e. the DRG

[5]These hospitals represent 99.98% of all cases in our full dataset, and all of the cases in the sample used for our regression analysis.

[6]In 2014, PCCL-levels of 5 and 6 were added in order to allow for a more fine-tuned differentiation for particularly severe cases.

fee is paid only once.

General Incentive Effects of PPS. The introduction of PPS created countervailing incentives for cost effectiveness and quality of treatment: Under FFS, all medically justifiable treatments were reimbursed, so that hospitals did not need to be too concerned about costs. In particular, the reimbursement for each additional day a patient spent in hospital exceeded marginal costs, so that hospitals with free capacities had incentives to keep patients in hospital. However, hospitals without free capacities may have had incentives for early discharges in order to take on patents with higher contribution margins. This latter point may be important as there was no financial incentive to reduce the risk of complications and readmissions under FFS, because additional treatment costs after readmission were covered by the health insurer.

Under PPS, there are lower bounds and upper bounds for the LOS for each DRG (and partially depending on secondary diagnoses). Within these two bounds, the reimbursement is independent of the LOS. Thus, when a patient is discharged one day earlier within these bounds, a hospital just saves the cost for this day. Similarly, hospitals reap the full financial benefits of saving other kinds of treatment cost. On the other hand, however, hospitals also have incentives to avoid complications and readmissions as they do not normally get any additional payment for patients readmitted within 30 days. Hence, from a purely financial perspective, the benefit of cost saving needs to be compared to the expected costs from a higher probability that the patient is readmitted within thirty days.

If a patient is discharged before the lower bound or after the upper bound, then deductions and surcharges apply, respectively. Thus, financial incentives are somewhat less pronounced beyond the lower and the upper threshold. Still, these additional payments or deductions are below the actual daily costs, so that there are no incentives to keep a patient unnecessarily long in hospital.[7]

[7]For specific diagnoses, the Medicare Hospital Readmission Reduction Program in the US sets additional quality incentives by imposing penalties on hospitals with readmission rates above average. Doyle Jr, Graves, and Gruber (2017) find that incentivizing quality measures has indeed positive quality

Summing up, hospitals under FFS have neither strong incentives for cost-cutting nor for avoiding readmissions, whereas they do have such incentives under PPS. Furthermore, the enhanced cost-cutting incentives under PPS create a countervailing effect to the enhanced quality incentives and vice versa, as cutting cost of medical treatment is likely to increase the readmission rate. It follows that the comparisons between the LOS and the readmission rate under PPS and FFS are theoretically ambiguous.

Incentive Effects Depending on Case and Hospital Characteristics. Although we cannot make clear-cut theoretical predictions regarding the average effect of adopting PPS, our discussion of the incentive structure allows us to draw the following three hypotheses on how this effect differs across case and hospital characteristics: First, if hospitals with free capacities under FFS want to extend a patient's stay in hospital, a patient with high opportunity costs is more likely to object. Thus, we hypothesize that the LOS of patients with low opportunity costs decreases after the system switch relative to patents with high opportunity costs, and we use an age above 65 as a proxy for low opportunity costs.

Second, we assume that the variance in capacity utilization decreases in the number of beds in hospital, so that the incentive to keep patients longer under FFS should be higher for small hospitals. Third, private hospitals are likely to respond stronger to financial incentives. However, as we do not have a clear-cut hypothesis about whether hospitals' incentives to keep patients in hospital are lower under PPS in order to save costs or higher in order to avoid readmissions, we cannot hypothesize about whether the treatment effect on LOS for private hospitals will be higher or lower than that for public hospitals either. However, as readmissions yield financial losses only under PPS, we might conjecture that the interaction term between private ownership and the system dummy is negative for the readmission rate.

effects including a reduction in mortality rates. Schreyögg (2017) discusses how incentives under PPS can be improved in Germany by conditioning payments on quality measures.

Documentation of Seondary Diagnoses. Another feature of FFS was that secondary diagnoses only were relevant for reimbursement in very specific cases where they made additional treatments necessary. As a consequence, there was no need to document them unless they were important information for subsequent treatments or they were needed to justify the costs of in-hospital treatments. Under PPS, however, the detailed case characteristics are taken into account to calculate the correct reimbursement for that case. Hence, all secondary diagnoses are relevant in every case, so that incentives for exact documentation are high. There is even reason to believe that hospitals (illegally) upcode the secondary diagnoses under PPS.[8]

This difference in the incentives to document secondary diagnoses under both systems has an important implication for our data: When using the secondary diagnoses actually documented by the hospital to calculate which PCCL a case reimbursed under FFS would have been assigned under PPS, this variable will systematically underestimate the severity of secondary diagnoses under FFS. We will, therefore, take this systematic measurement error into account when constructing such a variable. In Section 3, we will explain in more detail how we construct a measure for the severity of secondary diagnoses that is comparable between cases reimbursed under both systems.

3 Data

Our data set consists of all 193,943 in-hospital cases between 2001 and the end of 2008 covered by a German health insurer. 45,732 cases were reimbursed under the old system (FFS) and 148,211 cases under the new system (PPS). The patient files contain information on primary and secondary diagnoses, DRGs, PCCL (as part of the raw data under PPS, and calculated by us for cases under FFS by using the documented secondary diag-

[8]For instance, Abler, Verde, Stannigel, Mayatepek, and Hoehn (2011) and Jürges and Köberlein (2015) show that upcoding plays a large role in neonatology, where downscaling the birth weight by only a few gram yields a sharp increase in reimbursements, and where upgrading is hardly detectable since infants lose about 10% of their birth weight within the first few days. At an aggregated level, Schönfelder, Balázs, and Klewer (2009) estimate that more than three billion Euros paid as reimbursements to hospitals between 2004 and 2009 can be attributed to upcoding.

noses and the *Grouper* software applied by hospitals), dates of admissions and discharge, surgeries and treatments, discharge categories such as "regular" and "after hospital treatment", gender and age. Not all hospitals switched to the prospective payment system at the same time due to the aforementioned fact that there was a transition year (2003) where hospitals could choose between the two systems. The exact dates when the hospitals adopted PPS were taken from the official statistics of the national association of health insurance companies (AOK Bundesverband).

The advantage of using data from a health insurer instead of data from hospitals is that it contains readmissions to *any* hospital, so that we can safely use readmissions with the same main diagnosis as a proxy for treatment quality. As pointed out in the introduction, this is important as PPS sets incentives to trigger readmissions to other hospitals in order to avoid case consolidations. A disadvantage of the fact that just one health insurer is included is that we have only a few cases for many of the primary diagnoses in many of the 603 hospitals in our dataset. As we want to run regressions with hospital fixed effects in order to control for unobserved heterogeneity, we need to restrict attention to hospitals with sufficiently many observations for each DRG considered. We hence restrict attention to the 50 most frequent primary diagnoses, defined at the three-digit ICD level, and to hospitals which have, on average, at least ten cases per year for these primary diagnoses. This leaves us with 54 hospitals and 58,472 cases, of which 15,396 (26.3 %) were reimbursed according to the old FFS system. Table 9 in the Appendix provides an overview of the numbers of cases in these hospitals and the dates at which they switched to PPS.[9]

Table 1 structures information on the dependent variables (LOS and readmission rates) and on the explanatory variables used in our main regressions for the sample used in the empirical analysis. The average LOS has decreased from 7.2 to 6.3 days after PPS was introduced. This, however, needs to be interpreted with caution as there may be a time trend we need to control for. Recall that cases are consolidated when patients

[9]The average size is calculated as the average number of beds in the respective hospital during the observation period minus the average.

are readmitted to the same hospital with the same primary diagnosis. For those cases, hospitals report only the aggregated LOS of both stays, so that we need to exclude them when calculating the average LOS, and when analyzing the impact of the system change on LOS.

Insert Table 1 about here.

Next, table 1 shows that the average readmission rate within a year to any hospital is about 12.8% lower under the new system (0.360 compared to 0.413), and an even lager reduction of about 17% (0.082 compared to 0.099) is observed for the 30-day threshold within the same hospital, which is decisive for case consolidation. This lends some support for the hypothesis that hospitals strategically delay readmissions or trigger readmissions to other hospitals during the 30-day-period, while the fact that both readmission rates are lower under PPS is likely due to a time trend reflecting technical progress and the tendency to transfer after-hospital treatments to doctors' offices.

Recall from Section 2 that the severity of secondary diagnoses under PPS is summarized by the categorical variable PCCL reaching from PCCL0 (no additional complications expected) to PCCL4 (most severe secondary diagnoses). While PCCLs were not directly reported under FFS, we were able to calculate them from the secondary diagnoses reported in the patient files by using the software program 'Grouper' applied by all hospitals in Germany under PPS. However, for the reasons discussed in Section 2, there was no incentive for complete documentation under FFS, which implies that PCCLs calculated from the patient files are systematically higher under PPS compared to FFS (see Table 1).

In order to account for this problem, we proceed as follows: First, in order to keep the analysis straightforward, we treat PCCL as a binary variable by distinguishing only between cases where PCCL is zero ($P = 0$) or positive ($P = 1$). This seems to be justified as PCCL is zero in about 80% of all cases. We then implement a transformation of PCCLs observed under FFS that ensures that the average PCCL is the same under both systems. Specifically, we proceed as follows: First, we estimate a Probit model of

the probability that PCCL is positive (i.e., $P = 1$) using all cases under PPS. Table 2 shows the estimated parameters from this regression.

Insert Table 2 about here.

We then use the parameters from Table 2 to predict the probability that a case where PCCL is zero under FFS would have had a positive PCCL *if it had been treated under PPS.*[10] We denote this probability by $\widehat{P}$. Next, we assign $\widehat{P} = 1$ to all cases which have a positive PCCL (i.e., $P = 1$) even under FFS. The underlying assumption is that there are no cases where the PCCL calculated from the secondary diagnoses in the patient files is positive under FFS, but would have been zero under PPS. Given that there were no incentives for upcoding under FFS, this assumption seems reasonable. Furthermore, this procedure ensures that no case with $P = 1$ under FFS can have a lower $\widehat{P}$ than a case with $P = 0$ under FFS.

However, this procedure inevitably causes the average $\widehat{P}$ under FFS to be above the average P under PPS: Those cases for which PCCL=0 under FFS are assigned the probability of a positive PCCL predicted from the Probit estimation, so that the average $\widehat{P}$ for this group of cases is equal to the average P under PPS. At the same time, we leave cases with positive PCCL under FFS at $\widehat{P} = 1$, thus pushing the average $\widehat{P}$ for all cases under FFS above the average P of all cases under PPS. In order to balance the overall averages under both systems, we add the following final transformation $\widetilde{P}$: We leave $\widetilde{P} = \widehat{P} = P = 1$ for all cases with positive PCCL under FFS and multiply the $\widehat{P}$ for all cases with PCCL=0 under FFS with a fixed factor such that the average of the resulting $\widetilde{P}$ over all cases under FFS is the same as the average P over all cases under

[10] In other words, our procedure is based on the assumption that the correlation of PCCL with other case and hospital characteristics for a case reimbursed under PPS is an informative predictor of the probability that a case under FFS would have had a positive PCCL had it been reimbursed under PPS. In a robustness check, we will present results based on an alternative way of adjusting the PCCL for cases with $P = 0$ under FFS that does not rely on this assumption, assuming instead a fixed probability of a positive 'true' PCCL for all cases with $P = 0$ under FFS, independent of the observable case characteristics. We will explain this alternative transformation in greater detail before presenting the robustness check in Section 6.

PPS.[11] The line 'Severity of Secondary Diagnosis' in Table 1 displays descriptive statistics on this new variable $\widetilde{P}$.

Next, recall that under PPS, the hospital's reimbursement is fixed below the upper bound, so that hospitals have strong incentives to discharge patients before this threshold is reached. In line with this, table 1 show that only 5.7% are discharged after the threshold, while this holds for 16.3% of all cases under FFS.[12]

As readmissions may well depend on the specificities of the hospitals, we add hospital fixed effects and, in some regressions, the interaction of the system dummy with certain hospital characteristics such as the number of beds or the ownership structure.[13] In addition, we consider hospital-specific time trends in our robustness checks. The information on hospitals is taken from the official statistics at a federal level.

Table 1 furthermore shows that the patients' average age slightly decreased over time, and the percentage of females has increased due to the fact that, until 1996 all so-called industry health insurance companies insured only employees, but have been open for all people since then. Finally, the last row captures the different discharge categories expressed as percentages of all discharges. Apparently, the share of regular discharges has increased, while the share of discharges with a planned post-hospital treatment has decreased by almost 50%. This can be explained by the fact that, under the DRG system, post-hospital cases within the upper threshold have to be consolidated and jointly reimbursed with the inpatient case, which reduces the economic incentive for post-hospital treatments.

[11] Using N_P^{FFS} and N_P^{PPS} to denote the number of cases with observed P under FFS and PPS, respectively, and μ to denote the average $\widehat{P}$ conditional on $P = 0$ under FFS, this requirement is equivalent to $\frac{N_1^{FFS} + \widetilde{\widehat{P}} \mu N_0^{FFS}}{N_1^{FFS} + N_0^{FFS}} = \frac{N_1^{PPS}}{N_1^{PPS} + N_0^{PPS}}$, so that $\widetilde{P} = \frac{N_1^{PPS} N_0^{FFS} - N_1^{FFS} N_0^{PPS}}{(N_1^{PPS} + N_0^{PPS}) \mu N_0^{FFS}} \widehat{P}$ whenever $P = 0$ in cases under FFS.

[12] Under FFS, there was no upper bound as hospitals just needed to justify the LOS in order to get the additional costs reimbursed. Thus, we calculated how many patients under FFS were discharged after the upper threshold that applies under PPS.

[13] Using several proxies, Hull (2018) finds that larger hospitals and privately-owned hospitals provide on average higher quality. Seabury, Bognar, Xu, Huber, Commerford, and Tayama (2017) show that rural hospitals perform worse in several quality measures on stroke care.

4 Empirical Strategy

For identifying the impact of the system change on the average LOS and the readmission probability, comparing hospitals under FFS with hospitals under PPS is not sufficient as one could not tell which part of the difference is due to the system and which part due to differences in hospitals. Similarly, when comparing a cross section of hospitals before with the same cross section after the system change, the difference may be attributable to technical progress. To address these issues, we follow the general idea of the difference-in-differences approach which compares the performance difference of a treatment group and a control group in two different periods; after and before the treatment group is treated. In the simplest setting, this difference can be obtained in a regression with a period dummy (to distinguish between the period before and after treatment) and a treatment group dummy (to distinguish between the treatment group and the control group). The coefficient for the interaction of these two dummies is then the difference-in-differences estimator.

Our data differs from the simplest case by the important fact that all hospitals were eventually treated (i.e. switched from FFS to PPS), so that the distinction between a treatment and a control group cannot be applied at the hospital level. However, as we have sufficiently many pre-treatment and post-treatment observations for each hospital, we can consider all cases under PPS as treatment group and all cases under FFS as control group. The problem of unobserved heterogeneity at the hospital level can be solved by adding hospital fixed effects. Furthermore, our dataset includes information on the admission and release dates of each case, and each of these dates may have different sets of hospitals under each reimbursement system. Hence, we can replace the period dummy in the simplest difference-in-differences approach by a time trend to control for the fact that cases under PPS generally received medical care later.[14]

[14]We use a linear time trend throughout, which has the advantage that it can distinguish very finely between cases at different points in time while being very parsimonious in terms of degrees of freedom (as opposed to monthly dummies, for instance). This will be especially evident in the robustness check with hospital specific time trends. The downside is that it imposes a stronger assumption on the functional form than period dummies would. We, therefore, also looked at the main regression when using

A potential problem in our data is that hospitals had some discretion on when to adopt PPS. This casts some doubts on the interpretation of the system dummy as the average treatment effect as hospitals anticipating larger benefits from adopting PPS are likely to do so earlier. Figures 1 and 2 display mean LOS and readmission rates by year and reimbursement system, where the light grey bars indicate cases under PPS. The first group of bars represents cases in hospitals that adopted PPS in 2003, whereas cases in the last group of bars are hospitals that did so in 2005. While the patterns for LOS look similar across these groups, there seems to be some selection with regards to the readmission rate: Hospitals that seem to struggle with quality under FFS and anticipate to continue to do so after adopting PPS seem to delay the system change as long as possible. This is plausible as readmissions are costly under PPS but were fully reimbursed under FFS.

Insert Figures 1 and 2 about here.

While the figures support the intuition that the timing of the adoption of PPS is subject to selection, our estimation of the average treatment effect is independent of when each hospital chose to adopt PPS. The reason is that we have pre-treatment and post-treatment data for all hospitals, so that each hospital's treatment effect is considered in the same way when estimating the coefficient of the PPS dummy. Furthermore, in order to capture the impact of unobserved hospital characteristics on the time trend even more exactly, we present an additional regression as a robustness check in which we interact the hospital fixed effects with the time trend.

Another point that needs to be discussed is that the difference-in-differences method assumes that the dependent variables in the treatment and control groups follow a common pre-treatment trend. In our case, this means that hospitals which adopt PPS early follow the same trend as those that do so later. Comparing time trends of hospitals that adopted PPS in different years in Figures 1 and 2 does not indicate a violation of this

monthly, quarterly or annual dummies, and our main result that the effect of the reimbursement system is insignificant for length of stay and the readmission rate is confirmed, albeit with the opposite sign in the case of the readmission rate.

assumption. However, this may be due to unobserved heterogeneity in case or hospital characteristics. We, therefore, test more rigorously whether the common trend assumption is met by looking at regressions of our dependent variables LOS and readmission on the time trend and all variables included in our main regressions. In particular, we restrict attention to the time before the first hospital adopted PPS, i.e. to the years 2001-2002. In order to test the common trend assumption, we include the interaction of the time trend with a variable that indicates whether a hospital has adopted PPS earlier or later. An insignificant coefficient of this interaction then indicates that the common trend assumption is justified.

Insert Table 3 about here.

Table 3 shows that half of the hospitals adopted PPS in 2003. We hence test the common trend assumption by comparing hospitals which adopted PPS in 2003 with those that did so after 2003. Table 4 presents the results of these regressions. The interaction of the time trend with a dummy variable that indicates whether a hospital has switched to PPS after 2003 has an insignificant coefficient for both regressions for LOS (model (1)) and (model (2)). This supports the common trend assumption.

Insert Table 4 about here.

Furthermore, the fact that all hospitals had to adopt PPS eventually allows us to test whether early and late adopters had common trends *after* all hospitals had switched. As the last hospital in our dataset adopted PPS on 1st December 2005, we test for common trend after treatment by performing the same regressions from models (1)-(2) for the years 2006-2007. Models (3) and (4) of Table 4 show that there is no significant difference in the time trends between early and late adopters for these years either.

5 Results

Tables 5 and 6 show our main regression results. All regressions include dummies for primary diagnoses, a time trend, hospital fixed effects, and the calendar month in which

the patient was admitted to account for seasonal effects. Coefficients are estimated using OLS since we estimate readmission (models (2), (4) and (6)) within a linear probability model throughout.

Insert Table 5 about here.

Models (1) and (2) present the main estimations of LOS[15] and the readmission probability. The system dummy is insignificant both for LOS and the readmission rate. Thus, the observation from the descriptive statistics that patients are discharged earlier under PPS can be attributed to the significantly negative time trend rather than to the system change, and there is also no evidence for a reduction in treatment quality. This suggests that the countervailing effects discussed in section 2 offset each other, which contradicts the widespread view that the new system would lead to early discharges and potentially also to lower treatment quality. The controls for the observable case characteristics are in the expected direction: More severe cases as measured by our adjusted values for secondary diagnoses (PCCL>0; recall the procedure discussed in section 3) are kept in hospital longer and have higher readmission probabilities in both systems. Similarly, both LOS and readmission probability of patients older than 65 years are significantly larger than for younger patients.

While models (1) and (2) of Table 5 indicate that the average effect of PPS on all cases is not significantly different from zero, our main interest is how these treatment effects vary with certain case and hospital characteristics. A quick glance at Table 4 seems to support our hypothesis that the effect of adopting PPS on patients older than 65 is very different than for younger patients: When comparing model (1) of Table 4, which is the estimation of LOS based on a period in which all hospitals still used FFS, with model (3) of the same Table, which shows the result for the very same regression but based on a period in which all hospitals had already adopted PPS, the coefficient for the age dummy seems smaller in the latter case. Similarly, a positive PCCL seems to have a

[15]Recall that the LOS regressions (1), (3) and (5) do not include consolidated cases, which is why they have slightly less observations.

much smaller effect on LOS under PPS (model (3) of Table 4) than under FFS (model (1)), which would, in turn, mean that the effect of adopting PPS on LOS is smaller for cases with secondary diagnoses indicating more severe cases.

In order to test these effects more formally, we add the interaction terms between the system dummy and the age dummy, and between the system dummy and the adjusted value for PCCL, respectively, in models (3) and (5) of Table 5. Both interactions are significant at the 1-percent level.[16] The first interaction term supports our hypothesis that the LOS of patients who are less likely to have high opportunity costs reacts more to the adoption of PPS than that of younger patients. The same interaction term is insignificant for the readmission rate (see model (4)), which indicates that patients with low opportunity costs might in fact have been kept unnecessarily long in hospital under FFS. A comparable result is obtained for the interaction between the system dummy and our adjusted measure for PCCL: While the adoption of PPS reduced the LOS of more severe cases to a greater extent than that of less severe cases (model (5)), this had no impact on the readmission rate (model (6)).

Summing up, the results reported in Table 5 suggest that the new system has led to earlier discharges of elderly patients and patients with more severe secondary diagnoses without affecting the treatment quality measured by the risk of readmission within a year.

While Table 5 focused on the interaction of case characteristics (age and adjusted PCCL) with the system dummy, Table 6 reports interaction of hospital characteristics (time of switching to PPS, ownership and size) with the system dummy. Estimating system dummies separately depending on the year in which a hospital adopted PPS adds some analytical rigor to the selection issue discussed in Section 4. Model (1) in Table 6 supports the impression from Figure 1 that the pattern for LOS does not seem to depend on when hospitals switched to PPS. However, recall that Figure 2 indicated that hospitals may have self selected into different timings of adopting PPS according to the anticipated

[16]Recall that table 4 contains only the cases for years in which all cases were treated under either FFS or PPS, whereas table 6 contains all cases.

effect on readmission rates. Indeed, model (2) in Table 6 confirms that PPS reduced the readmission rate for early adopters but increased it for late adopters. This supports the intuition that hospitals expecting to do well under the new system were eager to adopt it as soon as possible, whereas hospitals that were likely to struggle with PPS sought to delay the system change.

Insert Table 6 about here.

Next, we argued in section 2 that, due to higher fluctuations in capacity utilization, smaller hospitals have more often incentives under FFS to keep patients unnecessarily long in the hospital. This hypothesis is supported by the positive interaction term of size and the system dummy in model (5) of Table 6.

Finally, models (3) and (4) can be used to analyze how ownership of a hospital affects the impact of adopting PPS. Recall that we argued in Section 2 that privately owned hospitals are likely to react stronger to financial incentives, but that there is no clear-cut theoretical prediction as to whether PPS increases or reduces the financial incentive to keep patients in hospital, due to the countervailing effects of the high-powered cost saving incentives and the negative consequences of readmission under PPS. Model (3) shows that the interaction between private ownership and the system dummy is insignificant, which gives some indication that these countervailing effects cancel out. Furthermore, as there were no financial incentives for higher quality under FFS, one would expect less quality effort in private hospitals under FFS compared to the new system. We indeed observe that the interaction of private ownership with the system dummy in model (4) is negative, but the result is to be treated with caution as the coefficient is only significant at the 10% level.

6 Robustness Checks

Re-adjusting PCCL. Recall that we accounted for the fact that PCCLs are relatively underestimated in FFS compared to PPS in a three-step-procedure: First, as there is

no incentive for upcoding under FFS, we assume that all cases with PCCL>0 under FFS would also have been assigned PCCL>0 under PPS. Second, we used the estimated probabilities of PCCL>0 under PPS to predict the probability that cases with PCCL=0 under FFS would have been assigned PCCL>0 had they been reimbursed under PPS. As these two steps together imply that the average values estimated for FFS are above the values actually observed under PPS, we multiplied the estimated values for all cases under FFS that had originally PCCL=0 with a factor that ensures that the average values are the same under both systems.

The underlying assumption of the second step of this procedure was that the probabilities for PCCL>0 estimated under PPS provide useful information also for predicting the PCCLs under FFS. A more radical and rather simple assumption is that the secondary diagnoses of FFS that would have been recorded had these cases been reimbursed under to PPS are uncorrelated with observable case characteristics.[17] If this is the case, then the best one can do is to assign all cases with PCCL0 under FFS the same probability $\widetilde{P}'$, and we get this value by ensuring that the average $\widetilde{P}'$ for all cases under FFS is equal to the ratio of cases with PCCL>0 under PPS.[18] Table 7 shows the coefficients that the regression in Table 5 would yield when using this alternative way of adjusting PCCLs. With the one exception discussed in the footnote,[19] all results are qualitatively confirmed. We also duplicated 6 with the alternative PCCL-adjustment. Here, all results are qualitatively confirmed.[20]

[17]We still assume that secondary diagnoses reported under FFS are never upcoded, so that all cases with PCCL>0 under FFS are assigned the probability $\widetilde{P}' = 1$.

[18]Recalling that N_P^{FFS} and N_P^{PPS} denote the number of cases with observed PCCL>0 ($P = 1$) and PCCL=0 ($P = 0$) under FFS and PPS, respectively, this requirement is equivalent to $\frac{N_1^{FFS}+\widetilde{P}'N_0^{FFS}}{N_1^{FFS}+N_0^{FFS}} = \frac{N_1^{PPS}}{N_1^{PPS}+N_0^{PPS}}$, so that $\widetilde{P}' = \frac{N_1^{PPS}N_0^{FFS}-N_1^{FFS}N_0^{PPS}}{(N_1^{PPS}+N_0^{PPS})N_0^{FFS}}$ for all cases with recorded PCCL0 under FFS.

[19]The only exception is that the interaction between the system dummy and the alternative measure for secondary diagnoses is now positive at the 5% level for readmissions. A reasonable explanation for this is that our procedure to upscale all cases with PCCL=0 under FFS at the same degree systematically underestimates (overestimates) PCCL for more (less) severe cases, whereas cases with higher probabilities of PCCL¿0 under PPS are in fact more severe. Thus, while the interaction term with system is not robust with respect to the adjustment of PCCL under FFS, it suggests that the procedure applied in our main regressions is more reliable.

[20]Regression table available on request.

Insert Table 7 about here.

30-day-readmission rate. Recall that the 30-day threshold is important under PPS as cases are consolidated if patients are readmitted to the same hospital within thirty days. We hence perform the same regressions for readmissions as in tables 5 and 6 for the 30-day threshold. Thereby, we distinguish between all readmissions within this period and readmissions to the same hospital only.[21] For PCCL, we use the same adjustments as in our main regressions.

Insert Table 8 about here.

Re-running the regression from table 5 with the 30-day threshold shows that all results are robust: the adjusted value for PCCL, gender and the age dummy are all highly significant in the same directions, while system itself and both interactions with system are insignificant for the readmission probability.[22] However, the impact of the years in which hospitals adopt the new system on the system's impact on readmissions is now different (cf. the interactions in model (6) in table 6 versus models (5) and (6) in table 8): In the regressions reported in table 6, we found that hospitals which switched later had significantly higher one-year-readmission rates, while the respective interaction terms are insignificant for the 30-day-threshold. This finding reinforces the importance of using a quality measure that is not directly linked to hospitals' financial incentives: As hospitals suffer financial losses only from patients readmitted within thirty days, they have strong incentives to avoid these readmissions, so that hospitals with quality issues will want to focus their efforts on avoiding readmissions particularly within this time frame. Using readmissions within 30 days as a proxy for quality would, in this case, conceal how late adopters of PPS struggle with quality and how hospitals self select into different timings of the system change.

[21] Note that, by contrast to tables 5 and 6, tables **??** and 8 restrict attention to readmissions as the readmission rate, and how it is measured, is irrelevant for regressions on LOS.

[22] Regression table available on request.

A potential objection to the results of our main regressions is that they do not account for the possibility that hospitals may follow different time paths. Hence, Tables 10 and 11 in the Appendix present regressions with hospital-specific time trends. All results remain qualitatively the same.

7 Conclusion

We have used a difference-in-differences approach to estimate the impact of introducing a high-powered incentive scheme on cost and quality provision in the healthcare sector. While we found no significant impacts of the system change on the average length of stay in hospital and the readmission rate, there are interesting impacts of patient and hospital characteristics, which are in line with the economic perspective on incentives and selection. Compared to younger patients, patients at an age above 65 are kept longer in the hospital under PPS, which we attribute to their lower opportunity costs. Patients with more severe secondary diagnoses are discharged relatively earlier under PPS compared to FFS, and the LOS increases under PPS for smaller hospitals. By contrast, all of the mentioned variables do not influence the system's change impact on readmission rates.

Turning to selection effects, our most interesting result is that hospitals that adopted PPS earlier have reduced their one year-readmission rates to a larger extent under PPS than those that switched later. This is in line with the fact that hospitals may suffer financially from readmissions under PPS, but not so under FFS. Furthermore, we do not find this effect when we substitute the one year-readmission rate by the 30-day-threshold which is decisive for case consolidations, i.e. for whether hospitals under PPS get additional money for treating patients after readmissions. This is likely to be attributable to the incentives to focus quality provisions on cases that might be readmitted within this period, so that selection becomes less important compared to incentives.

Our data allow for identifying these effects with a difference-in-differences approach, because we observe hospitals under both systems at the same time during the whole transition period, and cases under both systems for every hospital. We can then control

for unobserved heterogeneity by hospital fixed effects and for technological progress by a time trend. In contrast to most other papers analyzing quality effects, our data contain primary and secondary diagnoses, which enables us to calculate PCCLs as a measure for the severity of secondary diagnoses also for FFS. Another important advantage of our data is that we can observe readmission to any hospital, so that incentives to trigger readmissions to other hospitals under PPS in order to avoid case consolidations is no issue for our data.

Our study is subjects to a number of limitations: The first kind of limitation refers to the readmission rate as used in our analysis. While we did our best to account for the problem that incentives to report secondary diagnoses are lower under FFS, there seems to be no way to fully eliminate the problem. Furthermore, we have only data for readmissions to hospitals, but not for after discharge treatments in doctors' offices. If those treatments have systematically increased under PPS, then our empirical strategy overestimates the quality under PPS.

The second kind of limitations is that the readmission rate is the only reasonable quality proxy available from our data. Another quality measure often used for analyzing the impact of the system change is the mortality rate, for which all studies we are aware of find no significant impact of the system change (Rogers, Draper, Kahn, Keeler, Rubenstein, Kosecoff, and Brook (1990) and Kahn, Draper, and Keeler (1992)). Our data contain death as cause of discharge but we have no information on after-discharge mortality, so that using the mortality rate would not provide reliable results. This view is reinforced by findings of Cutler (1995) that the mortality rate in hospitals has increased after the system change but that there is no difference when considering longer time horizons. For Germany, the official concomitant research on the system change in Germany finds no impact but simply compares risk-adjusted mortality rates without extending to an econometric analysis (IGES (2013), pp. 401ff).[23] As the study points out, considering

[23]The study applies many additional quality indicators including e.g. institutionalized and standardized quality management, ratio of physicians with specific qualifications at all physicians, post-surgery complications at a very disaggregated level such as perioperative apoplectic strokes and patient satisfaction. Overall, it is concluded that there are no significant impacts of the system change. Patients are

only in-hospital mortality rates can be misleading as the importance of hospices may have changed over time.[24]

For post-stationary ambulant in-hospital treatment as quality measure we find a significantly negative system dummy. However, we do not view this as a reliable quality indicator as hospitals have financial incentives to avoid such treatment as there are no regular additional reimbursements. This suggests a substitution effect to medical offices and rehabilitation clinics which we cannot observe in our data.[25]

From an overall regulatory perspective, our results allow for two conclusions: First, the fact that we do not find a lower quality under PPS can most likely be attributed to the incentives to avoid case consolidations. Thus, whenever one implements a high-powered incentive scheme, one needs to think about smart devices for internalizing the externalities from lower quality provision. For instance, electricity providers in Germany pay high fines for blackouts and for less severe signals on quality reduction, while liabilities for delays in rail transportation are still so low that they hardly implement sufficient quality incentives. Second and more specifically, our results show that one needs to take case and firm characteristics carefully into account when predicting the impacts of regulatory transitions.

References

Abler, S., P. Verde, H. Stannigel, E. Mayatepek, and T. Hoehn (2011): "Effect of the introduction of diagnosis related group systems on the distribution of admission weights in very low birthweight infants," *Archives of Disease in Childhood – Fetal and Neonatal Edition*, 96(3), F186–F189.

overall equally satisfied in both systems (see also Geraedts (2017)) but satisfaction has decreased with regards to the information provided.

[24]Recent studies not dealing with the impact of a system change also emphasize the importance of longer time horizons, see Batty and Ippolito (2017) and Doyle Jr, Graves, and Gruber (2017).

[25]We replicated all specifications for readmission rates with after-hospital treatment. Regressions are available on request.

Armstrong, M., and D. E. Sappington (2006): "Regulation, competition, and liberalization," *Journal of Economic Literature*, pp. 325–366.

Batty, M., and B. Ippolito (2017): "Financial incentives, hospital care, and health outcomes: Evidence from fair pricing laws," *American Economic Journal: Economic Policy*, 9(2), 28–56.

Busato, A., and G. von Below (2010): "The implementation of DRG-based hospital reimbursement in Switzerland: A population-based perspective," *Health Research Policy and Systems*, 8(1), 1–6.

Busse, R., J. Schreyögg, and P. Smith (2008): "Variability in healthcare treatment costs amongst nine EU countries-results from the Health BASKET project," *Health Economics*, 17(S1), S1–S8.

Cheng, S.-H., C.-C. Chen, and S.-L. Tsai (2012): "The impacts of DRG-based payments on health care provider behaviors under a universal coverage system: A population-based study," *Health Policy*, 107(2), 202–208.

Cutler, D. (1995): "The Incidence of Adverse Medical Outcomes Under Prospective Payment," *Econometrica*, 63(1), 29–50.

Cutler, D. M., and R. J. Zeckhauser (2000): "The anatomy of health insurance," in *Handbook of Health Economics*, ed. by A. Culyer, and J. Newhouse, pp. 563–643. Elsevier.

Dafny, L. S. (2005): "How Do Hospitals Respond to Price Changes?," *American Economic Review*, 95(5), 1525–1547.

DesHarnais, S., E. Kobrinski, J. Chesney, M. Long, R. Ament, and S. Fleming (1987): "The early effects of the prospective payment system on inpatient utilization and the quality of care," *Inquiry*, 24(1), 7–16.

Doyle Jr, J. J., J. A. Graves, and J. Gruber (2017): *Evaluating Measures of Hospital Quality*. NBER Working Paper 23166.

Ellis, R., and T. McGuire (1996): "Hospital response to prospective payment: Moral hazard, selection, and practice-style effects," *Journal of Health Economics*, 15(3), 257–277.

Ellis, R. P., and M. M. Miller (2009): "Provider payment and provider incentives," in *Health Systems Policy*, ed. by G. Carrin, K. Buse, H. K. Heggenhougen, and S. R. Quah. Elsevier.

Faure, M., and H. Koziol (2001): *Cases on medical malpractice in a comparative perspective*. Springer.

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

Geraedts, M. (2017): "Strukturwandel und Entwicklung der Krankenhauslandschaft aus Patientensicht," in *Krankenhaus-Report*, ed. by J. Klauber, M. Geraedts, J. Friedrich, and J. Wasem. Stuttgart.

Hamada, H., M. Sekimoto, and Y. Imanaka (2012): "Effects of the per diem prospective payment system with DRG-like grouping system (DPC/PDPS) on resource usage and healthcare quality in Japan," *Health Policy*, 107(2), 194–201.

Hull, P. (2018): "Estimating hospital quality with quasi-experimental data," Discussion paper.

IGES (2013): *G-DRG-Begleitforschung gemäß §17b Abs. 8KGH, Endbericht des dritten Forschungszyklus*.

Jürges, H., and J. Köberlein (2015): "What explains DRG upcoding in neonatology? The roles of financial incentives and infant health," *Journal of Health Economics*, 43, 13–26.

Kahn, K., D. Draper, and E. Keeler (1992): "The effects of DRG-based prospective payment on quality of care for hospitalized Medicare patients: Final report," Discussion paper, RAND Corporation.

Kennedy, I., and A. Grubb (2000): *Medical law*, vol. 3. Oxford Univ Press.

List, J. A., and I. Rasul (2011): "Field experiments in labor economics," in *Handbook of Labor Economics Vol. 4 A*, ed. by D. Card, and O. Ashenfelter, pp. 103–228. Elsevier.

Louis, D., E. Yuen, M. Braga, A. Cicchetti, C. Rabinowitz, C. Laine, and J. Gonnella (1999): "Impact of a DRG-based hospital financing system on quality and outcomes of care in Italy," *Health Services Research*, 34(1 Pt 2), 405–415.

Lüngen, M., and K. Lauterbach (2003): *DRG in deutschen Krankenhäusern: Umsetzung und Auswirkungen; mit 44 Tabellen*. Schattauer.

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.

Rich, M., and K. Freedland (1988): "Effect of DRGs on three-month readmission rate of geriatric patients with congestive heart failure," *American Journal of Public Health*, 78(6), 680–2.

Rogers, W. H., D. Draper, K. L. Kahn, E. B. Keeler, L. V. Rubenstein, J. Kosecoff, and R. H. Brook (1990): "Quality of care before and after implementation of the DRG-based prospective payment system: a summary of effects," *Jama*, 264(15), 1989–1994.

Schönfelder, T., S. Balázs, and J. Klewer (2009): "Kosten aufgrund von DRG-Upcoding durch die Einführung der Diagnosis Related Groups in Deutschland," *Heilberufe*, 61(3), 77–81.

Schreyögg, J. (2017): "Vorschläge für eine anreizbasierte Reform der Krankenhausvergütung," in *Krankenhaus-Report*, ed. by J. Klauber, M. Geraedts, J. Friedrich, and J. Wasem. Stuttgart.

Seabury, S., K. Bognar, Y. Xu, C. Huber, S. R. Commerford, and D. Tayama (2017): "Regional disparities in the quality of stroke care," *American Journal of Emergency Medicine*, 35(9), 1234–1239.

Stargardt, T. (2008): "Health service costs in Europe: cost and reimbursement of primary hip replacement in nine countries," *Health Economics*, 17(S1), S9–S20.

Statistisches Bundesamt (2017): *Fallpauschalenbezogene Krankenhausstatistik (DRG-Statistik): Diagnosen, Prozeduren, Fallpauschalen und Case Mix der vollstationären Patientinnen und Patienten in Krankenhäusern.*

Ter-Martirosyan, A., and J. Kwoka (2010): "Incentive regulation, service quality, and standards in US electricity distribution," *Journal of Regulatory Economics*, 38(3), 258–273.

Yip, W., and K. Eggleston (2001): "Provider payment reform in China: the case of hospital reimbursement in Hainan province," *Health Economics*, 10(4), 325–339.

Zhang, J. (2010): "The impact of a diagnosis-related group-based prospective payment experiment: The experience of Shanghai," *Applied Economics Letters*, 17(18), 1797–1803.

Appendix

Insert Tables 9, 10 and 11 about here.

Tables and Figures

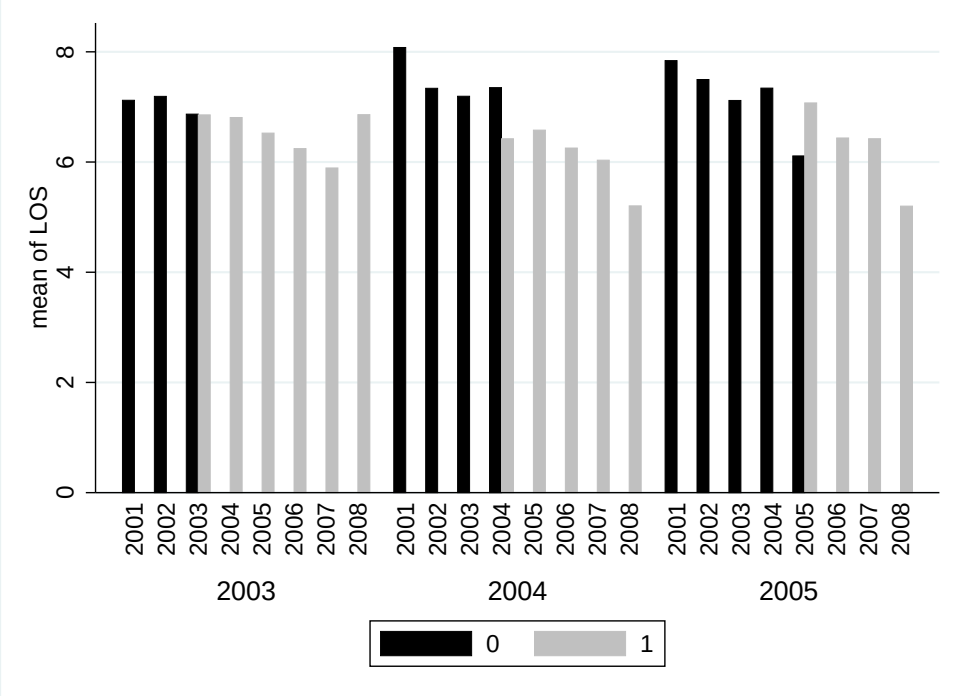

Figure 1: Mean LOS by system, year and timing of PPS adoption.

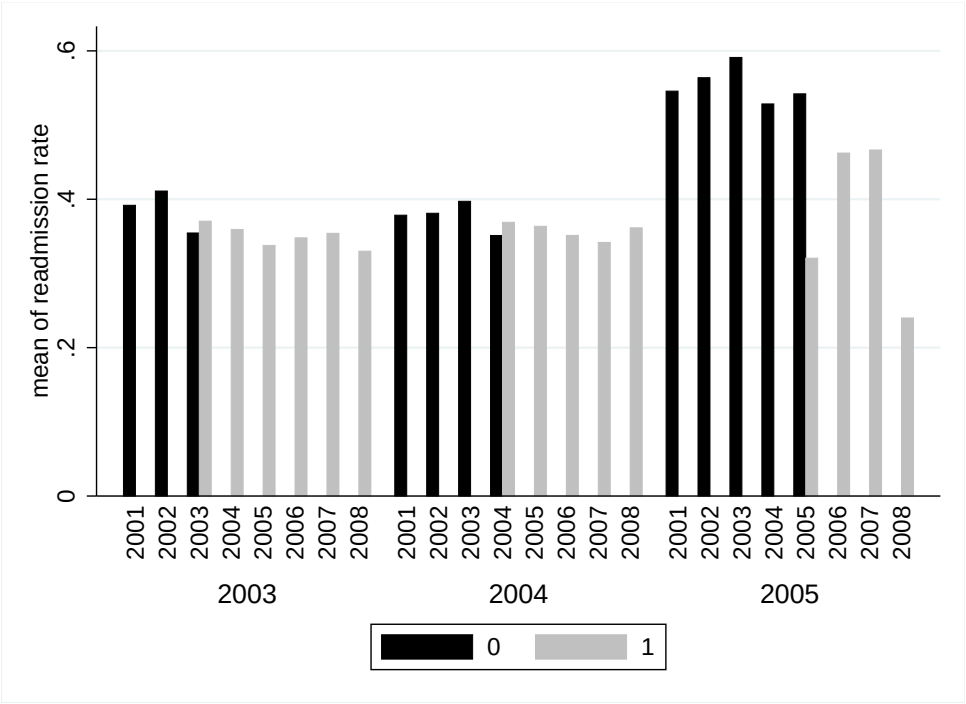

Figure 2: Mean readmission rates by system, year and timing of PPS adoption.

Table 1: Summary Statistics.

	Fee for service (FFS)		Prospective payment system (PPS)		
	Mean	Std. Dev.	Mean	Std. Dev.	t-value
Average Length of stay (LOS) (excl. consolidated cases)	7.241	8.908	6.348	7.255	11.177***
Readmission within 1 year to any hosp.	0.413	0.492	0.360	0.480	11.607***
Readmission within 30 days to same hosp.	0.099	0.298	0.082	0.275	6.066***
PCCL0	0.872	0.334	0.747	0.435	36.566***
PCCL1	0.005	0.067	0.008	0.091	-5.343***
PCCL2	0.064	0.244	0.114	0.318	-20.101***
PCCL3	0.045	0.206	0.087	0.282	-19.763***
PCCL4	0.015	0.123	0.044	0.205	-20.381***
Severity of Secondary Diagnosis	0.253	0.311	0.253	0.435	0.000
Patients discharged after upper limit	0.163	0.369	0.057	0.233	33.138***
Average number of beds	588.5	352.8	585.9	331.3	0.804
Private Ownership	0.394	0.489	0.368	0.482	-4.802***
Age (70 years = 0)	-25.7	19.5	-28.8	23.3	15.807***
Male	0.697	0.460	0.633	0.482	14.689***
Surgery	0.610	1.051	0.734	1.252	-11.899***
Discharge categories:					
Regular	0.869	0.338	0.895	0.307	-8.407***
Death	0.007	0.085	0.009	0.093	-1.694*
Rehabilitation	0.018	0.134	0.021	0.142	-1.921*
After Hospital Treatment	0.053	0.225	0.029	0.185	12.194***
Change of Hospital	0.034	0.180	0.023	0.149	6.735***
Other	0.019	0.136	0.024	0.154	-3.941***

Table 2: Probit Regression of PCCL> 0 under DRG.

Age	0.0147***
	(0.0009)
Male	0.1016***
	(0.0239)
Constant	0.3147***
	(0.0954)
N	43,071

Coefficients of Probit Regression, which includes fixed effects for hospitals, the calendar month and the primary diagnosis. Robust standard errors clustered at hospital level in parentheses. *, ** and *** denote significance at 5-percent, 1-percent and 0.1-percent levels, respectively.

Table 3: Time of Adoption of PPS by Number of Hospitals and Cases

PPS adopted in	Number Hospitals	Number Cases	Percentage of Cases
2003	27	29,491	50.44
2004	21	23,122	39.54
2005	6	5,859	10.02

Table 4: Test of the Common Trend Assumption.

	2001–2002		2006–2007	
	(1)	(2)	(3)	(4)
Dependent variable	Length of stay	Readmission	Length of stay	Readmission
Time trend	-0.0020***	0.0000	-0.0007	0.0000
	(0.0004)	(0.0000)	(0.0004)	(0.0000)
Time trend * Switch after 2003	0.0004	-0.0001	-0.0001	-0.0000
	(0.0009)	(0.0001)	(0.0004)	(0.0000)
Adjusted PCCL>0	4.4680**	0.0805*	2.4373***	0.1038***
	(1.2952)	(0.0321)	(0.1746)	(0.0093)
Older than 65	1.1706***	0.0807***	1.1485***	0.0924***
	(0.2931)	(0.0192)	(0.1593)	(0.0084)
Male	0.0777	-0.0178	0.0697	0.0290***
	(0.2958)	(0.0124)	(0.0890)	(0.0061)
Constant	3.9296***	0.4346***	5.6746***	0.1997***
	(0.9691)	(0.0770)	(0.5957)	(0.0470)
No. of observations	8,236	8,236	24,757	25,097
No. of hospitals	54	54	54	54
No. of pr. diagnoses	50	50	50	50
Primary Diagnosis FE	Yes	Yes	Yes	Yes
Calendar Month FE	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes

All regressions are OLS. Robust standard errors clustered at hospital level in parentheses. *, ** and *** denote significance at 5-percent, 1-percent and 0.1-percent levels, respectively.

Table 5: Benchmark Regressions and Impact of Case Characteristics.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Length of stay	Readmission	Length of stay	Readmission	Length of stay	Readmission
System	-0.1356	-0.0065	-0.0311	-0.0107	0.2913	-0.0099
	(0.1351)	(0.0090)	(0.1402)	(0.0096)	(0.2247)	(0.0105)
Adjusted PCCL>0	2.7613***	0.1026***	2.7688***	0.1023***	4.0904***	0.0917***
	(0.1837)	(0.0070)	(0.1842)	(0.0070)	(0.5327)	(0.0146)
System *					-1.5074**	0.0123
Adjusted PCCL>0					(0.4808)	(0.0159)
Older than 65	1.1928***	0.0903***	1.7549***	0.0676***	1.1941***	0.0903***
	(0.1264)	(0.0085)	(0.2256)	(0.0173)	(0.1252)	(0.0085)
System * Older than 65			-0.7304**	0.0294		
			(0.2148)	(0.0176)		
Male	0.0785	0.0162***	0.0826	0.0161***	0.0754	0.0163***
	(0.0743)	(0.0042)	(0.0743)	(0.0042)	(0.0741)	(0.0042)
Time trend	-0.0009***	-0.0000	-0.0009***	-0.0000	-0.0009***	-0.0000
	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Constant	6.1230***	0.2667***	6.0373***	0.2701***	5.8035***	0.2693***
	(0.2655)	(0.0264)	(0.2605)	(0.0259)	(0.2710)	(0.0256)
No. of observations	57,945	58,472	57,945	58,472	57,945	58,472
No. of hospitals	54	54	54	54	54	54
No. of pr. diagnoses	50	50	50	50	50	50
Primary Diagnosis FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes

All regressions are OLS. Robust standard errors clustered at hospital level in parentheses. *, ** and *** denote significance at 5-percent, 1-percent and 0.1-percent levels, respectively.

Table 6: Impact of Hospital Characteristics.

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Length of stay	Readmission	Length of stay	Readmission	Length of stay	Readmission
System	-0.0201	-0.0281*	-0.1583	0.0014	-0.1475	-0.0064
	(0.2077)	(0.0113)	(0.1570)	(0.0096)	(0.1352)	(0.0089)
System*Switch2004	-0.2188	0.0375**				
	(0.2305)	(0.0112)				
System*Switch2005	-0.1154	0.0383*				
	(0.3053)	(0.0158)				
System*Private			0.0653	-0.0224		
			(0.2489)	(0.0124)		
System*Size					0.0059**	-0.0000
					(0.0022)	(0.0001)
Adjusted PCCL>0	2.7583***	0.1030***	2.7604***	0.1029***	2.7630***	0.1026***
	(0.1837)	(0.0070)	(0.1839)	(0.0070)	(0.1831)	(0.0070)
Older than 65	1.1926***	0.0903***	1.1923***	0.0905***	1.1957***	0.0903***
	(0.1262)	(0.0085)	(0.1261)	(0.0085)	(0.1259)	(0.0085)
Male	0.0787	0.0162***	0.0785	0.0162***	0.0780	0.0162***
	(0.0743)	(0.0042)	(0.0743)	(0.0042)	(0.0744)	(0.0042)
Time trend	-0.0009***	-0.0000	-0.0009***	-0.0000	-0.0008***	-0.0000
	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Constant	6.2096***	0.2532***	6.1437***	0.2596***	6.0280***	0.2668***
	(0.2494)	(0.0266)	(0.2531)	(0.0267)	(0.2387)	(0.0263)
No. of observations	57,945	58,472	57,945	58,472	57,945	58,472
No. of hospitals	54	54	54	54	54	54
No. of pr. diagnoses	50	50	50	50	50	50
Primary Diagnosis FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes

All regressions are OLS. Robust standard errors clustered at hospital level in parentheses. *, ** and *** denote significance at 5-percent, 1-percent and 0.1-percent levels, respectively.

Table 7: Robustness Check: Alternative Measure for Secondary Diagnoses

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Length of stay	Readmission	Length of stay	Readmission	Length of stay	Readmission
System	-0.1237	-0.0061	0.0221	-0.0089	0.3220	-0.0153
	(0.1409)	(0.0087)	(0.1460)	(0.0094)	(0.2493)	(0.0103)
Adjusted PCCL>0	2.6307***	0.0954***	2.6610***	0.0948***	3.9684***	0.0678***
(alternative)	(0.1770)	(0.0073)	(0.1784)	(0.0073)	(0.5705)	(0.0143)
System *					-1.5488**	0.0318*
Adjusted PCCL>0					(0.5510)	(0.0156)
Older than 65	1.2167***	0.0915***	1.9948***	0.0765***	1.2257***	0.0913***
	(0.1291)	(0.0084)	(0.2220)	(0.0176)	(0.1288)	(0.0084)
System * Older than 65			-1.0139***	0.0194		
			(0.2100)	(0.0179)		
Male	0.0863	0.0166***	0.0914	0.0165***	0.0859	0.0166***
	(0.0746)	(0.0042)	(0.0745)	(0.0042)	(0.0746)	(0.0042)
Time trend	-0.0009***	-0.0000	-0.0009***	-0.0000	-0.0009***	-0.0000
	(0.0001)	(0.0000)	(0.0001)	(0.0000)	(0.0001)	(0.0000)
Constant	6.1456***	0.2682***	6.0218***	0.2705***	5.8195***	0.2749***
	(0.2682)	(0.0268)	(0.2628)	(0.0264)	(0.2729)	(0.0263)
No. of observations	57,945	58,472	57,945	58,472	57,945	58,472
No. of hospitals	54	54	54	54	54	54
No. of pr. diagnoses	50	50	50	50	50	50
Primary Diagnosis FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes

All regressions are OLS. Robust standard errors clustered at hospital level in parentheses. *, ** and *** denote significance at 5-percent, 1-percent and 0.1-percent levels, respectively.

Table 8: Robustness Check: Measures for Readmission

Dependent variable	(1) Readm. 30d	(2) Readm. 30d to same hosp.	(3) Readm. 30d	(4) Readm. 30d to same hosp.	(5) Readm. 30d	(6) Readm. 30d to same hosp.
System	-0.0007	0.0021	0.0079	0.0007	0.0057	0.0007
	(0.0073)	(0.0067)	(0.0071)	(0.0057)	(0.0058)	(0.0054)
System*Switch2004	0.0134	0.0006				
	(0.0077)	(0.0075)				
System*Switch2005	-0.0001	-0.0176				
	(0.0159)	(0.0208)				
System*Private			-0.0062	-0.0002		
			(0.0079)	(0.0102)		
System*Size					0.0000	-0.0001
					(0.0001)	(0.0000)
Adjusted PCCL>0	0.0535***	0.0354***	0.0534***	0.0354***	0.0534***	0.0353***
	(0.0074)	(0.0056)	(0.0074)	(0.0056)	(0.0074)	(0.0056)
Older than 65	0.0002	0.0067	0.0002	0.0067	0.0001	0.0066
	(0.0072)	(0.0059)	(0.0072)	(0.0059)	(0.0073)	(0.0059)
Male	0.0144***	0.0135***	0.0144***	0.0135***	0.0144***	0.0135***
	(0.0035)	(0.0022)	(0.0035)	(0.0022)	(0.0035)	(0.0022)
Time trend	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000	-0.0000
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Constant	0.0675***	0.0157	0.0714***	0.0172	0.0731***	0.0183
	(0.0109)	(0.0099)	(0.0108)	(0.0100)	(0.0112)	(0.0097)
No. of observations	58,472	58,472	58,472	58,472	58,472	58,472
No. of hospitals	54	54	54	54	54	54
No. of pr. diagnoses	50	50	50	50	50	50
Primary Diagnosis FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes

All regressions are OLS. Robust standard errors clustered at hospital level in parentheses. *, ** and *** denote significance at 5-percent, 1-percent and 0.1-percent levels, respectively.

Table 9: System Change Dates by Hospital.

nr	system change	private	Average size (mean=0)	no of cases
1.	01jan2003	1	.7814782	2609
2.	01jan2003	0	-221.3474	263
3.	01jan2003	1	-215.344	2386
4.	01jan2003	0	-221.844	951
5.	01jan2003	1	-452.098	697
6.	01jan2003	0	230.2076	2695
7.	01jan2003	0	-246.4475	95
8.	01jan2003	1	83.6284	2705
9.	01apr2003	1	-357.9182	1185
10.	01may2003	1	-420.5224	758
11.	01may2003	0	-226.6285	1083
12.	01jul2003	0	-375.5528	381
13.	01aug2003	0	-244.2047	747
14.	01aug2003	0	-233.1813	105
15.	01aug2003	0	-473.5528	120
16.	01aug2003	0	-56.90832	1499
17.	01aug2003	0	-299.732	173
18.	01aug2003	0	14.42551	184
19.	01aug2003	1	599.2391	4271
20.	01aug2003	1	-120.4715	160
21.	01aug2003	0	-230.3518	821
22.	01sep2003	0	-496.5528	301
23.	01sep2003	1	435.604	2934
24.	01sep2003	0	-355.3307	1522
25.	01oct2003	0	-162.239	102
26.	01oct2003	1	-199.345	207
27.	01oct2003	0	-472.3517	537
28.	01jan2004	0	-190.182	163
29.	01jan2004	0	-337.4944	137
30.	01jan2004	0	139.6812	718
31.	01jan2004	0	-336.5528	327
32.	01mar2004	0	-348.5528	904
33.	01apr2004	1	-433.9028	140
34.	01apr2004	0	786.3763	3084
35.	01apr2004	0	-378.4302	938
36.	01jun2004	0	-272.6604	1152
37.	01jun2004	0	-221.5822	1734
38.	01jun2004	0	-28.93138	2068
39.	01jun2004	0	-76.20374	3997
40.	01jul2004	0	-182.1454	1451
41.	01jul2004	0	-175.926	1369
42.	01jul2004	0	-122.784	1466
43.	01aug2004	0	-424.1227	386
44.	01sep2004	0	-205.3175	1836
45.	01sep2004	0	-192.1915	656
46.	01nov2004	0	-464.5628	199
47.	01nov2004	0	734.6019	291
48.	01dec2004	0	-490.4207	106
49.	01jan2005	1	-191.4406	107
50.	01jan2005	1	-291.588	369
51.	01jan2005	0	-100.5448	1750
52.	01jul2005	0	-75.19981	204
53.	01jul2005	1	-225.9949	190
54.	01dec2005	1	85.04559	3239

Table 10: Robustness Check: Hospital Specific Time Trends (1)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Length of stay	Readmission	Length of stay	Readmission	Length of stay	Readmission
System	-0.1387	-0.0086	-0.0381	-0.0130	0.2890	-0.0111
	(0.1580)	(0.0099)	(0.1606)	(0.0102)	(0.2492)	(0.0112)
Adjusted PCCL>0	2.7613***	0.1032***	2.7680***	0.1029***	4.0810***	0.0956***
	(0.1875)	(0.0069)	(0.1876)	(0.0069)	(0.5255)	(0.0151)
System *					-1.4976**	0.0086
Adjusted PCCL>0					(0.4728)	(0.0162)
Older than 65	1.1990***	0.0906***	1.7285***	0.0676***	1.1994***	0.0906***
	(0.1259)	(0.0085)	(0.2249)	(0.0177)	(0.1248)	(0.0085)
System * Older than 65			-0.6881**	0.0298		
			(0.2119)	(0.0182)		
Male	0.0814	0.0167***	0.0852	0.0165***	0.0781	0.0167***
	(0.0748)	(0.0042)	(0.0749)	(0.0042)	(0.0747)	(0.0042)
Constant	5.6051***	0.3493***	5.5579***	0.3513***	5.2879***	0.3511***
	(0.2615)	(0.0247)	(0.2620)	(0.0242)	(0.2793)	(0.0238)
No. of observations	57,945	58,472	57,945	58,472	57,945	58,472
No. of hospitals	54	54	54	54	54	54
No. of pr. diagnoses	50	50	50	50	50	50
Primary Diagnosis FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes
Hospital Specific time trends	Yes	Yes	Yes	Yes	Yes	Yes

All regressions are OLS. Robust standard errors clustered at hospital level in parentheses. *, ** and *** denote significance at 5-percent, 1-percent and 0.1-percent levels, respectively.

Table 11: Robustness Check: Hospital Specific Time Trends (2)

	(1)	(2)	(3)	(4)	(5)	(6)
Dependent variable	Length of stay	Readmission	Length of stay	Readmission	Length of stay	Readmission
System	-0.0379	-0.0346*	-0.2428	-0.0013	-0.1460	-0.0085
	(0.2132)	(0.0134)	(0.1871)	(0.0108)	(0.1513)	(0.0096)
System*Switch2004	-0.3437	0.0452*				
	(0.2862)	(0.0181)				
System*Switch2005	0.4548	0.0860***				
	(0.3544)	(0.0168)				
System*Private			0.2494	-0.0176		
			(0.2996)	(0.0210)		
System*Size					0.0064	-0.0001
					(0.0033)	(0.0003)
Adjusted PCCL>0	2.7623***	0.1039***	2.7606***	0.1033***	2.7627***	0.1032***
	(0.1897)	(0.0070)	(0.1874)	(0.0070)	(0.1879)	(0.0069)
Older than 65	1.1985***	0.0907***	1.1981***	0.0907***	1.1997***	0.0906***
	(0.1255)	(0.0085)	(0.1256)	(0.0085)	(0.1257)	(0.0085)
Male	0.0811	0.0168***	0.0812	0.0167***	0.0810	0.0167***
	(0.0747)	(0.0042)	(0.0748)	(0.0042)	(0.0749)	(0.0042)
Constant	5.5489***	0.3534***	5.5811***	0.3509***	5.6059***	0.3493***
	(0.2727)	(0.0245)	(0.2682)	(0.0243)	(0.2622)	(0.0247)
No. of observations	57,945	58,472	57,945	58,472	57,945	58,472
No. of hospitals	54	54	54	54	54	54
No. of pr. diagnoses	50	50	50	50	50	50
Primary Diagnosis FE	Yes	Yes	Yes	Yes	Yes	Yes
Calendar Month FE	Yes	Yes	Yes	Yes	Yes	Yes
Hospital FE	Yes	Yes	Yes	Yes	Yes	Yes
Hospital Specific time trends	Yes	Yes	Yes	Yes	Yes	Yes

All regressions are OLS. Robust standard errors clustered at hospital level in parentheses. *, ** and *** denote significance at 5-percent, 1-percent and 0.1-percent levels, respectively.