

The Effect of Phase 2 of the Premier Hospital Quality Incentive Demonstration on Incentive Payments to Hospitals Caring for Disadvantaged Patients

Andrew M. Ryan, Jan Blustein, Tim Doran, Marilyn D. Michelow, and Lawrence P. Casalino

Objective. The Medicare and Premier Inc. Hospital Quality Incentive Demonstration (HQID), a hospital-based pay-for-performance program, changed its incentive design from one rewarding only high performance (Phase 1) to another rewarding high performance, moderate performance, and improvement (Phase 2). We tested whether this design change reduced the gap in incentive payments among hospitals treating patients across the gradient of socioeconomic disadvantage.

Data. To estimate incentive payments in both phases, we used data from the Premier Inc. website and from Medicare Provider Analysis and Review files. We used data from the American Hospital Association Annual Survey and Centers for Medicare and Medicaid Services Impact File to identify hospital characteristics.

Study Design. Hospitals were divided into quartiles based on their Disproportionate Share Index (DSH), from lowest disadvantage (Quartile 1) to highest disadvantage (Quartile 4). In both phases of the HQID, we tested for differences across the DSH quartiles for three outcomes: (1) receipt of any incentive payments; (2) total incentive payments; and (3) incentive payments per discharge. For each of the study outcomes, we performed a hospital-level difference-in-differences analysis to test whether the gap between Quartile 1 and the other quartiles decreased from Phase 1 to Phase 2.

Principal Findings. In Phase 1, there were significant gaps across the DSH quartiles for the receipt of any payment and for payment per discharge. In Phase 2, the gap was not significant for the receipt of any payment, but it remained significant for payment per discharge. For the receipt of any incentive payment, difference-in-difference estimates showed significant reductions in the gap between Quartile 1 and the other quartiles (Quartile 2, 17.5 percentage points [$p < .05$]; Quartile 3, 18.1 percentage points [$p < .01$]; Quartile 4, 28.3 percentage points [$p < .01$]). For payments per discharge, the gap was also significantly reduced between Quartile 1 and the other quartiles (Quartile 2, \$14.92 per discharge [$p < .10$]; Quartile 3, \$17.34 per discharge [$p < .05$];

Quartile 4, \$21.31 per discharge [$p < .01$]). There were no significant reductions in the gap for total payments.

Conclusions. The design change in the HQID reduced the disparity in the receipt of any incentive payment and for incentive payments per discharge between hospitals caring for the most and least socioeconomically disadvantaged patient populations.

Key Words. Health economics, hospitals, incentives in health care, Medicare

INTRODUCTION

Hospital pay-for-performance (P4P) is scheduled for nationwide implementation in the United States as part of Medicare's Value-Based Purchasing Program in 2013 (United States Congress 2010). Despite the growing interest in P4P and its rapid expansion over the past several years, research on whether hospital-based P4P programs improve quality of care has been inconclusive (Mehrotra et al. 2009; Ryan and Blustein 2011). At the same time, concerns have emerged that P4P may have the unintended consequence of increasing the quality gap between rich and poor hospitals (Casalino et al. 2007). Lower performing hospitals tend to care for poorer patients (Gaskin et al. 2008) and are more likely to be located in economically disadvantaged areas (Blustein, Borden, and Valentine 2010). If these hospitals receive lower incentive payments in P4P programs, or face payment penalties, both of which could occur under Value-Based Purchasing, and they will be less able to fund quality improvement initiatives, potentially increasing race and income-related disparities in the delivery of care.

The Institute of Medicine and numerous analysts have suggested that P4P programs that reward both performance and improvement in performance may mitigate this unintended consequence of P4P (Casalino et al. 2007; Chien et al. 2007; Wharam et al. 2009; Ho, Moy, and Clancy 2010; Werner 2010). Rewarding quality improvement may create greater opportunity for hospitals that have lower initial quality, both because these hospitals

Address correspondence to Andrew M. Ryan, M.A., Ph.D., Assistant Professor, Department of Public Health, Joan and Sanford I Weill Medical College, 402 East 67th St., New York, NY 10065; e-mail: Amr2015@med.cornell.edu. Jan Blustein, M.D., Ph.D., is with the Robert F. Wagner Graduate School and Division of General Medicine, NYU Medical School, New York, NY. Tim Doran, M.D., is with the School of Health Sciences, University of Manchester, Manchester, UK. Marilyn D. Michelow, B.A., is with the Joan and Sanford I Weill Cornell Medical College, New York, NY. Lawrence P. Casalino, M.D., Ph.D., M.P.H., is with the Division of Outcomes and Effectiveness Research, Department of Public Health, Joan and Sanford I Weill Medical College, New York, NY.

have more room to improve and because the marginal cost of quality improvement likely increases as quality improves.

The effect of P4P on the distribution of incentive payments has important policy relevance given the impending implementation of Value-Based Purchasing. Value-Based Purchasing will redistribute hospital Medicare payments, without adding new money to the system, based on quality of care for acute myocardial infarction (AMI), heart failure, and pneumonia, creating winners and losers among hospitals. To date, no research has examined whether P4P designs that reward quality improvement, in contrast to attainment alone, reduce the gap in incentive payments between hospitals caring for populations experiencing different levels of socioeconomic disadvantage.

We address this question by exploiting a natural experiment in the Premier Hospital Quality Incentive Demonstration (HQID), a national P4P demonstration implemented by the Centers for Medicare and Medicaid Services (CMS) and Premier Inc. The HQID was implemented in two phases, changing from a design that rewarded only very high quality attainment in Phase 1 to a design that rewarded attainment over a wider range of quality performance and introduced incentives for quality improvement in Phase 2. The design used in Phase 2 is similar to the design of the upcoming Value-Based Purchasing program. We assess whether the change in the HQID incentive design reduced the gap in incentive payments received by hospitals caring for less socioeconomically disadvantaged patient populations compared with hospitals caring for more socioeconomically disadvantaged patient populations.

VARIATION IN PAYMENT INCENTIVES ACROSS THE TWO PHASES OF THE HQID

In the first phase of the HQID, from Q4 2003 to Q3 2006, the program rewarded quality attainment. A 2 percent bonus on Medicare reimbursement rates was paid to hospitals performing in the top decile of a composite quality measure for each clinical condition incentivized in the HQID (heart failure, AMI, community-acquired pneumonia, coronary-artery bypass grafting [CABG], and hip and knee replacement), and a 1 percent bonus was paid to hospitals performing in the second highest decile. During Phase 2, from Q4 2006 to Q3 2009, hospitals were eligible to receive three types of rewards: (1) a “Top Performer Award,” given to hospitals with scores in the top 20 percent of HQID hospitals in the current year; (2) an “Attainment Award,” given to

hospitals with composite scores exceeding the median from HQID hospitals 2 years prior; and (3) an “Improvement Award,” given to hospitals scoring above the median of HQID hospitals in the current year and also ranking within the top 20 percent in terms of quality improvement among HQID hospitals. Hospitals were able to receive Top Performer and Attainment Awards or receive Improvement and Attainment Awards, but they could not receive both Top Performer and Improvement Awards. In Phase 2, a fixed incentive pool was determined with 60 percent of incentive payments allocated to Top Performer and Improvement Awards and 40 percent allocated to Attainment Awards. Payments to hospitals were based on which awards they received and how many eligible discharges for these awards. The incentivized quality measures remained very similar across the two phases. In Phase 1, approximately \$8.2 million per year was disbursed in incentive payments (Premier Inc. 2008), increasing to \$12 million per year in Phase 2 (Premier Inc. 2011).

METHODS

Study Sample

Two hundred and sixty-five hospitals enrolled in the HQID at the start of Phase 1. Over the 5-year study period, 32 hospitals exited the HQID. All observations from hospitals that ultimately left the HQID were excluded from the analysis. An additional five hospitals had missing data on the Disproportionate Share Index. After these exclusions, the analysis included 229 hospitals, each with 5 years of data (3 years from Phase 1 and 2 years from Phase 2).

Data

Data on the size of incentive payments received by hospitals in the HQID are not publicly reported by Premier Inc. or CMS and therefore had to be estimated. To do this, we downloaded publicly available data from the Premier Inc. website (<http://www.premierinc.com>) on the hospitals in the HQID that received incentive payments. We used Medicare Provider Analysis and Review data (MedPar) from Q4 2003 to Q3 2006 to calculate the incentive payments for those hospital receiving awards in Phase 1. To obtain information about hospital characteristics (number of discharges, proportion of discharges from Medicaid patients, ownership status, number of beds, urban or rural location, and teaching status), we used American Hospital Annual

Survey data for financial years 2003 and 2005. We also used the CMS Impact File for financial year 2003 for information on hospitals' Disproportionate Share Index.

Measures

Our three outcome measures are (1) whether a hospital received payment for any of the incentivized conditions (AMI, heart failure, pneumonia, CABG, and hip and knee replacement); (2) total payments (in dollars) received by a hospital across all incentivized conditions; and (3) payments per discharge for the incentivized conditions. Both (2) and (3) varied across hospitals every year because hospitals had different numbers of discharges between and across years and received different yearly payments for each incentivized condition based on the type(s) of award they received during that year. We examined both total incentive payments and incentive payments per discharge because the two variables have different implications for hospitals' responsiveness to incentives. If considerations of fixed cost investments are central to hospitals' quality improvement decisions, then this would be reflected in responsiveness to total payments, as expected total payments would need to exceed the fixed costs of quality improvement should these investments be made. If marginal costs are more important to hospitals' decisions to improve quality, then payment per discharge would be more relevant, as expected payments per discharge would need to exceed the marginal cost per discharge of quality improvement activities.

The mechanics of our incentive estimation are as follows: To estimate incentive payments received by hospitals in Phase 1, we multiplied the sum of included Medicare hospital revenues (excluding Disproportionate Share, Indirect Medical Education, outlier, and pass through payments, as was done by in the HQID when calculating Phase 1 payments) for each incentivized condition by the 1 or 2 percent bonus for incentive-receiving hospitals. For Phase 2, following the procedure used for the HQID, we first identified the hospitals that received Top Performer and Improvement Awards for each condition in each year. Then, we summed the eligible discharges among hospitals receiving either of these awards across all incentivized conditions. We then calculated a dollar amount per discharge for the Top Performer and Improvement Awards by dividing \$7.2 million (the annual total designated for Top Performer and Improvement Awards) by the sum of all Medicare discharges for hospitals receiving these awards, and determined the total incentive payment for each hospital. To calculate dollar amount per discharge for Attainment

Awards, we divided \$4.8 million (the annual total designated for Attainment Improvement Awards) by the sum of eligible discharges among hospitals receiving Attainment Awards across all incentivized conditions. Eligible discharges among hospitals in the HQID that received incentive payments were identified from data downloaded from the Premier website. To calculate incentive payments per discharge, we divided total hospital incentive payments by eligible discharges across all the incentivized conditions.

We operationalize socioeconomic disadvantage with the Medicare Disproportionate Share Index (DSH). The DSH reflects the socioeconomic status of a hospital’s patients by accounting for both Medicaid admissions and admissions from patients receiving Supplemental Security Income. A higher DSH value indicates greater socioeconomic disadvantage. The DSH has been used in similar research to identify the gradient of hospital disadvantage (Jha, Orav, and Epstein 2010). We classified hospitals into quartiles based on their DSH in the first year of the HQID.

Statistical Analysis

To evaluate the presence of gaps in the study outcomes across the gradient of hospital disadvantage, we conducted joint Wald tests of the difference of each study outcome across the quartiles of the DSH, in each phase of the HQID. We then performed a difference-in-differences analysis to estimate the effect of the HQID design change on the distribution of incentive payments across the DSH quartiles. For hospital j at year t , we used generalized estimation equations (GEE) to estimate the following:

$$g\{E(Y_{jt})\} = b_0 + b_1 \text{DSH quartile}_j + b_2 \text{Phase } 2_t + b_3 (\text{DSH quartile}_j \cdot \text{Phase } 2_t),$$

where $g\{\}$ is the link function; Y is either the receipt of any incentive payment, total incentive payment, or incentive payment per discharge; DSH quartile is a vector of dummy variables representing hospitals’ quartile of the DSH in the first year of the HQID (omitting the least disadvantaged quartile); and Phase 2 is a dummy for Phase 2 of the HQID. The probability of receiving an incentive payment is binary (0/1) while the magnitude of incentive payments is characterized by a distribution with both many zero values and high values (Supporting Information, Appendix SA). Modeling a dependent variable with this distribution, similar to that of health care costs, has received extensive attention in the health services literature (Mullahy 1998). Following the

literature (Buntin and Zaslavsky 2004), we used a population averaged generalized estimating equation with a Gaussian distribution and log link to model total incentive payments and payments per discharge. A GEE model with a binomial distribution and a logistic link was used to model the probability of receiving any incentive payment.¹

The difference-in-differences estimates for Quartiles 2, 3, and 4 are the marginal effects of the coefficients contained in vector b_3 . We performed a joint Wald test of whether the difference-in-difference estimates are different from zero. This test indicates whether, overall, gaps in the outcomes were reduced in Phase 2. From the estimated models, we also generated predicted values for the dependent variable for each DSH quartile in both phases of the HQID. Because other hospital characteristics (ownership status, number of beds, urban or rural location, and teaching status) changed very little for hospitals across the two phases of the HQID, their influence on the outcomes is effectively canceled out when examining the differences in outcomes within hospitals over time. We therefore do not include other hospital characteristics in our models.

Next, we examined the extent to which differences in Phase 2 payments for hospitals with a higher DSH were a result of these hospitals receiving the new Improvement or Attainment Awards or instead were a result of these hospitals becoming more likely to receive incentive payments for top performance. Using only data from Phase 2, we modeled the probability that hospitals received each type of award (Top Performer, Attainment, and Improvement) as a function of DSH quartile by estimating separate GEE models (with a logit link) for each type of award. We then modeled the magnitude of the incentive payments received from each type of award in Phase 2 as a function of DSH quartile, again by estimating separate GEE models (with a log link) for each type of award. We then tested whether differences in these outcomes were significant in Phase 2.

In all analysis, standard errors were specified to be robust to clustering at the hospital level. All analysis was performed using *Stata* 11.0.

Sensitivity Analysis

Sensitivity analysis was performed to examine the robustness of our findings across different modeling specifications. We used the proportion of discharges from Medicaid patients as an alternative to DSH as a measure of socioeconomic disadvantage. To examine whether results were sensitive to the inclusion of major teaching hospitals—which frequently have a higher DSH

Table 1: Hospital Characteristics by Quartile of Disproportionate Share Index

<i>Characteristic</i>	<i>Quartile 1</i>	<i>Quartile 2</i>	<i>Quartile 3</i>	<i>Quartile 4</i>
<i>n</i>	58	57	57	57
Disproportionate Share Index (Mean)***	7.5	16.0	23.5	43.6
Ownership (%)				
Government-run	1.7	14.0	10.5	7.0
For-profit	0	0	0	0
Not-for-profit	98.3	86.0	89.5	93.0
Number of beds (Mean)***	223	258	322	405
Total discharges (Mean)***	11,681	11,763	15,141	18,719
Urban location (%)	84.5	73.7	78.9	84.2
Affiliated with medical school (%)	27.6	42.1	35.1	49.1
Member of council of teaching hospitals (%)**	6.9	8.8	12.3	26.3

Notes. Chi-square test is performed for categorical variables, and Wald test is performed for continuous variables. Table includes 229 hospitals. Quartile 1 hospitals have the lowest Disproportionate Share Index values, indicating the least disadvantage.

*** $p < .01$,

** $p < .05$ for test of equality across quartiles.

(Table 1) but may also have more resources to improve quality—we reran the analysis, excluding hospitals that were members of the Council of Teaching Hospitals. To examine whether our results were sensitive to the exclusion of hospitals that exited the HQID, we estimated different models that excluded only observations after these hospitals left the HQID and models with an intent-to-treat assumption, which set incentive payments to 0 after hospitals left the HQID. Finally, we examined whether our results were sensitive to classifying hospitals into DSH quartiles based on their values on the DSH Index each year, instead of their value on the DSH Index in the first year alone. We also estimated models in which the DSH Index was treated as a continuous variable, rather than grouping hospitals into quartiles based on DSH values.

RESULTS

Table 1 shows the descriptive statistics for the study hospitals in the four quartiles of the DSH. Hospitals in Quartile 1 of the DSH (least disadvantaged patient populations) had a mean DSH of 7.5 compared with 16.0, 23.5, and 43.6 for hospitals in Quartiles 2, 3, and 4, respectively. Hospitals in Quartiles 2, 3, and 4 of the DSH were more likely to be government-run and had more

beds, while hospitals in Quartile 4 had the highest proportion of membership in the Council of Teaching Hospitals.

Appendix SA shows box plots of incentive payments per discharge in Phase 1 and Phase 2 and the within-hospital change in incentive payments between the two phases of the HQID. In Phase 1, the median payment per discharge for hospitals in Quartile 1 of the DSH was \$30.21 compared with \$14.13, \$16.78, and \$6.43 for Quartiles 2, 3, and 4, respectively. However, in Phase 2, a much more uniform distribution of incentive payments per discharge was observed. The median change in incentive payments per discharge from Phase 1 to Phase 2 was $-\$2.58$ for Quartile 1 and $\$0.43$, $\$6.99$, and $\$14.85$ for Quartiles 2, 3, and 4, respectively. In other words, Phase 2 shifted incentive payments per discharge toward hospitals serving the most disadvantaged patients.

Figure 1 shows the receipt of any payment, total payments, and payment per discharge for the four quartiles of the DSH across the 5 years of the HQID. It shows that for the receipt of any payment and payment per discharge, the trends in Phase 1 were fairly flat for each of the quartiles, indicating that there were not strong trends toward higher or lower incentive payments preceding the start of Phase 2. For total payments, a modest downward trend, similar across the DSH quartiles, is observed during Phase 1. Between the end of Phase 1 and the beginning of Phase 2, there is a substantial jump for the receipt of any payment for each quartile of the DSH. Between the end of Phase 1 and the beginning of Phase 2, total incentive payments for the highest three quartiles of the DSH increased while total payments decreased for Quartile 1. Also, between Phase 1 and Phase 2, payment per discharge increased for hospitals in the two highest quartiles of the DSH but decreased for hospitals in the lowest DSH quartile. Figure 1 also shows that, during Phase 1, substantial gaps were observed for receipt of any payment, total payment, and payment per discharge across the quartiles of the DSH, but that these gaps decreased in Phase 2.

Table 2 shows the results from the GEE models evaluating the impact of the HQID design change for the receipt of any incentive payments, total payments, and payments per discharge. In Phase 1, there was a substantial gap in the receipt of any payments between Quartile 1 and the other quartiles of the DSH: Quartile 2 was 22.8 percentage points less likely to receive a payment ($p < .01$); Quartile 3 was 19.9 percentage points less likely ($p < .01$); and Quartile 4 was 32.8 percentage points less likely ($p < .01$). The difference in the receipt of any payment across the quartiles of the DSH was jointly significant ($p < .05$). There was a large, although nonsignificant gap, in the total

Figure 1: Trends in Receipt of Any Payment, Total Payment, and Payment per Admission by Disproportionate Share Index Quartiles. Note: dashed vertical line denotes period immediately preceding start of phase 2 of Hospital Quality Incentive Demonstration (HQID)

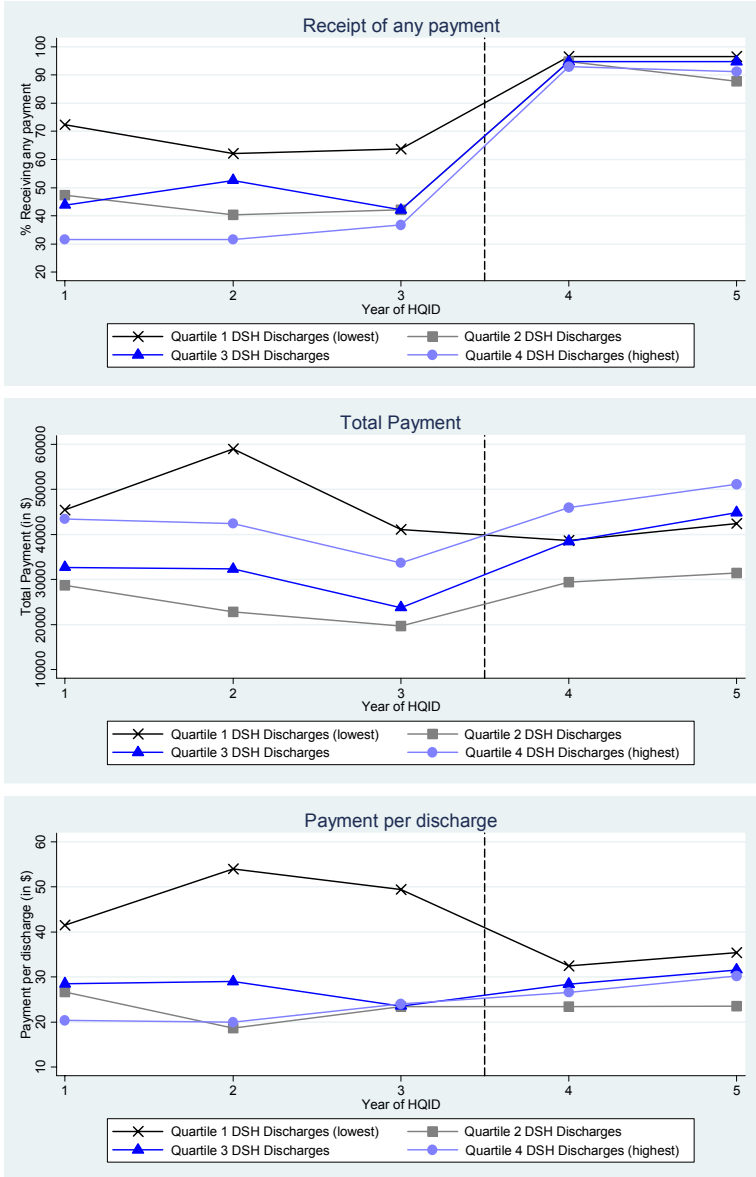


Table 2: Receipt of Any Payment, Total Payments, and Payment per Discharge by Disproportionate Share Index Quartile in Phase 1 and Phase 2 of Hospital Quality Incentive Demonstration

	Phase 1	Phase 1 Difference (Quartile _i - Quartile ₁)	Phase 2	Phase 2 Difference (Quartile _i - Quartile ₁)	Difference-in- Differences
Receipt of any payment ^{†,§}					
Quartile 1 (reference)	66.1 (4.5)	-	96.6	-	-
Quartile 2	43.3 (4.9)	-22.8 (6.7)***	91.2 (3.3)	-5.3 (3.9)	17.5 (7.0)**
Quartile 3	46.2 (4.8)	-19.9 (6.6)***	94.7 (2.7)	-1.8 (3.4)	18.1 (6.9)***
Quartile 4	33.3 (4.8)	-32.8 (6.6)***	92.1 (3.2)	-4.4 (3.8)	28.3 (7.0)***
Total payment [†]					
Quartile 1 (reference)	\$48,482 (13,024)	-	\$40,543 (5,041)	-	-
Quartile 2	\$23,761 (5,082)	-\$24,721 (13,980)*	\$30,450 (3,549)	-10,093 (6,165)	14,629 (12,199)
Quartile 3	\$29,617 (5,525)	-\$18,865 (14,147)	\$41,656 (4,090)	1,114 (6,491)	19,979 (12,541)
Quartile 4	\$39,843 (10,758)	-\$8,639 (16,892)	\$48,568 (5,762)	8,025 (7,656)	16,664 (14,978)
Payment per discharge ^{†,‡,¶}					
Quartile 1 (reference)	\$48.32 (6.49)	-	\$33.95 (3.03)	-	-
Quartile 2	\$22.92 (3.42)	-\$25.40 (7.34)***	\$23.47 (2.08)	-\$10.48 (3.67)***	\$14.92 (7.83)*
Quartile 3	\$27.03 (4.04)	-\$21.28 (7.65)***	\$30.00 (2.64)	-\$3.94 (4.02)	\$17.34 (8.32)**
Quartile 4	\$21.48 (4.35)	-\$26.84 (7.81)***	\$28.41 (2.78)	-\$5.53 (4.11)	\$21.31 (8.20)***

Notes: Table includes 229 hospitals each with 5 years of data. Quartile 1 hospitals have the lowest Disproportionate Share Index values, indicating the least disadvantage. Standard errors are shown in parentheses.

*** $p < .01$,

** $p < .05$,

* $p < .10$ for test of differences for levels and difference-in-differences estimates between a given quartile and Quartile 1.

† $p < .05$ for joint test of difference across quartiles of the Disproportionate Share Index in Phase 1.

‡ $p < .05$ for joint test of difference across quartiles of the Disproportionate Share Index in Phase 2.

§ $p < .05$ for joint test of whether difference-in-difference estimates are equal to 0.

¶ $p < .10$ for joint test of whether difference-in-difference estimates are equal to 0.

payments between Quartile 1 and the other quartiles in Phase 1. For payment per discharge, the Phase 1 gap was significant for all quartiles (Quartile 2, -\$25.40 per discharge [$p < .01$]; Quartile 3, -\$21.28 per discharge [$p < .01$]; Quartile 4, -\$26.84 per discharge [$p < .01$]) and was also jointly significant across the quartiles of the DSH.

Table 2 also shows that, in Phase 2, the gap in the receipt of any incentive payment was reduced between Quartile 1 and of the other quartiles of the DSH, such that the differences were no longer jointly significant. However, in Phase 2, the gap in total payments was jointly significant across the DSH quartiles and the gap in payment per discharge remained significant between Quartile 1 and Quartile 2, and also remained jointly significant across the DSH quartiles.

The difference-in-differences estimates in Table 2 show the reduction in the gap between Quartile 1 and each of the other quartiles between Phase 1 and Phase 2. For the receipt of any incentive payment, the reduction was large and significant between Quartile 1 and the other quartiles (Quartile 2, 17.5 percentage points [$p < .05$]; Quartile 3, 18.1 percentage points [$p < .01$]; Quartile 4, 28.3 percentage points [$p < .01$]) and the joint test that the difference-in-difference estimates equaled 0 was rejected ($p < .05$). For total payments, the reduction in the gap was not statistically significant between Quartile 1 and the other quartiles, nor was the joint test of the gap reduction rejected. However, the difference-in-difference estimates for payment per discharge were significant for Quartiles 3 and 4 (Quartile 3, \$17.34 per discharge [$p < .05$]; Quartile 4, \$21.31 per discharge [$p < .05$]), although a joint test that the difference-in-differences estimates equaled 0 was rejected only at $p < .10$.

Table 3 shows the results from GEE models evaluating the probability of hospitals receiving each type of incentive award and the magnitude of incentive payments for each type of award across the quartiles of the DSH in Phase 2 of the HQID. Hospitals in Quartile 1 were significantly more likely than hospitals in Quartiles 2 and 4 to receive the Top Performer Award. Hospitals in Quartile 1 also received larger incentive payments per discharge for Top Performer Awards (\$14.97 per discharge) than hospitals in all other quartiles: Quartile 2, \$7.60 per discharge ($p < .01$); Quartile 3, \$10.18 per discharge ($p < .10$); and Quartile 4, \$6.95 per discharge ($p < .05$). Table 3 also shows that, for the Attainment Award, although there was no difference in the receipt of any payment or total payments across DSH quartiles, hospitals in Quartile 1 received higher payments per discharge (\$12.21 per discharge) than hospitals in Quartile 2 (\$9.98 per discharge, $p < .05$) and Quartile 4 (\$10.13 per discharge, $p < .10$). For the Improvement Award, while the receipt of any

Table 3: Probability of Receiving Different Types of Incentive Payments and Magnitude of Incentive Payments (in \$) by Disproportionate Share Index Quartile in Phase 2 of Hospital Quality Incentive Demonstration

	Top Performer Award	Top Performer Award Difference (Quartile _i - Quartile ₁)	Attainment Award	Attainment Award Difference (Quartile _i - Quartile ₁)	Improvement Award	Improvement Award Difference (Quartile _i - Quartile ₁)
Receipt of any payment						
Quartile 1 (reference)	578 (5.7)	-	96.6 (2.1)	-	42.2 (5.4)	-
Quartile 2	36.0 (5.5)	-21.8 (7.9)**	91.2 (3.3)	-5.3 (3.9)	38.6 (4.7)	-3.6 (7.2)
Quartile 3	46.5 (5.8)	-11.3 (8.1)	94.7 (2.7)	-1.8 (3.4)	40.4 (5.2)	-1.9 (7.5)
Quartile 4	43.9 (5.6)	-13.9 (8.0)**	92.1 (3.2)	-4.4 (3.8)	51.8 (5.4)	9.5 (7.7)
Total payment						
Quartile 1 (reference)	\$17,499 (3,572)	-	\$15,728 (1,774)	-	\$7,316 (1,584)	-
Quartile 2	\$8,176 (1,839)	-\$9,322 (4,018)**	\$14,215 (1,472)	-\$1,513 (2,305)	\$8,059 (1,937)	\$743 (2,502)
Quartile 3	\$14,446 (3,109)	-\$3,053 (4,736)	\$17,070 (1,527)	\$1,341 (2,341)	\$10,141 (2,015)	\$2,825 (2,563)
Quartile 4	\$11,796 (2,711)	-\$5,702 (4,484)	\$18,601 (2,068)	\$2,873 (2,725)	\$18,170 (3,742)	\$10,855 (4,064)***
Payment per discharge ^{*,†,‡}						
Quartile 1 (reference)	\$14.97 (2.27)	-	\$12.21 (0.82)	-	\$6.77 (1.29)	-
Quartile 2	\$760 (1.61)	-\$7.37 (2.78)***	\$9.98 (0.50)	-\$2.23 (0.96)**	\$5.89 (1.21)	-\$0.88 (1.77)
Quartile 3	\$10.18 (1.66)	-\$4.79 (2.81)*	\$11.61 (0.68)	-\$0.60 (1.06)	\$8.21 (1.68)	\$1.44 (2.12)
Quartile 4	\$6.95 (2.14)	-\$8.02 (3.11)**	\$10.13 (0.71)	-\$2.09 (1.09)*	\$11.34 (1.60)	\$4.57 (2.06)**

Notes: Table includes 229 hospitals each with 2 years of data. Quartile 1 hospitals have the lowest Disproportionate Share Index values, indicating the least disadvantage. Standard errors are shown in parentheses.

*** $p < .01$,

** $p < .05$,

* $p < .10$ for test of differences for levels between a given quartile and Quartile 1.

† $p < .05$ for joint test of difference across quartiles of the Disproportionate Share Index for Top Performer Award.

‡ $p < .05$ for joint test of difference across quartiles of the Disproportionate Share Index for Attainment Award.

§ $p < .05$ for joint test of difference across quartiles of the Disproportionate Share Index for Improvement Award.

payment was similar across the quartiles of Medicaid discharges, hospitals in Quartile 1 (\$6.77 per discharge) had smaller total payments and smaller payments per discharge than hospitals in Quartile 4 (\$11.34 per discharge, $p < .05$).

Sensitivity analysis (1) using the proportion of Medicaid discharges instead of the Disproportionate Share Index to operationalize socioeconomic disadvantage; (2) excluding hospitals that were members of the Council of Teaching Hospitals; (3) using different exclusions and analysis assumptions for hospitals that ultimately exited the HQID; and (4) allowing DSH quartiles to be time varying and treating the DSH index as a continuous variable, resulted in qualitatively similar findings as the results shown in this paper (see Appendix SB).

DISCUSSION

Financial incentive schemes ultimately aim to improve the quality of care delivered to patients. Rewarding only the top performing providers, as was the approach in Phase 1 of the HQID, introduces competition and incentivizes providers to aspire to the highest level of performance. Quality of care, however, is partly dependent on the characteristics of patient populations and on historical levels of funding. Top performer awards have the potential to systematically disadvantage hospitals with more deprived patient populations and with lower baseline performance levels. Conventional incentive schemes therefore risk disincentivizing these hospitals and denying them resources to invest in quality improvement, thereby widening existing disparities in quality of care.

With the change in incentive design in Phase 2, a larger and more diverse group of hospitals received incentive payments in the HQID. From Phase 1 to Phase 2, there were statistically significant reductions in disparities in the receipt of any incentive payment and incentive payments per discharge between hospitals with the least disadvantaged and most disadvantaged patient populations. This is consistent with the aims of the change in the program made in Phase 2 (Premier Inc. 2010) and goes some way to meeting policy makers' concerns that hospital P4P disproportionately benefits providers serving less disadvantaged patient populations. Analysis of Phase 2 payments also showed that of the three classes of awards, incentive payments for improvement made up the largest share of total payments to hospitals serving the most disadvantaged patient populations. Again, this is consistent with the

aims of program sponsors and the wider policy making community. While the narrowing in the gap for total payments was not significant, this was largely a result of high variance in the payment data and the resulting limited power to detect effects. Nonetheless, despite the narrowing of incentive gaps, a significant gap persisted in Phase 2 for incentive payments per discharge across the gradient of hospital disadvantage.

To our knowledge, this is the first study to evaluate whether P4P designs that reward improvement and quality across a broader range of performance, as opposed to designs that reward high performance alone, benefit hospitals caring for socioeconomically disadvantaged patients. A prior study of the performance of hospitals in Phase 1 of the HQID found that hospitals that cared for more disadvantaged patient populations (as measured by the Disproportionate Share Index) reported worse quality at baseline but tended to improve quality more over time (Jha, Orav, and Epstein 2010). By the end of Phase 1, mean performance was statistically indistinguishable between high-DSH and low-DSH hospitals. Rather than evaluating performance scores, our study examined actual incentive payments and assessed whether the Phase 2 change in incentive design led to a change in the distribution of these payments. As we have argued, payment redistribution has important fiscal consequences, which in turn may influence a hospital's future ability to improve quality.

There are two reasons why the change in incentive design could have affected the distribution of incentive payments. First, the design change may have engendered a behavioral response toward greater quality improvement among hospitals with greater socioeconomic disadvantage, which frequently began the HQID with lower quality. Second, the design change may have not resulted in a behavior change, but by creating different criteria for incentive payments, redistributed a greater proportion of available incentive payments to hospitals with greater socioeconomic disadvantage. While not addressed directly in this study, the narrowing of the gap in incentive payments does not appear to be a result of a behavioral response among hospitals: recent evidence suggests that initially lower performing hospitals did not in fact improve more in response to the Phase 2 incentives than initially higher performing hospitals (Ryan, Blustein, and Casalino in press). Instead, the reductions in the incentive gaps were more likely the result of the change in the criteria that linked performance to payments.

Our study has important implications for Medicare's impending implementation of hospital-based P4P through the Value-Based Purchasing program. While the design for Value-Based Purchasing differs from Phase 2 of the

HQID—the HQID used discrete incentive thresholds to allocate incentives, while Value-Based Purchasing will use a system based on a continuous “exchange function” to translate quality into incentive payments—the two designs are similar in their emphasis on both quality attainment and improvement, making incentive payments attainable for most hospitals (Centers for Medicare & Medicaid Services 2011). Our findings suggest that under Value-Based Purchasing, there may be disparities in incentive payments that work against hospitals treating more disadvantaged patients. We estimate that hospitals with the least disadvantaged patients received approximately 20 percent more in incentive payments per discharge than hospitals with the most disadvantaged patients in Phase 2 of the HQID. However, our analysis suggests that this gap is much smaller than it would have been had payments been disbursed based on high-quality attainment alone.

This study has several limitations. The first of these is statistical power: with only 229 hospitals included in the analysis, there was limited ability to detect differences across quartiles defined by hospital disadvantage, particularly given the high variability of total payments and payments per discharge. This might explain our failure to detect a narrowing of the gap in total payments. Second, while data on whether a hospital received an incentive payment came directly from Premier Inc., our measure of the magnitude of payments was not reported by Premier Inc. Rather, we calculated this measure using published formulas from Premier Inc. and based on discharge data from the Medicare program. We estimated that \$25.3 million was distributed in bonuses in Phase 1 compared with the actual figure of \$24.1 million published by Premier Inc., which suggests that our calculations were reasonably accurate.

A third set of caveats pertains to generalizability. Participants in the Demonstration were self-selected and were likely more motivated to improve quality than the typical hospital in the United States. In addition, the incentive structure in Phase 2—the specifics of the way in which “top performance,” “attainment,” and “improvement” were defined—are unique to the HQID and differ from “improvement” and “attainment,” as defined under Medicare’s proposed hospital Value-Based Purchasing program (Centers for Medicare & Medicaid Services 2011). Medicare’s program, which will use a different incentive design and apply to all acute care hospitals in the United States, may have different motivational and distributional consequences across the gradient of socioeconomic disadvantage.

In addition, differences in the distribution of incentive payments between phases of the HQID may have resulted from another event occurring

at the same time as the change in incentive design or instead could reflect a continuation of an existing trend toward convergence in incentive payments across DSH index quartiles. However, Figure 1 shows that, in general, there were not strong trends toward convergence, or divergence, during Phase 1. It is not obvious what other event occurring at the same time as the start of Phase 2 of the HQID would affect the distribution of incentive payments.

Implications

Our findings lend support to the widely held belief that rewarding high performance leads to higher payments among hospitals serving less disadvantaged patient populations, but that adding incentives for improvement can yield a more diverse set of winners under P4P. Rewarding improvement may help mitigate the negative distributional impacts of P4P that threaten to widen the gap in quality of care between advantaged and disadvantaged patients. At the same time, rewarding improvement is not a panacea (Borden and Blustein 2012): gaps in incentive payments per discharge continued after the HQID's change in incentive design. It will be important to monitor future P4P efforts to understand how different reward structures differently impact performance, and—just as important—how those structures impact the distribution of incentive payments.

ACKNOWLEDGMENTS

Joint Acknowledgment/Disclosure Statement: The authors thank Xuming Sun for programming support and data management and Jayme Mendelsohn for research assistance. Funding for Dr. Ryan was supported by a career development award from the Agency for Healthcare Research and Quality (1 K01 HS018546-01A1).

Disclosures: None.

Disclaimers: None.

NOTE

1. When estimating a model with a balanced panel, one variable with only cross-sectional variation (DSH quartile), one variable with only time-series variation (Phase 2), and the interaction between these variables, linear regression, fixed effects,

random effects, and population averaged models (with GEE) all result in identical point estimates for the difference-in-differences estimates. The only difference between these models is in their estimates of standard errors, and the fact that the DSH quartile dummies drop out of fixed effects models because these variables are time-invariant.

REFERENCES

- Blustein, J., W. B. Borden, and M. Valentine. 2010. "Hospital Performance, the Local Economy, and the Local Workforce: Findings from a US National Longitudinal Study." *PLoS Med* 7 (6): e1000297.
- Borden, W. B., and J. Blustein. 2012. "Valuing Improvement in Value Based Purchasing." *Circulation: Cardiovascular Quality and Outcomes* DOI: 10.1161/CIRCOUTCOMES.111.962811.
- Buntin, M. B., and A. M. Zaslavsky. 2004. "Too Much Ado about Two-Part Models and Transformation? Comparing Methods of Modeling Medicare Expenditures." *Journal of Health Economics* 23 (3): 525–42.
- Casalino, L. P., A. Elster, A. Eisenberg, E. Lewis, J. Montgomery, and D. Ramos. 2007. "Will Pay-for-Performance and Quality Reporting Affect Health Care Disparities?" *Health Affairs (Millwood)* 26 (3): w405–14.
- Centers for Medicare & Medicaid Services. 2011. "Medicare Program; Hospital Inpatient Value-Based Purchasing Program. Final Rule." *Federal Register* 76 (88): 26490–547.
- Chien, A. T., M. H. Chin, A. M. Davis, and L. P. Casalino. 2007. "Pay for Performance, Public Reporting, and Racial Disparities in Health Care: How Are Programs Being Designed?" *Medical Care Research and Review* 64 (suppl 5): 283S–304S.
- Gaskin, D. J., C. S. Spencer, P. Richard, G. F. Anderson, N. R. Powe, and T. A. Laveist. 2008. "Do Hospitals Provide Lower-Quality Care to Minorities Than to Whites?" *Health Affairs (Millwood)* 27 (2): 518–27.
- Ho, K., E. Moy, and C. M. Clancy. 2010. "Can Incentives to Improve Quality Reduce Disparities?" *Health Services Research* 45 (1): 1–5.
- Jha, A. K., E. J. Orav, and A. M. Epstein. 2010. "The Effect of Financial Incentives on Hospitals That Serve Poor Patients." *Annals of Internal Medicine* 153 (5): 299–306.
- Mehrotra, A., C. L. Damberg, M. E. Sorbero, and S. S. Teleki. 2009. "Pay for Performance in the Hospital Setting: What Is the State of the Evidence?" *American Journal of Medical Quality* 24 (1): 19–28.
- Mullahy, J. 1998. "Much Ado about Two: Reconsidering Retransformation and the Two-Part Model in Health Econometrics." *Journal of Health Economics* 17 (3): 247–81.
- Premier Inc. 2008. "About the Hospital Quality Incentive Demonstration (HQID)" [accessed on June 13, 2011]. Available at <http://www.premierinc.com/quality-safety/tools-services/p4p/hqi/index.jsp>.
- Premier Inc. 2010. "Medicare Value-Based Purchasing Project Shows Any Hospital Can Achieve Positive Results, Regardless of Size, Location or Patient-Payer

- Mix" [accessed on June 15, 2011]. Available at <http://www.premierinc.com/about/news/09-aug/DSH-pressrelease080609.jsp>.
- Premier Inc. 2011. "Model Hospital Value-Based Purchasing Program Continues to Improve Patient Outcomes" [accessed on June 13, 2011]. Available at <http://www.premierinc.com/quality-safety/tools-services/p4p/hqi/index.jsp>.
- Ryan, A. M., and J. Blustein. 2011. "The Effect of the MassHealth Hospital Pay-for-Performance Program on Quality." *Health Services Research* 46 (3): 712–28.
- Ryan, A. M., J. Blustein, and L. P. Casalino. In press. "Medicare's Flagship Test of Pay-for-Performance Did Not Spur More Rapid Quality Improvement Among Low-Performing Hospitals." *Health Affairs*.
- United States Congress. 2010. "Patient Protection and Affordability Care Act." HR 3590, Public Law 111-148.
- Werner, R. M. 2010. "Does Pay-for-Performance Steal from the Poor and Give to the Rich?" *Annals of Internal Medicine* 153 (5): 340–1.
- Wharam, J. F., M. K. Paasche-Orlow, N. J. Farber, C. Sinsky, L. Rucker, K. J. Rask, M. K. Figaro, C. Braddock III, M. J. Barry, and D. P. Sulmasy. 2009. "High Quality Care and Ethical Pay-for-Performance: A Society of General Internal Medicine Policy Analysis." *Journal of General Internal Medicine* 24 (7): 854–9.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Appendix SA1: Author Matrix.

Appendix SA: Distribution of Incentive Payments.

Appendix SB: Sensitivity Analysis.

Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.