Do changes in reimbursement fees affect hospital prioritization?

Hans Olav Melberg
Department of Health Management and Health Economics,
University of Oslo

Kine Pedersen
Department of Health Management and Health Economics,
University of Oslo

UNIVERSITY OF OSLO
HEALTH ECONOMICS
RESEARCH NETWORK

Working paper 2015: 1

HERO

Do changes in reimbursement fees affect hospital prioritization?

January 2015

Hans Olav Melberg and Kine Pedersen

Department of Health Management and Health Economics, University of Oslo

Corresponding author, e-mail: h.o.melberg@medisin.uio.no

Keywords: DRG, activity based payment system, distortions, effect of financial incentives, priorities

JEL-Classification: I10, I12, I18

Abstract

It has been argued that activity based payment systems make hospitals focus on the diagnostic groups that are most profitable given costs and reimbursement rates. This article tests the hypothesis by exploring the relationship between changes in the DRG reimbursement rates and changes in the number of registered treatment episodes for all DRG-codes and hospitals in Norway between 2006 and 2013. The results show that hospitals increase the number of admissions in a diagnostic group when the reimbursement rate is increased, and decrease it (or increase it less) when DRG-rates are reduced. Across all treatments, during all years included in our primary analysis, the increase in volume is about four times as large for DRG-categories with increased reimbursement compared to DRGs with decreased reimbursement rates. However, the results also show that the number of treatment episodes in a diagnostic category is affected by many other factors in addition to the economic incentives.

Health Economics Research Network at the University of Oslo

ISSN 1501-9071 (print version.), ISSN 1890-1735 (online), ISBN 978-82-7756-245-2

Introduction

Many hospitals are paid partly using an activity-based system of fees for each treatment episode classified in a specific diagnostic related group (DRG). For instance, Medicare in the US pays hospitals according to a severity weighted version of the DRG (Vertrees, Averill, Eisenhandler, Quain, & Switalski, 2013). Similarly, many hospitals in Europe are partly financed using variants of payment based on volume measured by the DRG system (Busse R, 2011; Mikkola, Keskimäki, & Häkkinen, 2002), and several properties of these payment systems have been explored in a large literature (Barnum, Kutzin, & Saxenian, 1995; Jegers, Kesteloot, De Graeve, & Gilles, 2002). For example, it has been argued that the system will encourage up-coding of patients into DRG-groups which are most profitable (Kuhn & Siciliani, 2008), that it will lead hospitals to focus on patients which are relatively easy to treat or underprovision of services to patients with large needs (Ellis & McGuire, 1986; Martinussen & Hagen, 2009) and that hospitals will focus on the diagnostic groups that are most profitable given costs and reimbursement rates (Dafny, 2005; Ellis, 1998; Lægreid & Neby, 2012).

In this paper, we test the last hypothesis by estimating to what extent changes in the reimbursement for different DRGs are associated with changes in the number of treatment episodes. First, we develop a theoretical framework for the relevance of the DRG-payment to hospital decisions. Next, we specify the empirical strategy used to identify the causal effect in the theoretical framework, and estimate the effect using data on reimbursement rates and the number of hospital stays for all DRG-categories in Norway between 2006 and 2013. Finally, we report several tests of robustness and discuss potential complications of our findings.

The main result from our analysis is that hospitals increase the number of admissions within a DRG when its' reimbursement rate is increased, and decrease it (or increase it less) when DRG rates are reduced. Across all treatments, during all years included in our primary analysis, the increase in volume is about four times as large for DRG-categories with increased reimbursement compared to DRGs with decreased reimbursement rates. In absolute terms the effect is smaller: A ten percentage point increase in reimbursement for a DRG is predicted to increase volume by almost one percentage point.

1. Background

Even though the relationship between payment systems and hospital behavior has been extensively studied, the system-wide empirical effect of financial incentives on priority setting between different diagnostic groups has received less attention. Previous studies have often focused on a single diagnostic group, such as a large increase or reduction in the treatment of a specific diagnosis following a change in the reimbursement rate or the cost of treatment. For instance, one study showed that during the period of 1999-2002 the number of treatments for sleep apnea increased by 110% despite the fact that the proportion of patients with indication for surgery remained stable (Kvaerner, 2004). Another study explored changes in reporting practice in neonatology following the introduction of activity-based financing, and found significant changes in the diagnostic coding that indicated a change in coding practice and not a shift in disease pattern (Hansen, 2005). Although useful, these studies may exaggerate the effect of financial incentives because they focus on some striking cases and not the overall picture.

To avoid generalizing from single case studies, another approach in the literature has been to focus on upcoding, and in particular, the balance between admissions for diagnoses which are closely related (Carter, Newhouse, & Relles, 1990; Silverman & Skinner, 2004; Steinbusch, Oostenbrink, Zuurbier, & Schaepkens, 2007). For instance, some diseases have separate reimbursement codes for the same condition depending on whether it is classified as with or without complications, thereby providing an opportunity for the physicians and the coder to have some discretion in the choice of DRG.By examining how coding practice changes in relation to relative changes in reimbursement rates, it is possible to get a better overall picture of how hospitals react to incentives. In this case, the focus is mainly on changes in the coding practice, and the concern is that the hospitals are gaming the system by using the most profitable DRG. However, while changes in related DRGs are useful for identifying changes in reporting practice as a response to changes in DRG, it does not capture the extent to which different DRG reimbursements may change the actual priority given to one type of patients compared to another.

In contrast to the pure focus on changes in reporting practice, our aim is to also capture the extent to which the hospitals change the level of activity in different diagnostic groups as a response to changes in the reimbursement rates, and not just the coding practice. The extent to which changes in reimbursement rates influence priorities is relevant for the design of financial systems in general, but is of particular interest in systems that are officially intended to be neutral concerning priority-setting. This is the case in Norway, where the reimbursement fee is supposed to incentivize efforts to increase overall volume and stimulate costs reductions, but not to change medical priorities of whether one should provide more or less treatments in the different diagnostic groups. The Norwegian health authorities have stated that *"the main aim is to make the funding system as neutral as possible in terms of decisions regarding choice of form of treatment"* (Norwegian Directorate of Health, 2007, p. 57).Therefore, our aim was to investigate whether the Norwegian payment system functions according to policy-makers' explicit aims, or if hospitals prioritize treatment activities based on reimbursement rates.

2. Theory, Method and Data

2.1 Theoretical framework

Assume that the decision problem for a hospital is to maximize the production of health utility within a given budget constraint by determining the number of treatments to perform (x) within different diagnostic groups (i). Income is composed of a fixed transfer (k) and a variable income based on the number of interventions (x) performed in the different DRGs (i), each with a reimbursement rate of r_i. Formally we have:

(1) $\text{Max } \mathrm{U}(x_i) \text{ given } k + r_i x_i = c\ x_i$

In this framework, the price the hospital pays for a treatment episode in a specific diagnostic group is the difference between the cost of a treatment in that diagnostic group and the reimbursement:

(2) $p_i = c_i - r_i$

For the sake of simplicity we assume that all hospitals are identical in their costs structures and utility functions.[1] In such a system, the choice of how many DRGs to perform (i.e., the optimal level of a given DRG), is a function of the marginal importance the hospital attaches to performing the given procedure and the marginal price which is the difference between the marginal cost and the reimbursement rate:

$$x_i^* = f(U_i', c_i', r_i \ )$$

The reimbursement level is assumed to be calculated at regular intervals based on the estimated average cost of the intervention. In this sense the system is cost-neutral: The authorities do not use the reimbursement rates as a tool for influencing the hospitals' priorities between different treatments. All changes are based on changes in estimated costs, and not a desire to make the hospital increase production in a specific DRG. In this system, where the reimbursement is equal to the cost, financial incentives are cancelled out and only the marginal utility associated with the different interventions will matter when prioritizing between different disease groups. As long as reimbursements reflect costs, the optimal choice is then reduced to a function of the marginal utility of the various interventions:

$$x_i^{**} = f(U_i')$$

As long as the reimbursement is equal to the marignal cost, there should be no correlation between changes in reimbursement and the number of interventions in the different DRGs. However, reimbursements will in practice be imperfectly related to costs because the rates are updated with a lag: During a given year actual costs may change while the reimbursement rate stays constant. For instance, when the price of a pharmaceutical required for the treatment in a specific DRG increases in the middle of the year, the reimbursement rate will be lower than the cost-neutral reimbursement rate for part of that year. In other DRGs the costs may decrease during the year, such as when hospitals find ways to reduce the costs of an

[1] Non-identical hospitals would be an interesting extensions since if one knew the cost structure of different hospitals, one could exploit this to examine whether the hospitals are influenced by their cost structure to focus on financially the most favourable DRGs in each hospital. Similarly, more information about varying marginal costs, as opposed to average cost, of different DRGs would make it possible to examine whether those hospitals and those DRGs with most favourable marginal cost structures receive the largest increases in volume. We do not have detailed access to marginal costs for different hospitals or DRGs, so while intersting this approach for testing the effect of financial incentives is not explored in the current paper.

intervention. The imperfections caused by the lagged updating of reimbursements reintroduce the importance of financial incentives and make it tempting for hospitals to focus marginally more on the financially most favorable DRGs and correspondingly less on the unfavorable DRGs.[2]

Within the theoretical framework the question of whether hospitals in fact favor some DRGs more than others because of financial incentives, is reduced to the question of whether the mean change in treatment volume is different for DRGs with increased reimbursement compared to DRGs with reduced reimbursement.

Importantly, the question is not just whether the number of interventions changes when the reimbursement and the costs change. The number of interventions necessarily has to change when costs change since the hospital is operating within a budget constraint, and a higher cost must lead to fewer interventions. The question is not whether number of interventions change, but whether it changes more for the specific DRG which has become financially more favorable than the other DRGs. If the hospitals in fact use the reimbursement level to prioritize between different DRGs, one should observe a pattern in which favorable DRGs should have larger increases in treatment volume than unfavorable DRGs. To operationalize this, we will examine the contrast between changes in activity levels for DRGs which were more financially favorable with changes in the activity levels for DRGs which were less favorable.

2.2 Data and empirical operationalization

We collected data on rates of reimbursement for the different DRGs for each year between 2002 and 2013 from official governmental documents presenting the payment system (publicly available online) and compiled a database. For the same time period, we gathered the number of interventions in the different DRGs from the National Health Directorate, which in turn gets their information from the National

[2] Note that the motive for giving priority to the most profitable treatment need not be profit in itself. Even for non-profit hospitals focusing on the most profitable DRG-categories will be beneficial because it enables increased production of a good which we assume the hospital care about (overall health) and less altruistically it also makes it easier to balance the budget.

Patient Registry (NPR). The registry contains information for all hospital treatment episodes, inpatient and outpatient, from all hospitals in Norway.

The monetary reimbursement hospitals receive for each hospital stay is the product of three factors: the DRG weight, the bonus per DRG point, and the weight placed on bonus payment relative to the fixed per capita payment in the payment system. Since the introduction of Activity Based Financing (ABF) in Norway in 1997, there ABF-share of the budget has varied between 40% and 60%. However, since 2006 and until recent changes (2014), the share has remained constant at 40%. Thus, in order to avoid biases introduced by changes in the size of the overall bonus, we chose to investigate the time period from 2006 and onwards. Furthermore, although the price per DRG point increases every year, the price reflects the average treatment cost of a hospital stay and affects the absolute and not the relative price the hospital faces. Consequently, when investigating the effect of changes in reimbursement rates on hospital prioritization, the key remaining reimbursement component of relevance is the DRG-weight. Hence, in this paper, reimbursement rate refers to the DRG weight.

The theoretically important concepts financially favorable and unfavorable DRGs do not have a direct observable counterpart in the reimbursement database. The reimbursement for a DRG in a given year is known, but since true hospital costs are unknown we do not know in general which DRGs have the most or least favorable relationship between costs and reimbursements. However, by exploring information of changes over time, and the fact that these are lagged, it is possible to identify a set of diagnostic groups which are financially favorable and unfavorable for a limited period of time. If the DRG reimbursement is increased compared to the previous year, this demonstrates that the reimbursement rate was too low for at least part of the previous year. These changes in the DRG reimbursement rates are observable in the dataset and can be calculated.

In order to create a measure of the degree to which the reimbursement was favorable or unfavorable, we calculated the percentage change in the reimbursement rate for all DRG-categories for each year. An increase in the DRG reimbursement implies that the rate had been too low for part of the previous year. If the number of interventions is affected by the unfavorable reimbursement rate, one would expect the hospitals to have limited the number of interventions in the previous year and increased the number in the current year when the DRG reimbursement rates were increased. Similarly, DRG reimbursement rates which decrease indicate a DRG category with a favorable DRG rate in the previous year. If this rate influenced the hospital one would predict a relative high number of interventions in the previous year when the rate was high, and a decrease in the current year when the rate was decreased.

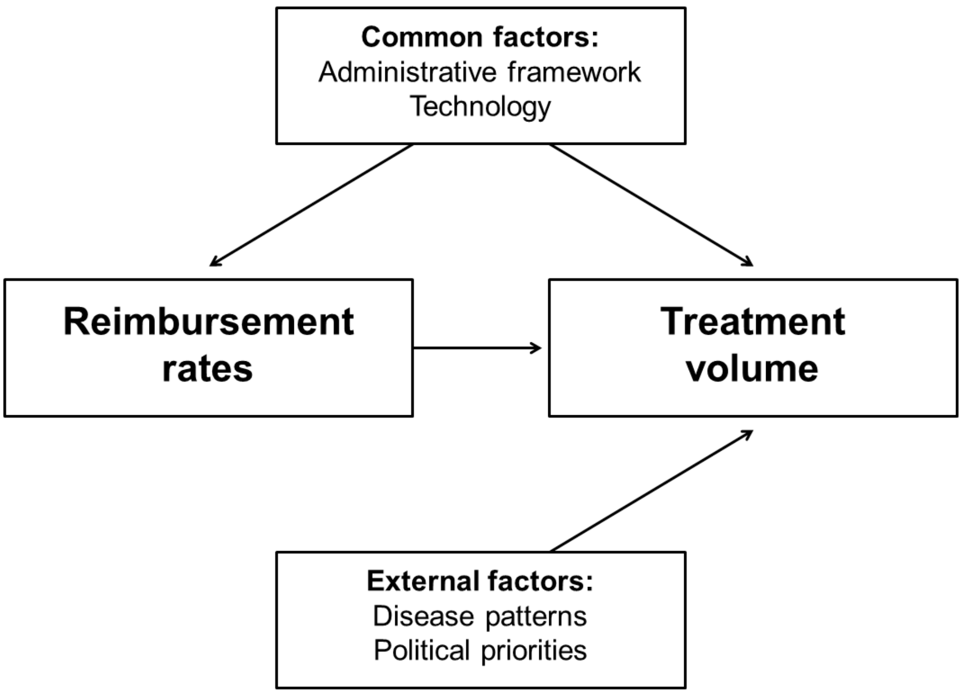
The general approach, then, is to use the percentage change in the reimbursement rate and the number of interventions for each DRG to identify favorable and unfavorable DRGs, and further, examine whether this is related to the percentage change in the number of interventions performed in the respective diagnostic related group. If the financial incentives influence priorities, the percentage increase in the number of treatments should be higher for DRGs with increased reimbursement than DRGs whose reimbursement is reduced.

2.3 Causality of financial incentives on priority setting

Although the data allows identification of financially favorable and unfavorable diagnostic groups, the causal relationship between financial incentives and prioritization between treatments may be confounded by many other factors. Figure 1 illustrates the problem of causal inferences about the effect of changes in reimbursement rates on treatment volume. First, treatment volume in different diagnostic groups is affected by several factors external to the financing system, such as disease patterns in the population, political priorities like the introduction of treatment guarantees for certain diseases, as well as other non-financial factors affecting changes in prioritization. Although important, these variables are only related to volume and not to reimbursement rates, and for this reason they do not confound the relationship between reimbursement rates and volume (Morgan and Winship, 2007). For instance, sudden events such as disease outbreaks may cause a large number of interventions to be registered in a particular DRG in one year, but as long as there is no systematic relationship between the outbreak and the reimbursement, these factors do not create any bias in the estimated relationship between financial incentives and treatments. Thus, the key challenge for causal identification is the second group of variables, those which affect both the reimbursement rate and the volume.

We assume that there are two main types of changes which affect both reimbursement rates and volume: Changes in the administrative framework and the introduction of new technology. Administrative changes in the financial system include the implementation of new DRGs, splitting of existing DRGs into new categories, and altering regulations for when a DRG is to be used. For instance, after splitting a DRG, both the reimbursement and the volume will change, but these changes in volume and reimbursement are not related to financial incentives. Second, the introduction of new technologies is likely to influence both the average treatment cost and the activity level, for instance by making a treatment more attractive due to increased effectiveness. In order to estimate the relationship between financial incentives and volume, it is important to reduce the influence of these types of variables and explore the type of bias they may create.

Figure 1. Causal relationships in identifying the effect of reimbursement rates on treatment volume.



The general strategy to reduce the confounding effect of administrative changes is to eliminate observations which are affected by these changes. This includes, first, eliminating all observations where the DRG reimbursement or the number of treatments was zero. A zero value is indicative of a code which was not in use and a large increase from zero, or a large decrease to zero, is mainly caused by administrative rules about coding and not a response to financial incentives which typically are more continuous. Furthermore, the introduction of a new DRG in one year will lead to an infinitely large (from zero) observed percentage change in the number of interventions in the relevant DRGs for that year. Moreover, it will affect the number of measured interventions in other DRGs, for instance leading to drastic reductions in another existing category for which most of the patients are now located in the new DRG.

To further eliminate the effect of paper changes, we have eliminated administrative and other special DRGs such as rehabilitation where additional reimbursements rules apply. For the years 2009 and 2010, major revisions of the reimbursement system took place, including the introduction of several new DRGs in both years and a comprehensive recalibration of all DRG weights in 2010. This means that the reimbursement and DRG categories in 2009 were very different from the system which was in place in 2010. In order to avoid the potential bias introduced by recalibrating and reorganizing the entire set of DRGs, we therefore excluded all observations from 2009 and 2010 (Figure 2).

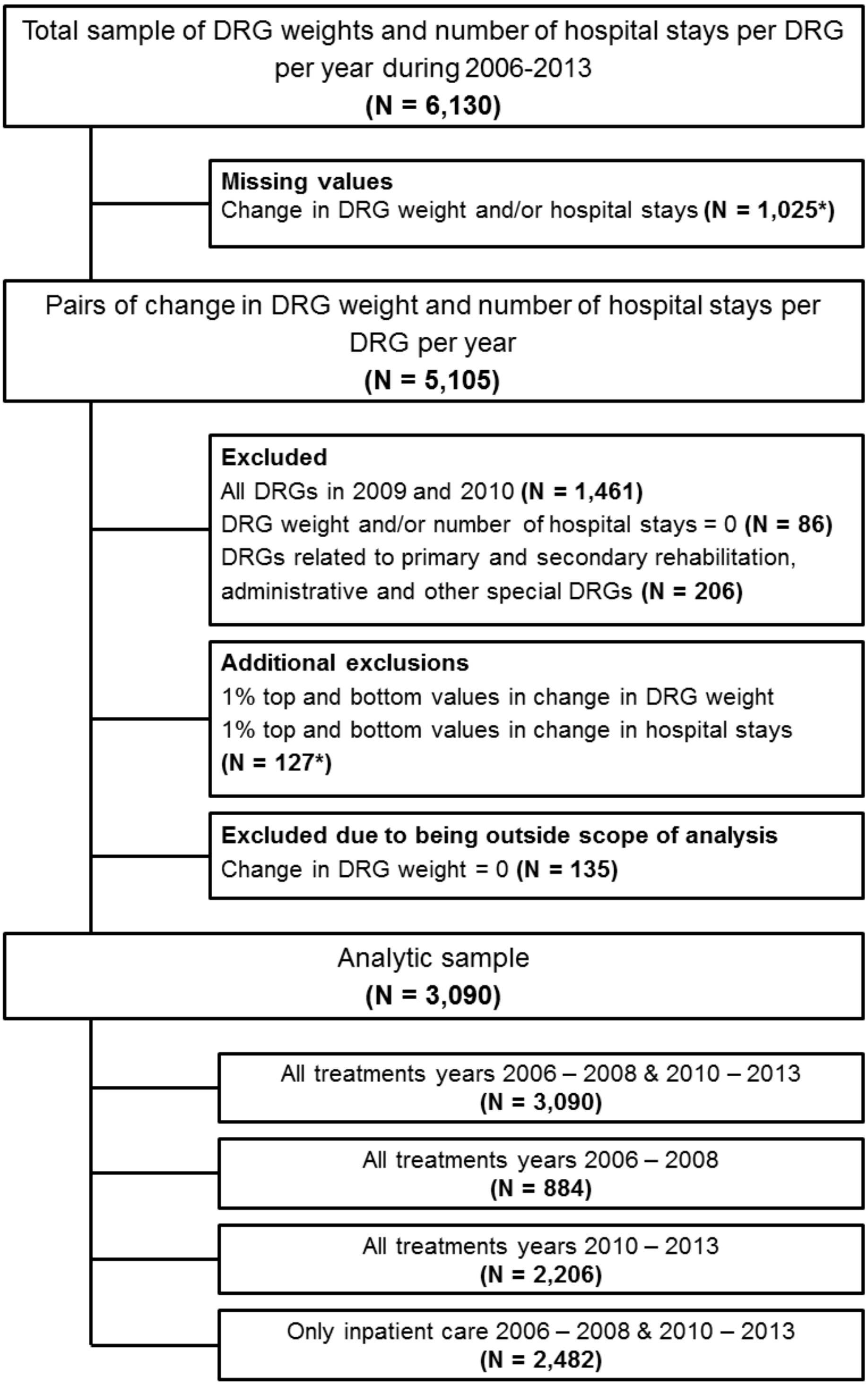
Another important issue relates to cases in which the percentage change in reimbursements rates or treatment volume is almost zero, but not quite zero. In this case, even minor absolute changes will cause large relative changes. We used two strategies to reduce this problem. First, extreme changes may be caused by financial incentives, but they may simply reflect administrative changes which need to be accounted for. For this reason, our first conservative strategy was to eliminate the most extreme changes, but we also provide the results with these very large changes included in the calculation. Second, we examined whether weighting the results by the budget share made a difference. Weighting the percentage change in a category by the number of interventions in that category will reduce the problem of the results potentially being dominated by many large and quite random changes in small DRG-categories.

To examine the robustness of results in different samples and sub-groups, we compared differences in means between favorable and unfavorable DRGs across a wide range of subgroups. The first subgroup contained the entire analytic sample, while the subsequent two subgroups distinguished between the years prior to and after the major administrative revisions in 2009 and 2010. In the final subgroup, we explored the degree to which including or excluding categories related to outpatient care biased the results by including only inpatient care through all years.

In addition to investigating both unweighted and weighted results for these four subgroups, we also explored the impact of budget share by means of the dichotomous characteristic “high” and “low” budget share. The prediction is that one would expect larger sensitivity to changes in categories which are important in the budget, but changes in small budget items may be less sensitive to changes in reimbursements. In this analysis, we compared all observations with the 75% lowest shares of the activity-based budget with the observations holding the highest 25% share of the budget. The subgroups are intended to serve as a tool to further eliminate the common factors affecting both reimbursement rates and treatment volume.

Finally, we conducted additional analyses based on two different sample selections. In the first analysis, we included information from 2002 to 2013, i.e., we did not exclude the observations from 2009 and 2010, nor did we exclude DRGs related to rehabilitation and other administrative or special DRGs. We did however exclude observations whose DRG weight and/or hospital stays were equal to zero, as well as unadjusted DRG weights and the 1% top and bottom values in change in DRG weight and change in hospital stays. In the second robustness analysis, we utilized our analytic sample, but included the 1% top and bottom extreme values in DRG reimbursement change and change in treatment volume.

Figure 2. Flow chart of sample derivation.



*The number of exclusions is presented jointly due to overlapping observations

2.4 Regression model

In addition to the comparison of average change, we also constructed two standard ordinary least squares (OLS) models to estimate the effect of changes in reimbursement fees on hospital treatment volume. The outcome variable is the number of treatments and the primary explanatory variable is the percentage change in the reimbursement fees from one year to another per DRG. Our first OLS model is thus represented by:

(i) $\Delta y_{ij} = \beta_0 + \beta_1 \Delta DRG_{ij} + \varepsilon_{ij}$

Where Δy_{ij} represents percentage change in the number of treatments in a diagnostic group *i* in year *j*, ΔDRG_{ij} represents percentage change in the DRG weight per DRG *i* in year *j*, and ε_{ij} is the error term. The estimated β_1 can thus be interpreted as the average percentage points increase in treatment volume associated with a one percentage point change in the reimbursement fee for the diagnostic groups.

In the second OLS model, we included a set of control variables in addition to our primary explanatory variable, in order to estimate the effect of reimbursement fees on treatment activity. The second OLS model can be written as:

(ii) $$\Delta y_{ij} = \beta_0 + \beta_1 \Delta DRG_{ij} + \beta_2 DRG_{ij} + \beta_3 \Delta DRGAbs_{ij} + \beta_4 y_{ij} + \beta_5 BS_{ij} + \beta_6 REIM_{ij} + \beta_6 LOS_{ij} + \beta_7 MED + \beta_7 DAY + \alpha_j YEAR_j + \varepsilon_{ij}$$

Where DRG_{ij} represents the absolute reimbursement rate (*i.e.*, the DRG weight), $\Delta DRGAbs_{ij}$ represents the absolute change in reimbursement rate, y_{ij} represents the number of hospital stays, BS_{ij} represents each DRG's share of the activity-based hospital budget, $REIM_{ij}$ represents the total monetary reimbursed amount, and

$LOSs_{ij}$ represents average length of stay, all variables per DRG *i* in year *j*. Additionally, the dummy variable MED represents a medical DRG as opposed to a surgical DRG, DAY represents a dummy for DRGs associated with day treatment or outpatient clinic procedures, and finally, $YEAR_j$represents year dummies for all years included in our analytic sample (years 2007, 2008, 2011, 2012 and 2013).

3. Results

3.1 Descriptive statistics

Across all DRGs for all periods, the DRG rate ranged from 0.01 to 39.13, with a mean value of 1.49 (Std. err. = 0.06) and a median of 0.73 (Table 1). The average change in DRG weights were 3.33% ranging from -51.92% to 102.90%. In the dataset, DRGs with a number of treatments less than four were registered as zero treatments. Thus, the number of hospital stays per DRG per year ranged from 4 to 299,229, with a mean value of 6,050 treatments. The mean change in activity was 2.04%, ranging from -44.14% to 72.73% (Figure 3 and Table 1).

Figure 3. Distribution of the change in in DRG rate and number of hospital stays.

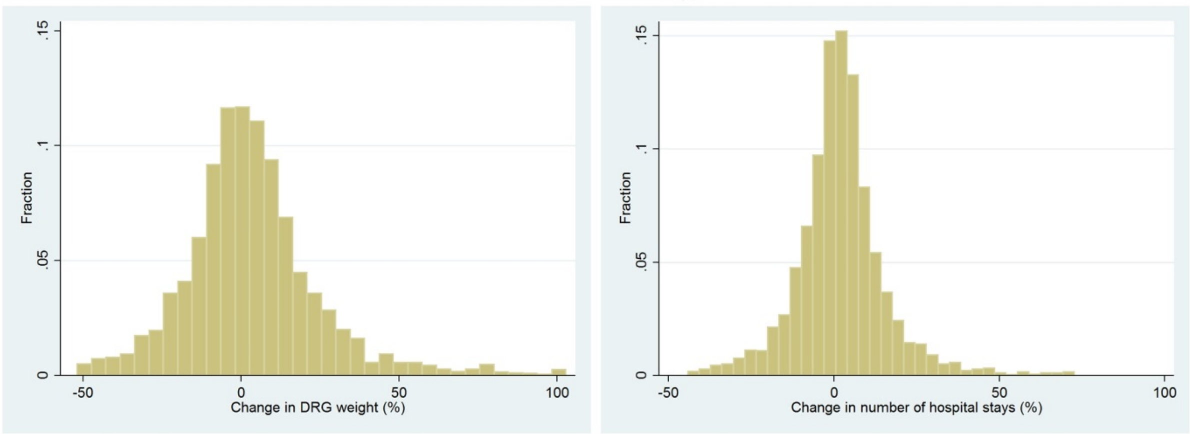
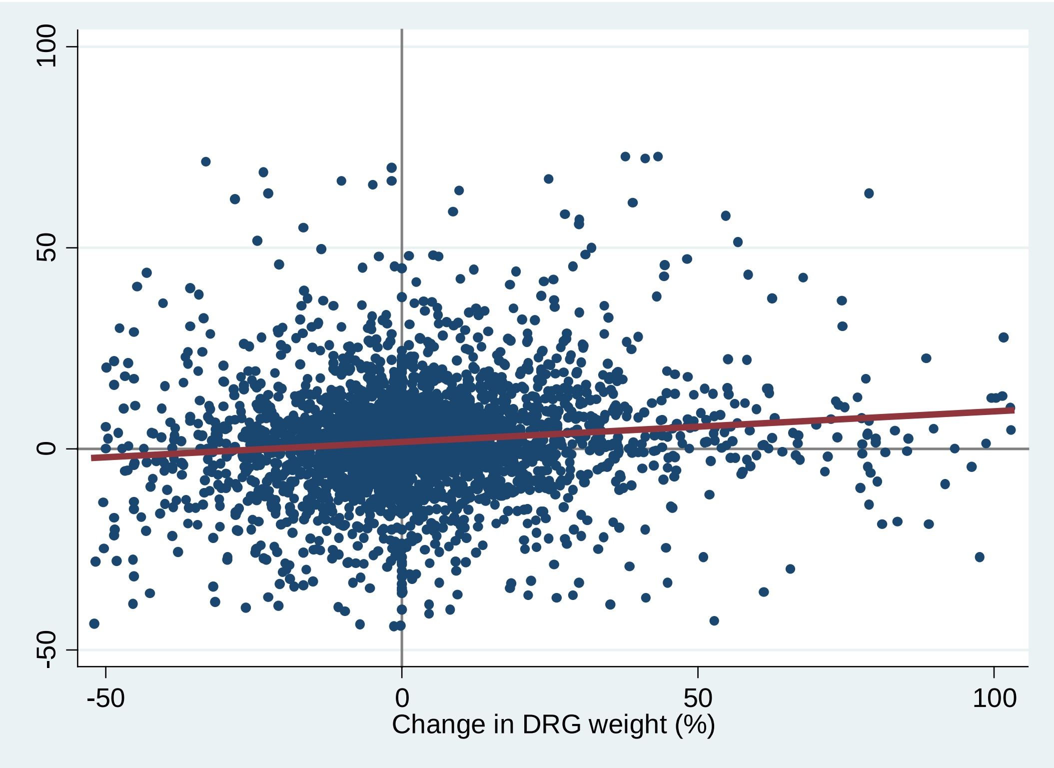


Table 1. Descriptive statistics for key variables.

Variable	**Mean**	**Std. Err.**	**Median**	**Min**	**Max**
DRG rate	1.49	0.06	0.73	0.01	39.13
Percentage change in DRG weight	3.33	0.38	1.71	-51.92	102.90
Number of hospital stays	6050	356	951	4	299229
Percentage change in number of stays	2.04	0.25	1.50	-44.14	72.73

In Figure 4, we have illustrated the distribution of all observations in the analytic sample including zero changes in the DRG weight between an increase and reduction in both reimbursement and the number of treatments performed. The analytic sample included 1,657 observations of increases and 1,433 reductions in the DRG weight. Among observations with an increased reimbursement, 61% also had an increase in registered number of treatments. Among observations with reduced reimbursement, on the other hand, 52% experienced an activity increase.

Figure 4. Scatter plot and regression line, change in DRG weight and number of hospital stays including zero changes in the DRG weight (main analytic sample)



3.2 Adjustments in reimbursement and the corresponding activity change

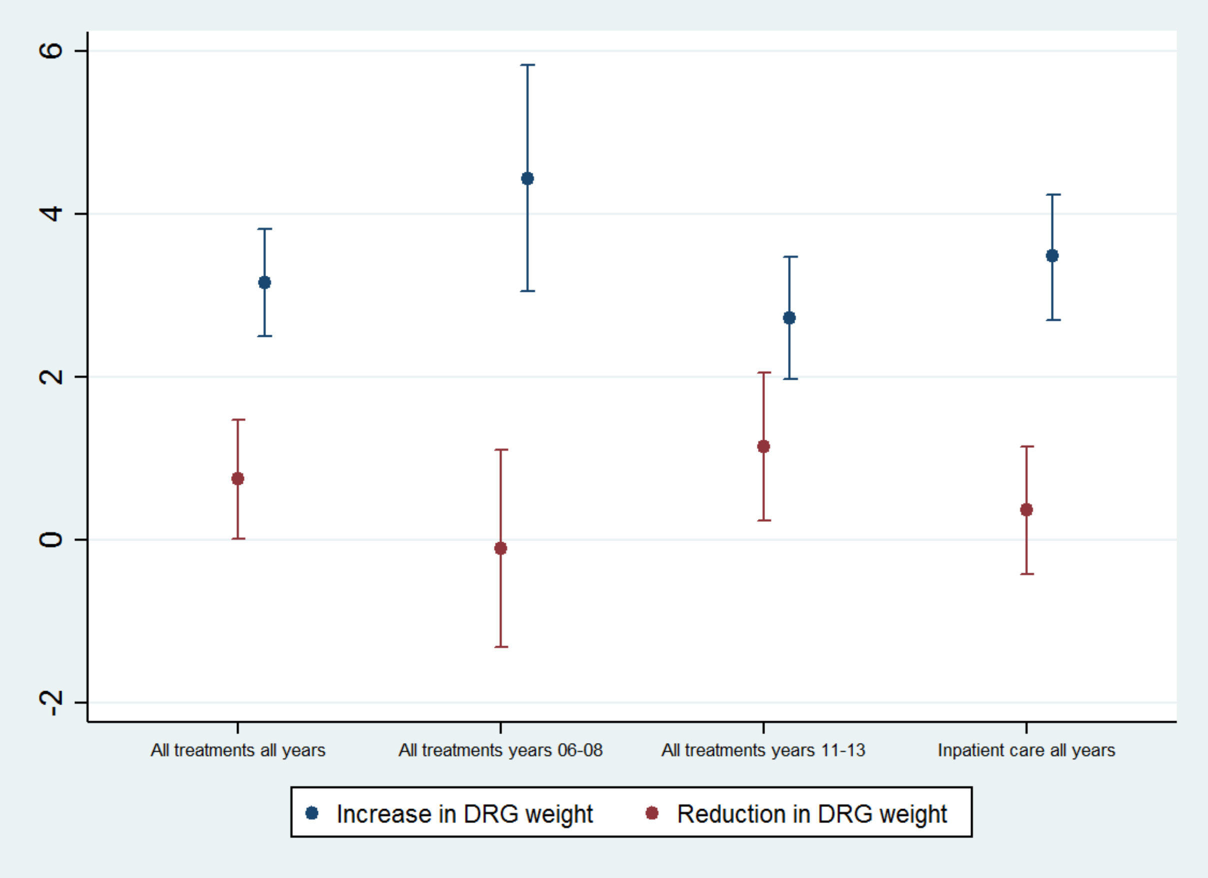
For inpatient and outpatient care across all years, a reduction in the reimbursement was associated with a 0.74% average increase in treatment activity, while treatment activity increased by 3.16% for procedures with increased reimbursement. The respective changes for all treatments during years 2006-2008 were -0.10% and 4.44%, while corresponding numbers were 1.14% and 2.72% for inpatient and outpatient care in years 2011-2013. Finally, for inpatient care exclusively across all years, a reduced reimbursement rate corresponded to an increase in activity at 0.36%, while the number of treatments increased by 3.49% for procedures with increased reimbursement. Thus, the difference between mean change in activity for increases and reductions in reimbursement for the four analytic groups, expressed by means of percentage points, were 2.42, 4.54 and 1.58 and 3.13, respectively. The differences were significant at a 1% level (Table 2 and Figure 5). In Table 2, the same results are also presented as means weighted by each procedure's share of the activity based budget. After weighting by the budget share, the percentage point difference between mean changes in hospital stays given an increase and decrease in reimbursement was 1.94 for inpatient and outpatient care across all years, 3.31 for all treatments in years 2006-2008, 1.12 for all treatments years 2011-2013, and 2.21 for exclusively inpatient care across all years.

Table 2. Unweighted and weighted mean change (%) in number of treatments by reduction and increase in the DRG weight.

Subgroup	Unweighted			Weighted by budget share		
	Reduction in DRG weight	Increase in DRG weight	Difference	Reduction in DRG weight	Increase in DRG weight	Difference
All treatments all years	0.74	3.16	2.42*	1.74	3.67	1.94*
	(0.36)	(0.33)		(0.46)	(0.47)	
All treatments years 06-08	-0.10	4.44	4.54*	1.81	5.12	3.31*
	(0.60)	(0.69)		(0.70)	(0.88)	
All treatments years 11-13	1.15	2.72	1.58*	1.69	2.81	1.12***
	(0.45)	(0.37)		(0.62)	(0.50)	
Only inpatient care all years	0.36	3.49	3.13*	1.53	3.74	2.21*
	(0.39)	(0.37)		(0.51)	(0.52)	

Standard error in parenthesis below the mean. *Significant at a 1% level, ***Significant at 10% level (Two-sample t-test assuming unequal variances).

Figure 5. Mean percentage change in number of treatments by subgroup and direction of change in the DRG weight.



In addition to calculating means weighted by budget share, we also investigated the impact of constituting a high share of the activity based budget by dichotomizing observations into respectively low and high budget share, defining high budget share as the observations with top 25% budget share (Table 3). For inpatient and outpatient care across all years, and the observations amongst the lowest 75% budget share, the mean change in number of treatments were 0.37% for DRGs with reduced reimbursement and 2.88% for procedures with increased reimbursement, while the respective numbers were 1.88 and 3.98 among the observations with a high budget share. Across all subgroups (outpatient and inpatient across all years, all treatments years 2006-2008, all treatments years 2011-2013, and inpatient care exclusively all years), the respective percentage point difference in mean change in activity between observations with an increase and a decrease in reimbursement were 2.51, 5.32, 1.57 and 3.36 for low budget share, while amounting to 2.10, 3.00, 1.44, and 2.37 for the high budget share observations. All differences were significant at either 1% or 5% level (Table 3).

Table 3. Mean change in number of treatments by reduction and increase in the DRG weight, categorized by low and high budget share.

Subgroup	Low budget share			High budget share		
	Reduction in DRG weight	Increase in DRG weight	Difference	Reduction in DRG weight	Increase in DRG weight	Difference
All treatments all years	0.37	2.88	2.51*	1.88	3.98	2.10*
	(0.45)	(0.41)		(0.51)	(0.51)	
All treatments years 06-08	-0.98	4.34	5.32*	1.64	4.65	3.00**
	(0.81)	(0.93)		(0.79)	(0.85)	
All treatments years 11-13	0.88	2.45	1.57**	2.03	3.46	1.44***
	(0.55)	(0.46)		(0.67)	(0.57)	
Only inpatient care all years	0.05	3.41	3.36*	1.34	3.72	2.37*
	(0.49)	(0.46)		(0.56)	(0.57)	

Low budget share indicate observations with the 75% lowest budget shares, while high budget share contains the remaining 25% observations with highest budget share. Standard error in paranthesis below the mean. *Significant at a 1% level, **Significant at 5% level, ***Significant at 10% level (Two-sample t-test assuming unequal variances).

3.3 Robustness of analysis

We scrutinized the robustness of our results by conducting the analysis on two additional sample selections, as previously described in section 2.3. For all years during the period 2002-2013 across the four subgroups, the difference in mean change in treatment volume between unfavorable and favorable DRGs ranged from 1.48 to 4.23 percentage points, while ranging from 0.71 to 5.58 when including extreme values in our initial analytical sample (Table 4). For years 2002-2013 the results were significant at 1% level. For the initial sample including extreme values, the results were significant at a 1% level for all subgroups except all treatments in years 2011-2013, where the difference of 0.71 percentage points were insignificant (p-value = 0.227).

Table 4. Mean change (%) in treatment volume by reduction and increase in the reimbursement fees, categorized by sample selection.

Subgroup	Years 2002-2013			Including extreme values		
	Reduction in DRG weight	Increase in DRG weight	Difference	Reduction in DRG weight	Increase in DRG weight	Difference
All treatments all years	-0.31	1.17	1.48*	1.62	3.70	2.09*
	(0.37)	(0.33)		(0.61)	(0.46)	
All treatments years 06-08	0.70	4.93	4.23*	-0.06	5.52	5.58*
	(0.54)	(0.66)		(0.78)	(1.07)	
All treatments years 11-13	1.07	2.74	1.68*	2.39	3.10	0.71
	(0.51)	(0.42)		(0.81)	(0.49)	
Only inpatient care all years	-0.55	1.16	1.71*	0.69	4.04	3.34*
	(0.37)	(0.36)		(0.50)	(0.52)	

The category 'Years 2002-2013' refers to a different sample including all years from 2002-2013, while 'Including extreme values' refers to the analytic sample including extreme values. Standard error in parenthesis below the mean. *Significant at a 1% level (Two-sample t-test assuming unequal variances).

DRGs with unadjusted reimbursement fees were not included in our initial analyses, due to our primary interest in DRGs with altered reimbursement fees. One may believe that if hospitals respond to financial incentives, the mean change in treatment volume for unchanged DRGs would lie somewhere between the mean change for DRGs with increased reimbursement and reduced reimbursement. Mean changes in treatment volume for unadjusted DRGs are presented in Table 5. The mean change lies between DRGs with an increase and decrease for the subgroups including all treatments in years 2006-2008 as well as only inpatient care all years, whereas the mean change is actually lower than the mean change for DRGs with reduced reimbursement for all treatments all years and all treatments years 2011-2013.

Table 5. Mean change in number of treatments for DRGs with unadjusted reimbursement fees by subgroup category.

Subgroup	N	Sample proportion (%)	Mean
All treatments all years	135	4.19	-0.03 (1.23)
All treatments years 06-08	60	6.36	1.95 (1.82)
All treatments years 11-13	75	3.29	-1.62 (1.67)
Only inpatient care all years	116	4.46	1.03 (1.32)

Standard error in parenthesis below the mean.

3.4 Regression results

Results from the bivariate and multivariate OLS regression models are presented in Table 6. We found a positive and significant effect of changes in reimbursement rates on change in number of treatments registered within one DRG from one year to another. The effect ranges from 0.0508 to 0.1303 percentage points change in activity for a one percentage point change in reimbursement, depending on subgroup and regression model. For each subgroup, the regression coefficient does not alter substantially when introducing multiple explanatory variables in the model.

Table 6. Regression results of the effect of percentage change in reimbursements on percentage change in treatment volume.

Subgroup	Model i) Bivariate		Model ii) Multivariate	
	Coefficient for drgChangePct	Adjusted R^2	Coefficient for drgChangePct	Adjusted R^2
All treatments all years	0.0755*	0.0135	0.0750*	0.0285
	(0.0140)		(0.0209)	
All treatments years 06-08	0.1290*	0.0297	0.1303*	0.0488
	(0.0337)		(0.0418)	
All treatments years 11-13	0.0606*	0.0093	0.0508*	0.02
	(0.0153)		(0.0238)	
Only inpatient care all years	0.1019*	0.0222	0.1092*	0.0403
	(0.0172)		(0.0224)	

Robust standard errors in parenthesis below the mean. *Significant at a 1% level.

4. Discussion and policy implications

Our analysis indicates that financial incentives have an impact on relative treatment volume between each DRG. The difference is consistent across a wide array of subgroups, and varies between a one to five percentage points difference depending on the subgroup. Furthermore, the regression analysis indicates a significant, positive relationship between changes in reimbursement rates and the respective change in treatment volume, a result that was robust to inclusion of additional variables in the model.

The results complements the findings of another study which investigated the effect of reimbursement changes, which estimated a price elasticity of a diagnostic procedure at 0.094, *i.e.*, a ten percent increase in the price for a treatment would yield an increase in treatment volume of almost one percent (Januelevicuite et al. (2013)). In comparison, our model uses data from a longer and more recent time period, and predicts that a ten percentage point increase in reimbursement would cause a 0.5-1.3 percentage points increase in treatment volume. Moreover, the regression coefficients from the bivariate models do not change substantially when introducing other variables.

Although we found evidence for a difference in means between DRGs with increased and reduced reimbursement rates, it remains unclear whether this difference is modest or substantial. On the one hand, the difference may not appear considerable since it is a small change measured in absolute percentages. However, the mean change in treatment volume across all DRG weights in the analytic sample is 2.04%. Measured against this average change in activity, the difference between the increase in the financially favorable DRGs (3.16%) and the financially unfavorable DRGs (0.74%) implies that DRGs with an increased reimbursement had a 55% larger increase in treatment activity than the overall mean increase in treatment activity, while DRGs with reduced reimbursement rates increased activity by 64% less than the overall mean activity increase.

A result indicating that financial incentives influence hospital prioritization should not be confused with the conclusion that hospitals care more about money than patients' health. On the contrary, it may be interpreted as a positive implication of clinicians who try to maximize provision of care to as many patients as possible. Faced with a budget constraint and externally set reimbursement rates, clinicians must take financial incentives into account if they want to maximize the number of treatments across different diseases. Despite the benefits associated with the financial incentives on hospital prioritization, precautions should be made to avoid excessive and harmful consequences of price differentials. Prioritization of procedures with relatively low price as compared to relatively expensive treatments may cause undesirable composition of health production in terms of over-treatment of patients on the verge of indication for an advantageous procedure for the hospital (who would otherwise be advised not to seek treatment), and the withholding of treatment for patients who are in strong need but who's procedure is too expensive for the hospital. In such a case, the hospital may maximize the production of treatments, but not maximize the production of health.

Our study has some limitations. First, the derivation of the analytic sample entailed exclusion of a large number of observations, amounting to almost 40% of our sample of non-missing observations. However, these cuts were necessary in order to isolate the effect, and the remaining observations comprised a considerable sample of DRG weights and activity level across multiple years, producing results that were stable across different samples and years.

In addition to limitations associated with sample derivation, another constraint in our analysis relates to the unobservable feature of technological change which is assumed to impact both reimbursement rates and treatment level. Innovations in medical treatments may cause both prices and effectiveness to increase or decrease, thus creating dependence between reimbursement rates and treatment volume, which may cause distortions in our results. Novel technologies present several different scenarios: i) increased effectiveness and reduced costs, ii) increased effectiveness and increased costs, iii) unaltered effectiveness and reduced costs. A potential limitation is that since we do not observe the cause of change in the DRG reimbursement, we may mistakenly attribute the change in volume to the change financial incentive, when in fact the causal mechanisms was that the new technology changed the utility of the treatment. However, exploring all these options show that the bias in all cases indicate that our estimates are not exaggerated. Although we do not know whether the change was caused by technology or something else, or what type of technological change it was, a detailed examination of all cases show that the problems always lead to an underestimation of the effect of financial incentives. In the first scenario a cheaper and better treatment technology appears. The reduced costs will be reflected in reduced reimbursement rates and hence provide a potential incentive for reducing the number of treatments. However, the price effect may be neutralized by the increased effectiveness which provides incentives for increasing the number of treatments. The net effect in the first scenario may therefore be that the pure price effect is dampened by the change in the utility of the treatment in which case the danger is that we underestimate the effect of financial incentives, not that we exaggerate its effects.

In the second scenario, better, but also more costly technology appears. When it appears, the net price will be higher, for a given reimbursement, than the old technology. For this reason the hospital will not use it as much as they would like to given a revised cost reimbursement. However, because the technology is better, the decline in the volume during the period of unrevised reimbursement rates will not be as large as it otherwise would be. Once again, the utility effect of the new treatment dampens the effect of financial incentives and the measured effect is a conservatie estimate of the true effect. In the last scenario, reduced costs will impact treatment volume only through reduced reimbursement. In this case the observed change in volume reflects the change in financial incentives. The size depends on the proportion of the period for which the reimbursement rate was wrong, which once again leads to an underestimation since the misaligned rate will not last for the whole time period, but none of the change in volume is affected by the constant utility of the technology. In sum, although technological advancements are unobservable, the different logical possibilities indicate that the estimates in the paper are conservative and if anything the various limitations indicate that the effect should be larger than our estimates.

Finally, we do not know to which degree hospitals can actually prioritize between treatments intentionally, or whether some procedures are more controllable than others. We utilized the number of treatments as our outcome variable to represent changes in prioritization between different procedures, and checked the robustness in our outcome variable by weighting the results by budget share as well as investigating differences in means across categories of high and low budget share, being left with consistent results. However, weighting also introduce new potential biases as the budget share is a composite of number of treatments and the reimbursement for the respective treatment, and the number of interventions in a DRG category may be systematically related to the degree of which hospitals can influence the interventions. For instance, one would not expect a change in the DRG to affect the number of births in a country, but the DRG rate for c-section may affect the balance between different kinds of interventions. Similarly, certain types of elective surgery, like meniscus surgery, are clinically easier to adjust based on the financial incentives, while life emergency treatment for broken legs is less malleable. Information on the degree to which hospitals have a discretionary influence on the number of interventions in a DRG code is not observed. However, if the large DRG-categories are less open to discretionary choice (birth, cancer treatment, heart attack), weighting the results by the budget share may create a downward bias and give large emphasis to precisely those categories where it is most difficult for the financial incentives to have an effect.

5. Conclusion

Understanding the mechanisms of the financing system is crucial for safe, efficient and desirable allocation of health care resources, and is essential for the design of future financing models. Our empirical analysis utilizing recent data from Norwegian hospitals provide supporting evidence for the association between financial incentives on hospital prioritization, while also suggesting that treatment volume is in fact a part of hospitals' utility function. More research is needed on to which degree hospitals can prioritize between different treatments, as well as the price elasticity of each procedure. In addition, more information about the causal relationship between reimbursement rates and activity level is needed to inform the analysis of financial incentives on prioritization. Although the health care sector is influenced by a range of other factors than financial incentives influencing the hospital production, the financing system provides an important framework for production and will influence output in one way or another. Thus, a financing system may never be 'neutral'. Ultimately, the decision makers will have to decide to which degree financial incentives should influence hospital production and the prioritization between different patient groups. In this respect, reimbursement fees may serve as a tool to steer prioritization and assist the achievement of health policy aims.

Acknowledgements: The authors wish to thank the Norwegian Health Directorate and seminar participants at the University of Oslo for comments on previous versions of this paper.

References

Barnum, H., Kutzin, J., & Saxenian, H. (1995). Incentives and provider payment methods. *The International Journal of Health Planning and Management, 10*(1), 23-45. doi: 10.1002/hpm.4740100104

Busse R, G. A., Quentin W, Wiley M, eds. . (2011). *Diagnosis related groups in Europe: Moving towards transparency, efficiency and quality in hospitals.* : Open University Press.

Carter, G. M., Newhouse, J. P., & Relles, D. A. (1990). How much change in the case mix index is DRG creep? *J Health Econ, 9*(4), 411-428.

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

Ellis, R. P. (1998). Creaming, skimping and dumping: provider competition on the intensive and extensive margins. *J Health Econ, 17*(5), 537-555.

Ellis, R. P., & McGuire, T. G. (1986). Provider behavior under prospective reimbursement. Cost sharing and supply. *J Health Econ, 5*(2), 129-151.

Hansen, T. W. (2005). Changes in the utilization of diagnostic codes in neonatology following the introduction of activity-based financing. *Health Policy, 74*(2), 218-223. doi: 10.1016/j.healthpol.2005.01.007

Janueleviciute, J., Askildsen, J. E., Kaarbøe, O., Siciliani, L., & Sutton, M. (2013). How do hospitals respond to price changes? Evidence from Norway. *Working papers in Economics (No. 15/12), Department of Economics, University of Bergen.*

Jegers, M., Kesteloot, K., De Graeve, D., & Gilles, W. (2002). A typology for provider payment systems in health care. *Health Policy, 60*(3), 255-273. doi: http://dx.doi.org/10.1016/S0168-8510(01)00216-0

Kuhn, M., & Siciliani, L. (2008). Upcoding and Optimal Auditing in Health Care (or The Economics of DRG Creep). *Discussion Paper No. DP6689, Centre for Economic Policy Research, London, UK*.

Kvaerner, K. J. (2004). Thought-provoking adaptation of otorhinolaryngologic surgery. *Tidsskr Nor Laegeforen, 124*(21), 2743-2745.

Lægreid, P., & Neby, S. (2012). Gaming the system and accountability relations: Negative side-effects of activity-based funding in the Norwegian hospital system. . *Working paper, Uni Rokkan Centre, The Stein Rokkan Centre for Social Studies*.

Martinussen, P. E., & Hagen, T. P. (2009). Reimbursement systems, organisational forms and patient selection: evidence from day surgery in Norway. *Health Econ Policy Law, 4*(Pt 2), 139-158. doi: 10.1017/s1744133109004812

Mikkola, H., Keskimäki, I., & Häkkinen, U. (2002). DRG-related prices applied in a public health care system—can Finland learn from Norway and Sweden? *Health Policy, 59*(1), 37-51. doi: http://dx.doi.org/10.1016/S0168-8510(01)00169-5

Norwegian Health Directorate (2007). *Activity-based funding of health services in Norway. An assessment and suggested measures*. Oslo, Norway.

Silverman, E., & Skinner, J. (2004). Medicare upcoding and hospital ownership. *Journal of Health Economics, 23*(2), 369-389. doi: http://dx.doi.org/10.1016/j.jhealeco.2003.09.007

Steinbusch, P. J., Oostenbrink, J. B., Zuurbier, J. J., & Schaepkens, F. J. (2007). The risk of upcoding in casemix systems: a comparative study. *Health Policy, 81*(2-3), 289-299. doi: 10.1016/j.healthpol.2006.06.002

Vertrees, J. C., Averill, R. F., Eisenhandler, J., Quain, A., & Switalski, J. (2013). Bundling post-acute care services into MS-DRG payments. *Medicare Medicaid Res Rev, 3*(3). doi: 10.5600/mmrr.003.03.a03