

*INCENTIVES IN PRIMARY CARE AND THEIR IMPACT ON POTENTIALLY
AVOIDABLE HOSPITAL ADMISSIONS*

Gianluca Fiorentini[§], Elisa Iezzi[#], Matteo Lippi Bruni*, Cristina Ugolini*

[§] Department of Economics, University of Bologna, Italy

[#] Department of Statistics, University of Bologna, Italy

* Department of Economics, University of Bologna and Centre for Household Income Labour and Demographic economics (CHILD), Italy

Address correspondence to:

Matteo Lippi Bruni, Piazza Scaravilli 2, 40126, Bologna, Italy

Tel. +39-051-2098145 - Fax + 39-051-2098040

e-mail: matteo.lippibruni2@unibo.it

Abstract

Financial incentives in primary care have been often introduced with the purpose of improving appropriateness of care and containing demand. We usually observe pay-for-performance programs, but alternatives have been also implemented, such as pay-for-participation in improvement activities and pay-for-compliance with clinical guidelines. Here, we assess the influence of different programs which ensure extra-payments to GPs, for containing episodes of avoidable hospitalisations. Our dataset covers patients and GPs of the Italian region Emilia-Romagna for year 2005, and we control for a wide range of factors potentially influencing GPs' behaviour. By separating pay-for-performance from pay-for-participation and pay-for-compliance programs, we estimate the impact on the probability of (inappropriate) hospitalisation of financial incentives included in contracts between GPs and the NHS. As dependent variable, we consider two different sets of conditions, for both of which timely and effective primary care should be able to limit the need of hospital admission. The first is based on 27 medical DRGs that Emilia-Romagna identifies as at risk of inappropriateness in primary care, while the second refers to the internationally recognised ACSCs (ambulatory care-sensitive conditions). We show that pay-for-performance schemes may have a significant effect over aggregate indicators of appropriateness, while the effectiveness of pay-for-participation schemes is adequately captured only by taking into account subpopulations affected by specific diseases. Moreover, the same incentive scheme has fairly different effects on the two sets of indicators used, with performance improvements limited to the target explicitly addressed by the policy maker (i.e. the list of 27 DRGs). This evidence is consistent with the idea that a "tunnel vision" effect may occur when public authorities promote and monitor specific sets of objectives, as proxies for more general improvements in health care practices.

Keywords

Primary care, preventable hospitalisations, financial incentives, organizational appropriateness, multilevel modelling.

JEL classification: I11, I18, C31

1. Introduction

In recent years, the efforts for cost containment in healthcare have increased the need to identify where resources can be most efficiently targeted and also primary care has received increasing attention. In order to promote higher quality of care, programs have been introduced worldwide to ensure extra financial rewards to GPs involved in care improvement activities. Such schemes frequently adopt a pay-for-performance structure, where payments are contingent on meeting indicators of provider effort. Alternatively, other strategies have been developed, each with its distinctive strengths and weaknesses, with physicians rewarded for participation in improvement activities and/or for compliance with clinical guidelines.

For instance, the Quality and Outcomes Framework signed in 2003 in the UK is a mix of pay-for-performance and pay-for-compliance schemes that makes a substantial part of GP's income (about 30%) contingent on attaining 146 quality indicators, covering clinical care for 10 chronic disease, organisation of care and patient experience. In the US, public and private payers have made available financial incentives (in the range of 1 to 10% of total revenues) to hospitals and other health care providers as well as to individual physicians participating in collecting information and/or meeting quality targets. In Italy, the experience of financial incentives is limited to primary care, but it refers to a small part of GP's total revenue.

Our study focuses on medical conditions where timely and effective provision of primary care should reduce the risk of hospitalization and examines the influence on (potentially inappropriate) hospitalisations of different sets of programs providing economic incentives to GPs for improving appropriateness of care. More precisely, we analyse the impact of the use of economic incentives by Regional and Local Health Authorities (LHAs) in contracts for primary care in the Italian region Emilia-Romagna for the year 2005, distinguishing pay-for-performance from pay-for-participation and pay-for-compliance programs. We test the hypothesis that, other things equal, patients cared by GPs that receive higher shares of their revenues by incentive based programs are less likely to experience “avoidable” hospitalisations.

The paper adopts a general perspective in terms of study population and programs considered. By taking as reference the whole adult population and the full set of programs that provide financial incentives, we move away from the disease based approach adopted in previous works (Lippi Bruni, Nobilio & Ugolini 2009), although this comes at the cost of merging together programs pertaining

to different areas. Through a comprehensive analysis, our aim is to evaluate the impact of this policy innovation, getting rid of the limitations implied by considering each program separately. For instance, a more narrow approach that focuses only on specific diseases, may fail to capture dysfunctional consequences such as the adoption of a “tunnel vision” attitude on part of GPs, as a response to performance monitoring, that induce a concentration of effort only on the areas included in the performance indicator scheme.

Taking up a broad viewpoint is particularly important as long as the implications of a widespread use of economic incentives in the Italian primary care system are potentially twofold. Not only, they are expected to improve results for the targets explicitly addressed, such as containing pharmaceutical prescriptions or reducing hospitalisation rates for specific diseases. They are also intended as tools for improving quality of care more generally by enforcing cooperation and inducing GPs, who are independent professionals contracted with the NHS, to take decisions in line with the general interest. In other words, providing extra money for clearly identified targets should (hopefully) improve quality of performances also on dimensions other than those strictly included in the contractual agreements, as a consequence of a more cooperative attitude of GPs.

Given these premises, we consider two indicators for avoidable admissions. The first measure refers to clinical conditions for which the Regional government has explicitly sought to promote an active participation of GPs in improving appropriateness, whereas the second indicator is an internationally accepted measure of avoidable hospitalisations. The first indicator of potentially inappropriate hospitalisations relates to policy targets explicitly addressed by healthcare authorities. It corresponds to a list of 27 medical DRGs selected by the Emilia-Romagna Region as at risk of inappropriateness in primary care. The second indicator is based on the Ambulatory Care Sensitive conditions (ACSCs) developed by Billings et al. (1993). Admissions for ACSCs are “avoidable” in the sense that they refer to conditions (*chronic conditions* such as diabetes and asthma, *acute conditions* such as ear/nose/throat infections, and *preventable illnesses* such as tetanus) that, if treated properly on an outpatient basis, should not require inpatient admission. Sometimes these admissions may be needed (ex-post), following a decline in patient’s health, but effective and timely ambulatory care should be able (ex-ante) to prevent such deterioration. Therefore, hospitalisations of this kind are used in the literature to identify possible deficiencies in the quality of primary care services.

To summarise, thanks to the inclusion of different incentive schemes as explanatory variables we obtain information on the extent to which the incentive structure influences the performances.

Moreover, by considering different dependent variables, we get insights also on the scope of such impact. In particular, the first measure employed here (27 DRGs) aims at capturing the effectiveness of the programs on targets that have explicitly attracted the policymakers' attention, while ACSCs can be taken as proxy for the gains in quality – interpreted in a very broad sense - favoured by enhanced cooperation between GPs and the other players of the system. Indeed, this latter measure displays several favourable features for this purpose. It is internationally validated as an indicator of good quality of primary care but it is not endorsed as policy target in the present institutional context. Moreover, it is defined according to very different criteria with respect to our first indicator (ICD-9-CM codes rather than DRGs), thus allowing for sufficient differentiation between the two variables,

The available dataset covers patients and GPs of the whole region and provides detailed information on health consumption of the population and on the different components of GP remunerations. In order to account for the hierarchical structure of the dataset where patients can be grouped within GPs and GPs within districts, our empirical strategy is based on multilevel modelling (Goldstein, 2003).

2. The role of incentives and motivation in health care

Our work contributes to the literature that studies how economic incentives can be used to improve the governance of institutions and organisations. Given the wide set of issues at stake and the extremely diversified situations that fall under such framework, both the theoretical and empirical literature provide mixed indications. Agency theory assumes that strong monitoring by the principal should motivate the agent to increase effort in order to reduce his penalty if caught shirking (Prendergast, 1999; Laffont & Martimort, 2002). Contrariwise, following the “crowding out” theory (Frey & Jegen, 2001), intensive monitoring could be considered as a signal of lack of trust and diminish effort by reducing pre-existing intrinsic motivations. The same result could hold in case of rewards. The premise of pay-to-perform schemes is to increase agents' marginal benefit of effort towards work outcomes relative to opportunity costs, but, introducing monetary rewards to compensate performance on a task that has been previously undertaken without any immediate monetary remuneration can actually reduce performance. Gneezy and Rustichini (2000) argue that this may be related to the size of the monetary incentive, with crowding out most likely to occur with modest payments. On the contrary, external interventions could also crowd “in” intrinsic motivation when there is a high degree of perceived autonomy, discretion and freedom in work

activity, personal relationships predominate and employees participate in decision making. In this context, the principal could signal his confidence in the agent's ability and strengthen the agent's intrinsic motivation in the long period.

Empirical studies of performance pay reforms in the private sector show that firms experience significant increases in productivity when switching from flat salary scheme to a pay for output scheme, where firms occasionally suffer decreases in quality when quality is not easily contractible (Lazear, 1996). Studies of performance pay reforms in the public sector are mostly focused on sporadic intervention programs and show mixed results (Heckman, Heinrich & Smith, 1996; Lavy, 2004). Skeptics of high-powered incentives in public organizations argue that using pay-for-performance schemes in the public sector could produce unintended consequences on workers' behaviour and induce extensive gaming by public employees that may lower their productivity (Dixit, 2000; Smith, 1995).

Examining the principal/agent relationship between purchaser and provider in healthcare, Goddard et al. (2000) stress the risks associated to the design of compensation mechanisms related to performance. The incompleteness of many measures may lead to the so called "tunnel vision", e.g. the concentration on areas included in the performance indicator scheme to the exclusion of other unmeasured areas, but also to some sort of "myopia" when the effort is concentrated on short-term issues, excluding long-term activities which may influence performance measures only in many years time. Besides, when outcomes depend on the joint effort of a number of agencies, it is very difficult to design adequate incentive schemes able to encourage cooperation and it becomes possible that only local, relatively narrow, objectives are pursued at the expense of the general organization's objectives. This could be relevant also if it is in the power of the agent to manipulate the reported data, giving rise to "mis-representation" of reported performance. Finally, when multi-period systems of targets and rewards are used, there is a danger of creating a "ratchet effect" when the good performance in one year is punished with higher future targets and the agent has a strong incentive to report persistently mediocre level of performance.

In healthcare pay-to-perform schemes have been increasingly adopted internationally. The focus of these initiatives ranges from primary to hospitals care to regional health care systems, and we observe a variety of methods to design the incentives and to assess their effects on quality. For the 1999 Australian pay-for-performance scheme for GP remuneration in chronic disease management, Scott et al. (2008) find a positive moderate effect on quality of diabetes care. In the US, Cutler (2006) reports that the empirical evidence to date is "more positive than negative", in the sense that

payment incentives result to produce some positive impact on the quality of care, but the effects are not all in the same direction. In the English pay-for-performance program, there is evidence of an increase in the quality of care during the first years, but also evidence of gaming and tunnel vision effect (Doran et al. 2006).

To overcome the unintended side effects, health economists suggest some broad strategies (Mannion et al, 2007; Godager, Iversen & Ma, 2009). The first strategy concerns the enhancement of the information set on which both principal and agent take their decisions, in order to reduce the scope for misdirected efforts, and the development of performance benchmarks which are independent of past activity. Another strategy has to do with the degree of dissonance between the objectives of principal and agent (Smith, 2002). In this case the idea is to create a “culture change”, that is to foster a closer alignment between their objective, promoting a favourable cultural environment and encouraging clinical professionals themselves to promote a common culture that might contribute to ameliorate the agency problem. Following this strategy, payers are exploring alternative solutions for improving quality such as pay-for-participation programs or pay-for-compliance programs. Both solutions generate a looser incentive structure with respect to pay-for-performance, since payments are conditional not on the accomplishment of precise, measurable targets, but on personal involvement of GPs in programs aimed at improving the quality of care for specific diseases and/or on the adherence to clinical protocols. The (still very limited) empirical literature on the topic, mainly based on surgical care, outlines that there are cases in which this kind of incentives may prove more effective in improving the quality of care (Birkmeier & Birkmeier, 2006). One of the potential advantages is physicians’ greater acceptance of this approach, together with significant improvements in providers’ adherence to evidence-based best practices. Nevertheless, incentive schemes that do not condition payments to the achievement of pre-defined individual targets may strongly attenuate the influence on physicians behaviour.

3. Primary care organisation and strategies for appropriateness

In the last fifteen years, significant changes introduced in the Italian NHS have widened regional powers in the organisation and provision of healthcare. Such process has gone together with an increase in direct financial responsibilities of regional governments according to the principles of fiscal federalism. As for primary care, a series of reforms implemented in the 90’s have enhanced the integration between district and primary care services and reinforced group practice. Moreover,

they have set the basis for introducing additional financial rewards to GPs for meeting specific policy goals (France, Taroni & Donatini, 2005).

The responsibility for organising primary care, and more generally non-hospital services, is attributed to health districts grouped into LHAs. The relationship between the NHS and GPs, who are independent professionals acting as gatekeeper to secondary care, is regulated through contractual agreements, which set, among other things, the remuneration profile. Capitation is the primary source of GPs' remuneration and money follows patients' decision to enrol with a particular GP. In recent years, a maximum limit of 1500 patients has been introduced and GPs already exceeding such threshold can keep their additional patients but they cannot add new ones. In addition to that, GPs receive also fee-for-service compensations for providing specific treatments such as minor surgery, preventive activities and post surgery follow-ups. Both the *fixed* (capitation) and the *variable* (fee-for service) components are negotiated between the GPs' trade unions and the central government and are uniform nationwide.

On top of them, each region can introduce a third component of the remuneration, typically aimed at promoting appropriateness of care and cost containment. Indeed, consistently with the policy guidelines and the institutional framework designed by each region, LHAs and health districts can autonomously negotiate with the local organisations of GPs *additional* payments for specific programs, that address priority targets identified at the regional and local level.

In Emilia-Romagna, the regional government has devoted substantial effort to enhance an appropriate use of hospital services. Using DRGs and Disease Staging, a classification was developed to identify cases of potentially inappropriate hospitalisations. To assist LHA managers in identifying the appropriate level of alternative, low intensity, treatment settings for patients who do not need acute care hospitalization, the Regional Health Authority has produced in 2004 an "Atlas of appropriateness of hospital use". As for primary care, since 2001 the LHAs have been involving GPs in policies aimed at substituting hospital with community-based care, for conditions typically associated to chronic pathologies such as diabetes, asthma, chronic obstructive pulmonary disease, congestive heart failure, hypertension and psychiatric disorders. Regional guidelines have been issued for encouraging cooperation between primary and secondary care providers, and the implementation of these measures was pursued through the introduction of several programs based on financial incentives at the district level.

Pay-for-performance programs are characterised by a strict link between financial transfer and targets achievement to be verified ex-post. In Emilia Romagna, they are negotiated between district or LHA and GPs trade unions. Examples are the financial rewards to GPs that meet targets for the prescription rate of generic relatively to non-generic drugs or extra-payments for GPs that keep hospitalization rates below predefined thresholds for a series of listed conditions. Usually, the financial rewards adjusts to the performance level, according to a graduated scale that begins at the attainment of a minimum threshold and increases up to a maximum level. As for primary care, GPs receive financial incentives to deliver high quality services that improve clinical and organizational appropriateness. The main target indicators refer to pathologies for which GPs have the stronger potential of influencing the rate of inappropriateness in the setting of care. For example, some local contracts provide financial incentives to contain the standardized ordinary hospitalization rate, to decrease the emergency admissions rate, to increase the number of protected hospital discharges for patients needing follow-up care or treatment after a hospital stay, to reduce hospitalization for elderly patients strengthening the capacity of the home and community care system.

Pay-for-participation schemes are an heterogeneous group of programs aimed at encouraging physician's participation in the management of specific conditions (e.g. diabetes, hypertension, oncology, asthma and dementia) which require additional efforts directly devoted to each diagnosed patient. These payments can be labelled "*pay-for-participation*" because the financial bonus does not depend on the achievement of specific targets as for pay-for performance. For instance, various districts – though not all - have developed programs that require physicians to take part in local diabetes management plans through which GPs are entitled to a financial bonus for the assumption of responsibility of each diabetic patient. GPs in charge of providing care to individuals expected to require higher than average effort - as diabetes type II patients - receive an extra payment that supplements standard capitation. A key feature of "pay-for-participation" is that the amount received by GPs is defined on a per-patient basis and, therefore, depends on the number of patients with the diagnosed disease falling under his responsibility and not on some measure of performance as it would be for the reduction in avoidable hospitalisations.

Finally, pay-for-compliance schemes introduce financial transfers to GPs who take part in various activities that promote cooperation with professionals dealing with other levels of care. Additional compensations are provided for attending audit meetings, for actively participating in the development of new protocols and guidelines or for implementing evidence-based best practices, such as increasing influenza immunisation uptake for the elderly and uptake of breast and cervical

cancer screenings or to increase the local rate of domiciliary care. In this case, differently from pay-for-participation, GPs' compensation is not related to the number of patients included in the list of the GP, but to the number of activities the physician is involved in.

4. Data and estimation issues

Our dataset covers year 2005 and encompasses the regional population aged between 18-74 (2.936.384 individuals), the 3.229 GPs operating in the region and the 39 healthcare districts belonging to 11 LHAs.

We consider two dependent variables, both identified through hospital discharge data. The first indicator is based on the list of 27 medical DRGs at risk of inappropriateness in primary care identified by the Region (Model A) and listed in the Regional resolution 319/2000 (Table 1). The total number of inappropriate admissions identified according to this criterion amounts to 11.552 (0,4% of the study population). In the second case, the dependent variable takes value 1 if the patient has been hospitalised for an episode classified as ACSCs (Model B) and 0 otherwise. We adopt here the list of ACSCs coded using the International classification of diseases (ICD-9-CM) and developed by Billings et al. (1993) and Caminal et al. (2004). We classify hospitalisations as inappropriate if at least one of the ICD-9-CM codes included in the ACSCs' list is recorded as the primary reason for admission (see Table 2). The total number of ACSCs admissions recorded in the dataset amounts to 16.924, which corresponds to the 0,58% of the study population .

INSERT TABLE 1 AND TABLE 2

Table 3 presents a cross tabulation of DRGs and ACSCs. Although both indicators proxy poor quality of primary care, the different criteria according to which they are defined lead to a limited overlap of events. In particular, only one third of the cases identified as 1 according to the DRG criterion are the same also under the ACSCs criterion (3918 out of 11552) and less than one fourth of the events classified as ACSCs are associated to one of the 27 DRGs (3918 out of 16924).

INSERT TABLE 3

We record detailed information at the patient, GP and district level (Table 4). Patient characteristics include gender, age and the presence of comorbidities that are summarised in the Charlson index (Charlson et al., 1987; Romano et al., 1993). The Charlson index is a weighted index of patient comorbidity, computed from the ICD-9-CM diagnostic and procedure codes available in administrative datasets according to their potential for influencing mortality. As regards GPs, we control for GP gender, age, practice location in urban areas and for type of practice, distinguishing single-handed from group practices. We include also information on list size and on the average age of the patients included.

INSERT TABLE 4

We aggregate the financial incentives received by GPs as variable part of their remuneration in three groups: pay-for-performance, pay-for-participation and pay-for-compliance. In order to get rid of scale effects due to list size, these variables are measured as share of GPs annual income. Table 5 summarises the size of the three groups of incentives across LHAs, while Figure 1 displays their distribution among the 39 regional districts. The share of total income received through the three schemes varies considerably among GPs. This is due, in the first place, to the variability in the financial size of the programs recorded across districts.

INSERT TABLE 5

INSERT FIGURE 1

As regards the district level, we investigate the influence of local conditions by including controls for supply side characteristics such as hospital beds and district hospitalization rates.

Patients are clustered according to the GP they are enrolled with and, in turn, GPs are clustered according to the district they work in. Grouping of observations may determine the presence of (unobserved) common factors among the units that belong to the same group. Having to deal with data characterised by such a hierarchical structure, our empirical strategy is based on multilevel modelling (Iversen & Luras, 2000). This estimation technique has the advantage of allowing the Var/Cov matrix of the error terms to account for possible correlation of observations belonging to the same layer. In this way, a correction is introduced in the size of the standard errors and appropriate confidence intervals can be computed (Guo & Zhao, 2000).

In our estimates three layers are identified: the patient (i), the GP (j) and the district (k). Given the dichotomous nature of the dependent variable, we estimate the following three-level logit model:

$$(1) \quad \text{logit}(\pi_{ijk}) = \log\left(\frac{\pi_{ijk}}{1 - \pi_{ijk}}\right) = \beta_0 + \beta_1 X_{ijk} + v_{0k} + u_{0jk} + \varepsilon_{ijk}$$

where π_{ijk} corresponds to $\Pr(y_{ijk} = 1 | X_{ijk}, v_{0k}, u_{0jk})$ and y_{ijk} is the realisation of a random variable Y_{ijk} assumed to follow a Bernoulli distribution with parameter π_{ijk} . Here, π_{ijk} represents the probability that a patient i falling under the responsibility of GP j in district k is hospitalised for an episode that is classified as avoidable admission. We estimate model (1) regressing two alternative dichotomous indicators, expressed by a list of DRGs at risk of inappropriateness and ACSCs, against the same set of explanatory variables. The random component of the estimated equation comprises three distinct parts, that represent the random errors for the patient (ε_{ijk}), GP (u_{0jk}) and district (v_{0k}) level respectively. Our distributional assumptions consider random components at different levels to be uncorrelated and normally distributed, while we permit observations referring to patients cared by the same physician or in the same district to be correlated (Rasbash et al, 2000). We compute also the intraclass correlation coefficients (ICCs) ρ_h as estimates of the share of the total variance that potentially refers to each layer h (Browne et al., 2005).

$$\rho = ICC = \frac{\text{population variance between macro - units}}{\text{total variance}} \quad (2)$$

For each dependent variable, we estimate a three-level logit model where GPs are nested within districts. Estimations are carried out with the SAS GLIMMIX procedure (SAS Institute, 2005), using the IGLS algorithm with a pseudo-likelihood procedure. We take deviance as measure of goodness of fit, and significance of the coefficients is assessed through the Wald statistic (Goldstein, 2003).

5. Empirical results

Table 6 presents estimates of equation (1) for the two dependent variables (Model A- DRGs and Model B- ACSCs). In both cases we consider a three level logit model.

INSERT TABLE 6

For both models, the estimated variances of the districts are significant and confirm the existence of substantial residual variability across local areas, which cannot be fully captured by controls for patients, physicians and districts characteristics. Consequently, despite the wide set of available information, potentially relevant unobservables affecting local areas and practices are still included in the error component.

Patients' characteristics, that adjust for differences in the case-mix across practices, are significant both in Model A and B. Interestingly, despite the use of different dependent variables not only the signs and significance levels, but also the coefficients are fairly comparable across specifications. In particular, male, older patients as well as those reporting more comorbidities, display a higher probability of incurring in avoidable hospitalisations. As regards physician characteristics, in both models the average age of the list and the urban location of the practice decrease the probability of hospitalisation for both dependent variables. Moreover, younger GPs and those working in single handed practices show a higher probability to have patients hospitalized for a DRG at risk of inappropriateness, while no significant difference is recorded when considering ACSCs.

One of the main policy question addressed in the paper concerns the role of financial incentives. Our results suggest that different programs produce a different impact. Pay-for-participation programs are never significant. On the contrary, in one of the estimated models, both pay-for-performance and pay-for-compliance, though the latter more weakly, influence the probability of avoidable hospitalisations. Moreover, financial transfers aimed at improving the appropriateness of hospital referrals through additional financial transfers to GPs are effective if we take the list of 27 DRGs as reference, while they are not, if one considers ACSCs. That is, we find a positive association between financial incentives and performances in primary care only if avoidable hospitalisations are measured according to conditions over which the healthcare authorities have explicitly drawn GPs' attention (i.e. the 27 DRGs).

In the policymaker's perspective two lessons can be drawn. First, the result from the DRGs model suggests that, when additional money is accompanied by active promotion of clearly identified policy objectives, such investments seem to be effective in inducing GPs to back the policymaker's effort for improving quality of care. At the same time, if we measure appropriateness of primary care through indicators widely used at the international level - such as ACSCs - but not explicitly endorsed in the present institutional context, no association is recorded between the share of

revenues received through incentive-based programs and quality of primary care. This result can be interpreted as suggestive of the fact that we cannot find support for the optimistic view according to which provision of extra-money to GPs for specific programs improves quality of care also on dimensions than go beyond the targets identified by policymakers. ACSCs can be seen as good proxy to capture such wider perspective since they are validated by the scientific community but here they are not identified as reference targets by public authorities and display a limited overlap with the DRGs on which public authorities have focused. In conclusion, our evidence indicates that, not only the way the incentive schemes is designed, but also the link between the target taken as reference and its institutional endorsement is crucial.

A further remark concerns pay-for-participation programs, whose coefficient is not significant in either model. As financial transfers for these programs represent the largest share of extra-payments to GPs, one could wonder why they are not as effective as the other ones. This empirical puzzle can be explained by keeping in mind that pay-for-participation schemes are usually associated to the management of specific diseases such as diabetes, asthma or hypertension. Consequently, their impact is likely to be loosely measured when we refer to the general population instead of focusing on the groups of patients targeted by each program.

In order to test this hypothesis, we consider a specific disease: diabetes type II. To extract the sub-population of type-2 diabetic patients we exploit identifiers previously developed for a related study (Lippi Bruni, Nobilio & Ugolini, 2009). Adjustments with respect to models A-B have to be made in order to account for disease specific features. In particular, outcomes are given by the ACSC codes listed in Table 2 that refer to acute complications for diabetes (comas) that should be avoided through an effective patient management. They are identified by hospital records in which ICD-9 codes 250.2 to 250.3 are documented as primary or most responsible diagnosis. Table 7 presents the explanatory variables that to a large extent corresponds to those employed in the general model. An exception being the substitution of the Charlson index with insulin dependence as proxy for severity, to get a disease specific indicator.

INSERT TABLE 7

Consistently with the more focused approach adopted here, we consider only payments aimed at improving diabetes care, which in Emilia Romagna typically take the form of pay-for-participation programs. Figure 2 presents the distribution of this payment scheme among the 38 regional districts as a fraction of GPs annual income (one district was excluded as its hospitalisation rates for diabetic

comas were over 10 times the regional average, probably due to a coding mistake). Finally, to account for supply side characteristics, we consider at the third level the presence of hospital diabetes specialised wards.

INSERT FIGURE 2

Despite using data for a different year (2003 vs. 2005) and a slightly different specification, the results confirm previous evidence (Lippi Bruni, Nobilio & Ugolini 2009). Patients' characteristics are again strong predictors of the probability of hospitalisation. As for financial incentives, we observe a significant association between the health outcome and the additional payments provided to GPs through pay-for-participation programs for diabetes care. This confirms that, given the design of incentives in the present institutional context, the role of pay-for-participation program can be better evaluated if one takes a disease-based perspective, instead of the more aggregate approach adopted throughout the paper.

6. Conclusions

In this paper we investigate the influence of economic incentives included in Regional and Local Health Authority contracts for primary care in the Italian region Emilia-Romagna (year 2005). We consider two different measures of avoidable admissions as dependent variable and three groups of programs: pay-for-performance, pay-for-participation and pay-for-compliance. The first dependent variable is obtained from the list of 27 medical DRGs that Emilia-Romagna identifies as at risk of organizational inappropriateness in primary care, while the second is the rate of preventable hospitalisations measured by ACSCs.

Data clustering across LHAs and GPs leads to the choice of multilevel modelling. Patients' characteristics emerge as the most important factors influencing the two dependent variables and all coefficients are fairly robust across specifications. Several physician characteristics display a significant impact, specially when we consider the 27 medical DRGs. In this case, younger GPs, those working in rural areas and in single handed practices show a higher probability to have their patients hospitalized for a condition at risk of inappropriateness.

In a policy perspective, the paper's main aim is to evaluate the impact of financial incentives introduced to promote good quality of primary care and, consequently, appropriate use of hospital resources. Two main results emerge from the analysis. First, taking two different sets of indicators as dependent variables produces fairly different results. When we refer to indicators, such as the list of 27 DRGs identified as important policy target by the Regional Health Authority we find that,

other things equal, additional financial transfers reduce the probability of episodes of inappropriateness of care. At the opposite, moving to indicators such as ACSCs, that are internationally accepted, but for which no specific action has been undertaken for raising GPs awareness in the present institutional context, we observe no relation between the amount of financial incentives received by a GP and the probability of experiencing avoidable hospitalisation for his patients. This first result suggests that explicit institutional support is crucial for making quality improvement activities successful.

A second important finding concerns the different impact produced by different types of programs. Our estimates show that, the larger the share of revenues provided to GPs through pay-for-performance programs, the lower the probability of their patients to experience hospitalisation for the 27 DRGs at risk of inappropriateness. A similar result holds also for pay-for-compliance programs, although the coefficient is significant only at the 10% level. At the opposite, pay-for-participation programs, despite providing a relatively larger amount of money, are never significant. We argue that the latter result is probably due to the fact that such scheme is mostly used for improving the management of particular diseases. Consequently, its impact is more precisely evaluated by choosing disease specific outcomes as dependent variable, as we document for diabetes care.

We conclude with a few cautionary remarks. First, our analysis bears the limitations implicit in the use of cross sectional data for the identification of strong causal relationship. Nonetheless, it must be reminded that incentive programs are contracted at the district level and, once an agreement is reached, GPs working in the district are part of it. This implies that, the largest share of variability in economic incentives is recorded between rather than within districts. At the same time, movements of physicians across districts are extremely rare. The propensity of citizens to change GP is limited, and mostly motivated by changes in residence. In this context, the time spam required to create a sufficiently large list in a new district is so long that the (opportunity) costs of moving becomes prohibitive. Given this, the possibility for a GP to participate in incentive based programs is crucially influenced by the district of reference and by the policy developed thereby but the choice of where to open the practice is typically taken once and for all and responds to criteria that are probably unrelated to the amount of incentives provided by each district. The largest majority of GPs took this decision well before the start of incentive based programs and also for the younger physicians other considerations such as local market conditions are likely to be more important.

Altogether, these institutional features greatly reduce the risk of endogeneity of economic incentives, when measured at the level of the individual GP.

A second limitation concerns the difficulty to assess the overall social and financial returns of the money invested for improving quality in primary care. Despite covering the general population and a wide set of conditions, the measure of benefits used in the paper is still narrow with respect to all potential improvements generated by incentive based programs of the kind considered here. They can stimulate activities that produce better outcomes of care, or financial savings, that fail to be captured by reductions in avoidable hospitalisations. As long as expected benefits pertain also to dimensions that go beyond a more appropriate use of hospital resources, to provide conclusive statements in terms of cost benefit analysis for these policy actions calls for more comprehensive studies and our work can be seen as first step towards this challenging direction.

Acknowledgments

This paper is part of the research projects: "The role of economic incentives in primary care", funded by the Italian Ministry of Health and the Strategic Research Project on "Health Economics: contracts, procurements and innovation", funded by the University of Bologna. Elisa Iezzi thanks "Fondazione Cassa dei Risparmi di Forlì" for the financial support. The authors thank participants to the ECHE (Rome), SIEP (Pavia) and AIES (Matera) conferences for useful suggestions and the Regional Agency for Health and Social Care for kindly providing the database used in this study. The opinions expressed here are the sole responsibility of the authors and do not represent the views of the Agency. Usual disclaimers apply.

References

- Billings, J., Zeitel, L., Lukomnik, J, Carey, TS, Blank, AE, & Newman, L. (1993). *Impact of Socioeconomic Status On Hospital Use In New York City*. Health Affairs, 12, 162-173.
- Birkmeier N.J., Birkmeier J.D. (2006). Strategies for improving surgical quality – should payers reward excellence or effort? *The New England Journal of Medicine*, 354, (8), 864-70.

- Browne, W.J., Subramanian, S.V., Jones, K., & Goldstein, H. (2005). Variance partitioning in multilevel logistic models. *Journal of the Royal Statistical Society, Series A* 168, 599-613.
- Caminal, J., Starfield B., Sanchez E., Casanova C., & Morales M. (2004). The role of primary care in preventing ambulatory care sensitive conditions. *European Journal of Public Health*, 14 (3): 246-51.
- Charlson, M., Pompei, P., Ales, K., & McKenzie, C. (1987). A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *Journal of Chronical Disease*, 40: 373-83.
- Cutler, D.M. (2006). The Economics of Health System Payment. *De Economist*, 154 (1): 1-18.
- Dixit, A. (2002). Incentives and organisations in the public sector: an interpretative review. *The Journal of Human Resources*, 37 (4): 696-727.
- Doran, T., Fullwood, C., Gravelle, H. et al. (2006). Pay-for-performance programs in family practices in the United Kingdom. *New England Journal of Medicine*, 355: 375-84.
- France, G., Taroni, F., & Donatini, A. (2005). The Italian health-care system. *Health Economics*, 14: S187-S202.
- Frey, B.S., & Jegen, R. (2001). Motivation crowding out theory: A survey of empirical evidence. *Journal of Economic Surveys*, 15: 589-611.
- Gneezy, U., & Rustichini, A. (2000). Pay enough or don't pay at all. *Quarterly Journal of Economics*, 115 (3): 791-810.
- Godager, G., Iversen T., & Ma A. (2009). Service motives and profit incentives among physicians. *International Journal of Health Care Finance and Economics* (in press).
- Goddard, M., Mannion, R., & Smith, P. (2000). Enhancing performance in health care: a theoretical perspective on agency and the role of information. *Health Economics*, 9: 95-107.
- Goldstein, H. (2003). *Multilevel statistical methods* (3rd edition). London: Kluwer Academic.
- Guo, G., Zhao, H. (2000). Multilevel modelling for binary data. *Annual review of sociology*. 26: 441-462.

- Heckman, J.J., Smith, J.A., & Taber, C. (1996). What Do Bureaucrats Do? The Effects of Performance Standards and Bureaucratic Preferences on Acceptance into the JTPA Program. NBER Working Papers 5535.
- Iversen, T., Luras, H. (2000). Economic Motives and Professional Norms: The Case of General - medical Practice. *Journal of Economic Behavior & Organization*. 43, 447-70.
- Josephson G.W., & Karcz A. (1997). The impact of physician economic incentives on admission rates of patients with ambulatory sensitive conditions: an analysis comparing two managed care structures and indemnity insurance. *The American Journal of Managed Care*, 3 (1): 49-56.
- Laffont J.J., & Martimort, D. (2002). *The Theory of Incentives: the principal-agent model*. Princeton: Princeton University Press.
- Lavy, V. (2004). Performance Pay and Teachers' Effort, Productivity and Grading Ethics, NBER Working Papers 10622.
- Lazear, E.P. (1996). Performance Pay and Productivity, NBER Working Papers 5672.
- Lippi Bruni, M., Nobile, L. & Ugolini, C. (2009). Economic incentives in general practice: the impact of pay-for-participation and pay-for-compliance programs on diabetes care. *Health Policy*, (in press).
- Mannion, R., Goddard, M., & Bate, A. (2007). Aligning incentives and motivations in health care: the case of earned autonomy. *Financial Accountability & Management*, 23 (4): 401-20.
- Prendergast, C. (1999). The provision of incentives in firms. *Journal of Economic Literature*, 37, 7-63.
- Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., & Healy, M. et al.(2000). Guide to MLWIN, VERSION 2.1. London: Institute of education.
- Rice, N., & Jones, A. (1997). Multilevel models and health economics. *Health Economics*, 6: 561-75.
- Romano, P.S., Roos, L.L., & Jollis, J.G. (1993). Adapting a clinical comorbidity index for use with ICD-9-CM administrative data: differing perspectives. *Journal of Clinical Epidemiology*, 46: 1075-1079.
- Scott, A., Schurer, S., Jensen, P.H., & Sivey, P. (2008). The Effects of Financial Incentives on Quality of Care: The Case of Diabetes. HEDG Working Paper, University of York.

Smith, P. (1995). On the unintended consequences of publishing performance data in the public sector. *International Journal of Public Administration*, 18: 277-310.

Smith, P. (2002). Measuring health system performance. *European Journal of Health Economics*, 3: 145-48.

Table 1: 27 medical DRGs at risk of inappropriateness in primary care

014 INTRACRANIAL HEMORRHAGE OR CEREBRAL INFARCTION
019 CRANIAL & PERIPHERAL NERVE DISORDERS W/O CC
025 SEIZURE & HEADACHE AGE >17 W/O CC
065 DYSEQUILIBRIUM
088 CHRONIC OBSTRUCTIVE PULMONARY DISEASE
089 SIMPLE PNEUMONIA & PLEURISY AGE >17 W CC
090 SIMPLE PNEUMONIA & PLEURISY AGE >17 W/O CC
091 SIMPLE PNEUMONIA & PLEURISY AGE 0-17
127 HEART FAILURE & SHOCK
131 PERIPHERAL VASCULAR DISORDERS W/O CC
133 ATHEROSCLEROSIS W/O CC
134 HYPERTENSION
142 SYNCOPE & COLLAPSE W/O CC
182 ESOPHAGITIS, GASTROENT & MISC DIGEST DISORDERS AGE >17 W CC
183 ESOPHAGITIS, GASTROENT & MISC DIGEST DISORDERS AGE >17 W/O CC
184 ESOPHAGITIS, GASTROENT & MISC DIGEST DISORDERS AGE 0-17
208 DISORDERS OF THE BILIARY TRACT W/O CC
243 MEDICAL BACK PROBLEMS
245 BONE DISEASES & SPECIFIC ARTHROPATHIES W/O CC
256 OTHER MUSCULOSKELETAL SYSTEM & CONNECTIVE TISSUE DIAGNOSES
294 DIABETES AGE >35
324 URINARY STONES W/O CC
395 RED BLOOD CELL DISORDERS AGE >17
426 DEPRESSIVE NEUROSES
427 NEUROSES EXCEPT DEPRESSIVE
429 ORGANIC DISTURBANCES & MENTAL RETARDATION
467 OTHER FACTORS INFLUENCING HEALTH STATUS

Table 2: Ambulatory Care Sensitive Condition ICD-9-CM Codes

Angina	411.1, 411.8, 413. Excludes cases with procedure codes [01-86.99]
Asthma	493
Bacterial pneumonia	481, 482.2, 482.3, 482.9, 483, 485, 486. Excludes cases with secondary diagnosis of sickle cell [282.6].
Cellulites	681, 682, 683, 686. Excludes cases with any procedure codes except 860 where it is the only procedure
Chronic obstructive pulmonary disease	491, 492, 494, 496, 466.0
Congestive heart failure	428, 402.01, 402.11, 402.91, 518.4
Dehydration - volume depletion	276.5
Diabetes	250.1, 250.2, 250.3, 250.8, 250.9, 250.0, 251
Gangrene	785.4
Gastroenteritis	558.9
Grand mal status and other epileptic convulsions	345, 780.3
Hypertension	401.0, 401.9, 402.00, 402.10, 402.90. Excludes cases with procedures 36.01, 36.02, 36.05, 36.1, 37.5, 37.7.
Hypoglycemia	251.2
Hypokalemia	276.8
Immunization-related and preventable conditions	032, 033, 037, 045, 055, 072, 320.0, 390, 391
Kidney/urinary infection	590, 599.0, 599.9
Pelvic inflammatory disease	614 (Excludes 68.3-68.8)
Peptic ulcer	[531, 532, 533]
Pulmonary tuberculosis and other tuberculosis	011, 012-018
Pyelonephritis	590
Ruptured appendix	540.0, 540.1
Severe ear, nose, and throat infections	382, 462, 463, 464, 465, 472.1
Skin grafts with cellulitis	DRG 263, DRG 264

Table 3: Cross tabulation of DRGs and ACSCs

ACSCs	DRGs		Total
	0	1	
0	2911926 (99.16%)	7634 0.26%	2919460
1	13006 (0.44%)	3918 (0.13%)	16924
Total	2924832	11552	

Table 4: Descriptive statistics. Patient and GP characteristics, year 2005

<i>Explanatory variable</i>	<i>Coding</i>	<i>Mean/ Proportion</i>	<i>Standard deviation</i>	<i>Min</i>	<i>Max</i>
<i>Patient level (n=2936384)</i>					
Patient gender	male=1	49.7			
Patient age	Continuous	46.4	12.8	18	74
Charlson index	Continuous	0.1	1.1	0	50
<i>Physician level (n=3229)</i>					
GP gender	male=1	72.1			
GP age	Continuous	52.0	5.4	72	32
Practice location urban	(if yes=1)	95.5			
Single handed practice	(if yes=1)	44.7			
Pay-for-performance	Continuous (% annual income)	0.4	1.4	0.0	39.6
Pay-for-participation	Continuous (% annual income)	5.4	4.4	0.0	56.8
Pay-for-compliance	Continuous (% annual income)	0.1	0.4	0.0	7.6
List size per GP	Continuous	1071.0	510.4	1	2521
List average age	Continuous	50.0	5.9	14	77
<i>District (n=39)</i>					
District total bed	Continuous	544.7	768.6	18	4290
Hospitalization rate	Continuous (% population)	19.4	9	4.8	33.5

Table 5: Economic incentives. Local Health Authorities, amounts in Euro, year 2005

		Pay-for-performance					Pay-for-participation					Pay-for-compliance				
LHA	GP	% GP	MIN	MAX	MEAN	STD	% GP	MIN	MAX	MEAN	STD	% GP	MIN	MAX	MEAN	STD
1	211	82%	0	4222	1977	1300	98%	0	49214	7755	6498	8%	0	3966	119	525
2	306	2%	0	4662	69	545	99%	0	12273	2264	2151	34%	0	3769	130	321
3	345	15%	0	3532	185	566	100%	0	21505	6781	2993	12%	0	8652	157	940
4	516	0%	0	878	3	55	98%	0	30333	6248	5118	10%	0	2066	96	334
5	639	83%	0	21780	628	1662	99%	0	16586	4331	3305	3%	0	5440	22	234
6	97	46%	0	5319	805	1636	100%	80	17287	3907	3417	5%	0	3130	120	580
7	297	0%	0	0	0	0	99%	0	26063	8038	4976	6%	0	176	6	28
8	294	11%	0	2436	47	197	99%	0	10964	2626	1516	55%	0	3906	204	420
9	148	0%	0	0	0	0	99%	0	30659	4289	4529	14%	0	1352	54	193
10	150	0%	0	0	0	0	97%	0	10705	3095	2051	13%	0	1963	109	313
11	226	0%	0	0	0	0	97%	0	11169	2364	1871	6%	0	1555	30	146

Figure 1. Economic incentives in % GP annual income. Districts, year 2005

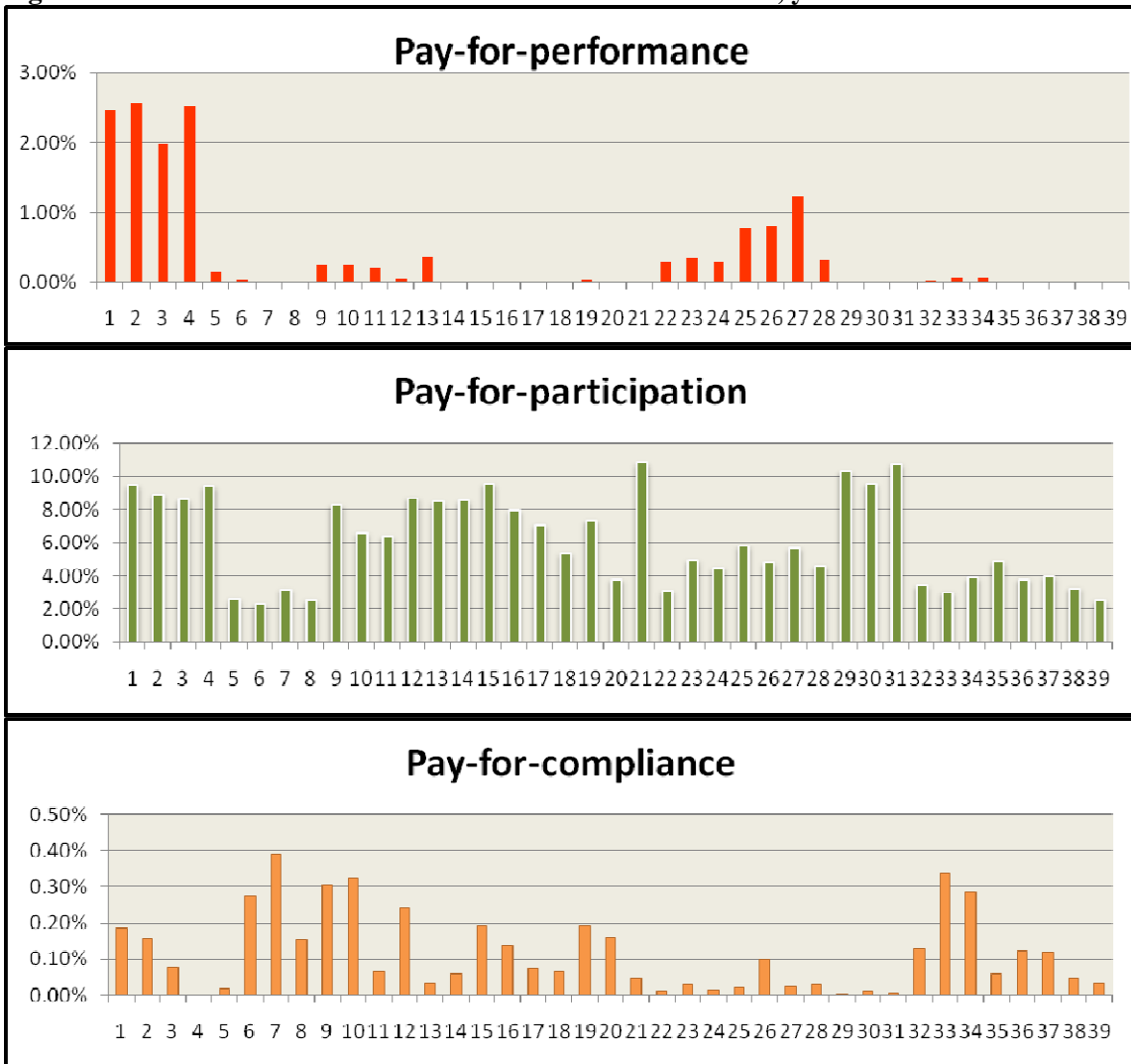


Table 6. GP and patient factors predicting an avoidable hospital admission, year 2005.

	MODEL A – 27 DRGs			MODEL B – ACSCs		
	Three level			Three level		
<i>Explanatory variables</i>	<i>Coefficient</i>	<i>SE</i>	<i>p ></i>	<i>Coefficient</i>	<i>SE</i>	<i>p ></i>
FIXED EFFECTS						
Constant	-7.084	(0.174)	***	-7.219	(0.173)	***
<i>Patient level</i>						
Patient gender	-0.108	(0.019)	***	-0.399	(0.016)	***
Patient age	0.049	(0.001)	***	0.062	(0.001)	***
Charlson index	0.003	(0.000)	***	0.004	(0.000)	***
<i>Physician level</i>						
GP gender	0.014	(0.024)		-0.002	(0.021)	
GP age	-0.006	(0.002)	***	-0.003	(0.002)	
List per GP	1.80E-05	(2.50E-05)		2.80E-05	(2.10E-05)	
List average age	-0.011	(0.003)	***	-0.017	(0.003)	***
Single handed practice	0.039	(0.017)	**	-0.002	(0.017)	
Practice location urban	-0.232	(0.041)	***	-0.197	(0.045)	***
Pay-for-performance	-0.022	(0.011)	**	-0.003	(0.010)	
Pay-for-participation	0.004	(0.003)		0.000	(0.003)	
Pay-for-compliance	-0.039	(0.024)	*	-0.014	(0.020)	
<i>District area level</i>						
Hospitalization rate	0.003	(0.001)		0.004	(0.004)	
District total beds	-2.72E-06	(9.10E-06)		-1.17E-06	(4.20E-05)	
RANDOM EFFECTS						
Level 2 - $\sigma^2 (u_{0jk})$	0.020	(0.007)	***	0.019	(0.005)	***
Level 3 - $\sigma^2 (v_{0k})$	0.029	(0.009)	***	0.037	(0.009)	***
ρ GP	0.006			0.006		
ρ districts	0.009			0.011		
Deviance [-2ln(L)]	23960851			23960418		

*** p-value \leq 0.01 ** p-value \leq 0.05 * p-value \leq 0.10

Table 7: Descriptive Statistics for Diabetes. Patient and GP characteristics, year 2005

<i>Explanatory variable</i>	<i>Coding</i>	<i>Mean/ Proportion</i>	<i>Standard deviation</i>	<i>Min</i>	<i>Max</i>
<i>Patient level (n=164574)</i>					
Patient gender	male=1	50.3			
Patient age	continuous	67.9	12.833	35	107
No insulin dependence	(if yes=1)	15.6			
<i>Physician level (n=2938)</i>					
GP gender	male=1	74.2			
GP age	continuous	50.8	5.539	35	71
Practice type	single-handed	31.4			
Practice location urban	(if yes=1)	94.2			
Postgraduate qualification	(if yes=1)	5.1			
Pay-for-participation	continuous (% annual income)	0.2	0.718	- 0.006	8.841
Diabetic list size per GP	continuous	56.8	20.354	4	129
List average age	continuous	50.8	5.533	35	71
<i>District (n=38)</i>					
Hospital beds in endocrinology	continuous	15.7	9.888	2	28

Figure 2. Pay-for-participation incentives for diabetes in % GP annual income. Districts, year 2005

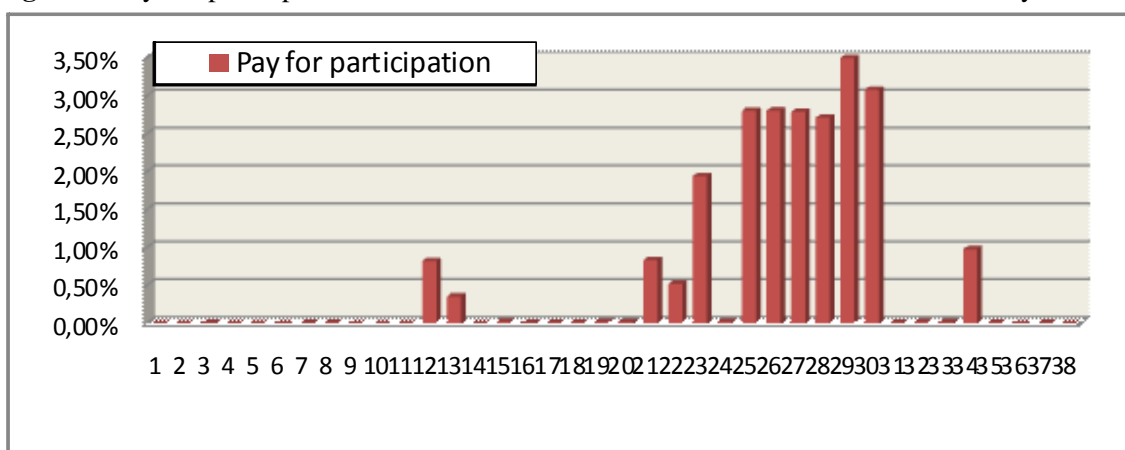


Table 8. Diabetes: GP and patient factors predicting an avoidable hospital admission, year 2005.

	MODEL A – DIABETES		
	Three level		
<i>Explanatory variables</i>	<i>Coefficient</i>	<i>SE</i>	<i>p ></i>
FIXED EFFECTS			
Constant	-6.473	1.065	***
<i>Patient level</i>			
Patient gender	0.017	0.003	***
Patient age	0.239	0.070	***
No insulin dependence	-1.232	0.070	***
<i>Physician level</i>			
GP gender	0.005	0.006	
GP age	-0.165	0.089	*
Single handed practice	0.126	0.076	*
Practice location urban	-0.315	0.127	***
List size	-6.00E-05	1.11E-04	
List average age	0.009	0.015	
Pay-for-participation	-0.045	0.025	**
<i>District area level</i>			
Hospital beds in endocrinology	0.006	0.004	
RANDOM EFFECTS			
Level 2 - $\sigma^2 (u_{0jk})$	1.27E-20		
Level 3 - $\sigma^2 (v_{0k})$	0.118		
ρ GP	3.73E-21		
ρ districts	0.035		
Deviance [-2ln(L)]	1356843		

*** p-value ≤ 0.01

** p-value ≤ 0.05

* p-value ≤ 0.10