The role of GP's compensation schemes in diabetes care: evidence from panel data.

Elisa Iezzi*, Matteo Lippi Bruni**, Cristina Ugolini**

* Department of Statistics, University of Bologna

**Department of Economics University of Bologna and CHILD

Preliminary Draft (this version 25th November 2009)

ABSTRACT

The design of incentive schemes that improve quality of care is a central issue for the healthcare sector. Nowadays we observe many pay-for-performance programs, where payment is contingent on meeting indicators of provider effort, but also other alternative strategies have been introduced, for example programs rewarding physicians for participation in diseases management plans. Although it has been recognised that incentive-based remuneration schemes can have an impact on GP behaviour, there is still weak empirical evidence on the extent to which such programs influence health outcomes. We investigate the impact of financial incentives in Regional and Local Health Authority contracts for primary care in the Italian Region Emilia Romagna for the years 2002-05. We focus on avoidable hospitalisations (Ambulatory Care Sensitive Conditions) for patients affected by type 2 diabetes mellitus, for which the assumption of responsibility and the adoption of clinical guidelines are specifically rewarded. We estimate a panel count data model using a Negative Binomial distribution to test the hypothesis that, other things equal, patients under the responsibility of GPs receiving a higher share of their income through these programs are less likely to experience avoidable hospitalisations. Our results support the hypothesis that financial transfers may contribute to improve quality of care, even when they are not based on the ex-post verification of performances.

Key words: primary care, quality, diabetes, avoidable hospitalizations, panel count data models.

JEL classification: I11, I18, C31

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

1 Introduction

Financial incentives have been extensively employed for improving performances of healthcare providers not only in insurance-based systems but also in countries with a National Health Service (NHS). Probably, the best known example is the Payment by Results program of the British NHS for financing hospital trusts, but the UK extended target payments also to primary care through the Quality and Outcomes framework (e.g. Campbell et al. 2009). The British experience is not unique as proved by the implementation in the Italian NHS of programs that provide General Practitioners (GPs) with extra-payments exceeding standard capitation. These transfers are usually paid to GPs for meeting targets of general interest or for taking part in care improvement activities promoted by public authorities.

In NHS systems, the attention to the design of an effective governance of primary care has increased as a consequence of the pivotal role attributed to GPs as providers of ambulatory care, as well as gatekeepers to secondary care. Consequently, public interest focuses not only in improving quality of primary care, but also in involving GPs in cost containment strategies conceived on a wider scale.

The common purpose of these programs is to promote physicians' internalisation of policymaker's objectives and the range of activities potentially involved is extremely diversified. Special bonuses may be provided for meeting pre-specified targets in order to promote the alignment of GPs' referral and prescription decisions to general health policy goals, usually associated to a more appropriate use of resources (e.g. prescription rate of generic drugs, containment of hospitalisation rates etc…). Nevertheless, financial incentives may reward also direct provision of treatments (e.g. immunisation uptake), assumption of responsibility for patient affected by diseases that require additional physician's effort (e.g. diabetes, hypertension), or the adoption of organisational routines aimed at improving cooperation among providers, such as participation in networks of GPs or adherence to evidence-based guidelines.

Another important distinction concerns performance monitoring. Pay-for-performance schemes typically associate financial transfers to the achievement of verifiable targets agreed in advance. Alternative approaches may choose to reward participation in care improvement activities, without necessarily linking additional funds to the attainment of specific objectives defined in terms of financial or epidemiological indicators to be verified ex-post. The first approach has the advantage of introducing a more stringent incentive structure, thus increasing its capacity to influence GPs behaviour in the desired direction. Whereas, the second approach is less intrusive of physicians professional autonomy and there are circumstances where it may prove effective in encouraging and promoting cooperation between independent providers, such as GPs, and the different actors of the public system.

Given the variety of possible institutional arrangements, the research agenda has suggested a number of relevant policy questions. Are primary care physicians' influenced by economic incentives in their referral and prescribing decisions? GPs' responses can be obtained only through large financial transfers or also with (relatively) small scale programs? To what extent are the results sensible to performance monitoring? Should financial incentives cover wide areas of intervention or do they work effectively also when focused on disease specific objectives? Does the target of the program (e.g. cost containment vs. quality enhancement) affect its effectiveness? What is the impact on quality of care? Do we observe differences across countries? Can they be attributed to institutional characteristics or are they associated to different practice style, local clinical culture and habits?

The array of experiences developed around the world provides a rich and promising area of investigation for evaluating the effectiveness of different incentive schemes based on financial transfers to GPs that add up to standard capitation. Our work contributes to the literature on program evaluation in primary care by analysing the case of diabetes in the Italian Region Emilia Romagna for the period 2002-2005. Emilia Romagna launched a Diabetes Project in 2003 for improving clinical appropriateness in disease management, through which GPs were assigned the home care responsibility of type 2 diabetic patients. In order to promote the accomplishment by GPs of a number of activities expected to improve quality of diabetes care, Local Health Authorities (LHAs) have been allowed to provide specific compensations to their GPs for activities such as regular reviews of patients, periodic measure of glycosylated haemoglobin, participation to the local diabetes management program. The identification of the activities to be incentivised and the extent of the corresponding financial transfers are bargained between health authorities and the representative organisations of GPs at the district level. Consequently, we record wide differences across districts in the amount of economic incentives.

The sources for this study are regional databanks including detailed information on the sources of GPs professional income and on the use of healthcare services by all regional patients (episodes of hospitalisation, prescription of drugs etc…). Information on drug utilisation and access to specialised centres allows to identify patients suffering from diabetes type 2. Then, the comprehensive information on economic incentives received by each GP permits us to investigate the relation between the share of funds received for special programs related to diabetes care and quality of care provided to registered patients. It is well documented in the clinical literature on diabetes mellitus type 2 that timely and accurate ambulatory care should be able to prevent deterioration of the patient's health status that can ultimately lead to emergency hospitalisation (ADA, 2002; Booth and Fang, 2003; Ices 2003; Fleming 2004, RACGP, 2009) In particular, we measure quality of diabetes care with the Ambulatory Care Sensitive Conditions (ACSCs) developed by Billings et al. (1993) for this chronic disease. According to the literature on ACSCs (Purdy et al. 2009), a high frequency of hospital admissions for such episodes is typically associated to deficiencies in disease management and

inadequate patient supervision, as appropriate outpatient care should be effective in preventing complications and therefore hospital admission.[1]

The paper tests the hypothesis that, other things equal, the higher the fraction of professional income a GP receives from special payments for diabetes care, the lower the number of avoidable hospitalisations (i.e. diabetic ACSCs) experienced by his type 2 diabetic patients. By doing so, we verify whether physician respond to economic incentives by improving the quality of care and patient supervision, measured by a reduction in (avoidable) adverse outcomes.

Our work contributes on several dimensions to the literature on the role of financial incentives on physician behaviour. First, differently from most of the experiences analysed so far, the incentive scheme implemented in this case is designed at compensating participation in caring improvement activities, such as assumption of responsibility of patients, adherence to guidelines etc, rather than rewarding high level of performances. A second important improvement is ensured by the use of a large dataset that covers the entire regional population for four years. The initial study population amounts to 2618087 inhabitants aged 35 or more, from which 164574 diabetic patients are extracted. The use of longitudinal data allows a more robust identification than it was done in previous works (Lippi Bruni, Nobilio, Ugolini 2009) of the link between the relative amount of economic incentives and outcomes of care. In particular, by exploiting variation across time and GPs, the effect of financial transfers associated to diabetes care can be isolated from confounding factors such as generalised improvement in clinical practice.

2. The role of financial incentives for quality improvement

Financial incentives aimed at improving quality of healthcare services are usually designed as pay-for-performance programmes where remuneration is conditional on achieving measurable targets, that reflect clearly identified policy goals. Whereas empirical evidence suggests that physicians respond to changes in compensation, studies where quality measures are introduced explicitly are at earlier stages of development. Numerous evaluations have focused on the Quality and Outcomes framework in UK using various methodologies to assess trends in quality indicators before and after introduction of the program in 2004. At this regards, Campbell, Reeves et al. (2009) find an improvement in quality for two of the three chronic conditions considered (diabetes and asthma but not for heart disease) in the short term but, once targets were reached, rates of quality improvement slowed considerably for all the three conditions, whereas continuity and the quality of care for conditions not linked to incentives declined significantly. In their conclusions, the pay-for-performance scheme introduced in the UK

[1] For example, ACSCs include short-term complications of diabetes mellitus such as diabetic ketoacidosis, hyperosmolarity, and coma. These life-threatening emergencies arise when a patient experiences an excess of glucose (hyperglycemia) or insulin (hypoglycemia) but with timely and high-quality outpatient care hospitalization for these -severe conditions should be preventable.

initially produced significant improvements in measurable dimensions of clinical performance but such acceleration was not sustained over the subsequent years.

Focusing on diabetes, Scott et al. (2009) find a positive but moderate impact on quality of care in diabetes management, measured by the probability of ordering an HbA1c test, as a consequence of the incentives introduced in Australia in 1999. Young et al. (2007) investigate the effect of a program conferring limited financial risk to primary care physicians in Rochester, US, between 1999 and 2004 and find no difference between the post- and pre-intervention trends indicating that the overall increase in performance was largely independent from the incentive program.

Some studies suggest that both organization- and physician-level incentives have a measurable impact on quality improvement but also that financial incentives are more influential when based on individual performance. Moreover, physicians' lack of responsiveness may be due to payments that are too small (Conrad and Perry, 2009). In a comprehensive survey, Christianson, Sutherland and Leatherman (2009) conclude that studies with stronger research design report no impacts of financial incentives on quality and so the evidence for justifying the growing use of pay-for-performance schemes is thin and inconclusive.

The empirical literature on payments for quality improvement faces three main challenges. First, incentives schemes are often complicated and nuanced, introduced on top of, or blended with, payment mechanisms aimed to pursue different scopes. This makes difficult to assess the independent effect of the financial scheme. Second, financial incentives can be manipulated owing to the incompleteness of many performance measures, leading to a concentration of efforts only on the areas included in the performance indicator scheme. Third, paying for performance could diminish intrinsic motivation and satisfaction to perform the task for its own sake. A precondition for crowding-out (Frey, 1993; Le Grand, 2003) is that individuals are sufficiently intrinsically motivated. The more important is the task— in the sense that the agent would perform it independently of reward — the more likely they become. Such relationships are particularly relevant for non-repetitive jobs with high discretion, e.g. physicians motivated by ethics and professionalism. In these contexts, if financial incentives are perceived to trivialize the non-financial motivations, their use could be harmful. This could be especially true in NHS systems where monetary incentives could devalue providers altruistic motivation (Siciliani, 2009).[2] Indeed, there are contradictory theories on the power of incentives in the public sector as public organizations are endowed with a set of unique characteristics - among others, the lack of markets for goods provided by public organizations and the non-monetary motivations of

[2] As regards the Quality and Outcome framework, McDonald, Harrison et al (2008) report that the financial incentives in the pay-for-performance program have not harmed the internal motivation of GPs; while Campbell, McDonald et al (2008) note that incentives have caused resentment on the part of some nurses about the distribution of bonus payments as they were given responsibility for achieving performance targets. Whalley, Gravelle et al (2008) in a survey of UK physicians find improvements in satisfaction relating to hours worked and remuneration.

many public sector employees - that may make it more difficult and less necessary to motivate public employees by paying them according to the value of their output (Andersen, 2009).

Contrariwise, financial incentives could also crowd "in" intrinsic motivation when agents perceive the external rewards as informative or supportive. Bénabou and Tirole (2003) formalize the distinction between intrinsic and extrinsic motivation indicating the importance of both while identifying the overall effect of external interventions on motivation. In their model the crowding out effect is more likely when the reward is offered to agents with limited ability and/or for unattractive tasks. In this case, the agent thinks the principal knows more about his abilities than he himself, so he takes the provision of incentives as a bad sign. In other words, rewards may be weak reinforcers in the short term but they could be counterproductive if they reduced the agent's motivation to undertake the tasks in the future and they become negative reinforcers when they are withdrawn. On the contrary, by offering low-powered incentives or non-contingent payments, the principal could signal his confidence in the agent's ability inducing stronger crowding-in effects. In this case the short-term implications of the principal's confidence-management motive would reduce the slope of the agent's compensation schedule but could strengthen the agent's intrinsic motivation in the long run. Striking a balance between wages, monetary rewards and monitoring is particularly important in the health sector, where production and delivery are highly labour intensive. Indeed, worker motivation is a key determinant of health sector performance where resource availability and worker skills are often necessary but not sufficient to ensure the desired performance.

In addition, theoretical work suggests that low-powered incentives can be optimal when agents allocate effort across multiple tasks some of which are observable only at high costs (Eggleston, 2005). These contracts usually combine an upfront payment (per patient or per unit of time) with a partial fee for service, trying to encourage the provision of features of care (e.g. quality) which are valuable for patients but not easily observable (Dumont et al., 2008). For example, to improve primary care, transfers of a fixed additional amount per chronic patient may help to reduce patients selection and contribute to align physicians' behaviour to the payer's and patient's objectives.

For these reasons, although interest in pay-for-performance mechanisms continues to grow, there are doubts about the ultimate effectiveness of these schemes in the presence of intrinsic motivations and multitasking. An emerging literature extends the general theory of economic incentives introducing concepts from the literature on Public Service Motivation (Andersen, 2009) such as identity with the organization goals, work ethic, non-pecuniary motivations and applying them to the health sector. Meanwhile, many countries such as US, Canada, Australia and New Zealand have been experimenting softer organisational drivers to improve performance, such as delegated autonomy (taking the form of decreased inspection and reporting requirements) and alternative schemes for quality improvements that try to align the interests of the principal with those of the agent by means of a looser incentive structure with respect to pay-for-performance (Smith, 2002; Birkmeier and Birkmeier, 2006; Mannion et al., 2007). The potential advantages of these schemes are providers adherence to evidence based

protocols, that help to reduce the risk of developing complications, but also increases in cooperation and coordination between the different actors of the caring process. However, such schemes present also potential shortcomings not to be overlooked, since the lack of binding connections between individual results and financial transfers may attenuate the influence on physicians behaviour. Systematic empirical literature on this topic is scarce but outlines that there are cases in which low-powered incentives may have a positive impact on the quality of care (Dumont et al., 2008).

3. The organisation of diabetes care

In Italy, primary care physicians work as independent professionals contracted with the NHS and are on charge of delivering primary care to the citizens registered with them. Moreover, GPs are gatekeepers to NHS-funded specialist and hospital care. LHAs are divided into Healthcare Districts (HDs) that have the responsibility of organising and coordinating outpatient specialist services, residential and primary care within their territory. Primary care services are free at the point of demand and citizens choose the GP with whom to register. Although the choice of the GP is reversible at any time, turnover rates are extremely limited and, in most cases, they are a consequence of a change in residence. In year 2000 it has been established a maximum list size of 1500 patients. GPs exceeding the limit when the new regulation was approved, were allowed to keep their additional patients but they cannot add new ones until their list has fallen behind the threshold.

Over the last two decades, the Italian NHS has been characterised by a devolution of powers to regional governments which has opened to regions new opportunities for experimenting institutional innovations also in the area of primary care. In Emilia Romagna - a north-eastern region with around 4 millions inhabitants - interventions have been carried out in order to further coordination between GPs and LHAs, using also economic incentives to complement the existing GPs' payment scheme based on capitation. In the policymaker's view these financial tools are aimed at promoting a closer alignment of GPs' decisions with general health policy goals. The objectives span from improvements in the quality and appropriateness of care, to the adoption of referral and prescribing decisions expected to back the public payer's effort for cost containment.

The GPs' remuneration scheme is defined every three years in a process of nationally centralized bargaining with GPs' trade unions. The capitation component, that represents the most relevant part of GPs' income paid by the NHS, is contracted at the national level. Nevertheless, regions have considerable autonomy in defining the additional part of GPs' income and geographical differences over this dimension are documented not only across but also within regions.

The Regional Government sets priorities for primary care but LHAs and HDs benefit of extensive degrees of freedom in defining the activities to be incentivised through additional payments. As regards diabetes care, the Regional Health Authority started in 2003 a Diabetes Project that defines

roles and responsibilities of LHAs, HDs and GPs for disease management, adopts clinical guidelines based on best practices and provides a general framework for introducing specific financial incentives[3]. Nevertheless, local agreements between GPs', LHAs and HDs define the extent of additional payments associated to diabetes care. Each agreement involves all GPs operating in a particular district, identifies the activities rewarded and the corresponding financial amount. According to the Regional Diabetes Project, the activities that each district may compensate range from adherence to clinical guidelines to the assumption of responsibility of diabetes patients[4] or participation in audit meeting. Local diabetes management plans introduce additional payments in a variety of ways. In some cases, as for the assumption of responsibility of patients, GPs receive a financial transfer that increases capitation for each diabetes patient registered in their list. In other cases, as for attendance to audit meetings or contribution to dissemination of new protocols and guidelines, transfers are associated to specific activities promoted at the local level. Consequently, they are not related to the number of diabetes patients followed by each GP.

The main purpose of all these programs is to compensate GPs for additional time and effort required for delivering care to diabetes patients and to foster coordination between GPs, HDs and (when necessary) also secondary care facilities. As a consequence of specific local objectives, the agreements signed in each HD display large variability in terms of remunerated activities and in the size of the incentives, but their common purpose is to promote better diabetes management. For this reason we have grouped in a single variable ("*financial incentives*") all payments received by GPs for the activities aimed at improving the delivery of medical services to diabetes patients. [5]

Our aim is to study the influence of this variable on quality indicators for diabetes care and the number of diabetic ACSCs represents a good reference for our purposes for several reasons. First, since the most important avoidable admissions for diabetes are hyperglycaemic emergencies, a reduction in the frequency of these life threatening episodes ensures substantial health gains. Moreover, regular and accurate reviews of patient's conditions together with delivery of appropriate care should be able to avoid such complication. Finally, a reduction in the frequency of diabetic ACSCs or diabetic comas is not explicitly contracted upon in any of the districts that we consider. This ensures that the outcome

[3] Before 2003, there was no formal shared coordination between primary and secondary care but some districts have been experimenting various forms of financial bonus for diabetes since the early 2000s.

[4] As a consequence of assumption of responsibility, GPs must comply with a range of requirements that include the regular reviews of their diabetes patients, consolidating the patient's knowledge and skills regarding eating plan, physical activity, home blood glucose self-monitoring, foot care etc., reviewing medication usage (oral hypoglycaemic agents or insulin), the periodic measure of their blood pressure and glycosylated haemoglobin (HbA1c), the timely referral to community and hospital based specialists.

[5] In a previous work (Lippi Bruni et al. 2009), for which data only for a single year were available, we consider two variables for financial incentives according to whether transfers were paid on per patient basis or not. Since in several districts the classification of the different items of GPs remuneration has not been fully consistent over time, in the present longitudinal study we prefer to use a unique indicator, that groups all financial items associated to local diabetes programs.

measure used in the analysis captures a very general dimension of quality and goes beyond the more narrowly defined sets of activities that are explicitly remunerated. Moreover, we limit the risk of using indicators affected by "tunnel vision effect", according to which professional effort tends to be concentrated only on the tasks that are explicitly remunerated, thus leading to possible overestimations of the effectiveness of the incentive-based programs.

An important feature of diabetes incentives is that in Emilia Romagna, differently from what the Region does elsewhere, additional payments are not linked to strict ex-post monitoring of performance. The main justification relies in the attempt to limit possible drawbacks due to motivation crowding out and reduced willingness to cooperate with public authorities that strict monitoring may induce, when physicians perceive supervision intrusive of the patient-physician relationship. Such concern appears particularly relevant in a context where GPs maintain a large professional autonomy in referral and prescribing. Nonetheless, the major shortcoming of such schemes is that the associated incentive structure may be too weak and fail to generate substantial improvements in the patterns of care.

Given these premises, it is a open issue whether programs with these characteristics can ensure health gains to the population of chronically-ill patients. For this reason, an empirical evaluation of their impact in the present case is useful for the design of effective programs also in other contexts.

4. Methods

Data

The study population consists of all regional patients affected by diabetes type 2 followed from 2002 to 2005. According to WHO criteria, patients are classified as having type 2 diabetes if they were aged 35 or more at the time of diagnosis or if they are not currently treated with insulin. Consequently, we include all individuals above 35 years who had at least one prescription for diabetes medications (oral agents or insulin) during 2002. As some patients managed through a diet and exercise alone can be missed with this strategy, we also include individuals who had at least one outpatient visit to a diabetic centre during 2002 or an hospital admission with a diabetic diagnostic code in the previous two years. During the period 2002-05 the average number of GPs active each year amount to 2960.

Our outcome indicators are the ACSCs that refer to diabetes developed by Billings et al. (1993). Hospitalisations are identified from hospital records when ICD-9 codes 250.1, 250.2, 250.3, 250.8, 250.9, 250.0, 251 are documented as primary or most responsible diagnosis. The total number of adverse outcomes in the observation period is 4357, averaging to 1089 hospitalizations per year. The dependent variable is measured as the yearly number of diabetic ACSCs recorded among the patients included in the list of each GP and its frequency distribution is shown in Table 1. As expected, the

sample displays a substantial fraction of zero outcomes: 71.18% zeros on the total amount of hospital admissions referred to our diabetic population.

TABLE 1

Table 2 displays descriptive statistics for the estimated sample. For GPs we include gender, seniority and the presence of postgraduate qualifications. We also control for rural practice location and for the type of practice, distinguishing single-handed from groups. As regards patients list, we consider size, average patients' age and the number of diabetic patients. Other characteristics of the list are the number of insulin dependent patients and number of visits to a diabetic outpatient clinic (DOC), which are expected to capture severity.

TABLE 2

Figure 1 shows the distribution of incentives across local areas as share of GPs annual income. They include direct financial incentives for assumption of responsibility for each diabetic patient and payments for participation in improvement activities or for compliance with regional and local guidelines.

FIGURE 1

Statistical analysis

In our study the count variable y –number of diabetic ACSCs recorded in each GP's list per year- is non-negative and follows a positively skewed distribution of the nonzero realizations (Figure 2).

FIGURE 2

We consider Poisson and Negative binomial regressions. The Poisson distribution is identified by a single parameter λ, equal for mean and variance. It is convenient to specify λ as a log-linear function of the explanatory variables x_i that account for observed sample heterogeneity, so that λ_i is positive for any combination of x_s and β_s:

$$E(y_i|x_i) = \lambda_i = \exp(\beta_0 + \sum_j x_{ij}\beta_j) = \exp(\mathbf{x}'_i \boldsymbol{\beta}) > 0 .$$

When *equidispersion* is not satisfied, the negative binomial regression (Hilbe, 2007) is the most commonly used alternative to the Poisson and can be obtained by introducing latent heterogeneity in the conditional mean of the Poisson distribution (Hilbe and Greene, 2008):

$$E(y_i|x_i) = \exp(\mathrm{x}'_i \beta + \varepsilon_i), \qquad Cov[\mathrm{x}, \varepsilon] = 0$$

To fully characterise the model, distributive assumptions over ε are required. In the standard negative binomial model (Greene, 2008) $h_i = \exp(\varepsilon_i) \xrightarrow{d} Gamma$ with mean 1, variance $1/\theta$ and

$f(h_i) = \frac{\theta^{\theta}}{\Gamma(\theta)} \exp(-\theta h_i) h_i^{\theta-1}$ with $h_i > 0$, $\theta > 0$, where $\Gamma(.)$ denotes the gamma integral that brings to an integer argument.

The marginal distribution of y is a Poisson-gamma mixture with a closed form and the marginal negative binomial (NB) distribution $NB(\lambda, \theta)$ is:

$$f(y_i \mid \mathbf{x}_i) = \frac{\Gamma(\theta + y_i)}{\Gamma(\theta)\Gamma(y_i + 1)} \left(\frac{\theta}{\lambda_i + \theta} \right)^{\theta} \left(\frac{\lambda_i}{\lambda_i + \theta} \right)^{y_i}$$

The moments are $E(y_i \mid \mathbf{x}_i) = \lambda_i$ and $Var(y_i \mid \mathbf{x}_i) = \lambda_i + \frac{1}{\theta}\lambda_i^2$, where $\frac{1}{\theta}$ is a dispersion parameter measuring the extent of *overdispersion* and the conditional variance is always greater than the conditional mean. The negative binomial model in previous equation is labelled NB2 (Hilbe, 2007) in reference to the appearance of quadratic term for λ_i in the conditional variance function. The generalized event count model is NBP for which NB1 and NB2 represent special cases with P=1 and P=2.[6]

We employ a likelihood ratio (LR) test of overdispersion parameter $1/\theta$ to evaluate the validity of the Poisson specification against the negative binomial model ($H_0 : 1/\theta = 0$). In order to choose between NB1 and NB2 (Greene 2008), as these models are not nested, we use the Vuong test (1989), whereas for selecting between NB1 and NBP or NB2 and NBP, we use a LR test since these models are nested.

Specifications for nonlinear panel data include pooled, fixed and random effects. A first possibility is simply to ignore time dependence. By pooling data together, the estimated standard errors will not be reliable for hypothesis testing. To address the problem, one may use a panel-robust estimate of variance that correct standard errors for dependence over time for any given individual.

Less naïve approaches assume correlation over time ($t = 1,...,T$) for observations pertaining to the same individuals i ($i = 1,...,n$). The fixed effect model incorporate individual heterogeneity with an individual-specific intercept term:

$$\log \lambda_{it} = \alpha_i + \boldsymbol{\beta}' \mathbf{x}_{it} + \varepsilon_{it},$$

where α_i can be interpreted as the coefficient of a binary variable taking value 1 for observations belonging to the *i-th* group. This model can be estimated by means of direct maximization of the full log-likelihood function (unconditional estimator) or with conditional maximum likelihood. In general, the unconditional estimator is preferred because it is consistent in T, where T is usually taken be fixed

---

[6] The conditional variance for the NBP model is $Var(y_i \mid \mathrm{x}_i) = \lambda_i + \frac{1}{\theta}\lambda^P$, whereas the conditional mean is the same in each NB specification.

and small. Simulation studies (Greene, 2006) demonstrate that scaling its standard errors by the deviance-based dispersion statistic produces standard error that are closer to the nominal values (Hilbe, 2007). The alternative is a random effect model such as the following:

$$\log \lambda_{it} = \boldsymbol{\beta}' \mathbf{x}_{it} + u_i$$

where u_i is an IID random effect for the *i*-th group, the same in every period, such that $\exp(u_i)$ has gamma distributions with parameters (θ_i, θ_i). Thus, $E[\exp(u_i)]$ has mean 1 and variance $1/\theta$. It is assumed that $\theta_i /(1+\theta_i)$ is distributed as Beta (a,b), which layers the random group effect onto the NB model. The random effect is added to the NB model by assuming that overdispersion parameter is randomly distributed across groups, with a log-likelihood function that is $L(\beta; y_{it}, a, b)$. This model can be estimated by a direct maximization of the full log-likelihood function.

Specification tests

First, we use a LR test to check whether a count model assuming equidispersion is an adequate specification. Under H_0, imposing restrictions in estimated parameters should make little difference to the maximized value of the likelihood function. The LR statistic is displayed in Table 3 and rejects the Poisson in favour of the Negative binomial (NB2) specification.

For choosing the more appropriate NB specification, we first compare NB2 and NB1, using the Vuong (1989) statistic: $V = \sqrt{n}\overline{m}/s_m$ where $m_i = \ln L_i(NB2) - \ln L_i(NB1)$ is the ratio of the logs of the fitted probabilities for the $i-$th observation under null hypotheses. The test statistic is the standard measure for testing whether a mean is zero and the limiting distribution of Vuong is normal (0,1). Large values favour H_0 against a broader alternative H_1, but the intermediate values are inconclusive. The statistic displayed in Table 3 falls in the inconclusive region (-1.96 to +1.96). It may be that NB1 and NB2 are not sufficiently different to enable a statistical choice on this basis. For this reason we use LR statistic for choosing NB1 vs. NBP and NB2 vs. NBP. The LR statistic (Table 3) rejects NB1 specification in favour of NBP, but does not reject the NB2 in favour NBP. At the end we prefer the NB2 model.

To test for the relevance of correlation across time, we use LR tests for the random and fixed effects against the pooled specification with robust standard error. The LR statistic (Table 3) rejects the pooled specification.

TABLE 3

Finally, a Hausman test is used to choose between fixed- and random-effects. Under the null, both specification are consistent, but the random effect is more efficient. Under the alternative, the random effect specification becomes inconsistent. Thus, if the null hypothesis is verified, the two estimators should be similar; divergence indicates rejection of the null. The statistic is:

$$H = (\hat{\beta}_{RE} - \hat{\beta}_{FE})' \left[\hat{Var}(\hat{\beta}_{RE}) - \hat{Var}(\hat{\beta}_{FE}) \right]^{-1} (\hat{\beta}_{RE} - \hat{\beta}_{FE})$$

Our test reported in Table 4 produces a χ^2 statistic of 6.97, which exceeds the critical value (0.9357). Consequently, random-effects are rejected in favour of the fixed-effects.

TABLE 4

5. Results

In Table 5 and 6 we report estimates for the different negative binomial model specifications. In particular, Table 5 reports the standard coefficients for (panel) count data models whereas Table 6 displays re-transformations of the coefficients in exponential forms to interpret them more directly as elasticities (incident-rate ratios).

TABLE 5 -6

Results are fairly consistent across specifications. Interestingly, the differences between fixed effects and the other specifications only marginally affect the most relevant policy variables, i.e. those related to financial incentives. Given the results of the Hausman test (Table 4), we focus our comment in particular on the results of the fixed effects estimates.

Individual characteristics of the GP, such as gender, specialisation and seniority are in general poorly significant. On the contrary, we obtain statistically significant effects when we consider the characteristics of the list of patients falling under the responsibility of each GP. List size as well as number of diabetic and insulin dependent patients, they all positively contribute to record an increase in the number of diabetic ACSCs among GP's patients. Our evidence indicates also that accessibility to primary and secondary care services plays an important role. For instance, rural location of a GP increases the risk of avoidable admissions among its patients, whereas a high frequency of visits to diabetic specialised wards reduces the numbers of adverse events. Quite interestingly the institutional arrangement of the practice has a limited impact, with patients followed by GPs operating in single handed practices displaying a higher frequency of diabetic ACSCs, once we control for the other covariates. Yet, this effect is only poorly significant (only at the 10% level).

An interesting insight is provided also by year dummies. The baseline case is year 2002, and all the year dummies display a negative sign, indicating that, other things equal, the number of adverse events has fallen over time. A more detailed check of the relative size of the coefficients suggests that a major improvement in the reduction of avoidable episodes was achieved between 2003-04 and between 2004-05, while no significant differences is observed between 2002 -03.

The major policy issue addressed in the analysis concerns the effectiveness of economic incentives in improving outcomes of diabetic care. More specifically, coefficients associated to financial incentives

are highly significant in all specifications and Table 6 suggests that elasticity is fairly stable and reaches values that are just below one. Overall, the empirical results on financial incentive indicate that they may be effective tools for improving outcomes in diabetes care. In particular, after controlling for a set of relevant factors such as GP individual characteristics and composition of the GP's list, an increase in the share of additional payments that a GP receives significantly reduces the expected number of avoidable hospitalisations experienced by patients included in its list.

6. Conclusions

In the present paper we have analysed the influence of economic incentives on the quality of primary care. We have focused on diabetes type 2 and we have considered financial transfers provided to GPs with the specific purpose of improving quality of treatment for such disease. Our dataset spans over 4 years (2002-05) and covers the population of diabetic patients and GPs operating in the Italian region Emilia Romagna. Thanks to a comprehensive dataset that links hospital records, pharmaceuticals prescriptions and provides details on GPs remuneration schemes, we have been able to associate the performance of each GP in terms an outcome indicator for diabetes care to the share of money he/she receives for participating in improvement activities in the same clinical area.

According to a well established epidemiological literature (Billings et al. 1993), our outcome indicator is expressed as number of avoidable hospitalisations per year for ICD-9CM codes recognized as diabetic ACSCs. As timely and good quality of treatment in a primary care setting should be able to prevent such hospitalisations, this adverse event is an indicator of poor quality of primary care.

We estimate a series of panel count data models, where the dependent variable consists of the number of diabetic ACSCs recorded in the list of each GP. We control for a set of covariates aimed at capturing differences in size and composition of the list, as well as in individual characteristics of the GP and results are fairly consistent across the different specifications.

The present work extends previous analysis to a more general setting, thanks to the availability of longitudinal data which allow to control for time-varying effects. We confirm and extend previous findings (Lippi Bruni et al. 2009) on the role of economic incentives in the Italian primary care system, and in particular in diabetes management. The panel structure of the dataset guarantees a more precise identification of the causal relationship between financial transfers and outcomes of care and provides more accurate estimations of the impact of the former on disease management. Our results stress the importance of the size and of the incentive structure induced by the different sets of programs in order to exert a positive influence on patient outcome: financial incentives are consistently negative and significant across specifications, thus implying that patients being followed by GPs that receive a higher share of their wage through the programs have a lower probability to experience an avoidable hospitalisation.

Acknowledgements

The study was financed by the Italian Ministry of Scientific Research (MIUR) and the Italian Ministry of Health. Elisa Iezzi gratefully acknowledges financial support from Fondazione Cassa dei Risparmi di Forlì. The authors thank the Regional Agency for Healthcare Services of Emilia Romagna for making the data available.

References

ADA Clinical Practice Recommendations. Hyperglycaemic crisis in patients with diabetes mellitus. Diabetes Care 2002; 25 (Suppl 1); S100-8.

Andersen LB. What determines the behaviour and performance of health professionals? International Review of Administrative Sciences 2009; 75; 79-97.

Benabou R, Tirole J. Intrinsic and extrinsic motivation. Review of Economic Studies 2003; 70 (3); 489-520.

Billings J, Zeitel L, Lukomnik J et al. Impact of socioeconomic status on hospital use in New York City. Health Affairs 1993; 12; 162-173.

Birkmeier NJ, Birkmeier JD. Strategies for improving surgical quality – should payers reward excellence or effort? The New England Journal of Medicine 2006; 354 (8); 864-70.

Booth GL, Fang J. Acute complications of diabetes, in Diabetes in Ontario: An ICES Practice Atlas, Hux J, Booth G, Slaughter P, Laupacis A. (eds), ICES, Ontario; 2003.

Campbell SM, McDonald R, Lester H. The experience of pay for performance in English family practice: a qualitative study. Annals of Family Medicine 2008; 6; 228–234.

Campbell SM, Reeves D, Kontopantelis E, Sibbald B, Roland M. Effects of Pay for Performance on the Quality of Primary Care in England. The New England Journal of Medicine 2009; 361; 368-78.

Conrad DA, Perry L. Quality-Based financial incentives in Health Care: Can we improve quality by paying for it? Annual Review of Public Health 2009; 30; 357-71.

Christianson J, Sutherland K, Leatherman S. Financial incentives, healthcare providers and quality improvements: a review of the evidence. QQUIP, The Health Foundation, London; 2009.

Dumont E, Fortin B, Jacquemet N. Shearer B. Physicians' multitasking and incentives: empirical evidence from a natural experiment. Journal of Health Economics 2008; 27; 1436-1450.

Eggleston K. Multitasking and mixed systems for provider payment. Journal of Health Economics 2005; 24 (1); 211-223.

Fleming DM. The prevalence of known diabetes in eight European countries. European Journal of Public Health 2004; 14; 10-14.

Frey B. Shirking or work morale? The impact of regulating. European Economic Review 1993; 37; 1523-32.

Greene W. LIMDEP Econometric Modeling Guide, Version 9, Plainview, NY:Econometric Software Inc. ; 2008

Greene W. Functional forms for negative binomial model for count data. Economic letters 2008; 99; 585-590.

Hilbe JM. Negative binomial regression. Cambridge University Press; 2007.

Hilbe JM , Greene WH. *Handbook of Statistics Vol.27,* Elsevier BV; 2008.

ICES (Institute for Clinical Evaluative Sciences) and Canadian Diabetes Association. Diabetes in Ontario: an ICES practice atlas; 2003.

Le Grand J. Motivation, Agency and Public Policy: of Knights and Knave, Pawns and Queen. Oxford University Press, Oxford; 2003.

Lippi Bruni M, Nobilio L, Ugolini C. Economic incentives in general practice: the impact of pay-for-participation and pay-for-compliance programs on diabetes care. Health Policy 2009; 90 (2-3); 140-148.

Mannion R, Goddard M, Bate A. Aligning incentives and motivations in health care: the case of earned autonomy. Financial Accountability & Management 2007; 23 (4); 401-20.

McDonald R, Harrison S, Checkland K. Incentives and control in primary health care: findings from English pay-for-performance case studies. Journal of Health Organization and Management 2008; 22; 48–62.

Purdy S, Griffin T, Salisbury C, Sharp D. Ambulatory care sensitive conditions: terminology and disease coding need to be more specific to aid policy makers and clinicians. Public Health 2009; 123 (2); 169-173.

RACGP (Royal Australian College of General Practitioners). Diabetes Management in General Practice (15th edition); 2009.

Scott A, Schurer S, Jensen PH, Sivey P. The effects of an incentive program on quality of care in diabetes management. Health Economics 2009; 18 (9); 1091-1108.

Siciliani L. Paying for performance and motivation crowding out. Economic Letters 2009; 103; 8-71.

Smith P. Measuring health system performance. European Journal of Health Economics 2002; 3; 145-48.

Vuong Q. Likelihood ratio tests for model selection and non-nested hypothesis. Econometrica 1989: 57;307-344

Whalley D, Gravelle H and Sibbald B. Effect of the new contract on GPs' working lives and perceptions of quality of care: a longitudinal survey. British Journal of General Practice 2008; 58; 8–14.

Winkelmann R. *Econometric analysis of count data* - 4. ed. – Springer, Berlin; 2003.

Young GJ, Meterko M, Beckman H et al. Effects of paying physicians based on their relative performance for quality. General Internal Medicine 2007; 22; 872-876.

Table 1 - **Frequency distribution of the adverse outcome (diabetic ACSCs in GP's list)**

Number of adverse outcome	Freq.	Percent	Cum.
0	8428	71.18	71.18
1	2674	22.58	93.77
2	591	4.99	98.76
3	108	0.91	99.67
4	30	0.25	99.92
5	3	0.03	99.95
6	2	0.02	99.97
7	3	0.03	99.99
9	1	0.01	100
Total	**11,840**	**100**	

Table 2 - **Descriptive Statistics, GP characteristics year 2002-2005**

Variable	Coding	Mean	Std. Dev.	Min	Max
Adverse outcome	Continuos	0.37	0.67	0	9
Postgraduate qualification	(if yes=1)	0.05	0.22		
GP gender	Male=1	0.77	0.42		
Practice rural location	(if yes=1)	0.06	0.23		
Practice type	Associated	0.68	0.46		
GP seniority	continuous (yrs)	18	8	0	47
List average age	continuous (yrs)	47	5	12	76
List size	continuous (nr. of patients)	1162	383	10	1941
List diabetics size	continuous (nr. of patients)	53	23	1	137
Insulin patients	continuous (nr. of patients)	8	5	0	29
Specialist visits	continuous (nr. of patients)	48	40	0	255
Financial incentives	continuos (% annual income)	0.67	1.31	0	14

Figure 1 - **Financial incentives in % GP annual income. LHAs, years 2002-2005**

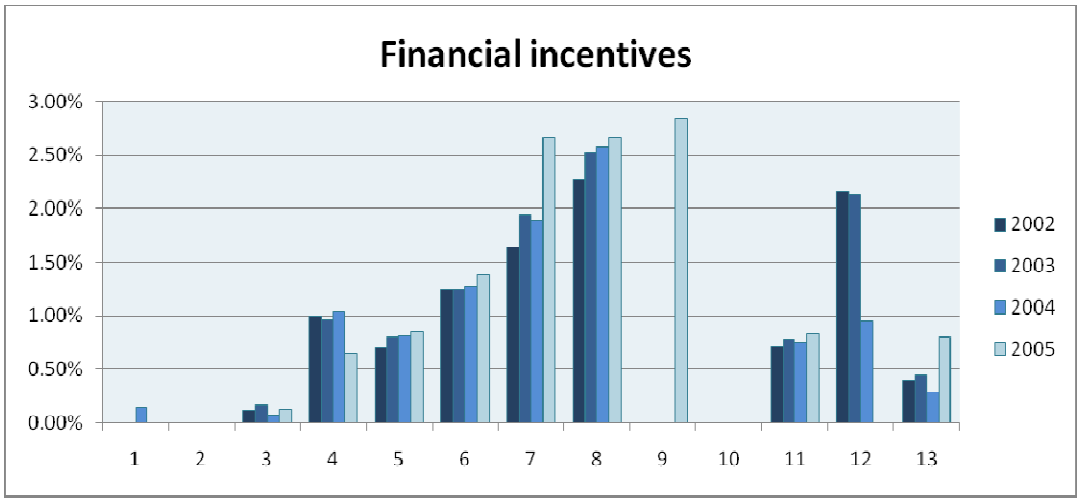

Figure 2 **- Density histogram of dependent variable with normal line. Years 2002-2005**

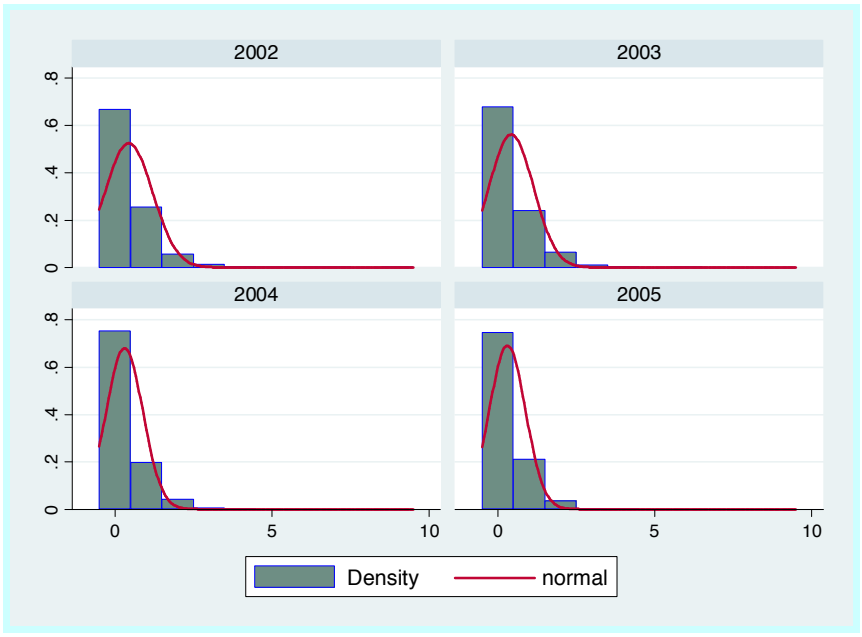

Table 3 – **Likelihood-ratio and Vuong test**

Compare	LR test	Prob>chi2	Result
POIS vs NB2	78.95	6.37E-19	Reject Poisson in favour NB2
NB1 vs NBP	3.702	0.054	Reject NB1 in favour NBP
NB2 vs NBP	0.294	0.588	Reject NBP in favour NB2
	VUONG	**|Prob|>NormSt**	**Result**
NB1 vs NB2	0.445	0.672	INCONCLUSIVE REGION

Compare	LR test	Prob>chi2	Result
FE vs POOL	60.04	9.29E-15	Reject POOLED in favour NB_FE
RE vs POOL	102.39	4.75E-22	Reject POOLED in favour NB_RE

Table 4 **- Hausman test**

	COEFFICIENTS			
	(b)	(B)	(b-B)	sqrt(diag(V_b-V_B))
	NB_RE	NB_FE	Difference	S.E.
Postgraduate qualification	0.023326	0.065659	-0.0423332	0.0394576
GP gender	0.045512	0.059878	-0.0143659	0.0203386
GP seniority	2.84E-05	0.000304	-0.0002752	0.0013285
List average age	0.003639	0.003605	0.0000338	0.0021948
List size	0.000231	0.000244	-0.0000128	0.0000352
List diabetics size	0.010398	0.009529	0.0008686	0.0007775
Insulin patients	0.046521	0.044499	0.002022	0.0028099
Specialist visits	-0.00219	-0.00213	-0.0000617	0.0002692
Practice rural location	0.253378	0.268803	-0.0154252	0.0383911
Practice type	-0.07045	-0.05657	-0.0138834	0.0188318
Financial incentives	**-0.04585**	**-0.0477**	**0.0018501**	**0.0063467**
year(2003)	-0.06515	-0.06582	0.0006724	0.0076073
year(2004)	-0.38532	-0.38612	0.0008025	0.0086071
year(2005)	-0.30232	-0.30963	0.0073116	0.008802

Test: Ho: differences in coefficients are not systematic

chi2(14) = (b-B)'[(V_b-V_B)^(-1)](b-B)=6.97
Prob>chi2 =0.9357

Table 5 – **Count data estimations and statistics**

Dependent variable: number of diabetic ACSCs in GP's list. Year 2002-2005.

Variable	NB_Pool_Rob			NB_FE			NB_RE		
Postgraduate qualification	0.0649	(0.073)		0.6569	(0.065)		0.0233	(0.079)	
GP gender	0.0575	(0.424)		0.0599	(0.040)		0.0455	(0.045)	
GP seniority	0.0002	(0.002)		0.0030	(0.002)		0.0001	(0.003)	
List average age	0.0036	(0.004)		0.0036	(0.004)		0.0036	(0.005)	
List size	0.0002	(7.30E-05)	***	0.0002	(6.69E-05)	***	0.0002	(7.80E-04)	***
List diabetics size	0.0099	(0.001)	***	0.0095	(0.001)	***	0.0104	(0.001)	***
Insulin patients	0.0449	(0.004)	***	0.0445	(0.004)	***	0.0465	(0.005)	***
Specialist visits	-0.0022	(4.28E-04)	***	-0.0021	(4.01E-04)	***	-0.0022	(4.83E-04)	***
Practice rural location	0.2671	(0.063)	***	0.2688	(0.058)	***	0.2534	(0.070)	***
Practice type	-0.0585	(0.036)		-0.0566	(0.034)	*	-0.0704	(0.039)	*
Financial incentives	**-0.0478**	**(0.013)**	***	**-0.0477**	**(0.012)**	***	**-0.0459**	**(0.014)**	***
year(2003)	-0.0652	(0.043)		-0.0658	(0.400)	*	-0.0651	(0.041)	
year(2004)	-0.3879	(0.045)	***	-0.3861	(0.043)	***	-0.3853	(0.044)	***
year(2005)	-0.3106	(0.046)	***	-0.3096	(0.044)	***	-0.3023	(0.045)	***
Constant	-2.1486	(0.197)	***	-2.1320	(0.188)	***	1.5770	(0.604)	***
ln_alpha	-1.3184		***						
ln_r							5.2608	(0.519)	
ln_s							1.5216	(0.126)	
STATISTICS									
N	11840			11840			11840		
ll	-8982			-9013			-8932		
df	15			15			17		
aic	17993			18055			17898		
bic	18104			18166			18023		

*Legend: * p<.1; ** p<.05; *** p<.01*

Table 6 - **Count data estimations**

Dependent variable: Number of diabetic ACSCs in GP's list. Year 2002-2005. Coefficients in exponentiated form.

Variable	NB_Pool_Rob		NB_FE		NB_RE	
Postgraduate qualification	1.0670		1.0679		1.0236	
GP gender	1.0592		1.0617		1.0466	
GP seniority	1.0002		1.0003		1.0000	
List average age	1.0037		1.0036		1.0036	
List size	1.0002	***	1.0002	***	1.0002	***
List diabetics size	1.0099	***	1.0096	***	1.0105	***
Insulin patients	1.0459	***	1.0455	***	1.0476	***
Specialist visits	0.9978	***	0.9979	***	0.9978	***
Practice rural location	1.3062	***	1.3084	***	1.2884	***
Practice type	0.9432		0.9450	*	0.9320	*
Financial incentives	**0.9534**	***	**0.9534**	***	**0.9552**	***
year(2003)	0.9369		0.9363	*	0.9369	
year(2004)	0.6785	***	0.6797	***	0.6802	***
year(2005)	0.7330	***	0.7337	***	0.7391	***
Constant	0.1166	***	0.1186	***	4.8403	***
ln_alpha	-1.3184	***				
ln_r					192.6390	
ln_s					4.5797	

*Legend: * p<.1; ** p<.05; *** p<.01*