Does the extension of primary care practice opening hours reduce inappropriate use of emergency services?

An IV approach.

Matteo Lippi Bruni, Irene Mammi and Cristina Ugolini

Department of Economics, University of Bologna

Correspondence to:

Cristina Ugolini, Piazza Scaravilli 2, 40126, Bologna, Italy.

Tel. +39-051-2098139 - Fax +39-051-2098040.

cristina.ugolini@unibo.it

Acknowledgements

The paper is part of the research project "*The role of economic incentives in the governance of primary care*" financed by the Health Department of Emilia Romagna, Italy, that also provided the data. We are also grateful to Andrea Donatini of the Health Department of Emilia Romagna for useful suggestions and valuable assistance throughout the research. The work reflects exclusively authors' opinions and does not involve the above mentioned institution. Other usual disclaimers apply.

4 June 2013

ABSTRACT

Inappropriate Emergency Department (ED) attendance generates serious inefficiencies in the allocation of health care resources. Several studies have focused on the role of patient characteristics, whereas the role of primary care physicians is less frequently investigated, although it could provide relevant insights in a policy perspective if avoidable ED attendance was directly linked with primary care characteristics. To this aim, we investigate the impact of regional policies aimed at increasing primary care accessibility on the appropriateness of ED visits in the Italian region Emilia Romagna using administrative data for year 2009. We focus on ED patients identified as potentially non-urgent according to the Italian hospital triage system and classified as white codes, i.e. patients with no priority in the ED who could have been dealt with safely in primary care. First, we allow for alternative specifications of a count data model for ED visits, we test for over-dispersion and then estimate a Negative Binomial model for over-dispersed count data. The dependent variable is the number of inappropriate ED attendances for patients registered with each GP. Our aim is to test whether extending GP's practice opening hours up to 12 hours/day is effective in reducing the inappropriate utilization of ED. We also account for the potential endogeneity of the extended opening variable by adopting two alternative instrumental variable strategies: a two-stage residual inclusion (2SRI) approach and a generalized method of moments (GMM)/Non linear instrumental variable (NLIV) approach. As endogeneity comes out to be a crucial issue here, in order to draw reliable policy implications from the results, we give preference to IV estimation. After controlling for patient characteristics and for a set of confounding factors, our results support the hypothesis that improving primary care organization in terms of accessibility favours a more appropriate use of ED services.

Key words: primary care, access to services, avoidable emergency department attendance, panel count data models.

JEL classification: I11, I18, C31

1. Introduction

Overcrowding in Emergency Departments (EDs) due to intense utilisation by non-urgent patients is well-documented worldwide (Sempere-Selva et al., 2001; Tsai et al. 2011; Flores-Mateo et al. 2012). Inappropriate ED attendance attracts increasing interest by policymakers as it may generate serious inefficiencies in resource allocation between community and hospital care and it may limit the capacity of the emergency services to provide timely and accurate responses to patients affected by severe and urgent conditions. Moreover, the risk of disrupting continuity of care and of impairing proper therapy for chronic diseases may cause additional adverse effects on health outcomes and reduce the quality of community care.

The phenomenon displays a strong variability across OECD countries owing partly to the different approaches used to identify patients that could have been effectively treated in alternative settings, such as primary or specialist care (Bezzina et al., 2005). A frequently held view equates appropriateness with urgency so that non-urgent patients are considered inappropriate users on clinical grounds. The percentage of ED visits triaged as non-urgent - a not life or limb threatening situation for which the recommended time frame to see a physician is greater than 2 hours - in the USA varies between 5% and 13% (McCraig et al., 2006; Gao, 2009) and shows similar rates among the uninsured, Medicaid insured and those with private insurance (Garcia et al., 2010). In Canada the percentage of non-urgent patients is estimated around 25% (Redelmeier and Fuchs, 1993; Afilalo et al., 2004; Howard et al., 2008). As for Europe, for France Lang et al. (1996) provide estimates of the share of non-urgent attendances around 30% and similar results are found also for Sweden (Hansagi et al., 1987) and for Spain (Sempere-Selva et al., 2001). In Italy a seminal analysis investigating the impact of non-urgent patient visits to ED was conducted by Bianco et al. (2003) for the Calabria Region: they estimate the rate of non urgent visits to EDs at 20%, in line with the literature based on European data.

However, non-urgent attendances may not fully coincide with inappropriate ones, since a proportion of ED visits triaged as non urgent typically require to be treated in a hospital setting, and therefore patient’s self-referral to his General Practitioner (GP) would not ensure an adequate response. This argument has led to alternative approaches which try to single out the share of non-urgent ED attendances that are inappropriately seeking response at the hospital level, by considering disease severity assessed through specific coding procedures implemented in the ED either by nurses at the moment of the admission (prospective assessment) or by physicians after the patient has been treated (retrospective assessment). Using such more restrictive criteria, the percentage of inappropriate ED visits in different countries is reported to be between 20% and 80% of non-urgent attendances (Afilalo et al., 2004).

Regardless of the country considered and of the approach used to identify inappropriate visits, the unifying indication emerging from these analysis suggests that a substantial fraction of emergency visits not followed by hospital admission should have been avoided because patient could have been treated equally - if not more- effectively in an outpatient setting. In particular, in countries like Italy where access to hospital care is largely filtered by GP, this calls into question the effectiveness of gate-keeping activities and suggests that more accurately designed policies in the area of primary care, including improved access to family physicians, may contribute to contain inappropriate utilisation of ED care.

In this paper we analyse the determinants of inappropriate ED attendance in a NHS based system using a large administrative data set for the year 2009. We investigate in particular the impact produced by policies aimed at increasing primary care accessibility – mainly in terms of extending practice opening hours - on the appropriateness of ED attendance in the Italian region Emilia-Romagna. The available regional databanks include information on the different characteristics of the GPs and on patients' use of healthcare services. Moreover, the link between the GP and patients registered in his list allows to study to what extent the practice characteristics influence ED use by rostered patients.

We first estimate by Maximum Likelihood alternative specifications of a count data model for ED visits under the hypothesis of exogeneity of all the regressors. After testing for over-dispersion, we focus on a Negative Binomial model for over-dispersed count data. The dependent variable, measured at the GP level, is the number of inappropriate ED attendances for patients registered in the GP list. Our aim is to assess whether GP organizational characteristics and differences in accessibility to primary care services are associated with different rates of inappropriate utilisation of ED services. In particular, we analyse whether coordinated opening hours which extend coverage of primary care services available for patients up to 12 hours/day are effective in reducing the inappropriate utilization of ED. At this scope we include in our empirical analysis among the regressors a dummy variable aimed to signal the extension of opening hours that represents our variable of interest for policy implications.

From a methodological point of view, we then allow for the potential endogeneity of the extended opening dummy and address the issue by adopting two alternative instrumental variable strategies to test and to control for endogeneity: a two-stage residual inclusion (2SRI) specification and a GMM approach.

Our paper contributes and extends the existing literature on several dimensions. First, we are able to disentangle inappropriate ED attendances from the remaining urgent and non-urgent ED visits by exploiting exogenous information provided by the Italian hospital triage system. Such system classifies attendances not followed by hospitalisation according to a coding criterion based on four categories: red, yellow, green and white codes. These labels correspond to conditions of decreasing severity/urgency and patients classified as *white codes* are treated as lowest priority cases. It must be remarked that *white codes* are identified on institutional grounds as episodes of inappropriate utilisation of ED services and, for this reason, patients falling under this category are charged a service fee because they should have sought care through a different channel of care within the NHS. They correspond not only to non-urgent conditions, for which also a *green code* could have been attributed, but they are also characterised by low severity. In this way, our analysis gets rid of the limitations originated by the adoption of often subjective and questionable definitions of inappropriate attendances to ED.

Second, and most relevant, the institutional setting provided by the Italian NHS allows us to investigate the influence of organisational features of primary care on the utilisation of ED services. While several studies have analysed the influence of patient characteristics on

ED attendance rates (Lang et al., 1996; Lee et al., 2000), the role of the institutional context, and in particular the link between the organisation of primary care and the use of emergency services has been more rarely addressed. In this perspective, our work contributes to the strand of literature that studies the role of primary care policies in ensuring a more appropriate utilisation of hospital services, a strategy which is expected to contribute to cost containment and/or improvement in health outcomes (e.g. Dusheiko et al. 2011; Scott et al. 2011; Wilson, 2013; Emmert et al. 2012; Eijkenaar, 2013).

Third, most studies that analyse the association between primary care characteristics and the use of emergency services (Lowe et al. 2005 for the US; Howard et al., 2008 and McCusker et al., 2010, for Canada) make use of survey data which inevitably suffer the problem of covering a limited share of total number of ED users and misses potential spillover effects in utilisation rates across groups of patients. On the contrary, the information for our study is drawn from administrative data sets which cover the entire regional population (around 4 million inhabitants) that benefits of universal NHS coverage, therefore we can control for overall ED utilisation.

Finally, from a methodological standpoint, thanks to the identification of relevant instruments that allow to control for endogeneity in the variable of interest, we are able to control for the potential biases due to the self-selection into the incentive programs by GPs that extend practice opening hours of the practice.

In a policy perspective, our work aims at filling a gap in our understanding of how to design effective policies based on coordination between primary and secondary care, in a moment where several ongoing reforms attempt to reduce inappropriate use of hospital care in general, and emergency services in particular, by improving effectiveness and extending availability of community care services. Indeed, assessing to what extent primary care initiatives which encourage cooperation among GPs and extend daily coverage of primary care practices are effective in limiting the inappropriate utilisation of ED services is crucial for the overall efficiency of the system.

Our main findings indicate that GP organizational characteristics are strongly correlated with the inappropriate ED attendance. The extended opening of the ambulatories and the consequent higher availability of primary care services are effective in answering the demand for non-urgent care that, when not addressed, flows inappropriately into the EDs.

2. Background Literature

An early stream of research has examined the influence of patient characteristics on the utilisation of ED services for conditions that should have been effectively treated in a different setting. Poverty, minority status and lack of a personal physician are shown to increase the probability for patients to attend ED for non-urgent care (Lang et al., 1996), even if it has been also documented that families of low socio-economic status and elderly patients were more likely to contact their GP first (Klijakovic et al., 1981; Sempere-Selva et al. 2001). Other studies have shown also that inappropriate ED users more often belong to middle and upper classes (Shah et al., 1996; Lee et al., 2000) and that in the US the use of EDs is influenced by the insurance status of the patient. The presence of cost sharing significantly reduced ED admissions, even if the absolute size of the copayment seems to produce little effect on utilisation (Roberts and Mays, 1998).

Several studies have also paid specific attention to patients' motive to skip GPs and self-refer to ED. Primary reasons include the perception of a need to receive immediate care and relief through hospital care and difficulty in accessing primary care practices. Lee et al. (2000) use data from a telephone survey on a sample of 2410 patients randomly selected in four EDs in Hong Kong and find that the primary reason for high ED attendance was the closure of the GP practice on public holidays or at night, reflecting scarce accessibility to GPs services and suggesting the need for GPs to set up a network system able to provide out of hours services. More recently, a study conducted through a postal questionnaire sent to 339 Dutch patients reports that shorter waiting time and travelling distance compared to GP practices are the principal motivations for ED use (Moll van Charante et al. 2008). Overall, these results confirm previous indication from Puig-Junoy et al. (1998) who estimate a nested multinomial logit model on a cross-sectional Spanish National Health Survey and find that self-referred emergency visits are substitute for general practitioner visits, with a demand highly elastic with respect to the waiting time needed to see a community physician. Instead of accessibility of primary care, Carlsen et al. (2007) focus on the quality measured by patient satisfaction of primary care services and use data from an extensive survey conducted in 1998 in Norway to construct an indicator of user satisfaction with GP services. They detect a robust negative relationship between patient satisfaction of primary care services and the probability of hospital admissions, both for total hospital access and for emergency care. Similar results are highlighted by Sempere-Selva et al. (2001), Gutman et al. (2003), McCusker et al. (2010) stressing that primary reasons for self-referral to EDs are a general perception of unmet healthcare needs in primary care, the frustration with scheduling appointments and long waiting times, the perception of long waits before gaining access to other secondary services and greater trust in hospital care. To the best of our knowledge, very few studies have addressed the Italian case. An interesting exception is provided by Lega and Mengoni (2008) who analyse data from a structured questionnaire administered to 527 patients of the province of Macerata (Marche region), finding that what most appeals to ED users is the possibility to rapidly receive a specialist consultation with no waste of time, especially in presence of lack of trust in GPs and a general dissatisfaction in their appointment hours.

As regards the influence of physician organisation, most of the empirical research is based on US data (Weinberger, Oddone and Henderson, 1996; Lowe et al., 2005) and its results are not easily applicable to national health systems. For example, using a cohort study for one year of 57.850 Medicaid patients assigned to 353 primary care practices, Lowe et al. (2005) calculate that overall ED use would decrease by 13% if patients in all practices used the ED at a rate observed for practices with 12 or more evening hours a week, or would decrease by 5% if all practices had office hours also in the week end. Besides, studies of the experiences of Kaiser Permanente and the Veterans Administration suggest that improving the integration of primary care with hospitals may reduce utilization of ED services (Feachem et al, 2002; Armstrong et al, 2006). For UK the influence of practice organisations on patients ED use seemed marginal, as patient's factors - such as recent migrants or socioeconomic conditions of unskilled population- account for most of the variation in ED use (Saxena et al, 2006; Calderòn-Larranaga et al, 2011). Harris, Patel and

Bowen (2011) conducted an observational, cross-sectional ecological study of 68 GPs in one PCG in north London for which they considered several independent variables explaining GP access characteristics, population characteristics and health status, aggregated to the level of GP. They tested the hypothesis that variation in ED attendance is explained by the variation in the degree of access to GP practices - expressed in term of total number of opening hours per week - but none of the indicators that proxy access to GP services turns out significant, thus suggesting that in this case inappropriate ED use is probably driven by underlying patients' characteristics, including social deprivation, whereas in Lee et al. (2000) a substantial proportion of the higher socio-economic group utilized ED more inappropriately than the socially disadvantaged. Focusing on the provision of out-of-hours primary care services, Thompson et al. (2010) found for a UK district Hospital in West Suffolk examined in the period 1999-2006 little change in ED attendances attributed to the introduction of out-of-hours primary medical care in 2004, as the number of patients attending the ED with non-traumatic conditions out-of-hours rose after the changes were implemented. More successful appears to have been the Dutch reform that reorganised primary care towards large-scale GP cooperatives able to provide out-of-hours assistance: examining a population of 62,000 people during two four-month periods in a five-year interval, Moll van Charante et al. (2007) showed that the GP cooperatives are able to deal with the large majority of out-of-hours requests, handling 88% of all out-of-hours contacts in the second period of analysis, whereas self-referrals to ED represent a small group of patients that attend hospitals mainly for appropriate reasons.

3. Primary care in the Italian National Health Service

The Italian National Health Service (NHS) was established in 1978 to replace a Bismarckian social insurance health care system with a Beveridgian model based on the principles of universalism, comprehensiveness and equity. Health expenditure is relatively low according to international standards: public health expenditure amounts to 7.2% of GDP in 2011, whereas total health expenditure reaches 9.5% of GDP, slightly above the average of 8.9% in OECD countries. Despite relatively low public health spending, in the nineties the high level of public debt lead to the introduction of a series of reforms aimed at efficiency enhancement and cost containment. One of the most intense changes in the Italian NHS is its progressive regionalisation, leading to the introduction of fiscal federalism. This evolution gave the twenty Italian regions political, administrative and financial responsibility regarding the organisation and delivery of publicly financed health care (Fiorentini, Lippi Bruni, Ugolini, 2008).

Family physicians, who are independent contractors with the NHS, deliver primary care to registered citizens and play a gatekeeping role to more complex treatments such as specialist and hospital care. Health Districts (HDs) organise outpatient specialist services, residential and primary care and they are aggregated into Local Healthcare Authorities (LHAs) which have direct responsibilities over inpatient care and whose managers are appointed by the Regional Governments according to a top-down model of governance. Primary care services are free of charge at the point of need, registration with a family physician is compulsory and citizens can freely choose the GP to be enrolled with. Individuals may easily change their GP but, in practice, there is a very low turnover rate as long as most of list variations are induced by change in residence rather than by an unsatisfactory patient-physician relationship. Each physician has a maximum list size of 1500 registered patients, although exceptions were allowed for GPs who exceeded such limit when it was first introduced.

Capitation is the most important source of income for family physicians, nationally contracted every three years together with other relevant features of GPs' remuneration. The payment scheme is organised in three parts, with a variable and an additional component topping up capitation, The variable part is regulated by the national contract as a fee-for-service compensation for specific treatments, including minor surgery, preventive care and post-surgery follow-up. The additional part is intended as a reward for providing high quality, appropriate care or to adhere to cost containment policies and it can be designed either as a pay-for-performance mechanism or as low powered schemes (Fiorentini et al. 2011). In this way policymakers intend to favour the alignment of physicians' incentives to the general goals outlined at the system level. Following the 1992 and 1999 NHS reforms, regions have attained a large autonomy in designing this additional part of GPs remuneration, an autonomy usually devolved at the local level, which has led to a substantial heterogeneity of the incentive schemes between and within regions (Lo Scalzo et al., 2009). The areas covered by specific incentive programs differ across regions and LHAs and may involve, for instance, containment of referrals to hospital and specialist care, increase in prescription rates of generic drugs but also assumption of responsibility for chronic patients (Lippi Bruni et al. 2009).

Innovations in primary care have been extended also to the possibility of establishing medical associations or partnerships among GPs for improving quality and comprehensiveness of care. Such initiatives have taken different forms in the different regions (Fattore et al. 2009; Shaw and Meads, 2012), but they are all motivated by the view that single handed practices are considered to be relatively less effective compared to coordinated team activity in providing high quality care because of limited possibilities to ensure continuity of care, in particular to chronic patients (Fantini et al. 2012), and to acquire diagnostic equipment.

As regards the inappropriate utilization of emergency department, patients who attend an ED must pay a fixed fee if their condition is evaluated as deferrable to primary care, although low income individuals and patients affected by chronic conditions may be exempted from the payment. Over the years such fee was subject to several changes according also to specific regional policies, which led in some cases to a temporary suspension of its implementation. Nevertheless in the last decade, with the increasing pressure of demand for ED services and the scarcity of financial resources, it has consistently been applied throughout the country. Despite all these interventions, a commonly held view suggests that emergency departments are still intensively used as a substitute for GP care. In the health policy debate it is frequently argued that if practice opening hours covered a longer time span, this would reduce the demand for (inappropriate) emergency admissions. With these policy guidelines, the organization of primary care has been changing also in Italy and there are also some experiments of new models of extending opening hours and of out-of-hours care such as GPs groups and cooperatives, primary care centres integrated into hospitals EDs or hospital minor injury or illness ambulatories designed to provide care to non-urgent patients. The evidence of these interventions is still scarce.

In Emilia-Romagna major emphasis has been placed on improving the organisation of primary care and increasing cooperation between primary and hospital care, assuming that this could reduce inappropriate referrals to emergency care. For this purpose, since the years 2000 the Regional Health Department has been using different schemes of financial incentives to encourage GPs to adhere to the targets of health policies. In order to obtain efficiency and efficacy gains in the relationship between community and secondary care, a two-phase policy has been promoted. The first step, started in 2000, was aimed to formally organise the GPs activity in networks or groups towards a progressive increase of coordinated associations in primary care. Association's types made available to GPs range from coordination of practice opening hours and substitutions in case of illness for networks to a more intense cooperation in the case of groups, which are expected also to share the same ambulatory and staff (Fattore et al. 2009). Furthermore, incentives are paid to GPs for the recruitment of administrative assistants and nurses.

Once that a high number of GPs associations was reached, the second phase, launched at different times across LHAs from the second half of the 2000s, is intended to incentivise GPs with additional remuneration for coordinating the opening hours of the practices included in networks or groups in order to ensure an overall daily coverage exceeding 6 hours and up to a maximum of 12 hours. This policy does not force GPs to organize an out-of-hours service but is aimed to ensure that patients - especially people with chronic conditions and those needing routine or preventive care - can access services conveniently during the day to fit around work, school and other caring responsibilities, without being forced to attend EDs inappropriately. The first extensions of opening hours have interested mostly urban practices and groups with an higher number of partners, pointing out that the location of the practice and the number of GPs involved in the turnover could be crucial for the initial successful implementation of this new organizational model. Assessing the impact of these organisational features on hospital ED attendances is the main goal of the present work, since a positive evaluation could be useful to the regional policy makers in order to proceed encouraging GPs to extend their opening hours also for smaller groups and those located in rural areas.

4. Data and estimation issues

4.1 The data

The initial study population consists of all regional patients registered in the lists of GPs active in the Region in 2009. For this group of individuals we observe all visits to ED not followed by hospitalisation that took place in any regional ED during the year 2009. The data on ED flows are routinely collected for administrative reasons and report the triage code attributed to each episode. Each patient is characterised by a unique encrypted identification number which is consistent across all the datasets provided by the Regional Department of Healthcare Services. Such patient identification code was used to link the medical-use records for ED services to his/her GP.

As the main aim of the work is to investigate whether the organisational characteristics of primary care practices, and in particular arrangements for extending their opening hours, are effective restraints for the inappropriate use of ED, we adopt the GP as unit of the analysis and we aggregate data at the GP level. The count dependent variable is the number of white code attendances to ED (not followed by hospitalization) by patients rostered in the list of each GP operating in the region over the year 2009.

The total number of white codes in 2009 is 236335, which represent about 18% of total visits to EDs not followed by hospitalisation (over 1.3 million). In our analysis we consider only the GPs who had at least 200 registered patients (3113 GPs in 2009), as the involvement in primary care can be considered as a part time activity below such threshold. The total number of white codes for such sample of GPs amounts to 198738, about the 18% of the total 1112272 ED visits.

Each GP can organize himself/herself a in single-handed practice but different collaborative forms among GPs are also possible. Even when GPs participate in a collaborative arrangement, patients are always registered with a specific GP and not with the network or group: it is therefore still possible to assign to each GP only the ED visits of patients actually rostered in her/his list.

The choice of the organisational form is voluntary, though financial incentives are provided for enhancing the creation and maintenance of networks and groups. Moreover, additional regulated bonuses are paid to GPs if they ensure coverage of their patients' need beyond 6 hours per day. We focus here in particular on the agreements that ensure practice opening between 10 up to 12 hours per day, because these organisational arrangements are explicitly designed (and financially incentivised), among other things, also for containing access to ED. Ensuring similar increases in the accessibility to primary care services through the extension and coordination of the practice's opening hours is possible essentially only for GPs that are organised in network and group, the two most intense forms of cooperation. To the extent that physicians other than in networks or groups do not have the possibility to ensure such long-lasting daily service, we include in our estimating sample only the GPs who belonged to one of these organisational arrangements in 2009.

The total number of white codes ED visits considered in the sample of interest, that includes 2370 associated GPs, is 157005, which represents again about the 18% of the total 887445 ED episodes not followed by hospitalization, similarly to what happens if one considers the entire population of GPs.

We estimate an exponential conditional mean count data model allowing for alternative specifications in order to check the robustness of our findings. Given the count variable of interest y, that measures the number of white code ED visits by the patients with the same GP, we model the conditional mean of y according to an exponential form $E(y|\mathbf{x})=\mu=\exp(\mathbf{x}'\boldsymbol{\beta})$ and allow for different specification of the conditional variance $Var(y|\mathbf{x})$.

The vector **x** includes in the first place GP's characteristics such as gender, age, seniority and rural location of the practice. We allow for non-linear effects of age and seniority by including also quadratic terms among the regressors. Additional important controls are the characteristics of the list. GP's exposure to the risk of experiencing white code visits among her/his patients is mainly captured by list size included in logarithmic form. With respect to the demographic composition of the list, we consider the share of male patients and the age distribution of rostered patients. As long as cultural differences and health literacy may influence patients' decision to use of emergency vis-à-vis primary care services, we also include the share of non-native patients in the list.

In regard to organizational characteristics, for each GP we know whether her/his network or group practice organizes its activity over a daily schedule that exceeds either 6 or 9 hours respectively. As we focus on the extension over the 9 hours, we are therefore include a dummy variable for the extended opening of the practice ambulatory for more than 9 hours up to 12 hours. Finally we include a dummy for the presence of nursing staff in the practice.

Local fixed effects are controlled for by geographical dummies for the three macro-areas “Vast Areas” in which the Region is divided and that are in charge of coordinating healthcare policies within the regional borders.

TABLE 1

In Table 1 we report the main descriptive statistics for the variables included in the model relative to the whole sample of 2370 GPs, to the sub-sample of 1651 GPs that do not have an extension of the opening hours and to the sub-sample of 719 GPs (about the 30% of the total) that belong to practices that organize the activities over the 9 hours per day.

With respect to the whole sample, it is worth noting that about the 70% are male and that the average list size is well above 1200 patients, though this variable displays a remarkable variability over the sample, as captured by the high standard deviation. The GPs list presents on average a relevant proportion (7.5%) of non-native patients which is in line with the regional average reported by census data. Yet, there are GPs for whom such percentage is much higher, as the variable displays a high variability. It is therefore important to control for the share of non-native patients registered with each GP since this subgroup of individuals often presents different utilisation patterns of healthcare services for cultural and socio-economic reasons. With respect to the organizational characteristics of interest, around the 12.4% of the GPs have nursing staff employed in the practice. Finally, the 21.5% of the practice premises is located in rural municipalities, namely in totally or prevalently mountainous districts.

The extensions of opening hours up to 10-12 hours involves 30.3% of GPs working in group of network: it is therefore worth looking at disaggregated statistics for the two sub-groups of doctors. We note that the statistics relative to the individual characteristics (gender, age, seniority) as well as those about the characteristics of the list are mirror-like for the two sub-samples. Though, not surprisingly we have significant differences between the two groups with respect to the location of the practice premises and to the presence of nursing staff. The ambulatories of the practices that have the extension are prevalently located in urban municipalities (more than 86% of them vs. only the 75% for non-extension GPs): it is pretty intuitive that doctors are more likely to join a group or a network and thus organize the activities of the practice over the 9 hours if they work in urbanized areas where the number of GPs is higher and the facilities are closer one to the other; in general, the connections among GPs are easier than in rural municipalities. Among group and network practices, the 20% of the GPs with the opening extension avails itself of nursing staff, while only the 9% of non-extension GPs does: in order to guarantee an extended opening, it seems reasonable that the practice hires nursing personnel that collaborates with the physicians.

As regards the dependent variable, i.e. the number of white codes ED visits per GP, in Figure 1 and 2 we graph the distribution of the white-code ED visits per GP respectively for the non-extension and the extension sub-groups. The problem of zero inflated data can be ruled out as the minimum number of white code visits to ED for a GP amounts to 1, with an average of 66 visits and a standard deviation of about 53: this implies that the variance of the outcome *y* is about 42 times larger than the average, suggesting that the presence of severe over-dispersion of the dependent variable.

In Figure 3 we plot the observed proportions of the variable y versus Poisson and Negative Binomial probabilities fitted taking into account the sample average and the sample variance of the white code visits. Graphical inspection provides further evidence of over-dispersion which requires modelling the outcome with a Negative Binomial model.

3.2 Estimation of the count model

3.2.1 Specification of the model

Given the non-negative and discrete nature of the variable of interest, i.e. the number of white codes per GP, we face nonlinearities that cannot be dealt with by means of standard linear methods. We therefore fit a count data model where the conditional mean for ED visits is an exponential function of a dummy for the extending opening and a set of additional regressors:

$$E(y|\mathbf{x})=\mu=\exp(\mathbf{x}'\boldsymbol{\beta})=\exp(\beta_0+\beta_1 \textit{extended\_opening} +\beta_2 x_2+\beta_3 x_3+ \ldots + \beta_p x_p). \qquad (1)$$

The vector of regressors x_1, x_2,..,x_p includes the variables previously discussed and presented in Table 1. In the case that the *extended_opening* variable and the set of control variables are exogenous, the model in equation (1) can be consistently and efficiently estimated by Maximum Likelihood.

Assuming the exogeneity of all the regressors, we therefore focus on likelihood-based estimation strategies first. We start by estimating by ML a Poisson model for (1) which assumes that the count variable *y* is independently Poisson distributed (that is $\Pr[Y=y]=\frac{\mu^y e^{-\mu}}{y!}$) and relies on the hypothesis of equality of mean and variance, namely the equi-dispersion property of the Poisson distribution $E(y|\mathbf{x})=Var(y|\mathbf{x})=\mu$.

The Poisson model however generally imposes too restrictive assumptions, as most health economics data show over-dispersion (namely conditional variance that exceeds the conditional mean) and often an excess of zeroes in the distributions. If the restrictions are not met, the Poisson MLE still provides consistent estimates of the coefficients but not of the standard errors. In order to get robust standard errors and to deal someway with over-dispersion, we actually estimate the model by Poisson Pseudo-MLE that is consistent under weaker assumptions and does not requires the data to have a Poisson distribution [see Cameron and Trivedi (2005)].

If over-dispersion is detected in the data, the estimation of a Negative Binomial is a preferable strategy to tackle over-dispersion and to fit the model in (1). We therefore test for the possible over-dispersion of the count. Following the auxiliary regression-based approach suggested by Cameron and Trivedi (1990, 2005, 2010) we test the null hypothesis of equi-dispersion $Var(y|\mathbf{x})=\mu$ against two alternative hypotheses that $Var(y|\mathbf{x})=E(y|\mathbf{x})+\alpha^2 E(y|\mathbf{x})$, which is the variance equation that is assumed by the NB2 model and $Var(y|\mathbf{x})=E(y|\mathbf{x})+\alpha E(y|\mathbf{x})$. We then estimate a NB2 model where we allow for a constant quadratic form of the conditional variance that accounts for over-dispersion. Namely, assuming that $Var(y|\mu, \alpha)= \mu(1+ \alpha \mu)$, where α is the variance parameter of a Gamma distribution, we fit a Negative Binomial model for (1) by ML and also test the hypothesis that α is not significantly different from zero, in which case the NB distribution would reduce to the Poisson model.

The results for the Poisson and NB2 MLEs are reported in Table 2.

3.3 Instrumental variable estimation

As the decision of extending the opening hours of the practice is a voluntary choice of the members of groups and networks, one should seriously consider the possibility that the associated indicator is endogenous. This is a potential problem as the estimates given by MLEs are consistent only if the regressors are exogenous. As a consequence, the policy implications over the effects of extending opening hours of primary care practice on the (appropriate) use of emergency services are vulnerable to potential endogeneity of the associated control variable.

It is therefore crucial to investigate the role played by unobserved characteristics and, in particular, the potential correlation between the choice of extending opening hours and unmeasured latent factors that affect the outcome variable. A way to tackle this problem is to test for the presence of endogeneity and to control for it in the estimates by exploiting proper instrumental variables (IV). The validity of the instruments relies on the fact that they are correlated with the potentially endogenous variable but, at the same time, they are exogenously determined with respect to dependent variable, which in our case is the number of visits to EDs in 2009 coded as "white".

Count data specifications present peculiar challenges in the implementation of IV methods. Terza, Bradford and Dismuke (2008) argue that addressing endogeneity by applying conventional linear IV methods, that ignore the non-linear specification of the relationship of the count variable y with an endogenous regressor x_e, and a set of additional confounders, can lead to biased estimates of the causal effects of interest.

Mullahy (1997), Windmejier and Santos-Silva (1997), Terza (1998), Terza et al. (2008) and Wooldridge (1997, 2002), among the others, have suggested alternative estimators to tackle endogeneity in count/exponential regression models.

In line with the approaches proposed by the contributions mentioned above, we employ two IV strategies to estimate a model in which a (binary) regressor is allowed to be endogenous to the outcome variable. In order to tackle endogeneity, we exploit first a 2-stage residual inclusion (2SRI) strategy that involves the estimation of a reduced form model in the first-stage and a count regression in the second. Second, we consider a moment-based GMM/NLIV (Generalized method of moments/Non-linear instrumental variable) estimation procedure. Both estimators yield consistent estimates of the coefficients in presence of endogeneity, allow unobservable confounders to be correlated with the included regressors and require only weak structural assumptions on the data generating process.

3.3.1 Exclusion restrictions

Identifying potential candidates for valid instruments is often a challenging task, and even more so in the present context where information has to be linked consistently across different data banks, with the GP exerting a pivotal role.

For the purpose of identifying valid instruments, we exploit the evolution in the organisational models adopted in the Italian NHS for primary care, whose evolvement has emphasised the importance of strengthening connections and collaborative agreements among GPs, in the belief that increased coordination would have improved efficiency and overall quality of outpatient care. In Emilia Romagna this has produced a sustained trend of transitions from single handed to coordinated practices initiated in the mid-nineties.

In more recent years, the creation of a formal agreement such as network or group in the policymaker's view has potentially represented also a first step towards more intensive forms of cooperation including the coordination of opening hours of the practice to ensure coverage of primary care services for up to 10-12 hours per day.

Since the latter agreement is complex to manage and extremely demanding in terms of joint professional effort, it is usually implemented only after that the GPs belonging to a network or group have successfully run their coordinated activity for a certain number of year. Given this premise, we argue that an important determinant for the fact that GPs have opted for coordinating the extension of opening hours by 2009 is that the network or group they belong have already been operating for some years. Therefore, we use the information on the number of years since the GPs joined a network or group for the first time, as an instrument for the presence of extension in 2009.

In addition to it, the coordination of practice opening hours extended beyond the basic contractual standard is more likely to occur in urban settings with higher density of practices. Agreements of this kind can be seen as providing substantial benefits to registered physicians only if the location of collaborating practices are sufficiently close to each other, and most effective when GPs share the same facilities. Consequently, one may figure out that density of GPs influence the probability of extending opening hours but at the same time is uncorrelated with the (inappropriate) use of emergency services. We construct a concentration index expressed as number of GPs active in the districts per 100,000 residents.

To summarise, we consider two possible instruments for the dummy indicator of interest represented by the coordinated extension of the opening of the practice between 10-12 hours per day: the first one is a continuous variable measured as the number of years during which the GP has been operating in network or group; the second one is represented by the density of GPs in the district where each physician operate. The average number of GPs per 100,000 inhabitants amounts to 73.6 while the average number of years in network or groups to 6.4. The correlation among the two variables is 3%, an indication that excludes the risk of collinearity across instruments.

3.3.2 Two-stage residual inclusion

From the methodological point of view, we cannot rule out the possibility that the dummy variable for extended opening is correlated with unobservable confounders that also affect the dependent variable. We assume that such unobserved heterogeneity is uncorrelated with the other regressors in the model, that are still treated as exogenous variables.

The 2SRI suggested by Terza, Basu and Rathouz (2008) is a version suitable for non-linear models of the Hausman (1978) endogeneity test and draws from the 2SRI strategies suggested among others by Rivers and Vuong (1988), Smith and Blundell (1986) for specific nonlinear models; it is a version of the standard control function approach and was first developed for count data models by Wooldridge (1997, 2002).

When dealing with a potentially endogenous binary variable, we can test for its exogeneity through an asymptotically efficient Wald test on the coefficients of the first-stage residuals in the second-stage outcome model (Smith and Blundell (1986)). The test on the first stage residuals actually allows to determine the extent to which unobserved latent factors affect the outcome variable (Pizer (2009)). In brief, the 2SRI procedure allows to test and to correct for endogeneity by estimating a first-stage reduced form for the endogenous variable, obtaining consistent estimates of the residuals, including them in the second-stage outcome model and testing their significance.

Following Cameron and Trivedi (2010) and Terza et al. (2008), we now discuss the 2SRI strategy in detail.

Denoting y the number of white codes, d the binary variable for extended opening, $\mathbf{x}$ the set of additional confounders presented above and u an error term, we specify the conditional exponential mean model for y as follows:

$$\mu = E(y|\mathbf{x}, d, u) = \exp(\gamma d + \mathbf{x}'\boldsymbol{\beta} + u) \tag{2}$$

where u accounts for unobserved heterogeneity due to unobservable (omitted) latent variables and is assumed to be correlated with d, this correlation being the source of endogenity and induced over-dispersion in the model, but uncorrelated with the variables in $\mathbf{x}$. It is obvious that, in order to solve such endogeneity problem and for identification of the model in (2), one or more instruments for d need to be identified and excluded from the model. We then specify a reduced form equation for the extension variable d as follows:

$$d = \mathrm{r}(\mathbf{x}'\boldsymbol{\xi} + \mathbf{z}'\boldsymbol{\kappa}) + \varepsilon \tag{3}$$

where r() is a known potentially nonlinear function, $\mathbf{z}$ is a vector of exogenous (instrumental) variables that, in our framework, includes the two variables discussed above and ε is an unmeasured latent factor that affect both d and y and that is, once we control for observable variables, the only source of correlation between the extension dummy and the count outcome.

In order to better specify such correlation, we assume that the error terms u and ε are linked according to the following equation:

$$u = \rho\varepsilon + \upsilon \tag{4}$$

where υ is a stochastic term independent of ε and such that $E(\exp(\upsilon))$ is constant.

We can thus rewrite the model in (2) as:

$$\mu = E(y|\mathbf{x}, d, \varepsilon) = \exp(\gamma d + \mathbf{x}'\boldsymbol{\beta} + \rho\varepsilon). \tag{5}$$

If ε were observable, we would include it among the regressors, but, as it is not, we need to replace it by a consistent estimate. To do so, we adopt the two-step procedure that involve in the first stage the estimation of the reduced form in (3). We estimate the model in (3) in two alternative ways. We first estimate a linear probability model and obtain the first stage residuals: a LPM in the first-stage allows to compute the F-test for the joint relevance of the instrumental variables and it is the safest approach to estimate a LDV model in the first-stage when the distribution of the dependent variable is unknown (Angrist (2000)). However, as the reduced-form for the extension is likely to be nonlinear and a LPM can give many out-of- sample predicted probabilities, we also fit a Probit model for the model in (3) thus estimating the relation:

$$\Pr(\textit{extension}=1) = \Phi\ (\textit{concentration index; years of association; } \mathbf{x}'\boldsymbol{\beta}) + u \tag{6}$$

where the link function Φ is the probit one and u is the unobserved heterogeneity component. Similarly to the LPM case, a χ^2 test for the joint relevance of the IVs can be performed. After predicting the probability of an extended opening through both LPM and Probit regressions as a function of the instrumental variables and the exogenous set of regressors, we compute the residuals by subtracting the predicted probabilities from the extended opening dummy and we thus obtain consistent estimates for ε in equation (5).

In the second stage, we fit the model in (5) and model the count variable y as a function of the endogenous extension dummy, the set of observable exogenous regressors and the residuals from the first stage regression. We estimate such model by negative binomial MLE. The residuals from the first stage substitute for the unobserved confounders correlated with the extended opening and the count variable.

The inclusion of the residuals from the reduced form in the second stage regression has a twofold advantage: on the one hand, it allows to control for endogeneity of the extended opening hours caused by the correlation with unobserved factors; on the other hand, it provides a simple Wald exogeneity test for the potentially endogenous variable in a nonlinear framework. If ρ is statistically different from zero, the extension variable should be better considered endogenous as we have evidence of the presence of underlying unobserved factors that affect both the opening variables and the outcome variable.

To account for the fact that we include among the regressors in the second-stage residuals from a first-stage estimation, standard errors for the outcome model coefficients need to be bootstrapped.

Tables 2 and 3 presents respectively first and second stage estimates for the 2SRI method.

3.3.3 GMM/NLIV estimation

Endogeneity issues in econometrics are often dealt with through GMM estimation methods since Hansen [1982]. The GMM approach is particularly appealing also in a nonlinear context, such as in the exponential conditional mean regression model for count data, as, through the exploitation of proper moment conditions, it allows to get consistent and efficient estimates of the coefficients also when the regressors are correlated with unobserved heterogeneity and when heteroskedasticity is present. In particular, the GMM estimator does neither require assumptions about equi-dispersion in a Poisson model or overdispersion in NB nor assumptions about the reduced form for the endogenous variables. Furthermore, GMM can be applied also in count models with binary endogenous variables and still gives consistent estimates.

Windmeijer and Santos-Silva (1997) and Mullahy (1997) develop alternative GMM/NLIV estimators for count models with (dummy) endogenous variables and discuss two alternative formulations of the errors in the model, an additive or a multiplicative specification, that hence imply different orthogonality conditions between the excluded instruments and the error term.

When the exponential conditional mean model is specified with zero-mean additive errors ξ we have:

$$y=\exp(\mathbf{x}'\boldsymbol{\beta})+\xi \tag{7}$$

from which

$$E(y|\mathbf{x})=\mu=\exp(\mathbf{x}'\boldsymbol{\beta}) \tag{8}$$

such that the residuals are u=$\boldsymbol{y}$-$\exp(\mathbf{x}'\boldsymbol{\beta})$. When one or more regressors are endogenous, we have that $E[\mathbf{u}|\mathbf{x}] \neq 0$ and that ML estimates of the model are inconsistent. If valid and relevant instruments $\mathbf{z}$ are available such that $E[\mathbf{u}|\mathbf{z}]=0$ we can exploit such orthogonality conditions and estimate the model by GMM/NLIV.

Alternatively, if the exponential conditional mean model is specified with zero-mean multiplicative errors we have:

$$y=\exp(\mathbf{x}'\boldsymbol{\beta}+\tau)= \exp(\mathbf{x}'\boldsymbol{\beta})v \tag{9}$$

from which, in case of endogeneity, we have:

$$E(\mathbf{v}|\mathbf{x}) \neq 1. \tag{10}$$

Mullahy (1997) shows that, if instruments are available such that $E[\mathbf{v}|\mathbf{z}]=1$, such moment conditions can be exploited to estimate the coefficients of the model consistently by GMM.

Windmeijer and Santos-Silva (1997) warn about the fact that, under endogeneity, the same set $\mathbf{z}$ of IVs in general is not orthogonal to both the formulations of the residuals and argue that which specification of the error has to be preferred is an empirical matter. We estimate the count exponential model by GMM under both the formulations of the orthogonality conditions for the same IVs. As we have two instrumental variables at our disposal and assume that only one regressor is endogenous, we can estimate an over-identified model and test the validity of the orthogonality conditions through the standard Hansen test for overidentifying restrictions. The Hansen test can hopefully provide indications on which set of orthogonality conditions should better be exploited.

Table 5 reports the GMM estimations and the Hansen tests for both the additive error and multiplicative error formulations.

3.3.4 Additional specification tests

In addition to the test on endogeneity of the extension variable performed through the 2SRI procedure, we can also test the endogeneity assumption by a standard Hausman test. As argued in Staub (2009), the NLIV/GMM in the formulation of Mullahy (1997) is consistent under both exogeneity and endogeneity of the regressors but it is non-efficient. On the contrary, the MLE is consistent only under exogeneity but it is efficient. The estimates of the coefficients obtained by NLIV and MLE can therefore be used to construct an Hausman test based on the estimated difference in the coefficients for NLIV and ML.

5. Results

Table 2 presents two alternative specifications of our count data model: the central columns of the table display results obtained by fitting a Poisson model, whereas the columns on the right are obtained from a Negative Binomial model. The regression-based tests for over-dispersion (available on request) based on Cameron and Trivedi (1990) always reject the null hypothesis of equi-dispersion. Consistently, the likelihood-ratio test on the Negative Binomial parameter α, reported in Table 2, highlights the presence of over-dispersion and confirms that the Negative Binomial specification fits the data better than the Poisson distribution. Even though our tests suggest that Negative Binomial estimates are to be preferred, the findings are fairly consistent across both specifications and, most importantly, the differences between the two models do not affect significance and size of the coefficient for extending opening hours which is our main variable of interest.

TABLE 2

Actually, a striking result comes out from the estimates of the coefficients associated to this dummy variable. The impact of the extension is highly significant and positive across both specifications. At first glance, the interpretation of such evidence turns out to be problematic as the identified effect goes in the opposite direction than one could expect. Indeed, improving primary care organization in terms of extended opening hours of the practice seem to raise the inappropriate use of ED services, after controlling for a set of relevant covariates including physician characteristics and composition of the list.

One possible explanation for this empirical puzzle is that, as argued above, the results may be biased due to the potential correlation between the choice of extending opening hours and unmeasured latent factors that also affect the outcome variable. As previously discussed, the selection into the program that extends practice opening is in fact not random as it reflects a voluntary choice by GPs. We tackle this potential endogeneity issue by adopting the IV estimation strategies discussed in detail in previous sections.

5.1 Instrumental variable estimation

We can briefly recall here the basic idea behind our two-stage IV approach. We estimate first the probability of recording an extension in practice opening hours (10-12 hours per day) against a set of regressors where two instruments are added to the controls included in Table 2. The instruments chosen are the number of years the GP has worked in network or group before 2009 and the district physicians' concentration index of the district where the GP operates. In the second stage, we use the residuals from the first stage equation as additional covariate for the count data equation. The residuals are expected to control for unobserved latent factors that affect both the outcome variable and the dummy for opening extension.

TABLE 3 -4

Table 3 and 4 presents the 2SRI estimates. In table 3 the LPM and Probit results for the first-stage reduced form model for the extended opening presented in equation (3). Table 4 shows the second-stage Negative Binomial estimates for the model in equation (5) where the LPM and Probit residuals from the first-stage are respectively included together with the set of regressor used in equation (1). Our exclusion restriction necessary for identification requires that the instruments affect the decision to extend opening hours but are not related to the (inappropriate) use of emergency services.

A first important result is that, although not directly comparable in terms of marginal effects, the evidence of the first stage LPM and Probit estimates is consistent across specifications and shows that, even after controlling for a set of relevant confounding factors, the two instrumental variables are significant predictors of the probability to have extended opening hours in 2009. Having worked in association with colleagues for a longer period increases the probability that GPs agreed to extend opening hours by 2009. Similarly, a higher density of GPs in the district facilitates the extension of practice opening. The standard F-test on the joint relevance of the instruments in the LPM is well above the standard threshold of 10 (Staiger and Stock (1997)), a result which indicates that the instruments are jointly relevant and good predictors of the extension. Similarly, the χ^2-test on the IV relevance for the Probit estimates fails to reject the null hypothesis that the instruments are jointly relevant. From the analysis of the estimates of the reduced-form for the extension, we have evidence that supports our choice of the instruments.

The second-stage ML NB estimates includes the residuals from the first-stage estimation and provide first a direct test of exogeneity for the extension variable. In addition to it and provided that the instruments are valid ones, the estimates account now for the potential endogeneity of the variable of main policy interest. In the second stage estimation standard errors are bootstrapped in order to account for the fact that the Negative Binomial specification includes a regressor obtained from first stage estimates which substitutes for true unobservable latent factors of interest. The corrected standard errors are reported in the table.

The estimates of the coefficient ρ of the first-stage residuals are significantly different from zero with residuals obtained from both a LPM and a Probit first stage estimates. Our evidence seems to confirm that the extended opening variable is not exogenously determined and is correlated through unobservable factors that also affect the outcome variable y.

Once we account for such correlation and allow the policy variable to be endogenous, we observe a substantial change in the estimated effect of the extending opening hours of the practice. The coefficients are still statistically significant but with a negative sign and point out that an improvement in daily accessibility of primary care practices reduces the inappropriate use of ED services.

As for the remaining set of covariates, our estimates provides fairly similar results across the two specifications and the indication about the effects of these determinants on the use of emergency services are the very same.

Individual characteristics of the GPs affect the pattern of (inappropriate) utilisation of ED services by listed patients. Age and seniority are always significant and have a non linear impact, as is shown by significant quadratic terms, while gender is significant at 5% only in the first specification. In general, male GPs and GPs with a longer professional experience in primary care appear to have a lower frequency of inappropriate visits to the ED by their patients. Also the coefficient for rural practice location is significant and captures a higher propensity of attending inappropriately EDs for patients treated by GPs located in non-urban areas. This may be due to the fact that residents in low densely-populated areas may face longer distance and higher accessibility problems to their GPs and to other outpatient facilities with the consequence of making relatively more attractive the option of attending hospital emergency rooms instead of outpatient facilities. The presence of nursing staff is highly significant and has the expected negative sign. The practices that avail themselves of nursing personnel come out to be better at preventing inappropriate ED attendance of the GPs involved.

Statistically significant effects emerge also when we consider the characteristics of the list. In particular, a higher proportion of male and non-native patients in the list significantly increases the probability of attending ED inappropriately. Patient age has an effect only relative to the younger age groups: the propensity of inappropriate access to ED first increases with age and reaches the highest probability for patients aged from 36-50; then the coefficients for older age classes loose significance. This could signal that patients who enjoy regular access to their GPs, typically the elderly, are less likely to go to an emergency room than patients who are unable or do not prefer to see a doctor regularly because of a higher opportunity cost of time.

TABLE 5

Table 5 reports the GMM estimation for the outcome model under the two alternative formulations of the error term: the left hand side of the table presents the estimates for a model with a multiplicative error term [equation (9)], while in the right hand side the error term is specified as additive [equation (7)].

The instrumental variables used in the estimation are the same proposed for 2SRI estimation: number of years in an association form and the concentration of GPs, but the orthogonality conditions vary according to the specification of the error term in the model, as discussed above.

First of all, it is worth considering the outcome of the Hansen's J test for over-identifying restrictions in order to have indications on the valid set of moment conditions in this framework, as not necessarily both sets hold for the same model. The instruments can be safely excluded from the model only if they are orthogonal to the residuals. In the multiplicative error framework, the Hansen p-value is about 0.11 so that we fail to reject the null hypothesis of joint validity of the exclusion restrictions. On the contrary, in the additive error formulation, given a p-value of less than 0.02, we reject the null hypothesis suggesting that the orthogonality conditions do not hold. We thus opt for the multiplicative error formulation and estimate the model by the GMM approach developed by Mullahy (1997).

In line with the findings of the 2SRI strategy, the coefficient for the extended opening is negative and highly significant and confirms that higher time accessibility to primary care practices can be effective at reducing the number of inappropriate ED episodes.

The effects of the covariates are generally in line with the results discussed above, with only a few exceptions. GP age loses significance both in its linear and quadratic term, while seniority the maintains only a linear impact on the outcome variable. Among the characteristics of the list, only the share of non-native patients in the list diverges from the previous evidence as it is no more significant.

In order to test for the exogeneity of the policy variable also in a moment-based framework, an Hausman test can be performed to compare ML estimates with all the regressors assumed to be exogenous and the estimates obtained by GMM/NLIV in the formulation with multiplicative error. Since we systematically reject the null hypothesis of no significant differences between the two estimates, we have again evidence of the presence of endogeneity of the policy variable.

To summarise, all our findings point out that the policy variable of interest is endogenous to the outcome variable due to a likely self-selection into the extension program by GPs. Once this issue is addressed and controlled for, our results consistently indicate a significant impact of the extension of the opening hours on inappropriate utilisation of ED services and the negative sign of the coefficient points to an effect leading the reduction of inappropriate visits to ED. We show here that improved accessibility to primary care, as proxied by the extension of the opening hours, acts as a relevant restraint for the inappropriate use of ED being able to answer more effectively to the demand for non-urgent care. Such evidence is robust to the choice of different IV estimation strategies that tackle endogeneity.

5.2 Robustness checks

Finally, we examine some possible extensions of the analysis in order to test the robustness of our results. At this scope we re-estimate the GMM model in the multiplicative error formulation using two different specifications of the dependent variable.

TABLE 6

First, we isolate from the administrative dataset a subsample of white code visits that are registered by the triage nurses at the moment of admission as “self-referred” visits. There are not strict guidelines by the Regional Department of Health for the compilation of the field relative to the sender (GP, specialist physician, self-referral…) of the patient to the ED. This field is not compulsory so this information is not always thoroughly filled in the hospital admission form: the variable therefore presents many missing values and it is subject to measurement errors. We therefore do not exploit it in our main analysis as outcome variable but, though not very accurate and reliable, we think it can still provide a useful additional check. Except for the rural practice location that remains significant, with an higher coefficient than in Table 5, and that captures a higher propensity of a self-referral to EDs for patients treated in non-urban areas, the other covariates signs and their significances are not always consistent with the previous estimates and do not seem very reliable. We attribute this evidence to the weakness of this dependent variable. However, the extension of opening hours continues to be highly significant in decreasing the probability that a patient self-refers to EDs, supporting the validity of our main conclusion.

Second, we drop from the estimating sample the GPs with more than 200 white code accesses in the year 2009 which reduces our observations to 2317. From Figure 1 and 2 we can identify several outliers in the distribution of white codes per GP: we drop those observations in order to check the robustness of the results and to ensure that the finding in the main analysis are not driven by the outliers. The presence of these GPs who present outlier values in the outcome variable could be due both to the existence in the list of frequent flyers patients - an heterogeneous group of patients (often with chronic medical, mental health, alcohol and drug problems, as well as other psychosocial diseases) that tend to be persistent heavy users of ED – and/or to a relatively higher propensity by the GP to send recurrently her/his patients to the ED, no matter the severity of the disease. The GMM estimation for this new subsample confirms previous results and supports our conclusion of the positive effects produced by extending opening hours in reducing inappropriate access to ED. The evidence in the main analysis therefore comes out to be robust to the presence of outliers in the distribution of the dependent variable.

6. Conclusive remarks

Improving the accessibility to primary care services has been a recurring headline topic in the agenda of Italian policymakers in more recent years. This is an important target not only nationwide but also for the region we consider, Emilia Romagna, where increasing efforts have been provided to extend opening hours of the GPs associated in networks or groups in order to ensure an overall daily coverage up to a maximum of 12 hours. The main goal of the present work was to assess the impact of these innovative organisational features on hospital ED attendances, in order to gain better insights on the effectiveness of a policy aimed at increasing efficiency and appropriateness within the healthcare sector.

To this aim, we focused on ED patients classified as white codes by the Italian hospital triage system as they are cases that, according to both urgency and severity criteria, should have been treated in a primary care setting. Our dependent variable was the number of inappropriate ED attendances for patients registered in the GP list. We estimated a Poisson and a Negative Binomial model for over-dispersed count data and accounted for the potential endogeneity of our variable of interest, the extended opening hours up to 12 hours/day, by adopting a two-stage residual inclusion (2SRI) strategy and a generalized method of moments (GMM) approach. Whereas standard estimations provide a counterintuitive evidence, and extension of opening hours seems to be associated to higher frequency in of inappropriate attendance to ED, once we account for the potential endogeneity of the variable of main policy interest through an IV approach, our results support the hypothesis that improving primary care organization in terms of accessibility favours a more appropriate use of ED services.

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. As the opening hours of general practices seem to influence the behaviour of patients, the opportunity to exploit longitudinal data could help derive more conclusive policy implications. A second limitation concerns the fact that our data sets do not register time of arrival at the ED. To provide more robust evidence on the fact that we are genuinely measuring the impact produced by extended hours in primary care practices on ED attendances, one promising way to extend further research will be focused on different specifications of the dependent variable, in order to distinguish, for example, patterns occurring during week days from those of the weekends, when GP services are typically not available and the two care setting cannot be seen as substitute.

References

Afilalo J, Marinovich A, Afilalo M, et al. (2004), *Nonurgent emergency department patient characteristics and barriers to primary care,* Academic Emergency Medicine, 11 (12): 1302-10.

Angrist JD (2000), *Estimation of limited-dependent variable models with dummy endogenous regressors: simple strategies for empirical practice,* NBER Technical Working Paper 248.

Armstrong B, Levesque O, Perlin JB et al. (2006), *Reinventing Veterans Health Administration: focus on primary care.* Healthc Q, 9: 80-85.

Bezzina AJ, Smith PB, Cromwell D, Eagar K (2005), *Primary care patients in the emergency department: who are they? A review of the definition of the "primary care patient" in the emergency department,* Emergency Medicine Australasia, 17: 472-9.

Bianco A, Pileggi C, Angelillo IF (2003), *Non-urgent visits to a hospital emergency department in Italy,* Public Health, 117: 250-55.

Calderòn-Larranaga A, Carney L, Soljak M et al. (2011), *Association of population and primary healthcare factors with hospital admissions rates for chronic obstructive pulmonary disease in England,* Thorax, 66 (3): 191-96.

Cameron, AC and PK Trivedi. (1990), *Regression-Based Tests for Overdispersion in the Poisson Model*." Journal of Econometrics, 46:347-364.

Cameron AC and PKTrivedi (1998), *Regression Analysis of Count Data*, Econometric Society Monograph No.30, Cambridge University Press.

Cameron AC and PKTrivedi (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press.

Cameron AC and PKTrivedi (2010), *Microeconometrics Using Stata*, Revised Edition, Stata Press books, StataCorp LP.

Carlsen F, Grytten J, Kjelvik J, Skau I (2007), *Better primary physician services lead to fewer hospital admissions*, European Journal of Health Economics, 8:17-24.

Dusheiko, M. Gravelle, H. Martin, S. Rice, N. Smith, P. (2011). *Does better disease management in primary care reduce hospital costs? Evidence from English primary care*. Journal of Health Economics, 30, pp. 919-932.

Eijkenaar F. *Key issues in the design of pay for performance programs.* European Journal Health Economics 2013; 14; 117-131.

Emmert M, Eijkenaar F, Kemter H, Esslinger AS, Schöffski O. (2012). Economic evaluation of pay-for-performance in health care: a systematic review. European Journal Health Economics, 13 (6): 755-67.

Fantini, M.P.; Compagni, A.; Rucci, P.; Mimmi, S.; Longo, F. (2012), *General practitioners' adherence to evidence-based guidelines: A multilevel analysis.* Health Care Management Review, 37, 67–76.

Fattore G., Forini F., Salvatore D.. Tozzi V. (2009), S*ocial network analysis in primary care: the impact of interactions on prescribing behaviour*. Health Policy, 92, 141-148.

Feachem RG, Sekhri NK, White KL (2002), *Getting more for their dollar: a comparison of the NHS with California's Kaiser Permanente.* British Medical Journal, 324: 135-41.

Fiorentini G., Iezzi E., Lippi Bruni M., Ugolini C. (2011), *Incentives in primary care and their impact on potentially avoidable hospital admissions.* European Journal of Health Economics, 12: 297-309.

Fiorentini, G., Lippi Bruni, M., Ugolini, C. (2008). Health Systems and Health Reforms in Europe: the Case of Italy. Intereconomics, 43: 205-212.

Flores-Mateo G., Violan-Fors C., Carrillo-Santisteve P., Peiro S., Argimon JM. (2012) Effectiveness of organizational interventions to reduce emergency department utilization: a systematic review. PLoS ONE; 7(5):e35903.

GAO, General Accounting Office (2009), Hospital emergency departments. Publication 09-347, Washington DC, April.

Garcia T.C, Bernstein A.B., Bush M.A (2010), Emergency Department Visitors and Visits: Who Used the Emergency Room in 2007? NCHS Data Brief, no. 28.

Guttman N., Zimmerman D.R., Nelson M.S. (2003), The many faces of access: reason for medically nonurgent emergency department visits. Journal of Health, Politics and Law, 28 (6): 1089-120.

Hansagi H, Carlsson B, Olsson M, Edhag O (1987), *Trial of a method of reducing inappropriate demands on a hospital emergency department*, Public Health, 10: 99-105.

Harris MJ, Patel B, Bowen S (2011), *Primary care access and its relationship with emergency department utilisation,* British Journal of General Practice, December, e787-93.

Hausman JA (1978), *Specification Tests in Econometrics*, Econometrica, 46(6):1251-1271.

Howard M, Goertzen J (2008), *Emergency Department and Walk-in-Clinic use in models of primary care practice with different after-hours accessibility in Ontario*, Healthcare Policy, 4, 1: 73-88.

Klijakovic M, Allan BC, Reinker J (1981), *Why skip the General Practice and go to the Accident and Emergency Department?,* NZ Medical Journal, 93: 49-52.

Lang T., David, A, Diakitè B, Agay E, et al. (1996), *Non-urgent care in the hospital medical emergency department in France: how much and which health needs does it reflect?*, Journal of Epidemiol Community Health, 50: 45-62.

Lee A, Lau FL, Hazlett CB et al. (2000), *Factors associated with non-urgent utilization of Accident and Emergency services: a case-control study in Hong Kong*, Social Science & Medicine, 51: 1075-85.

Lega F., Mengoni A. (2008), *Why non-urgent patients choose emergency over primary care services? Empirical evidence and managerial implications*, Health Policy, 88: 326-38.

Lo Scalzo A, Donatini A, Orzella L et al. (2009), Italy – Health System Review, Health System in Transition, European Observatory on Health Systems and Policies, 11 (6).

Lowe RA, Localio AR, Schwarx DF et al. (2005), *Association between primary care practice characteristics and emergency department use in a Medicaid managed care organization*, Medical Care, 3(8): 792-800.

McCusker J, Roberge D, Lévesque JF et al (2010) *Emergency Department visits and primary care among adults with chronic conditions*, Medical Care, 48, 11: 972-80.

McGraig L., Burt CW (2006) *National hospital ambulatory medical care survey: 2004 emergency department summary.* Advances in Data, 372: 1-5.

Moll van Charante EP, ter Riet G, Bindels P JE (2008), *Self-referrals to the A&E department during out-of-hours: patients'motive and characteristics*, Patient Education and Counselling, 70: 256-65.

Moll van Charante EP, van Steewikj-Opdam P CE, Bindels P JE (2007), *Out-of-hours demand for GP care and emergency services: patients'choice and referrals by general practitioners and ambulance services,* BMC Family Practice, 8: 46.

Mullahy J (1997), *Instrumental-Variable Estimation Of Count Data Models: Applications To Models Of Cigarette Smoking Behavior*, The Review of Economics and Statistics, vol. 79(4): 586-593.

Pizer SD (2009), *An intuitive review of methods for observational studies of comparative effectiveness,* Health Serv Outcome Res Method, 9:54-68.

Puig-Junoy J, Saez M, Martìnez-Garcìa E (1998), *Why do patients prefer hospital emergency visits? A nested multinomial logit analysis for patient-initiated contacts,* Health Care Management Science, 1: 39-52.

Redelmeier DA, Fuchs VR (1993), *Hospital expenditures in the United States and Canada,* New England Journal of Medicine, 328: 772-78.

Rivers D, Vuong Q (1988), *Limited Information Estimators and Exogeneity Tests for Simultaneous Probit Models*, Journal of Econometrics, 39: 347-366.

Roberts E, Mays N (1998), *Can primary care and community-based models of emergency care substitute for the hospital A&E department?* Health Policy, 44: 191-214.

Saxena S, George J, Barber J et al. (2006), *Association of population and practice factors with potentially avoidable admission rates for chronic diseases in London*, J R Soc Med, 99 (2): 81-89.

Scott A, Sivey P, Ait Ouakrim D, Willenberg L, Naccarella L, Furler J, Young D. (2011) *The effect of financial incentives on the quality of health care provided by primary care physicians* (Review).The Cochrane Collaboration. Published by JohnWiley & Sons, Ltd.

Sempere-Selva T, Peirò S, Sendra-Pina P, Martinez-Espin C, Lòpez-Aguilera I (2001), *Inappropriate use of an accident and emergency department: magnitude, associate factors and reasons*, Ann Emerg Med, 37: 568-79.

Shah NM, Shah MA, Behbehani J (1996), *Predictors of non-urgent utilization of hospital emergency services in Kuwait,* Social Science & Medicine, 42: 1313-23.

Shaw SE, Meads G (2012), *Extending primary care: potential learning from Italy,* Primary Health Care Research & Development, 13 (4): 289-93.

Smith RJ, Blundell RW (1986), *An Exogeneity Test for a Simultaneous Equation Tobit Model with an Application to Labor Supply*, Econometrica,54: 679-685.

Staiger, D and JH Stock (1997), *Instrumental Variables Regression with Weak Instruments,* Econometrica, 65: 557-586

Staub KE, *Simple tests for exogeneity of a binary explanatory variable in count data regression models*, Communication in Statistics – Simulation and Computation, 38:9, 1834-1855.

Terza, JV (1998), *Estimating count data models with endogenous switching: Sample selection and endogenous treatment effects*, Journal of Econometrics 84: 129–154.

Terza JV, Bradford WD, Dismuke CE (2008), *The use of linear instrumental variables methods in health services research and health economics: a cautionary note*, Health Services Research, 43:3, 1102-1120.

Terza JV, Basu A, Rathouz PJ (2008), *Two-stage residual inclusion estimation: Addressing endogeneity in health econometric modeling*, Journal of Health Economics, 27:531-543.

Thompson C, Hayhurst C, Boyle A (2010), *How have changes to out-of-hours primary care services since 2004 affected emergency department attendances at a UK District General Hospital? A longitudinal study,* Emergency Medicine, 27: 22-25.

Tsai J.C-H, Chen W-Y, Liang Y-W (2011), *Nonemergent emergency department visits under the National Health Insurance in Taiwan,* Health Policy, 100: 189-95.

Weinberger M, Oddone EZ, Henderson WG (1996), Does increased access to primary care reduce hospital readmissions?, New England Journal Medicine, 334(22): 1441-47.

Wilson K.J. Pay-for-performance in health care: what can we learn from international experience. Quality Management in Health Care 2013; 22 (1); 2-15.

Windmeijer FAG and Santos Silva JMC (1997), *Endogeneity in count data models: an application to demand for health care*, Journal of Applied Econometrics, Vol. 12, 281-294.

Wooldridge JM (1997), *Quasi-likelihood methods for count data*, in Pesaran H, Schmidt P, eds Handbook of Applied Econometrics. Volume II: Microeconometrics, Blackwell Publishers Ltd, Malden, MA.

Wooldridge JM (2002), *Econometric analysis of cross section and panel data,* MIT Press, Cambridge, MA.

Table 1. Descriptive statistics for the variables included in the model

	Whole sample		*No Extension*		*Extension*	
	N=2370		N=1651		N=719	
Variable	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**	**Mean**	**Std. Dev.**
Male GP	0.707	0.455	0.701	0.458	0.719	0.450
GP age	54.886	4.891	54.949	4.804	54.741	5.084
GP seniority	21.450	8.283	21.389	8.330	21.588	8.178
Rural GP practice	0.215	0.411	0.250	0.433	0.135	0.342
Nursing staff	0.124	0.330	0.091	0.287	0.202	0.402
List size	1251.001	328.477	1251.028	329.540	1250.9374	326.251
Age group 14-20 (% list)	0.054	0.020	0.053	0.020	0.056	0.020
Age group 21-35 (% list)	0.194	0.045	0.192	0.045	0.199	0.045
Age group 36-50 (% list)	0.280	0.037	0.280	0.037	0.281	0.036
Age group 51-65 (% list)	0.215	0.036	0.215	0.036	0.215	0.036
Age group > 65 (% list)	0.257	0.066	0.260	0.067	0.249	0.063
Foreign patients (% list)	0.075	0.067	0.075	0.067	0.074	0.066
Male patients (% list)	0.478	0.039	0.477	0.039	0.482	0.039

Figure 1. Distribution of white codes per GP for the non-extension sub-sample

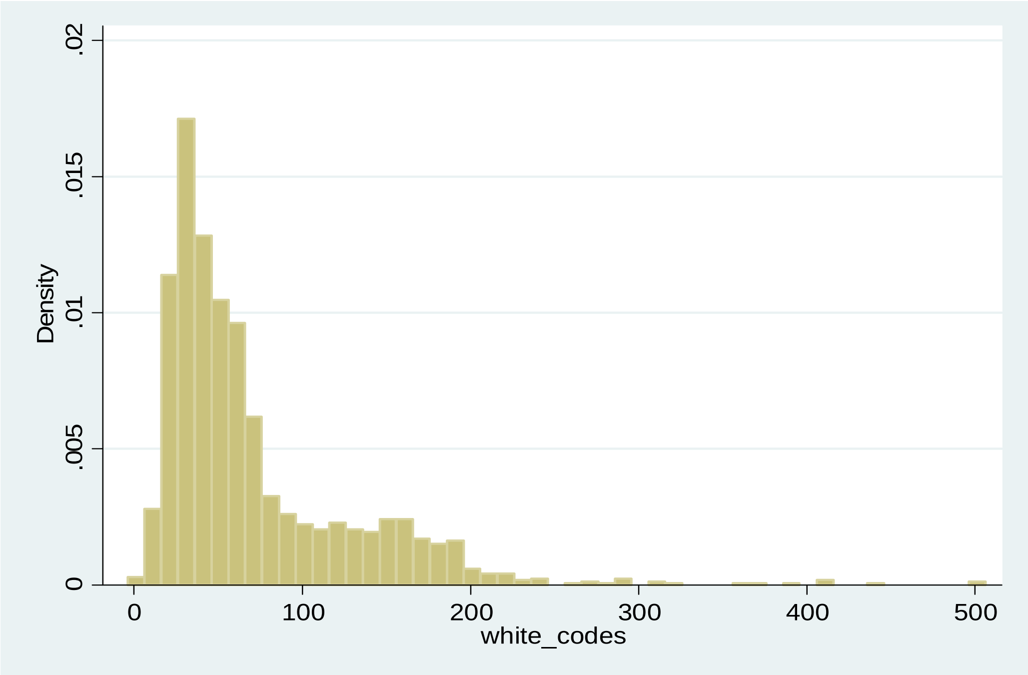

Figure 2. Distribution of white codes per GP for the extension sub-sample

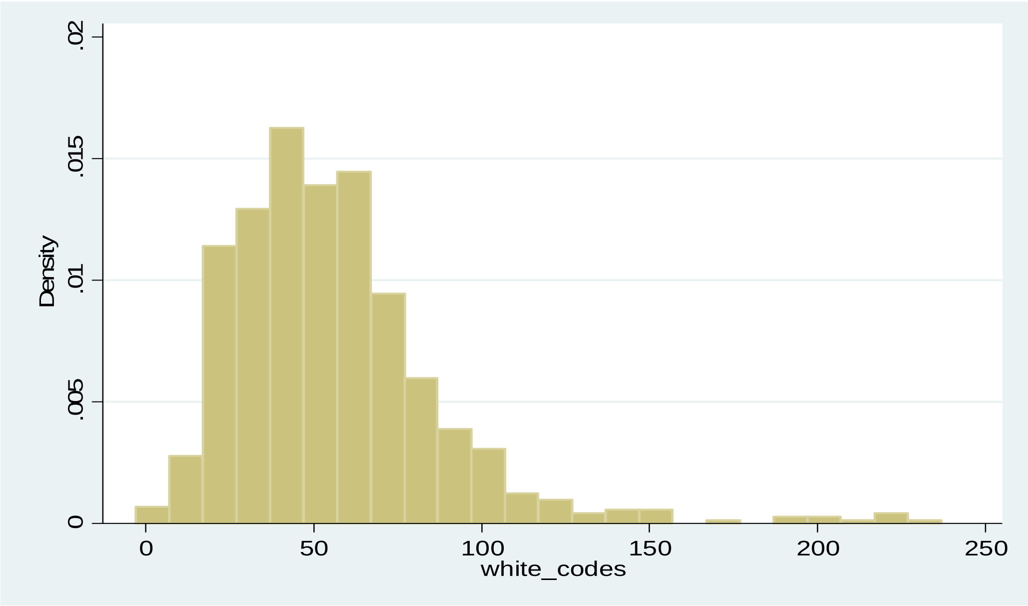

Figure 3. Observed proportions for white codes vs Poisson and Negative Binomial probabilities

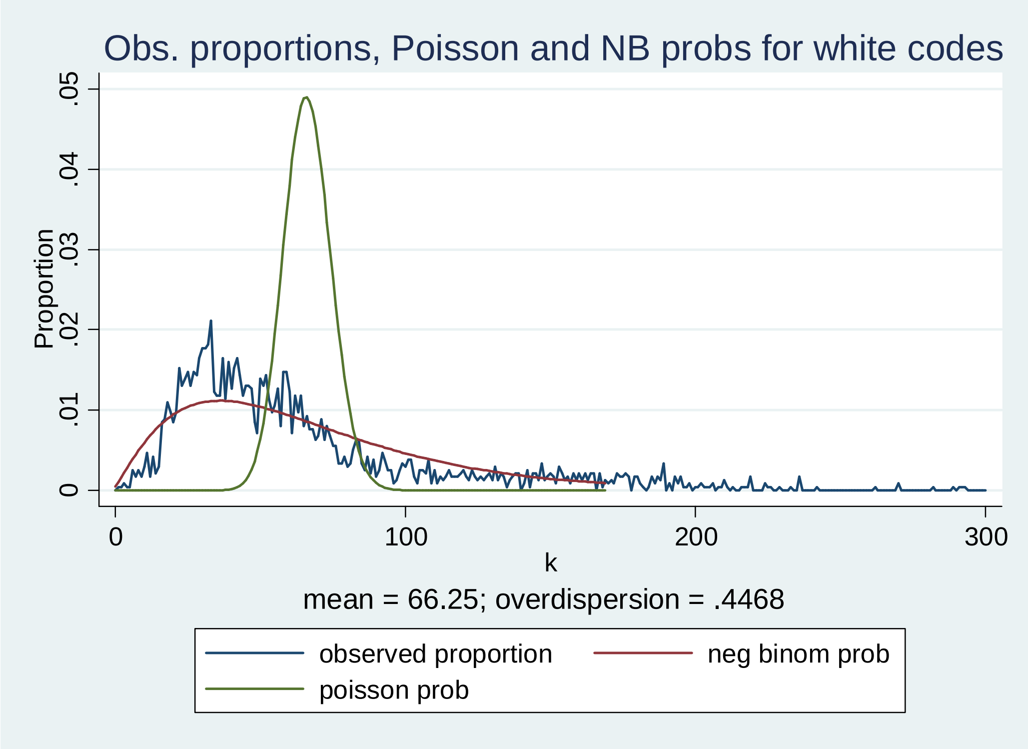

Table 2. White codes for GP, year 2009.

White codes for GP	*Poisson model*			*Negative binomial model*		
	Coefficient (SD)	*p value*	*IRR*	*Coefficient (SD)*	*p value*	*IRR*
Extended opening hours 10-12	**0.0709 (0.245)**	**0.004**	**1.0734**	**0.1089 (0.218)**	**0.000**	**1.1151**
Male GP	-0.0580 (0.034)	0.089	0.9437	-0.0427 (0.031)	0.169	0.9582
GP age	0.0916 (0.033)	0.006	1.0960	0.0719 (0.029)	0.012	1.0745
GP age squared	-0.0008 (0.000)	0.011	0.9992	-0.0006 (0.000)	0.022	0.9994
GP seniority	-0.0350 (0.008)	0.000	0.9656	-0.0309 (0.008)	0.000	0.9696
GP seniority squared	0.0006 (0.000)	0.003	1.0006	0.0006 (0.000)	0.003	1.0006
Rural GP practice	0.2779 (0.031)	0.000	1.3203	0.2334 (0.030)	0.000	1.2629
Nursing staff	-0.2436 (0.036)	0.000	0.7838	-0.2130 (0.032)	0.000	0.8081
List size	1.12083 (0.045)	0.000	3.0674	1.0700 (0.043)	0.000	2.9154
Age group 21-35 (% list)	0.5986 (0.955)	0.531	1.8195	1.8637 (0.783)	0.017	6.4473
Age group 36-50 (% list)	2.6193 (1.012)	0.010	13.726	2.8698 (0.826)	0.001	17.633
Age group 51-65 (% list)	0.2409 (0.942)	0.798	1.2724	0.5866 (0.773)	0.448	1.7979
Age group > 65 (% list)	-0.0664 (0.799)	0.934	0.9357	0.3942 (0.666)	0.554	1.4832
Foreign patients (% list)	1.2150 (0.245)	0.000	3.3703	1.0492 (0.237)	0.000	2.8553
Male patients (% list)	1.63644 (0.407)	0.000	5.1368	1.1825 (0.352)	0.001	3.2627
Vast Area 2	0.75961 (0.030)	0.000	2.1374	0.7831 (0.033)	0.000	2.1883
Vast Area 3	-0.1240 (0.022)	0.000	0.8834	-0.0948 (0.022)	0.000	0.9095
Constant	-8.0752 (1.192)	0.000	0.0003	-7.5057 (1.011)	0.000	0.0005
Alpha				0.1964 (0.007)		

Table 3. Two-stage residual inclusion estimates: FIRST STAGE

First stage LPM			**First stage Probit**		
Extended opening 10-12 hours	***Coefficient (SD)***	***p value***	***Extended opening 10-12 hours***	***Coefficient (SD)***	***p value***
Years in Group or network	0.01617	0.000	Years in Group or network	0.05987	
	(0.002)			(0.008)	0.000
Concentration index	0.01842	0.000	Concentration index	0.05395	
	(0.002)			(0.006)	0.000
Male GP	-0.05767	0.030	Male GP	-0.21588	
	(0.002)			(0.093)	0.020
GP age	-0.02880	0.300	GP age	-0.10082	
	(0.028)			(0.088)	0.254
GP age squared	0.00026	0.311	GP age squared	0.00089	
	(0.000)			(0.000)	0.265
GP seniority	-0.00374	0.589	GP seniority	-0.01190	
	(0.007)			(0.024)	0.624
GP seniority squared	0.00006	0.713	GP seniority squared	0.00019	
	(0.000)			(0.001)	0.749
Rural GP practice	-0.06382	0.005	Rural GP practice	-0.24326	
	(0.023)			(0.087)	0.005
Nursing staff	0.15220	0.000	Nursing staff	0.46639	
	(0.029)			(0.084)	0.000
List size	-0.01274	0.645	List size	-0.04942	
	(0.028)			(0.100)	0.622
Age group 21-35 (% list)	0.03186	0.965	Age group 21-35 (% list)	-0.73702	
	(0.717)			(2.463)	0.765
Age group 36-50 (% list)	-0.80914	0.285	Age group 36-50 (% list)	-3.82297	
	(0.756)			(2.569)	0.137
Age group 51-65 (% list)	-0.89899	0.199	Age group 51-65 (% list)	-3.72467	
	(0.670)			(2.357)	0.114
Age group > 65 (% list)	-1.02903	0.082	Age group > 65 (% list)	-4.20066	
	(0.590)			(2.017)	0.037
Foreign patients (% list)	-0.95718	0.000	Foreign patients (% list)	-3.20721	
	(0.211)			(0.741)	0.000
Male patients (% list)	1.38344	0.000	Male patients (% list)	5.25807	
	(0.317)			(1.114)	0.000
Vast Area 2	-0.39954	0.000	Vast Area 2	-1.36244	
	(0.025)			(0.100)	0.000
Vast Area 3	-0.27460	0.000	Vast Area 3	-0.77733	
	(0.024)			(0.077)	0.000
Constant	0.09192	0.925	Constant	-0.12075	
	(0.970)			(3.179)	0.970
F test (2, 2353) on the IVs	78.52	0.000	Chi2 (2) on the IVs	121.36	0.000

Table 4. Two-stage residual inclusion estimates: SECOND STAGE

Second stage NB (first-stage LPM residuals included)			**Second stage NB (first-stage PROBIT residuals included)**		
White codes for GP	***Coefficient (Bootstrap SD)***	***p value***	***White codes for GP***	***Coefficient (Bootstrap SD)***	***p value***
Extended opening 10-12 hours	**-0.50957 (0.088)**	**0.000**	**Extended opening 10-12 hours**	**-0.14713 (0.714)**	**0.039**
Residuals from first-stage LPM	0.67036 (0.095)	0.000	Residuals for first-stage Probit	0.28186 (0.081)	0.001
Male GP	-0.06617 (0.031)	0.034	Male GP	-0.05318 (0.030)	0.076
GP age	0.05970 (0.029)	0.040	GP age	0.06693 (0.028)	0.018
GP age squared	-0.00052 (0.000)	0.053	GP age squared	-0.00057 (0.000)	0.030
GP seniority	-0.02724 (0.008)	0.000	GP seniority	-0.02929 (0.007)	0.000
GP seniority squared	0.00048 (0.000)	0.012	GP seniority squared	0.00053 (0.000)	0.003
Rural GP practice	0.23330 (0.030)	0.000	Rural GP practice	0.22951 (0.032)	0.000
Nursing staff	-0.11432 (0.035)	0.001	Nursing staff	-0.17183 (0.036)	0.000
List size	1.04006 (0.041)	0.000	List size	1.05794 (0.419)	0.000
Age group 21-35 (% list)	1.81181 (0.796)	0.023	Age group 21-35 (% list)	1.79987 (0.827)	0.029
Age group 36-50 (% list)	2.24901 (0.843)	0.008	Age group 36-50 (% list)	2.55133 (0.882)	0.004
Age group 51-65 (% list)	0.45573 (0.841)	0.575	Age group 51-65 (% list)	0.47174 (0.763)	0.536
Age group > 65 (% list)	0.13793 (0.679)	0.839	Age group > 65 (% list)	0.12387 (0.711)	0.862
Foreign patients (% list)	0.61378 (0.247)	0.013	Foreign patients (% list)	0.85281 (0.244)	0.000
Male patients (% list)	1.71747 (0.359)	0.000	Male patients (% list)	1.43722 (0.353)	0.000
Vast Area 2	0 .55529 (0.049)	0.000	Vast Area 2	0.69535 (0.041)	0.000
Vast Area 3	-0.22554 (0.030)	0.000	Vast Area 3	-0.14614 (0.027)	0.000
Constant	-6.52155 (1.017)	0.000	Constant	-7.06948 (0.991	0.000
alpha	0.190765 (0.007)		alpha	0.19546 (0.007)	
Likelihood-ratio test of alpha=0:	Chibar2(01)= 2.4e+04	0.000	Likelihood-ratio test of alpha=0:	Chibar2(01)= 2.5e+04	0.000

Table 5. GMM estimation for white codes for GP, year 2009.

White codes for GP	*Multiplicative error*		*Additive error*	
	Coefficient (SD)	*p value*	*Coefficient (SD)*	*p value*
Extended opening hours 10-12	**-0.52573**	**0.000**	**-1.13955**	**0.000**
	(0.113)		**(0.245)**	
Male GP	-0.06382	0.089	-0.09202	0.027
	(0.037)		(0.042)	
GP age	0.04404	0.176	0.04640	0.249
	(0.032)		(0.040)	
GP age squared	-0.00039	0.201	-0.00040	0.279
	(0.000)		(0.000)	
GP seniority	-0.02115	0.021	-0.02278	0.023
	(0.009)		(0.010)	
GP seniority squared	0.00035	0.118	0.00034	0.177
	(0.000)		(0.000)	
Rural GP practice	0.20112	0.000	0.23839	0.000
	(0.033)		(0.035)	
Nursing staff	-0.12917	0.001	-0.08106	0.114
	(0.037)		(0.051)	
List size	1.03883	0.000	1.06465	0.000
	(0.044)		(0.049)	
Age group 21-35 (% list)	2.19686	0.017	0.07575	0.945
	(0.920)		(1.108)	
Age group 36-50 (% list)	2.08861	0.030	1.08058	0.368
	(0.961)		(1.201)	
Age group 51-65 (% list)	0.36740	0.678	-0.38163	0.727
	(0.886)		(1.095)	
Age group > 65 (% list)	0.10245	0.893	-1.18754	0.203
	(0.763)		(0.932)	
Foreign patients (% list)	0.44597	0.158	0.56196	0.092
	(0.316)		(0.333)	
Male patients (% list)	1.71713	0.000	2.52252	0.000
	(0.440)		(0.504)	
Vast Area 2	0.50303	0.000	0.47947	0.000
	(0.067)		(0.049)	
Vast Area 3	-0.28180	0.000	-0.30591	0.000
	(0.041)		(0.036)	
Constant	-6.09407	0.000	-5.46732	0.000
	(1.160)		(1.453)	
Hansen's J chi2(1) test	2.50753	0.113	-5.63760	0.018

Table 6. GMM estimation (multiplicative error) for robustness checks, white codes for GP, year 2009.

	Self-referred white codes for GP		*GPs with < 200 white accesses*	
	Coefficient (SD)	*p value*	*Coefficient (SD)*	*p value*
Extended opening hours 10-12	**-0.64021 (0.126)**	**0.000**	**-0.50918 (0.100)**	**0.000**
Male GP	-0.16435 (0.045)	0.000	-0.05951 (0.036)	0.099
GP age	-0.00562 (0.005)	0.222	0.04523 (0.032)	0.156
GP age squared	-0.00284 (0.003)	0.357	-0.00039 (0.000)	0.186
GP seniority	0.24908 (0.040)	0.000	-0.02907 (0.009)	0.001
GP seniority squared	-0.07553 (0.043)	0.080	0.00055 (0.000)	0.011
Rural GP practice	1.01908 (0.047)	0.000	0.16181 (0.030)	0.000
Nursing staff	1.97511 (1.146)	0.085	-0.10966 (0.036)	0.003
List size	0.03513 (1.220)	0.977	1.01248 (0.045)	0.000
Age group 21-35 (% list)	-0.31313 (1.129)	0.781	2.89820 (0.888)	0.001
Age group 36-50 (% list)	-0.32291 (0.973)	0.740	2.96707 (0.930)	0.001
Age group 51-65 (% list)	0.48543 (0.388)	0.211	0.92821 (0.856)	0.278
Age group > 65 (% list)	2.74197 (0.575)	0.000	0.69612 (0.739)	0.347
Foreign patients (% list)	-0.11591 (0.064)	0.070	0.28526 (0.301)	0.344
Male patients (% list)	-0.64475 (0.045)	0.000	1.44495 (0.420)	0.001
Vast Area 2	-4.17040 (0.984)	0.000	0.41664 (0.062)	0.000
Vast Area 3	-0.64021 (0.126)	0.000	-0.27300 (0.038)	0.000
Constant	-0.16435 (0.045)	0.000	-6.41858 (1.139)	0.000
Hansen's J chi2(1) test	1.52442	0.217	1.82369	0.177