Spatial disparities in hospital performance*

Laurent Gobillon (INED and PSE)† Carine Milcent (PSE)‡

November 19, 2012

*We are grateful to Chong-en Bai, Marteen Lindeboom, Thierry Magnac, Francesco Moscone and Jordan Rappoport for very useful discussions, as well as the participants of the 16th EU Health Econometrics Workshop in Bergen and the 54th NARSC in Savannah. This project was carried out with the financial support of the French Direction of Research, Studies, Evaluation and Statistics (DREES). We are responsible for all remaining errors.

†Corresponding author. INED, PSE, IZA and CEPR. INED, 133 boulevard Davout, 75980 Paris Cedex. Email: laurent.gobillon@ined.fr. Webpage: http://laurent.gobillon.free.fr.

‡PSE (CNRS-EHESS-ENPC-ENS), 48 boulevard Jourdan, 75014 Paris. Email: milcent@pse.ens.fr. Webpage: http://www.pse.ens.fr/milcent/index.html.

Abstract

Using a French exhaustive dataset, this paper studies the determinants of regional disparities in mortality for patients admitted to hospitals for a heart attack. These disparities are large, with an 80% difference in the propensity to die within 15 days between extreme regions. They may reflect spatial differences in patient characteristics, treatments, hospital characteristics, and local healthcare market structure. To distinguish between these factors, we estimate a flexible duration model. The estimated model is aggregated at the regional level and a spatial variance analysis is conducted. We find that spatial differences in the use of innovative treatments play a major role whereas the local composition of hospitals by ownership does not have any noticeable effect. Moreover, the higher the local concentration of patients in a few large hospitals rather than many small ones, the lower the mortality. Regional unobserved effects account for around 20% of spatial disparities.

Keywords: spatial health disparities, economic geography, stratified duration model

JEL code: I11, C41

1 Introduction

In many countries, spatial disparities between local markets are large and raise major policy concerns. Whereas the focus of attention is often the labour market (Combes and Overman, 2004), disparities also occur on other markets such as housing or health. This paper develops a new descriptive approach to explore the factors related to spatial disparities in healthcare quality. In the health literature, some studies quantify the international variations in healthcare reimbursment and utilization (Wagstaff and van Doorslaer, 2000) and the interregional variations in health care delivery (Sutton and Lock, 2000). Others assess the effect of local healthcare structure on health outcome (see Gaynor, 2006 for a survey). However, they do not examine the extent to which spatial variations in patients' characteristics, hospital composition or local healthcare structure are related to spatial disparities in quality.

In fact, evaluating the marginal effect of certain factors on mortality and assessing how spatial variations in these factors may explain spatial disparities in mortality are two related but different exercises. For instance, it is usually found that sex significantly affects mortality. However, if there is no variation in the share of females across the territory, the differences in the local sex composition do not explain the disparities in mortality across locations. The same arguments apply when considering local determinants such as local competition indices. Local competition may have a significant marginal effect on mortality but only small spatial variations. In that case, it is largely unrelated to the spatial disparities in mortality across the territory.

Spatial disparities in health outcomes have become a major concern in France. A driver of these disparities is often thought to be the missallocation of resources caused by a lack of information on local needs. As a consequence, a reform called "Plan Juppé" was introduced in 1996 to decentralize the organization and funding of the healthcare system at regional level. Since then, a global national budget has been allocated across regions. Some public regional agencies allocate

the regional budgets between public and private not-for-profit hospitals located in their region after some bilateral bargaining. Regions can thus be considered to some extent as some local healthcare markets.

In this paper, we explore the factors related to regional disparities in mortality by acute myocardial infarction (AMI) in France. Our approach is descriptive and we do not claim to establish causal effects. We focus on three types of factors in line with the health literature. First, spatial disparities in mortality can be related to differences in the local composition of patients (case-mix) if there is some spatial sorting according to individual attributes related to the propensity to die (such as age or sex). Second, they can be related to spatial differences in hospital attributes such as ownership status which is usually found to affect hospital performances. McClellan and Staiger (2000) show that within specific markets in the US, elder care is of better quality in for-profit hospitals than in not-for-profit hospitals. Milcent (2005) finds that patients in for-profit hospitals in France have a lower probability of death from a heart attack than patients in public hospitals.[1] Hospitals also exhibit some variations in equipment, innovative treatments, physician skills and activities that can be related to differences in health outcome (Tay, 2003). Third, spatial disparities in mortality can be related to spatial differences between local healthcare markets. In particular, local competition is often investigated but its effect on mortality has not been clearly demonstrated. Whereas Kessler and McClellan (2000) find for the US that local competition has a significant negative impact on mortality, the effect of local competition is an open public debate in the UK (Bloom et al., 2011). Other mechanisms related to the organization of local markets were also investigated such as the effect of local specialization in surgical treatments on mortality (Chandra and Staiger, 2007).

[1] Other references include Hansman (1996), Newhouse (1970), Cutler and Horwitz (1998), Gowrisankaran and Town (1999), Silverman and Skinner (2001), Sloan et al. (2001), Kessler and McClellan (2002), Shortell and Highes (1998), Ho and Hamilton (2000).

Many articles use the mortality rate within the hospital adjusted by the severity of illness to measure hospital quality (Geweke et al., 2003). However, their approach does not take into account the possible correlation between the mortality rate and the patients' in-hospital length of stay (Hamilton and Hamilton, 1987). This issue can be addressed by considering a duration model which takes into account patients' discharge through censorship as shown by Milcent (2005). The hospital unobserved heterogeneity in capacity, efficiency and management is taken into account though hospital fixed effects. However, such an approach does not control for variations in discharge practices across hospitals at each day of patients' stays. In the current paper, we use a very flexible econometric specification building on Ridder and Tunali (1999) and Gobillon et al. (2011). We estimate a Cox duration model stratified by hospital allowing for censorship. This specification is flexible enough to take into account hospital discharge specificities on each day of a stay and to recover a hospital-specific mortality profile for the duration of stay. At the individual level, the propensity of patients to die during their stay in hospital is specified as a function of three types of explanatory factors: demographic characteristics, diagnoses and comorbidities, and procedures. In a second step, we try to relate the hospital-specific death profiles to hospital characteristics and local healthcare market structure. We finally average the estimated model at the regional level and conduct a spatial variance analysis to assess the importance of each group of factors in explaining regional disparities in mortality in the spirit of the literature in labour economics (Abowd et al., 1999) and in economic geography (Combes et al., 2008).

Estimations are conducted on a unique matched patient-hospital dataset from exhaustive French administrative records over the 1998-2003 period. This original dataset contains some information on the demographic characteristics of patients, their diagnoses, their comorbidities and their treatments. It also provides details on the hospitals where the patients are treated, such as the location, the ownership status and the capacity.

We find that regional disparities in mortality are quite large, with an 80% raw difference in the

propensity to die within 15 days between extreme regions. Our spatial variance analysis shows that the regional use of innovative treatments plays a major role. Demographic characteristics (age and sex) which significantly affect mortality have a lesser role because spatial variations in patients' composition are not large. Interestingly, the regional composition of hospital ownership has a negligible role. This finding is explained by the fact that, when the use of innovative procedures is held constant, the propensity to die is similar in for-profit hospitals and public hospitals. Several results also suggest the existence of learning-by-doing or the concentration of skilled physicians in the same place. Indeed, we find that the concentration of patients in fewer and larger hospitals is correlated with a lower mortality. Mortality is also lower when the proportion of patients in the hospital treated for a heart attack is higher. Finally, regional unobserved effects account for around 20% of spatial disparities.

The rest of the paper is organized as follows. In the second section, we present the factors which may be related to spatial differences in healthcare quality in France and review the corresponding literature. The third section describes our dataset and presents descriptive statistics on regional disparities. The fourth section details the methodology used to assess the importance of factors related to regional disparities in mortality. The fifth section summarizes the results obtained with our approach.

2 French hospitals and healthcare market

We now describe the French healthcare system during the 1998-2003 period studied in our paper and highlight how its specificities may be linked to regional disparities in healthcare and hospital quality.

2.1 Reimbursement rules

Hospitals reimbursement rules vary depending on their ownership status. For-profit (FP) hospitals are financed *via* a fee-for-service system and are managed as profit-maximizing firms. By contrast, public and not-for-profit (NFP) hospitals are funded under a global budget system. If these hospitals do not spend all their budget on healthcare during a year, the remaining money is taken back by the public regulator and the hospital budget is reduced accordingly the following year. The global budget system does not provide any financial incentive to make any profit or to attract patients.

2.1.1 Innovation and reimbursement rules

When patients having an AMI are admitted to a hospital, there is no waiting list for treatment with consumable innovative medical devices such as angioplasty or stent.[2] However, the propensity to be treated with an innovative procedure depends both on demand factors such as the characteristics and diagnoses of the patients, and supply factors such as the ability of hospitals to perform the procedure and the hospital budget constraint at the time patients need to be treated.

Because of the reimbursement rules, hospitals do not have the same budget constraints to treat patients with consumable innovative medical devices. FP hospitals are reimbursed for every device used in care. By contrast, public and NFP hospitals are reimbursed for every stay at a flat rate which depends on the type of stay, and additional consumable devices used for care are not reimbursed. Therefore, FP hospitals have more incentives to perform innovative procedures than public and NFP hospitals. Spatial sorting of FP hospitals can thus create regional disparities in the use of innovative treatments and, consequently, regional disparities in mortality.[3]

[2] See below for a definition of the angioplasty and stent.

[3] Alternatively, spatial disparities in innovative treatments and mortality may also be linked to regional differences in patient composition.

2.1.2 Selection of patients and reimbursement rules

In France, there is no segmentation of the healthcare market by insurance status interfering with the effect of the hospital status.[4] Hospital in-patient expenditures are fully reimbursed by a unique public compulsory insurer funded by taxes and patients can freely choose their hospital within the region. The public health insurance covers in-patient expenditures in all hospitals whatever their ownership status. It is thus possible to identify the effect of hospital status without any interference with insurance status. However, in for-profit hospitals, patients have to pay "out-of-pocket" fees for catering, accommodation and, in some cases, physicians' top-up fees which are marginal compared with the total medical cost but are quite onerous for the patients and are not covered by the public health insurance.

Contrary to public hospitals which admit everyone, FP hospitals can select patients.[5] The selection is usually designed to maximize profit by taking into account the patients' health status and their medical needs. When hospitals invest in high-tech equipment and recruit physicians who are able to perform innovative treatments, they have strong incentives to admit patients with comorbidities and diagnoses that justify the use of innovative treatments.

In our study, we disentangle the role of hospital status and patients' attributes by introducing in our regressions an extensive list of patients' comorbidities and diagnoses that may justify the use

[4] The context is similar in Taiwan (see Lien et al., 2008). It contrasts with the case of the US where people with a managed care insurance must choose an establishment in a given subset of hospitals with a specific status.

[5] According to the French administrative literature, NFP hospitals are not supposed to select patients as they are required to take care for all individuals as a duty of public service. However, they do receive private donations. The amount of donations depends on the reputation and size of the hospital. Moreover, part of these donations can be used to pay the staff (managers, physicians, nurses, ...) as mentioned by Besley and Ghatak (2005), and Ghatak and Mueller (2011). NFP thus have incentives to attract skilled surgeons with wage bonuses paid on donations, as well as patients with high-quality services, and to select patients to some extent to avoid bad health outcomes. NFP hospitals thus have incentives to select patients, but to a lesser extent than FP hospitals.

of innovative treatments. It would be useful to control for the financial situation of patients as it determines whether they can afford the "out-of-pocket" fees in FP hospitals. As this information is not available in our data, we control instead for average income in the patients' municipality of residence.

2.2 Competition between hospitals

There is a growing body of literature in the US and the UK on the effect of local competition between hospitals on their performance. It is shown that in the US, the local competition between hospitals tends to decrease the patients' propensity to die (see Gaynor, 2006, for a survey), whereas in the UK this is still an open public debate (Bloom et al., 2011). Some recent scientific papers show the benefits of hospital competition on quality (see Propper et al. 2004, 2008; Cooper et al., 2011). We now assess the extent of competition between hospitals in the French healthcare system and whether it affects regional disparities in mortality.

2.2.1 Competition and ownership

In France, prices are regulated in both the public and private sectors as there is a unique public health insurance system. Hence, competition between hospitals can only be based on quality. The goal of FP hospitals is to make profit. They thus have incentives to attract patients needing expensive healthcare. This can be done by providing some services of better quality than other FP hospitals, public hospitals and NFP hospitals.

At a first glance, hospitals in the public sector do not have incentives to compete with each other or with FP hospitals by providing a better quality of care. In fact, the situation is more complex as there are some incentives for physicians in the public sector to improve their efficiency.[6] Indeed,

[6] This is made possible by the public sector allowing research activities and the use of high-tech equipment in large hospitals.

in the public sector, physicians are salaried civil servants and their wages do not depend on their performance. One day a week, though, physicians can work outside their hospital, in particular in a FP hospital where they are self-employed. Their income in FP hospitals is far higher than in public hospitals.[7] Moreover, physicians receive additional fees when performing innovative procedures.[8] In parallel, a patient choosing a hospital for treatment is attracted by the physicians who can provide the best care. As FP hospitals want to attract patients to make profit, they try to get the physicians with the best reputation. Hence, physicians in public and NFP hospitals have some incentives to compete with each other by providing a higher quality of care in order to increase their reputation and get hired part-time in FP hospitals. This is a specific form of competition based on quality.

2.2.2 Competition and innovation

Interestingly, the form of competition in public hospitals has an effect on medical practices. As physicians want to work in FP hospitals, they have an incentive to perform innovative procedures to increase their skills and reputation through learning-by-doing. Dormont and Milcent (2005) show that in public hospitals, the proportion of patients treated with innovative procedures is larger than expected when considering that the reimbursement rules do not create any incentive to perform innovative treatments.

There is another consequence of the relationship between competition and innovative treatments. Some hospital characteristics such as physicians' behaviours are not well captured by usual hospital variables, yet they may be correlated both with the performance of innovative treatments and mortality. In that case, when only the observed characteristics of hospitals are taken into account

[7] Physicians working in NFP hospitals are also salaried but their wages are far higher than in public hospitals.

[8] There is an abundant literature on the effect of payment rules on physicians' incentives (Hart and Holmstrom, 1987; Pauly, 1990; Blomqvist, 1991; Milgrom and Roberts, 1992; Newhouse, 1996; McGuire, 2000).

in the analysis, innovative treatments are not exogeneous. However, we deal with this issue by including hospital unobserved heterogeneity in our empirical specification.

Urban areas differ in their market structure as some of them concentrate patients in a few large hospitals, while others admit patients in many small ones.[9] The market structure has an effect on hospital efficiency because of the competition in quality between physicians. Indeed, when doctors work in a large hospital, they can increase their efficiency through learning-by-doing in the use of innovative procedures. In some papers in the literature, a Herfindahl index capturing the local concentration of patients in a few large hospitals rather than many small ones, is expected to have a positive effect on mortality because it is negatively correlated with competition between hospitals (Kessler and McClellan, 2000; Town and Vistnes, 2001; Gowrisankaran and Town, 2003). However, in the French context, the local concentration of patients may also be related to the concentration of skilled physicians in the same place. In that case, the effect of the Herfindahl index is expected to be negative. As long as the Herfindahl index is found to have a sizable effect (whatever the sign), regional disparities are related to variations in the value of the index across the territory.

2.3 The regional scale

Since the "Plan Juppé" in 1996, the French healthcare system has been decentralized and regional regulators have a significant influence in the organization of regional healthcare markets during our period of study. The government determines the national global budget every year and allocates it across regions. The regional budget is shared between NFP and public hospitals depending on the budget of the previous year, on bilateral bargaining between the regional regulator and the hospital managers, and on the local level of poverty. The regional regulators also decide the distribution of bed capacities for the public and private sectors in their region, and the level of hospital investment in the public sector. Investments in the private sector mostly depend on hospital managers. The

[9] This heterogeneity is due to historical reasons, regional planning and competition.

efficiency of regional regulators in organizing regional healthcare organization and the budget devoted by each hospital to the treatment of heart attack are not observed and will enter the hospital unobserved heterogeneity in our analysis.

Hospitals usually accept patients on the basis of their region of residence. However, patients are sometimes transferred to a neighbouring hospital in another region. Over the 1998-2003 period, a very high proportion of AMI patients (92.9%) were treated within their region of residence. This proportion is slightly lower for FP hospitals (91.4%) than for public hospitals (93.1%) and NFP (95.8%). These statistics support the assumption that regions can be viewed as local healthcare markets for heart attack. Our spatial variance analysis will thus be conducted at the regional level although we also assess the robustness of our results at the city level.

It is possible to check empirically that healthcare supply varies between regions. In 1999, 80% of beds were in public hospitals in the West and in Franche-Comté (in the East), versus only 46% in the PACA region (the south-eastern French Riviera). The proportion of NFP hospitals is highest in some Eastern regions at the German border (Alsace and Lorraine) for historical reasons. Conversely, the proportion of beds in FP hospitals is higher in the South-East (around the French Riviera region) where the population is older and richer. These regions are characterized by a substantial use of innovative procedures like stents. In fact, the rank correlation between the regional proportion of stents and the regional proportion of patients in FP hospitals is .61. When considering NFP hospitals instead of FP hospitals, the correlation is still quite high at .44.

2.4 Treatment of heart attack

We measure healthcare quality in a hospital by the death outcome of patients admitted for a serious disease. As the evidence shows that the effect of characteristics on mortality is disease-specific (Wray et al., 1997), we focus on one single disease, acute myocardial infarction also called heart attack. In line with the literature, we select all patients whose main pathology is coded I21

or I22 under the Tenth International Classification of Disease (ICD-10). These codes correspond only to emergency AMI admissions (WHO, 2009). Due to its clinical definition, AMI is a well-defined pathology with only a few re-admissions. Besides, AMI belongs to the ischemic-disease group which is one of the main causes of mortality in France. AMI mortality is an event frequent enough to yield reliable statistical results.

Mortality from AMI has been widely studied in the literature to assess the quality of hospital care in the US and the UK (see Goworisakaran and Town, 2003, for the US, and Propper et al., 2004, for the UK). Moreover, risk-adjusted AMI mortality is a commonly used measure of clinical performance. The same hospital features that lead to high healthcare quality for AMI are common to treatments for other conditions. These features include care coordination, speed of treatment and timely access to surgical interventions (Cooper et al., 2011).

Before age 35, heart attacks are often related to a cardiac disfunction. In line with the WHO standards, we limit our analysis to patients above age 35. Heart attacks occur when arteries or veins which irrigate the heart are clogged. In hospitals, patients can benefit from various treatments and procedures to improve the blood flow in clogged arteries. These include bypass surgery, cardiac catheters, percutaneous transluminal coronary angioplasty (PTCA) and stent.[10] A catheter is a thin flexible tube which is inserted in a vein. Bypass surgery involves taking a vein or a artery from the patient's body and using it to divert blood from the coronary arteries. In some

[10]Intravenous thrombolytic or clot-busting drugs that break down the blockage causing the heart attack are alternative treatments to restore blood flow. However, these drugs are used as both pre-hospital and in-hospital treatments. Our data only contain information on in-hospital treatments. In the 1990s, the MONICA investigators studied the effect of some drug therapies. They collected information on both pre-hospital and in-hospital treatments in order to compute a reliable index of clot-busting drug use (Tunstall-Pedoe et al., 1988 and 2000). As we do not have the necessary pre-hospital information to compute such an index, we do not include clot-busting drugs in our analysis. In any case, these drugs do not follow the same logic as surgical procedures as they are much cheaper.

cases, stent and angioplasty are alternative procedures to bypass which yield a better quality of life after home return. An angioplasty consists in inflating a balloon in a vessel to crush a blockage and create a channel. This procedure is costly as it raises the cost of a stay by 30% to 60% (Dormont and Milcent, 2002). The stent is a spring-shaped prosthesis used as a complement to angioplasty. The use of stent with an angioplasty significantly improves the results. Angioplasties and stents are innovative treatments that were first introduced over the period 1998-2003.

In this article, the term *stent* refers to an angioplasty together with one or more stents, the term *angioplasty* refers to an angioplasty without stent, and the term *catheterization* refers to a catheterization without angioplasty or stent.

3 The dataset

3.1 Data sources on patients, hospitals and areas

We use the PMSI dataset (*Programme de Médicalisation des Systèmes d'Information*) which provides records of all patients discharged from all French acute-care hospitals over the 1998-2003 period. It is compulsory for hospitals to provide these records on a yearly basis.[11] Three nice features of this dataset are that it provides some information at the patient level, it keeps track of hospitals across time, and it is exhaustive for both the public and private sectors.[12]

The dataset contains information on the demographic characteristics of patients, as well as very detailed information on diagnoses, comorbidities and treatments. A limit of the data is that the day on which treatments are performed during stays is not recorded. Likewise, certain comorbidity factors are not recorded, and this may seen as a drawback. However, McClellan and Staiger (1999)

[11] With the exception of local hospitals, for which it is not compulsory. This does not affect our study since these hospitals do not treat AMI patients.

[12] It should be mentioned however that only 90% of the private sector was covered in 1998 and 95% in 1999.

show that more detailed medical data on disease severity and comorbidity do not add much when taking patient heterogeneity into account. The dataset also provides us with the type of entry (whether the patients come from their place of residence, another care service in the same hospital or another hospital) as well as the type of exit (death, home return, transfer to another hospital or transfer to another care service). A limit of the data is that patients cannot be followed across time if they come back later to the same hospital or if they change hospital. As we cannot keep track of patients when they are transferred to another hospital, we restrict our sample to patients who come from their place of residence. After deleting observations with missing values for the variables used in our study (very few in number), the data contain 341,861 stays for patients distributed across 1,105 hospitals.

We match our dataset with the hospital records from the SAE survey (*Statistiques Annuelles des Etablissements de santé*) that was conducted every year over the 1998-2003 period. The SAE survey contains information on the municipality where each hospital is located, the number of beds (in surgery and in total) and the number of days that beds are occupied (in surgery and in total). The matching rate is very good and covers 97% of the patients.

The municipality code in the SAE survey also allowed us to match our dataset with wealth variables at the municipality level taken from other sources: the municipal unemployment rate computed from the 1999 population census, the median household income from the 2000 Income Tax dataset and the existence of a disadvantaged area in the municipality (disadvantaged areas being defined by a 1997 law under the label *zones urbaines sensibles*). These variables are intended to control for the sorting of patients across hospitals according to their financial constraints and are also used as proxies for the spatial differences in the funding of public and NFP hospitals based on the local level of poverty.

With the municipality code, we identify the urban area in which hospitals are located.[13] A

[13] An urban area (*aire urbaine*) is defined as an urban centre (which includes more than 5,000 jobs) and the

Herfindahl index of concentration is computed using the number of patients in hospitals within each urban area.[14] As in our empirical specification, hospital variables need to be time-invariant (see Section 4), all hospital and geographic variables are averaged across years.

3.2 Preliminary statistics

As we are interested in disparities across regions, we compute the average probability of death in every region (see Appendix A1 for details). Graph 1 displays the probability of death within 15 days for all regions. This probability is quite low in the Paris region, the East and South-East. It is higher in the West and South-West. Graph 2 represents the probability of death as a function of the length of stay for *i)* the region where the chances of dying are the lowest (Alsace: 8.51% after 15 days), *ii)* the region where the chances of dying are the highest (Languedoc-Roussillon: 15.31%), and *iii)* the most densely populated region (Ile-de-France which is the Paris region: 9.93%). The three death profiles are significantly different. Overall, descriptive statistics suggest some significant regional disparities in mortality.

[*Insert Graphs* 1 *and* 2]

We also report in Table 1 some indices of regional disparities (max/min ratio, Gini index and coefficient of variation) for the probability of death within 1, 5, 10 and 15 days. Global indices like the Gini index (.07) and the coefficient of variation (.22) remain quite small and suggest that

municipalities in its catchment area. There are 359 urban areas in mainland France and they do not cover the whole territory (as some municipalities are excluded and remain rural).

[14] The Herfindahl index for an urban area u is $H_u = \sum_{j \in u} \left(\frac{N^j}{\underline{N}^u} \right)^2$ where j indicates the hospitals, N^j is the number of patients in hospital j, and $\underline{N}^u = \sum_{j \in u} N^j$ is the total number of patients within the urban area u. H_u increases from $\frac{1}{n_u}$ to 1 as the concentration of patients increases, where n_u is the number of hospitals in the urban area u. When $H_u = \frac{1}{n_u}$, the patients are equi-distributed between the n_u hospitals. When $H_u = 1$, they are all treated within one hospital.

disparities are not systematic. However, the max/min ratio shows that regional disparities between extremes are large. Indeed, the difference in the probability of death within 15 days between the maximum (Languedoc-Roussillon) and the minimum (Alsace) is 80%. Interestingly, disparities are a bit larger for the probability of death within 1 day (Max/Min ratio of 94%). This may be due to different behaviours across regions in terms of transfers and home returns in the early days of AMI stays.

[*Insert Table* 1]

We also compute indices of regional disparities for factors potentially related to regional disparities in mortality, whether these factors are measured at the patient, the hospital, the municipality or the urban area level. We focus on Gini indices which are global measures of disparities contrary to the Max/Min which contrasts extremes. Alternatively, we could also comment the results obtained with the coefficients of variation which are similar. We consider in the sequel that disparities are small when the Gini index is below .1, they are moderate for an index from .1 to .2, they are large for an index from .2 to .3, and they are very large for an index above .3. Disparity indices are reported in Tables 1 and 2.

Interestingly, disparity indices computed from the regional averages of demographic individual variables (case-mix) are quite small, as shown in Table 1. The case-mix is usually found to have a strong effect on mortality, but as variations in the local composition of patients across the territory are small, the case-mix does not necessarily play an important role. There are some moderate regional disparities for comorbidities and secondary diagnoses, the Gini index reaching .18 for obesity, .17 for excessive smoking, and .16 for a history of vascular diseases.

Regional disparities in the use of procedures are much larger. Widespread procedures like angioplasty and stent have a large Gini index, which takes the values .31 and .21, respectively.

Table 2 shows that there are large disparities across regions in the average size of hospitals where patients are treated, whether the hospital size is measured with the total number of patients (Gini:

.17) or the number of AMI patients (.20). Disparities are even larger for the number of beds (.35) and the number of beds in surgery (.34). These disparities indicate some sorting of hospitals by size. Finally, disparities are smaller, but still large, for the hospital status and more specifically for being a FP hospital (.27).

Regional disparities are moderate for geographic variables such as the presence of a disadvantaged area in the municipality (.14) and the Herfindhal index of concentration computed at the urban area level (.14).

[*Insert Table* 2]

In summary, the case-mix, innovative procedures, hospital characteristics and local healthcare market structure can all influence regional disparities in mortality. We now present our empirical methodology to assess their respective roles.

4 Econometric method

We briefly describe the econometric model explaining the propensity to die with patients' characteristics, treatments, hospital characteristics and spatial variables. We also explain how the specification can be aggregated at the regional level to perform a spatial variance analysis. Technical details on the model and the estimation method are given in Appendices A2 and A3.

4.1 Specification of the propensity to die

4.1.1 Model

The naive approach to study the determinants of death would be to regress an indicator of mortality over a given period, for instance 15 days, on some explanatory variables potentially related to mortality. However, in our data, patients are not followed when they are discharged before 15

days, so we do not know whether they die after leaving hospital. We could consider that there is a risk of death only while patients are in the hospital where they were admitted and compute the indicator of mortality during the truncated duration of stay in that hospital. This approach is likely to yield biased results as it does not capture the specific behaviour of some hospitals which may discharge certain patients (in particular by transferring them to another hospital) to limit the number of longer stays and deaths. Put differently, this approach does not take into account the correlation between the in-hospital mortality rate and the length of stay. We therefore opt for an alternative strategy that is flexible enough to incorporate hospital unobserved heterogeneity that might be correlated with patients' characteristics affecting the propensity to die and that takes into account the possibility of a discharge.

We consider a Cox duration model at the patient level stratified by hospital. The hazard function of a patient i in a hospital $j(i)$ is:

$$\lambda\left(t\left|X_i, j\left(i\right)\right.\right) = \theta_{j(i)}\left(t\right)\exp\left(X_i\beta\right) \tag{1}$$

where $\theta_j(\cdot)$ is the instantaneous hazard function for hospital j, X_i are the patient-specific explanatory variables (demographic shifters, comorbidities, diagnoses and treatments) and β are their effect on death. Each hospital has its own hazard function which is left unspecified and can thus capture some behaviour specificities and be correlated with individual explanatory variables. With this specification, *i)* spatial effects related to hospital locations are captured by the hospital hazard functions and *ii)* correlations between spatial factors and individual observable characteristics are allowed through the correlation between hospital hazard functions and patients' characteristics.

The coefficients of patients' variables can be estimated by maximizing the stratified partial likelihood (see Ridder and Tunali, 1999; Gobillon et al., 2011). The contribution to likelihood of a patient i who dies after a duration t_i is his probability of dying conditionally on someone at risk

in his hospital dying after this duration. It is given by:

$$P_i = \frac{\exp\left(X_i\beta\right)}{\sum\limits_{i\in\Omega_{j(i)}(t_i)} \exp\left(X_i\beta\right)} \tag{2}$$

where $\Omega_j\left(t\right)$ is the set of patients at risk at day t in hospital j, i.e. the set of patients that are still in hospital j after staying there for t days. This procedure is similar to estimating a linear panel data model projected in the within dimension: a within transformation makes the hospital heterogeneity disappear. This means that the coefficients of patients' variables are identified thanks to the variations in the characteristics of patients at risk and the variation in their propensity to die within each hospital for each duration. The hazard function of each hospital can be recovered in a second step from the estimated coefficients of patients' variables.

We also want to characterize the hospital hazard functions in a descriptive way using hospital and geographic variables. We first consider that each hospital hazard function can be decomposed multiplicatively into a hospital fixed effect and a baseline hazard function. The variation of hospital fixed effects across regions will be used to characterize the importance of hospital and geographic effects in explaining the spatial disparities in mortality without resorting to specific indicators that may not be exhaustive enough to fully capture differences between regions. The estimated hospital fixed effects are then regressed on hospital and geographic variables to assess the importance of hospital composition and spatial location. More formally, we specify the hazard function of each hospital j multiplicatively as:

$$\theta_j\left(t\right) = \theta\left(t\right)\alpha_j \tag{3}$$

where $\theta\left(\cdot\right)$ is a baseline hazard function common to all hospitals and α_j is a hospital effect capturing the hospital specificities which is itself specified as $\alpha_j = \exp\left(Z_j\gamma + \eta_j\right)$ with Z_j some hospital variables (status, size, occupation rate, etc.) and some geographic variables (concentration index, municipality income variables, regional fixed effects, etc.), and η_j a random error capturing the hospital and geographic unobserved heterogeneity.

The baseline hazard and hospital fixed effects are estimated using moment conditions exploiting the multiplicative structure of hospital hazard functions given by (3). In particular, the estimated hospital fixed effects capture the average behaviour of hospitals over the first 30 days of stay.[15] The coefficients of hospital and geographic variables are estimated from the specification of hospital fixed effects in logarithm $\ln \alpha_j$, where the true value of hospital fixed effects has been replaced by its estimator $\ln \widehat{\alpha}_j$:

$$\ln \widehat{\alpha}_j = Z_j \gamma + \eta_j + \phi_j \tag{4}$$

where $\phi_j = \ln \widehat{\alpha}_j - \ln \alpha_j$ is the sampling error on the hospital fixed effects which comes from the replacement of their true value by their estimator. Equation (4) can be estimated using weighted least squares where the weight is the number of patients in the hospital.[16] The sampling error on the estimated hospital fixed effects is taken into account when computing the standard errors and the R-square (see Gobillon et al., 2011, for more details on the computations). Note that this specification takes into account unobserved heterogeneity at the hospital and geographic level through the term η_j and the standard errors are robust to aggregate shocks which often create large biases on the estimated standard errors when they are not taken into account (see Moulton, 1990).

4.1.2 Discussion

It would be tempting to estimate directly the effect of all patient, hospital and geographic variables, introducing all of them directly in a simple Cox duration model where there is a baseline hazard

[15] The estimators of hospital fixed effects are biased but consistent when the number of patients in each hospital tends to infinity. The median number of patients per hospital is 84.

[16] Note that for a given hospital, equation (4) is well defined only when there is at least one patient who dies in the hospital over the $1998-2003$ period (otherwise the quantity $\widehat{\alpha}_j$ from which we take the log would be zero). This condition may not be verified for hospitals that have only a few patients. In fact, these hospitals have a negligible weight in the sample and they are discarded.

function common to all hospitals and thus no stratification by hospital. There are, however, four limits to that approach: (i) the estimated coefficients of patients' variables may be biased because the correlation between patients' variables and hospital/geographic unobservables is not taken into acccount; (ii) the estimated coefficients of patients' variables may be biased because the specific discharge behaviour of each hospital is not taken into account; (iii) the estimated standard errors may be biased because the hospital/geographic unobservables are not taken into account in their computation; (iv) the overall importance of hospital and geographic factors cannot be assessed by computing the spatial variations of hospital fixed effects as they are not introduced in the specification. Our approach is robust to these four issues.

It would also be tempting to use an alternative two-stage approach where a Cox duration model with patients' variables and hospital fixed effects is first estimated, and the resulting estimated hospital fixed effects are then regressed on hospital and geographic variables (taking into account the hospital/geographic unobservables in the computation of standard errors). However, this approach is not as flexible as ours because hospital-specific hazard functions and not only hospital fixed effects must be used to properly take into account the hospital-specific discharge behaviour. Moreover, it is unfeasible in practice to estimate such a model with current computer resources because it requires maximizing a likelihood with more than a thousand parameters. Convergence is not granted as some hospitals only admit a very few patients for an AMI which makes it difficult to estime their hospital fixed effects. Finally, even if it were possible to maximize the likelihood, the estimated parameters would be biased because of the incidental parameter problem arising from the small number of patients in some hospitals (Lancaster, 1990). Our approach is robust to these issues and the estimation procedure used to recover the coefficients of patients' characteristics and the hospital fixed effects is quite fast.

4.2 Spatial variance analysis

We then average the model at the regional level to perform a spatial variance analysis of mortality. Taking the logarithm of the hazard function of a patient i in a hospital $j(i)$ given by (1) under the multiplicative assumption (3), and computing the average for any region r gives:

$$\frac{1}{\underline{N}^r} \sum_{i|j(i)\in r} \ln \lambda \left(t \left| X_i, j\left(i\right)\right.\right) = \overline{X}^r \beta + \overline{\ln \alpha}^r + \theta\left(t\right) \tag{5}$$

where $\underline{N}^r$ is the number of patients in region r, $\overline{X}^r$ is the regional average of individual explanatory variables and $\overline{\ln \alpha}^r$ is the regional average of hospital fixed effects weighted by the number of patients in the hospitals. This equation states how, at the regional level, the average hazard at t days for patients entering a hospital for an AMI relates to their average characteristics, the average hospital and geographic effects, and the baseline hazard at t days. We qualitatively assess the strength of these relationships from the variance of right-hand side terms in (5) (in a manner similar to Abowd, Kramarz and Margolis, 1999). In fact, the larger the variance, the stronger the relationship. In practice, as β and $\overline{\ln \alpha}^r$ are not observed, they are replaced by their estimators $\widehat{\beta}$ and $\widehat{\overline{\ln \alpha}}^r$ (the latter being defined as the regional weighted average of $\ln \widehat{\alpha}_j$). An estimator of the left-hand side term is obtained from the sum of right-hand side terms. Using the same approach, we also assess the strength of the relationship between the regional average hazard rate and $\overline{X}^r_s \widehat{\beta}$ for some sub-groups $\overline{X}^r_s$ of explanatory variables. Importantly, this procedure measures the strength of relationships *ex ante* before any filtering of patients through transfers or home returns. We can further assess the strength of relationships for hospital and geographic variables. Taking the log of the expression of hospital fixed effects and averaging at the regional level, we get:

$$\overline{\ln \alpha}^r = \overline{Z}^r \gamma + \overline{\eta}^r \tag{6}$$

where $\overline{Z}^r$ and $\overline{\eta}^r$ are the regional averages of explanatory variables and unobserved terms, respectively. We can assess the strength of relationships between the regional average hazard rate and

$\overline{Z}^r\gamma$ or $\overline{Z}^r_s\gamma$, for some sub-groups $\overline{Z}^r_s$ of explanatory variables, in the same way as for individual variables (replacing γ by its estimator). Here again, for any effect, we can compute the variance to get an idea of the strength of relationships.

5 Results

Table 3 reports the estimation results of the first-stage equation (1).[17] Demographic characteristics have the expected effect on mortality. The propensity to die increases with age and females are more likely to die than males. This is consistent with procedures being better adapted to males than to females (Milcent et al., 2007).

Some secondary diagnoses and comorbidities have a positive effect as they adversely affect health: renal failure, stroke, heart failure, conduction disease and alcohol. Others have a negative effect: diabetes, obesity, excessive smoking, vascular disease, peripheral arterial disease, previous coronary artery disease and hypertension. An interpretation of this result is that patients exhibiting these diagnoses and comorbidities are monitored more carefully before and after having a heart attack (Milcent, 2005).

All the treatments (bypasses, catheters, PTCA, other dilatations and stents) have the expected

[17] Estimated coefficients obtained with our approach for innovative procedures differ from those obtained using a Cox model without stratification. In particular, using a Cox model we obtain without stratification: -1.139 for catheter, -.481 for angioplasty and -.797 for stent. Estimated coefficients obtained for demographic variables, co-morbidities and secondary diagnoses when using Cox models with and without stratification are quite close.

negative effect on the propensity to die and the effect is large.[18]

[*Insert Table* 3]

We also assess to what extent regional disparities in mortality remain after controlling for the effect of patient-specific variables. Regions at the extremes are the same as when studying the raw data: Alsace (at the German border) exhibits the lowest probability of death and Languedoc-Roussillon (in the South-East) the highest. Graph 3 represents the probability of death as a function of the duration of stay for these two extreme regions and the Paris region. The difference between the extreme regions is smaller but still significant.

[*Insert Graph* 3]

Corresponding regional disparity indices are reported in Table 4. In particular, the difference in regional probability of death within 15 days between the extreme regions has decreased from 80% to 47% (this corresponds to a 41% decrease). More systematic disparity indices like the coefficient of variation and the Gini index also decrease, but to a lesser extent (by 19% and 17%, respectively). Overall, results suggest that some regional disparities remain and that hospital and geographic factors have a role to play in these disparities.

[*Insert Table* 4]

We then decompose multiplicatively the hospital hazard functions into some hospital fixed effects and a baseline hazard function common to all hospitals, and regress the logarithm of hospital fixed effects on a set of hospital and geographic variables. Regression results are reported in Table 5

[18] A limit of our approach is that we cannot take into account the date at which treatments are performed and we have to assume that patients experience the effect of their treatment during their entire stay. This assumption may create a bias in the results. However, as surgical treatments usually take place during the first days of stays, our assumption thus seems reasonable.

for three specifications including successively the hospital variables, the geographic variables, and both. The adjusted-R^2 is higher, but only moderately so, when both sets of variables are included in the specification rather than only one set of variables. This suggests that hospital variables and geographic variables are quite correlated but each of these two sets of variables plays a significant role.

We now comment on the sign of the estimated coefficients for the full specification including both hospital and geographic variables (Column 3). As regards the effect of hospital characteristics, we find that the propensity to die is nearly the same in FP hospitals and public hospitals. This result may look surprising but it comes from the fact that we control for the use of innovative treatments (mainly angioplasty and stent) as shown by Gobillon and Milcent (2012). If we remove the variables related to innovative treatments from the first-stage specification, the propensity to die becomes higher in public hospitals than in FP hospitals. Hence, the higher efficiency of FP hospitals may be attributable to wider use of innovative treatments. We also find that the propensity to die is lower in a NFP hospital than in a public or a FP hospital.

Besides, three results on hospital variables suggest the existence of either some learning-by-doing or some sorting of heart surgeons across hospitals according to their performance. Indeed, *i)* the proportion of patients in the hospital treated for an AMI has a negative and significant effect on mortality. It is thus likely that hospitals concentrating AMI patients have specialized in cardiac pathologies and perform better. *ii)* The propensity to die is lower in hospitals where there is a higher proportion of beds in surgery, and it is plausible that these hospitals have more highly qualified staff. *iii)* The propensity to die decreases with the occupation rate of beds in surgery (significantly at the 10% level). It is possible that hospitals attracting more patients per bed capacity are those attracting the most skilled surgeons or are those where surgeons simply end up doing more learning-by-doing, which brings down mortality.

Focusing on geographical variables, the municipality indicators of wealth do not have any sizable

effect. Still, the presence of a disadvantaged area in the municipality has a positive effect on mortality significant at the 10% level. This suggests the existence of missing variables related to income that may affect mortality.

The Herfindahl index which measures the local concentration of patients in a few large hospitals rather than many small ones has a significant negative effect. This result can be explained once again by learning-by-doing or the sorting of efficient surgeons in large hospitals.

[*Insert Table* 5]

As shown in Table 6, regional dummies have a significant negative effect compared to the reference (Languedoc-Roussillon). Remaining regional disparities in mortality can be explained by regional differences in hospital budgets, in the use of thrombolytic drugs in the pre-hospital or/and in-hospital stays, or in the propensity to transfer patients when they are more likely to die. Whereas hospital budgets and transfers are related to the regional organization of healthcare, spatial differences in drug use mostly depend on practices learnt in medical school .

[*Insert Table* 6]

We finally report in Table 7 the results of our spatial variance analysis. Patients' variables are more strongly related than hospital fixed effects to regional disparities in mortality, as shown by the respective variances of their effects.[19] Regional disparities in innovative treatments play a major role as the effect of innovative treatments has a far larger variance than the effect of the usual demographic determinants (age and sex) of mortality.

The role of hospital variables is very limited as the effects of hospital variables have a rather small variance. In particular, this is the case for ownership status. This is consistent with the

[19]Note that the sum of variances is larger than 100%. This occurs because the effect of individual variables and log-hospital fixed effects are slightly negatively correlated. The correlations between the effects of variables, or groups of variables, is available upon request.

coefficients obtained when regressing hospital fixed effects on ownership dummies as the effect of being treated in a public hospital rather than in a for-profit hospital is not significant. This does not mean, however, that the sorting of for-profit hospitals across space is not related to spatial differences in mortality as for-profit hospitals more often perform innovative procedures which decrease mortality.

Interestingly, among geographic variables, the Herfindahl index plays a significant role as the variance of its effect is quite large. Hence, spatial disparities in the concentration of patients in a few large hospitals rather than many small ones are related to spatial disparities in mortality. By contrast, the effect of municipality variables related to local wealth has a negligible variance.

Finally, regional fixed effects, which capture all unobserved regional effects and the regional average of unobserved hospital effects, have a large variance which is comparable with that of the effect of innovative treatments. This means that some unobserved factors have not been taken into account in the analysis. We believe that regional disparities in budgets and local healthcare organization are some plausible candidates.

[*Insert Table* 7]

6 Robustness checks

In our analysis, we conduct the spatial variance analysis at the regional level but regions are administrative units and cities usually have a more functional role in economic activities. As a robustness check, we replicate our spatial variance analysis at the city level. Table 8 shows that our main results are robust to considering cities instead of regions. Indeed, the use of innovative procedures is still the main factor related to spatial disparities in mortality. Demographic factors (age and sex) once again are less strongly related to these disparities than innovative procedures. Spatial differences in hospital and geographic factors seem to play a larger role than at the re-

gional level as shown by the variance of hospital fixed effects. However, the effects of hospital and geographic explanatory variables have a negligible variance except, once again, the Herfindahl index.

[*Insert Table* 8]

We also conducted our analysis using hospital fixed effects in aggregate regressions. These hospital fixed effects are summaries of hospital and geographic effects over a duration of thirty days. Alternatively, one can regress hospital integrated hazards for a given duration t to study mortality at t days. We run these regressions for the "short-run" and "long-run" durations of 5 days and 15 days. Results of the spatial variance analysis are reported in Table 9.[20] While results are similar for patient variables, there are some slight variations for hospital and geographic explanatory variables. Importantly, our results again show that the Herfindahl index plays a significant role and the hospital status plays no role.

[*Insert Table* 9]

7 Conclusion

In this paper, we explore the factors related to regional disparities in mortality for patients admitted to hospital for a heart attack. Our empirical study uses a unique matched patients-hospitals dataset constructed from exhaustive administrative records for the 1998-2003 period.

We show that regional disparities are fairly large. The difference in mortality rate between the extreme regions reaches 80%. We determine the factors related to mortality using a Cox duration model stratified by hospital which allows a flexible modelling of the hospital heterogeneity. The estimated specification is averaged at the regional level and a spatial variance analysis is conducted to determine which factors are most related to spatial disparities in mortality.

[20] Results on the effect of patient, hospital and geographic variables are available upon request.

Our results show that regional variations in the use of innovative treatments play a major role. Interestingly, the regional composition of hospital ownership has a negligible role. This finding is due to the fact that, when holding constant the use of innovative procedures, the propensity to die is similar in for-profit hospitals and public hospitals. As a result, the spatial sorting of for-profit hospitals is related to spatial disparities in mortality only through the greater use of innovative procedures in for-profit hospitals than in public hospitals. In addition, mortality is lower in hospitals where the proportion of patients treated for a heart attack is higher, suggesting the existence of learning-by-doing. Also, the higher the local concentration of patients in a few large hospitals rather than many small ones, the lower the mortality. Finally, regional unobserved effects account for around 20% of spatial disparities.

A limit of our analysis is that patients are not tracked in the data when they are transferred to another hospital. For patients who are transferred, we considered that the length of stay is censored. An interesting extension of our work would be to study how hospitals interact through transfers and to what extent the transfer of patients to another hospital affects their propensity to survive. Space may play a major role in transfers as some hospitals are isolated and others are close to an establishment specialized in heart surgery.

8 Appendix: technical aspects of our approach

8.1 Appendix A1: regional probability of death

For each hospital, we compute a gross survival function for exit to death using the Kaplan-Meier estimator, other exits (home returns and transfers) being treated as censored. For a given region, we then recover a survival function weighting the survival functions of hospitals in that region by the number of patients still at risk in these hospitals.

We could have directly computed a survival function for each region. However, we believe that the

relevant unit of patient treatment is the hospital. Also, our approach at the hospital level parallels our model presented in the next two sub-sections.

When the length of stay increases, the number of patients in a given hospital decreases. Above a given length of stay, there are no longer any patients at risk and the hospital is not included in the computation of the survival function. Hence, there is a selection of hospitals as the length of stay increases. We limited our analysis to lengths of stay below fifteen days to avoid ending up with too few observations in most hospitals of a given region.

The probability of death in a region for any given length of stay is one minus the survival function of that region.

8.2 Appendix A2: model

For each patient, we observe the length of stay in the hospital and the type of exit (death, home return or transfer). In the sequel, we only study exit to death. All other exits are treated as censored. We specify the hazard function of a patient i in a hospital $j(i)$ as:

$$\lambda\left(t \left| X_i, j\left(i\right)\right.\right) = \theta_{j(i)}\left(t\right) \exp\left(X_i \beta\right) \tag{7}$$

where $\theta_j(\cdot)$ is the instantaneous hazard function for hospital j, X_i are the patient-specific explanatory variables and β are their effect on death. The model is estimated maximizing the stratified partial likelihood. The contribution to likelihood of a patient i who dies after a duration t_i is his probability of dying conditionally on someone at risk in the same hospital dying after this duration. It can be written as:

$$P_i = \frac{\exp\left(X_i \beta\right)}{\sum\limits_{i \in \Omega_{j(i)}(t_i)} \exp\left(X_i \beta\right)} \tag{8}$$

where $\Omega_j(t)$ is the set of patients at risk at day t in hospital j, i.e. the set of patients that are still in hospital j after staying there for t days. The partial likelihood to be maximized can then be

written as $L = \prod_i P_i$. Denote $\widehat{\beta}$ the estimated coefficients of patient-specific explanatory variables. It is possible to compute the integrated hazard function $\Theta_j(t)$ of any hospital j using the estimator proposed by Breslow (1974). It can be written as:

$$\widehat{\Theta}_j(t) = \int_0^t \frac{I(N_j(s) > 0)}{\sum_{i \in \Omega_j(s)} \exp\left(X_i \widehat{\beta}\right)} dN_j(s) \tag{9}$$

where $I(\cdot)$ is the indicator function, $N_j(s) = card\ \Omega_j(s)$, and $dN_j(s)$ is the number of patients exiting from hospital j between the days s and $s+1$. From the Breslow's estimator, we compute a survival function for each hospital j as $\exp(-\widehat{\Theta}_j(t))$ (an estimator of its standard error is recovered using the delta method). The hospital survival functions are averaged at the regional level weighting by the number of patients at risk to give the regional survival functions. The regional probability of death is computed as one minus the regional survival function. As the hospital hazards are left completely unspecified, the study of regional disparities in death using regional probabilities of death remains very general.

We then study the factors related to hospital disparities by specifying the hospital hazard rates in a multiplicative way:

$$\theta_j(t) = \alpha_j \theta(t) \tag{10}$$

where α_j is a hospital fixed effect and $\theta(t)$ is a baseline hazard common to all hospitals. We show in appendix A3 how to estimate the parameters using empirical moments derived from (3).[21] Note that we need an identifying restriction since α_j and $\theta(t)$ can be identified separately only up to a

[21] In doing so, we depart from the log-linear estimation method proposed by Gobillon et al. (2011). Our approach is better suited when exits are rare, as in our case. Indeed, Gobillon, Magnac and Selod split the timeline into K intervals denoted $[t_{k-1}, t_k]$. Introduce $\theta_k = \int_{t_{k-1}}^{t_k} \theta(t)\, dt / (t_k - t_{k-1})$ and $y_{jk} = [\Theta_j(t_k) - \Theta_j(t_{k-1})] / (t_k - t_{k-1})$. Integrating (3) over each interval and taking the log, they get: $\ln y_{jk} = \ln \alpha_j + \ln \theta_k$. y_{jk} is not observed but can be replaced by a consistent estimator: $\widehat{y}_{jk} = \left[\widehat{\Theta}_j(t_k) - \widehat{\Theta}_j(t_{k-1})\right] / d_k^j$ where d_k^j is the amount of time in interval $[t_{k-1}, t_k]$ where at least one patient is at risk. The equation to estimate is then: $\ln \widehat{y}_{jk} = \ln \alpha_j + \ln \theta_k + \psi_{jk}$ where

multiplicative constant. We impose for convenience that: $\frac{1}{N}\sum_t N_t\theta(t) = 1$ where N_t is the number of patients still at risk at the beginning of day t and $N = \sum_t N_t$. After some calculations, we get:

$$\theta(t) = \left(\frac{1}{N^2}\sum_{j,t} N^j N_t \theta_j(t)\right)^{-1}\left(\frac{1}{N}\sum_j N^j\theta_j(t)\right) \quad (11)$$

$$\alpha_j = \left(\frac{1}{N^j}\sum_t N_{jt}\theta(t)\right)^{-1}\left(\frac{1}{N^j}\sum_t N_{jt}\theta_j(t)\right) \quad (12)$$

where N_{jt} is the number of patients at risk at time t in hospital j, $N^j = \sum_t N_{jt}$, and the sum on t, $\sum_t$, goes from $t = 1$ to $t = T$ (here, we fixed $T = 30$ as most patients have already left their hospital after that duration). An estimator of $\theta(t)$ denoted $\widehat{\theta}(t)$ can be obtained, replacing $\theta_j(t)$ by the estimator $\widehat{\theta}_j(t) = \widehat{\Theta}_j(t) - \widehat{\Theta}_j(t-1)$ on the right-hand side of equation (11). An estimator of α_j denoted $\widehat{\alpha}_j$ can then be derived, replacing $\theta_j(t)$ and $\theta(t)$, respectively by $\widehat{\theta}_j(t)$and $\widehat{\theta}(t)$, on the right-hand side of equation (12). We show in appendix A3 how to compute the covariance matrices of $\widehat{\theta} = \left(\widehat{\theta}(1), ..., \widehat{\theta}(T)\right)'$ and $\widehat{\alpha} = (\widehat{\alpha}_1, ..., \widehat{\alpha}_J)'$. We then explain the hospital fixed effects with some hospital and geographic variables denoted Z_j. We specify: $\alpha_j = \exp\left(Z_j\gamma + \eta_j\right)$ where γ are the effects of hospital and geographic variables on death, and η_j includes some unobserved hospital and geographic effects. For a given hospital j, taking the log and replacing the hospital fixed effect with its estimator, we get:

$$\ln\widehat{\alpha}_j = Z_j\gamma + \eta_j + \phi_j \quad (13)$$

$\psi_{jk} = \ln\widehat{y}_{jk} - \ln y_{jk}$ is the sampling error. This equation can be estimated with standard linear panel methods. The authors use weighted least squares where the weights are the number of individuals at risk at the beginning of the interval. A limit of this method is that $\ln y_{jk}$ can be replaced by its estimator $\ln\widehat{y}_{jk}$ only if $\widehat{y}_{jk} \neq 0$. When such is not the case, observations should be discarded from the sample. When implementing this approach in our case, this could be an issue as exits are rare and a significant number of observations would have to be discarded when the time spent in the hospitals becomes long. In practice however, the results obtained with the two approaches are quite similar.

where $\phi_j = \ln \widehat{\alpha}_j - \ln \alpha_j$ is the sampling error on the hospital fixed effect. Equation (4) can be estimated using weighted least squares where the weight is the number of patients in the hospital. The standard errors and R-square (adjusted to take into account the sampling error), are computed as proposed by Gobillon et al. (2011).

8.3 Appendix A3: second-stage estimation

In this appendix, we explain how to construct some estimators of the baseline hazard and hospital fixed effects. We first average equation (3) across time, weighting the observations by the number of patients at risk at each date. We obtain:

$$\frac{1}{N}\sum_t N_t \theta_j(t) = \alpha_j \frac{1}{N}\sum_t N_t \theta(t) \tag{14}$$

where N_t is the number of patients at risk at the beginning of period t, $N = \sum_t N_t$ with $\sum_t$ the sum from 1 to T days (with $T = 30$ in the application). A natural identifying restriction is that the average of instantaneous hazards equals one: $\frac{1}{N}\sum_t N_t \theta(t) = 1$. We obtain:

$$\alpha_j = \frac{1}{N}\sum_t N_t \theta_j(t) \tag{15}$$

An estimator of hospital fixed effects could be constructed from this formula, but weights (namely: N_t) are not hospital-specific and thus do not reflect hospital specificities. Hence, we propose another estimator of hospital fixed effects in the sequel which we believe better captures hospital specificities.

We also average equation (3) across hospitals, weighting by the number of patients at risk (summed across all dates) in each hospital. We get:

$$\frac{1}{N}\sum_j N^j \theta_j(t) = \frac{1}{N}\left(\sum_j N^j \alpha_j\right)\theta(t) \tag{16}$$

where $N^j = \sum_t N_{jt}$ with N_{jt} the number of patients at risk in hospital j at the beginning of date t (such that $N = \sum_j N^j$). Replacing α_j with its expression (15), we obtain: $\theta(t) =$

$\left(\frac{1}{N^2}\sum_{j,t}N^jN_t\theta_j(t)\right)^{-1}\left(\frac{1}{N}\sum_{j}N^j\theta_j(t)\right)$. An estimator of the hazard rate at date t in hospital j can be constructed from Breslow's estimator such that $\widehat{\theta}_j(t)=\widehat{\Theta}_j(t)-\widehat{\Theta}_j(t-1)$. A natural estimator of the baseline hazard is then:

$$\widehat{\theta}(t)=\left(\frac{1}{N^2}\sum_{j,t}N^jN_t\widehat{\theta}_j(t)\right)^{-1}\left(\frac{1}{N}\sum_{j}N^j\widehat{\theta}_j(t)\right) \tag{17}$$

We then construct an estimator of a given hospital fixed effect α_j averaging equation (3) across time for this hospital and weighting by the number of patients at risk at the beginning of each day in this hospital. We obtain:

$$\frac{1}{N^j}\sum_{t}N_{jt}\theta_j(t)=\alpha_j\frac{1}{N^j}\sum_{t}N_{jt}\theta(t) \tag{18}$$

An estimator of the hospital fixed effect is then:

$$\widehat{\alpha}_j=\left(\frac{1}{N^j}\sum_{t}N_{jt}\widehat{\theta}(t)\right)^{-1}\left(\frac{1}{N^j}\sum_{t}N_{jt}\widehat{\theta}_j(t)\right) \tag{19}$$

We also computed the asymptotic variances of $\widehat{\theta}=\left(\widehat{\theta}(1),...,\widehat{\theta}(T)\right)'$ and $\widehat{\alpha}=(\widehat{\alpha}_1,...,\widehat{\alpha}_J)'$, denoted V_θ et V_α, with the delta method. Indeed, the covariance matrix of $\widehat{\theta}_J=\left(\widehat{\theta}_1(1),...,\widehat{\theta}_J(T)\right)'$ can be estimated from Ridder et Tunali (1999). Its estimator is denoted $\widehat{V}_{\theta_J}$. We can then compute the estimators: $\widehat{V}_\theta=\left(\frac{\partial\widehat{\theta}}{\partial\widehat{\theta}_J'}\right)\widehat{V}_{\theta_J}\left(\frac{\partial\widehat{\theta}'}{\partial\widehat{\theta}_J}\right)$ and $\widehat{V}_\alpha=\left(\frac{\partial\widehat{\alpha}}{\partial\widehat{\theta}_J'}\right)\widehat{V}_{\theta_J}\left(\frac{\partial\widehat{\alpha}'}{\partial\widehat{\theta}_J}\right)$. The vectors $\frac{\partial\widehat{\theta}}{\partial\widehat{\theta}_J}$ and $\frac{\partial\widehat{\alpha}}{\partial\widehat{\theta}_J}$ are given by:

$$\frac{\partial\widehat{\theta}(t)}{\partial\widehat{\theta}_k(\tau)} = \frac{NN^k}{\sum_{j,t}N^jN_t\widehat{\theta}_j(t)}1_{\{t=\tau\}}-\frac{NN^kN_\tau}{\left[\sum_{j,t}N^jN_t\widehat{\theta}_j(t)\right]^2}\sum_{j}N^j\widehat{\theta}(t) \tag{20}$$

$$\frac{\partial\widehat{\alpha}_j}{\partial\widehat{\theta}_k(\tau)} = \frac{N_{k\tau}}{\sum_{t}N_{j,t}\widehat{\theta}(t)}1_{\{k=j\}}-\widehat{\alpha}_j\frac{\sum_{t}N_{j,t}\frac{\partial\widehat{\theta}(t)}{\partial\widehat{\theta}_k(\tau)}}{\sum_{t}N_{j,t}\widehat{\theta}(t)} \tag{21}$$

In practice, to simplify the computations, we neglected the second term on the right-hand side of (21). This is only a slight approximation that does not have much impact on the estimated

variance of $\widehat{\alpha}_j$. It amounts to neglecting in (19) the variations of $\frac{1}{N^j}\sum_t N_{jt}\widehat{\theta}(t)$ and to considering only the variations of $\frac{1}{N^j}\sum_t N_{jt}\widehat{\theta}_j(t)$. Put differently, $\widehat{\theta}(t)$ is supposed to be non-random in (19).

References

[1] Abowd M., Kramarz F. and D. N. Margolis (1999), "High wage workers and high wage firms", *Econometrica*, 67(2), pp. 251-333.

[2] Besley, T. and M. Ghatak (2005), "Competition and Incentives with Motivated Agents", *American Economic Review*, 95(3), pp. 616-636.

[3] Bloom N., Cooper Z., Gaynor M., Gibbons S., Jones S., McGuire A., Moreno-Serrae R., Propper C., Van Reenen J. and Seiler S. (2011), "In defence of our research on competition in England's National Health Service", *Lancet*, 378(9809), pp. 2064-2065.

[4] Blomqvist A. (1991), "The doctor as double agent: Information asymmetry, health insurance, and medical care", *Journal of Health Economics*, 10(4), pp. 411-432.

[5] Breslow N.E. (1974), "Covariance Analysis of Censored Survival Data", *Biometrics*, 30, pp. 89-99.

[6] Chandra A. and D. Staiger (2007), "Productivity Spillovers in Health Care: Evidence from the Treatment of Heart Attacks", *Journal of Political Economy*, 115(1), pp. 103-140.

[7] Combes PPh., Duranton G. and L. Gobillon (2008), "Spatial Wage Disparities: Sorting Matters", *Journal of Urban Economics*, 63(2), pp. 723-742.

[8] Combes PPh. and H. Overman (2004), "The Spatial Distribution of Economic Activities in the European Union", in *Handbook of Urban and Regional Economics, vol. 4*, pp. 2845-2910, Henderson V. and J.F. Thisse (eds.), Elsevier-North Holland, Amsterdam.

[9] Cooper Z., Gibbons S., Jones S. and A. McGuire (2011) "Does Hospital Competition Save Lives? Evidence From The English NHS Patient Choice Reforms ", *The Economic Journal*, 121(554), pp. F228-F260.

[10] Cutler D.M. and J.R. Horwitz (1998), "Converting Hospitals from Not-for-Profit to For-Profit States: Why and What Effects?", NBER Working Paper 6672.

[11] Dormont B. and C. Milcent (2002), "Quelle régulation pour les hôpitaux publics français?", *Revue d'Economie Politique*, 17(2), pp. 117-142.

[12] Dormont B. and C. Milcent (2005), "Innovation Diffusion under Budget Constraints: Microeconometric Evidence on Heart Attack in France", *Annales d'Economie et de Statistique*, 79/80, pp. 697-726.

[13] Ghatak M. and H. Mueller (2011), "Thanks for nothing? Not-for-profits and motivated agents", *Journal of Public Economics*, 95(1-2), pp. 94-105.

[14] Gaynor M. (2006), "What do we know about competition and quality in health care markets?", NBER Working Paper 12301.

[15] Geweke J., Gowrisankaran G. and R. Town (2003), "Bayesian Inference For Hospital Quality in a Selection Model", *Econometrica*, 71, pp. 1215-1238.

[16] Gobillon L., Magnac T. and H. Selod (2011), "The effect of location on finding a job in the Paris region", *Journal of Applied Econometrics*, 26(7), pp. 1079-1112.

[17] Gobillon L. and C. Milcent (2012), "Innovative procedures: the key factor for hospital quality", Working Paper.

[18] Gowrisankaran G. and R. Town (1999), "Estimating quality of Care in Hospitals Using Instrumental Variables", *Journal of Health Economics*, 18, pp. 747-67.

[19] Gowrisankaran G. and R. Town (2003), "Competition, Payers, and Hospital Quality", *Health Services*, 38, pp. 1403-1422.

[20] Hamilton B. and V. Hamilton (1997), "Estimating surgical volume-outcome relationships applying survival models: accounting for frailty and hospital fixed effects", *Health Economics*, 6, pp. 383-395.

[21] Hansman H.B. (1996), *The ownership of Enterprise*, Cambridge, Havard University Press.

[22] Hart O.D. and B. Holmstrom (1987), "The theory of contracts", in *Advances in Economic Theory*, Bewley T. (ed.), Cambridge University Press, Cambridge.

[23] Ho V. and B. Hamilton (2000), "Hospital mergers and acquisitions: Does market consolidation harm patients?", *Journal of Health Economics*, 19(5), pp. 767-791.

[24] Kessler D. and M. McClellan (2000), "Is hospital competition socially wasteful?", *Quaterly Journal of Economics*, 115(2), pp. 577-615.

[25] Kessler D. and M. McClellan (2002), "The effects of hospital ownership on medical productivity", *RAND Journal of Economics*, 33(3), pp. 488-506.

[26] Lancaster T. (1990), *The Econometric Analysis of Transition Data*, Cambridge University Press: Cambridge.

[27] Lien H., Chou S. and J. Liu (2008), "Hospital ownership and performance: Evidence from stroke and cardiac treatment in Taiwan", *Journal of Health Economics*, 27, pp. 1208–1223.

[28] McClellan M. and D. Staiger (1999) "The Quality of Health Care Providers", NBER Working Paper 7327.

[29] McClellan M. and D. Staiger (2000), "Comparing Hospital Quality at For-Profit and Not-for-Profit Hospitals", in *The Changing Hospital Industry: Comparing Not-for-Profit and For-Profit Institutions*, D.M. Cutler ed., University of Chicago Press.

[30] McGuire T.G. (2000), "Physician agency", in *Handbook of Health Economics*, pp. 461-536, Culyer A.J. and J.P. Newhouse (eds.), Elsevier, Amsterdam.

[31] Milcent C. (2005), "Hospital Ownership, Reimbursement Systems and Mortality rates", *Health Economics*, 14, pp. 1151–1168.

[32] Milcent C., Dormont B., Durand-Zaleski I. and P.G. Steg (2007), "Gender Differences in Hospital Mortality and Use of Percutaneous Coronary Intervention in Acute Myocardial Infarction", *Circulation*, 115(7), pp. 823-826.

[33] Milgrom P. and J. Roberts (1992), *Economics, Organization and Management*, Prentice Hall, Englewood Cliffs.

[34] Moulton B. (1990), "An Illustration of a Pitfall in Estimating the Effects of Aggregate Variables on Micro Units", *The Review of Economics and Statistics*, 72(2), pp. 334-338.

[35] Newhouse J. (1970), "Toward a Theory of Nonprofit Institutions", *American Economic Review*, 60(1), pp. 64-74.

[36] Newhouse J.P. (1996), "Reimbursing health plans and health providers: efficiency in production versus selection", *Journal of Economic Literature*, 34(3), pp. 1236-1263.

[37] Pauly M.V. (1990), "The rational nonpurchase of long-term-care insurance", *Journal of Political Economy*, 98(1), pp. 153-168.

[38] Propper C., Burgess S. and K. Green (2004), "Does competition between hospitals improve the quality of care? Hospital death rate and the NHS internal market", *Journal of Public Economics*, 88(7-8), pp. 1247-1272.

[39] Propper C., Burgess S. and D. Gossage (2008), "Competition and Quality: Evidence from the NHS Internal Market 1991-1996", *The Economic Journal*, 118, 138-170.

[40] Ridder G. and I. Tunali (1999), "Stratified partial likelihood estimation", *Journal of Econometrics*, 92(2), pp. 193-232.

[41] Shortell S.M. and E.F.X. Highes (1998), "The effects of Regulation, Competition, and Ownership on Mortality Rates Among Hospital Inpatients", *New England Journal of Medecine*, 318, pp. 1100-1107.

[42] Silverman E.M.and J.S. Skinner (2001), "Are For-Profit Hospitals Really different? Medicare Upcoding and Market Structure", NBER working paper 8133.

[43] Sloan F., Picone, G., Taylor D. and S. Chou (2001), "Hospital Ownership and Cost and Quality of Care: Is There a Dime's Worth of Difference?", *Journal of Health Economics*, 20(1), pp. 1-21.

[44] Sutton M. and P. Lock (2000), "Regional Differences in Health Care Delivery: implications for a national resource allocation formula", *Health Economics*, 9, pp. 547-559.

[45] Tay A. (2003), "Assessing competition in hospital care markets: The importance of accounting for quality differentiation", *RAND Journal of Economics*, 34(4), pp. 786-814.

[46] Town R. and G. Vistnes (2001), "Hospital Competition in HMO networks", *Journal of Health Economics*, 20, pp. 733-753.

[47] Tunstall-Pedoe H. for WHO MONICA Project Principal Investigators (1988), "The World Health Organization MONICA Project (Monitoring Trends and Determinants in Cardivascular Disease): A Major International Collaboration", *Journal of Clinical Epidemiology*, 41, pp. 105-114.

[48] Tunstall-Pedoe H., Vanuzzo D., Hobbs M., Mähönen M., Cepaitis Z., Kuulasmaa K. and U. Keil (2000), "Estimation of the Contribution of Changes in Coronary Care to Improving

Survival, Event Rates, and Coronary Heart Disease Mortality across the WHO MONICA Project Populations", *Lancet*, 355, pp. 688-700.

[49] Wagstaff A. and E. van Doorslaer (2000), "Equity in Health Care Finance and Delivery", in *Handbook of Health Economics*, pp. 1723-1760, Culyer A.J. et J.P. Newhouse (eds.), Elsevier.

[50] WORLD HEALTH ORGANIZATION (2009), "International Classification of Diseases 10 Code Framework - ", *http://www.who.int/classifications/icd/en/*. Access November 1, 2009.

[51] Wray N., Hollinghsworth J., Petersen N. and C. Ashton (1997), "Case-Mix Adjustment Using Administrative Databases: A Paradigm to Guide Future Research", *Medical Care Research and Review*, 54, pp. 326-356.

Table 1: disparity indices computed from regional averages of individual variables

	Mean	Min	Max	Max/Min	Std. Dev.	Coeff. of variation	Gini
Proba. of death within 1 day (KM)	0.019	0.012	0.023	1.940	0.003	0.159	0.086
Proba. of death within 5 days (KM)	0.056	0.038	0.066	1.721	0.008	0.136	0.073
Proba. of death within 10 days (KM)	0.090	0.061	0.107	1.749	0.011	0.119	0.062
Proba. of death within 15 days (KM)	0.129	0.085	0.153	1.800	0.016	0.125	0.065
Female, 35-55 years old	0.024	0.015	0.032	2.145	0.005	0.220	0.123
Female, 55-65 years old	0.026	0.021	0.034	1.609	0.003	0.134	0.073
Female, 65-75 years old	0.073	0.060	0.089	1.475	0.006	0.088	0.048
Female, 75-85 years old	0.112	0.093	0.134	1.435	0.009	0.085	0.046
Female, over 85 years old	0.088	0.059	0.110	1.852	0.014	0.163	0.090
Male, 35-55 years old	0.181	0.135	0.239	1.771	0.027	0.152	0.084
Male, 55-65 years old	0.135	0.116	0.158	1.372	0.014	0.102	0.057
Male, 65-75 years old	0.178	0.145	0.195	1.343	0.012	0.066	0.035
Male, 75-85 years old	0.137	0.105	0.159	1.510	0.017	0.122	0.067
Male, more than 85 year old	0.046	0.027	0.062	2.259	0.010	0.209	0.115
Excessive smoking	0.124	0.062	0.196	3.160	0.038	0.310	0.171
Alcohol problems	0.012	0.004	0.017	4.148	0.003	0.276	0.151
Obesity	0.067	0.018	0.111	6.273	0.022	0.323	0.176
Diabetes mellitus	0.155	0.092	0.208	2.254	0.026	0.170	0.085
Hypertension	0.301	0.203	0.373	1.833	0.041	0.136	0.074
Renal failure	0.049	0.028	0.078	2.760	0.011	0.216	0.112
Conduction disease	0.197	0.134	0.247	1.843	0.026	0.131	0.069
Peripheral arterial disease	0.063	0.036	0.109	3.019	0.015	0.243	0.122
Vascular disease	0.044	0.025	0.078	3.109	0.013	0.289	0.149
History of coronary artery disease	0.041	0.017	0.070	4.000	0.012	0.295	0.158
Stroke	0.031	0.020	0.048	2.448	0.006	0.202	0.103
Heart failure	0.158	0.128	0.204	1.598	0.020	0.126	0.069
Cabbage or Coronary Bypass surgery	0.008	0.001	0.036	36.312	0.008	0.946	0.434
Catheter	0.188	0.130	0.271	2.081	0.037	0.197	0.107
Percutaneous transluminal coronary Angioplasty (PTCA)	0.047	0.010	0.106	10.914	0.028	0.588	0.312
Dilatation other than PTCA	0.001	0.000	0.005	\\	0.002	1.301	0.629
Stent	0.219	0.107	0.411	3.836	0.086	0.395	0.210

Source: computed from the PMSI dataset (1998-2003).

Note: variables considered here are initially defined at the patient level. We construct regional variables as the averages of patient variables by region. Indices are computed from these regional variables.

Table 2: disparity indices computed from regional averages of hospital and geographic variables

	Mean	Min	Max	Max/Min	Std. Dev.	Coeff. of variation	Gini
Number of patients	3786	2363	9644	4.081	1524	0.402	0.172
Number of AMI patients	324	173	968	5.585	163	0.503	0.204
Proportion of AMI patients	0.086	0.061	0.220	3.581	0.032	0.374	0.125
Public	0.780	0.590	0.935	1.584	0.107	0.137	0.076
Not-for-profit	0.039	0.000	0.261	\\	0.063	1.594	0.696
For-profit	0.181	0.060	0.367	6.129	0.089	0.490	0.267
Unemployment rate	0.159	0.126	0.225	1.789	0.027	0.169	0.090
Poor area in the municipality	0.700	0.363	0.947	2.612	0.174	0.249	0.136
Municipality median income	13559	11552	17455	1.511	1198	0.088	0.043
Proportion of beds in surgery	0.393	0.323	0.451	1.395	0.033	0.083	0.046
Number of beds in surgery	503	243	3172	13.062	618	1.229	0.339
Proportion of occupied surgery beds	0.857	0.781	0.901	1.153	0.034	0.040	0.021
Number of beds	1267	595	8481	14.253	1665	1.315	0.348
Proportion of occupied beds	0.824	0.774	0.865	1.118	0.025	0.030	0.017
Number of beds in the urban area	4275	1107	47033	42.475	9838	2.301	0.572
Herfindahl index for hospitals in the urban area	0.675	0.130	0.893	6.874	0.182	0.269	0.135

Source: computed from the PMSI, the SAE, and the municipality datasets (1998-2003).

Note: variables considered here are initially defined at the hospital level. We construct regional variables as the averages of hospital variables by region, weighting observations by the number of AMI patients in the hospitals. Indices are computed from these regional variables.

Table 3: estimated coefficients for the individual variables, death

Variable	Estimate
Female, 35-55 years old	< reference >
Female, 55-65 years old	0.546***
	(0.111)
Female, 65-75 years old	1.040***
	(0.096)
Female, 75-85 years old	1.378***
	(0.094)
Female, over 85 years old	1.742***
	(0.094)
Male, 35-55 years old	-0.352***
	(0.101)
Male, 55-65 years old	0.231**
	(0.099)
Male, 65-75 years old	0.813***
	(0.095)
Male, 75-85 years old	1.274***
	(0.094)
Male, over 85 years old	1.653***
	(0.095)
Excessive smoking	-0.478***
	(0.041)
Alcohol problems	0.342***
	(0.066)
Obesity	-0.247***
	(0.041)
Diabetes mellitus	-0.058***
	(0.018)
Hypertension	-0.576***
	(0.016)
Renal failure	0.369***
	(0.018)
Conduction disease	0.875***
	(0.013)
Peripheral arterial disease	-0.025
	(0.024)
Vascular disease	-0.444***
	(0.028)
History of coronary artery disease	-0.225***
	(0.029)
Stroke	0.298***
	(0.024)
Heart failure	0.061***
	(0.014)
Cabbage or Coronary Bypass surgery	-0.499***
	(0.080)
Cardiac catheterization	-1.279***
	(0.030)
Percutaneous Transluminal Coronary Angioplasty	-0.683***
	(0.039)
Dilatation other than PTCA	-0.602***
	(0.216)
Percutaneous revascularization using coronary stents (PCI – stenting)	-1.032***
	(0.026)

Source: computed from the PMSI dataset (1998-2003). Note: ***: significant at 1%; **: significant at 5%; *: significant at 10%. Number of observations: 341,861.

Note: the coefficients can be interpreted as follows. Females, aged 55-65 are 100*(exp(0.546)-1)=72.6% more likely to die than Females aged 35-55.

Table 4: disparity indices computed from
the regional probability of death obtained from the model

	Mean	Min	Max	Max/Min	Std. Dev.	Coeff. of variation	Gini
Probability of death within 1 day	0.019	0.015	0.024	1.558	0.002	0.108	0.057
Probability of death within 5 days	0.056	0.049	0.073	1.476	0.005	0.091	0.044
Probability of death within 10 days	0.085	0.074	0.108	1.453	0.008	0.095	0.049
Probability of death within 15 days	0.114	0.099	0.145	1.465	0.013	0.116	0.062

Source: computed from the PMSI dataset (1998-2003).

Note: the probability of death within a given duration of stay is computed for every region as follows. We first compute the survival function for each hospital as the exponential of minus the integrated hazard computed from the model using Breslow's estimator. The probability of death in a hospital is then given by one minus the survival function. For a given region, the probability of death is finally defined as the average of the probabilities of death of all hospitals located in that region, weighting the observations by the number of AMI patients in the hospitals.

Table 5: regression of hospital fixed effects on hospital and geographic variables, exit to death

Variable	Regression (1)	Regression (2)	Regression (3)
Constant	-5.895***	-6.899***	-7.040***
	(0.215)	(1.406)	(1.496)
Public hospital	< reference >		< reference >
For-profit hospital	0.286***		0.058
	(0.040)		(0.051)
Not-for-profit hospital	0.030		-0.113**
	(0.071)		(0.073)
Proportion of AMI patients in the hospital	-1.047***		-0.688**
	(0.174)		(0.211)
Number of beds (in log)	0.105***		0.034
	(0.016)		(0.020)
Occupation rate of beds	0.116		0.222
	(0.221)		(0.223)
Proportion of beds in surgery	-0.136		-0.290***
	(0.089)		(0.090)
Occupation rate of beds in surgery	-0.246		-0.243
	(0.160)		(0.156)
Median municipality income		0.202	0.249*
		(0.139)	(0.150)
Presence of a poor area in the municipality		0.097***	0.073**
		(0.030)	(0.031)
Municipality unemployment rate		0.025	0.225
		(0.546)	(0.573)
Herfindahl index for the healthcare structure		-0.448***	-0.427***
		(0.058)	(0.070)
Regional dummies	Non	Oui	Oui
Number of hospitals	789	834	789
Corresponding number of patients	332,827	333,810	332,827
Adjusted-R^2	0.132	0.226	0.281

Source: computed from the PMSI, the SAE, and the municipality datasets (1998-2003).

Note: ***: significant at 1%; **: significant at 5%; *: significant at 10%. We introduced a dummy for the municipality not to be in a urban area (*dummy for rural area*), and a dummy for the municipality to be related to several urban areas (*dummy for multi-polarized municipality*). The coefficients can be interpreted as follows. In regression (3), patients staying in a not-for-profit hospital are -100*(exp(-.113)-1)=10.7% less likely to die than patients in public hospitals.

Table 6: regional dummies obtained from the hospital fixed-effect regression

Region code	Name	Coefficient	Ranking on raw data
91	Languedoc-Rousillon	< reference >	(1)
41	Lorraine	-0.162*	(19)
		(0.087)	
25	Basse-Normandie	-0.171*	(3)
		(0.089)	
53	Bretagne	-0.175**	(4)
		(0.079)	
22	Picardie	-0.180**	(2)
		(0.086)	
72	Aquitaine	-0.197***	(7)
		(0.075)	
43	Franche-Comté	-0.209**	(16)
		(0.099)	
83	Auvergne	-0.214**	(9)
		(0.088)	
93	Provence-Alpes-Côte-d'Azur	-0.216***	(11)
		(0.070)	
74	Limousin	-0.219**	(18)
		(0.101)	
21	Champagne-Ardenne	-0.219**	(10)
		(0.089)	
26	Bourgogne	-0.220**	(4)
		(0.088)	
54	Poitou-Charentes	-0.232***	(12)
		(0.086)	
82	Rhône-Alpes	-0.237***	(17)
		(0.073)	
24	Centre	-0.241***	(8)
		(0.082)	
73	Midi-Pyrénées	-0.248***	(5)
		(0.077)	
52	Pays de la Loire	-0.256***	(6)
		(0.078)	
31	Nord-Pas-de-Calais	-0.275***	(13)
		(0.069)	
42	Alsace	-0.283***	(21)
		(0.098)	
23	Haute-Normandie	-0.320***	(15)
		(0.086)	
11	Ile-de-France	-0.431***	(20)
		(0.082)	

Source: computed from the PMSI, the SAE, and the municipality datasets (1998-2003).

Note: in the last column, the ranking of the regions obtained from raw data is reported in brackets. The coefficients can be interpreted in as follows. Patients staying in a hospital located in Lorraine are -100*(exp(-.162)-1)=15.0% less likely to die than patients located in a hospital in Languedoc-Roussillon.

Table 7: variance analysis for the probability of death at the regional level

Group of variables from which we consider the effect	Variance
Integrated hazard	100.0%
Individual variables (averaged at the regional level)	80.7%
Innovative treatments	26.1%
Non-innovative treatments	0.0%
Diagnoses	5.9%
Demographic variables (age x sex)	14.0%
Log-hospital fixed effects (averaged at the regional level)	20.0%
Hospital and geographic variables (averaged at the regional level)	17.0%
Hospital variables	1.2%
Status and mode of reimbursement	0.2%
Proportion of AMI patients	1.3%
Beds (capacity and occupation rate)	0.7%
Geographic Variables	17.9%
Municipality variables	0.7%
Income-related variables	0.3%
Dummies for the municipality to be rural or multi-polarized	0.3%
Herfindahl index for healthcare structure	14.6%
Regional dummies	16.7%

Source: computed from the PMSI, the SAE, and the municipality datasets (1998-2003).

Note: we compute some regional variables from patient and hospital variables. When a variable is defined at the patient level, the corresponding regional variable is the regional average. When a variable is defined at the hospital level, the corresponding regional variable is the regional average, weighting observations by the number of AMI patients in the hospitals. The effect of a regional variable is defined as the variable times its coefficient, and the effect of a group of regional variables is defined as the sum of variables times their coefficients. We are interested in the variance of a regional variable or a group of regional variables. In the second column, we report this variance as a fraction of the variance of the average integrated hazard for the region. The higher the variance, the larger the explanatory power.

Table 8: variance analysis for the probability of death at the city level

Group of variables from which we consider the effect	Variance
Integrated hazard	100.0%
Individual variables (averaged at the regional level)	72.6%
Innovative treatments	24.8%
Non-innovative treatments	0.0%
Diagnoses	3.1%
Demographic variables (age x sex)	12.0%
Log-hospital fixed effects (averaged at the regional level)	38.5%
Hospital and geographic variables (averaged at the regional level)	4.8%
Hospital variables	0.6%
Status and mode of reimbursement	0.1%
Proportion of AMI patients	0.1%
Beds (capacity and occupation rate)	0.5%
Geographic Variables	4.1%
Municipality variables	0.5%
Income-related variables	0.4%
Dummies for the municipality to be rural or multi-polarized	0.0%
Herfindahl index for healthcare structure	2.1%
Regional dummies	1.0%

Source: computed from the PMSI, the SAE, and the municipality datasets (1998-2003).

Note: we compute some city variables from patient and hospital variables. When a variable is defined at the patient level, the corresponding city variable is the average for the city. When a variable is defined at the hospital level, the corresponding city variable is the average for the city, weighting observations by the number of AMI patients in the hospitals. The effect of a city variable is defined as the variable times its coefficient, and the effect of a group of city variables is defined as the sum of variables times their coefficients. We are interested in the variance of a city variable or a group of city variables. In the second column, we report this variance as a fraction of the variance of the average integrated hazard at for the city. The higher the variance, the larger the explanatory power.

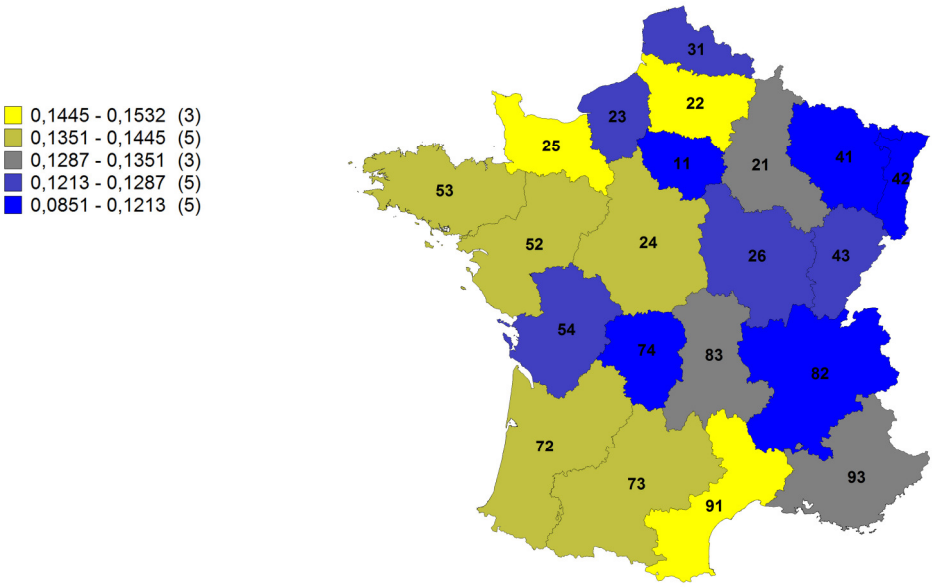
Table 9: variance analysis for the probability of death at the regional level, regression of the integrated hazard at 5 days and 15 days

Regression on:	Integrated hazard at 5 days	Integrated hazard at 15 days
Group of variables from which we consider the effect	Variance	Variance
Integrated hazard	100.0%	100.0%
Individual variables (averaged at the regional level)	78.7%	84.9%
Innovative treatments	25.4%	27.4%
Non-innovative treatments	0.0%	0.0%
Diagnoses	5.8%	6.2%
Demographic variables (age x sex)	13.6%	14.7%
Log-hospital fixed effects (averaged at the regional level)	15.4%	25.2%
Hospital and geographic variables (averaged at the regional level)	13.4%	21.3%
Hospital variables	1.1%	1.2%
Status and mode of reimbursement	0.2%	0.4%
Proportion of AMI patients	2.0%	0.8%
Beds (capacity and occupation rate)	1.1%	0.7%
Geographic Variables	14.9%	21.1%
Municipality variables	0.3%	1.7%
Income-related variables	0.1%	1.1%
Dummies for the municipality to be rural or multi-polarized	0.2%	0.4%
Herfindahl index for healthcare Structure	11.9%	18.5%
Regional dummies	18.2%	22.4%

Source: computed from the PMSI, the SAE, and the municipality datasets (1998-2003).

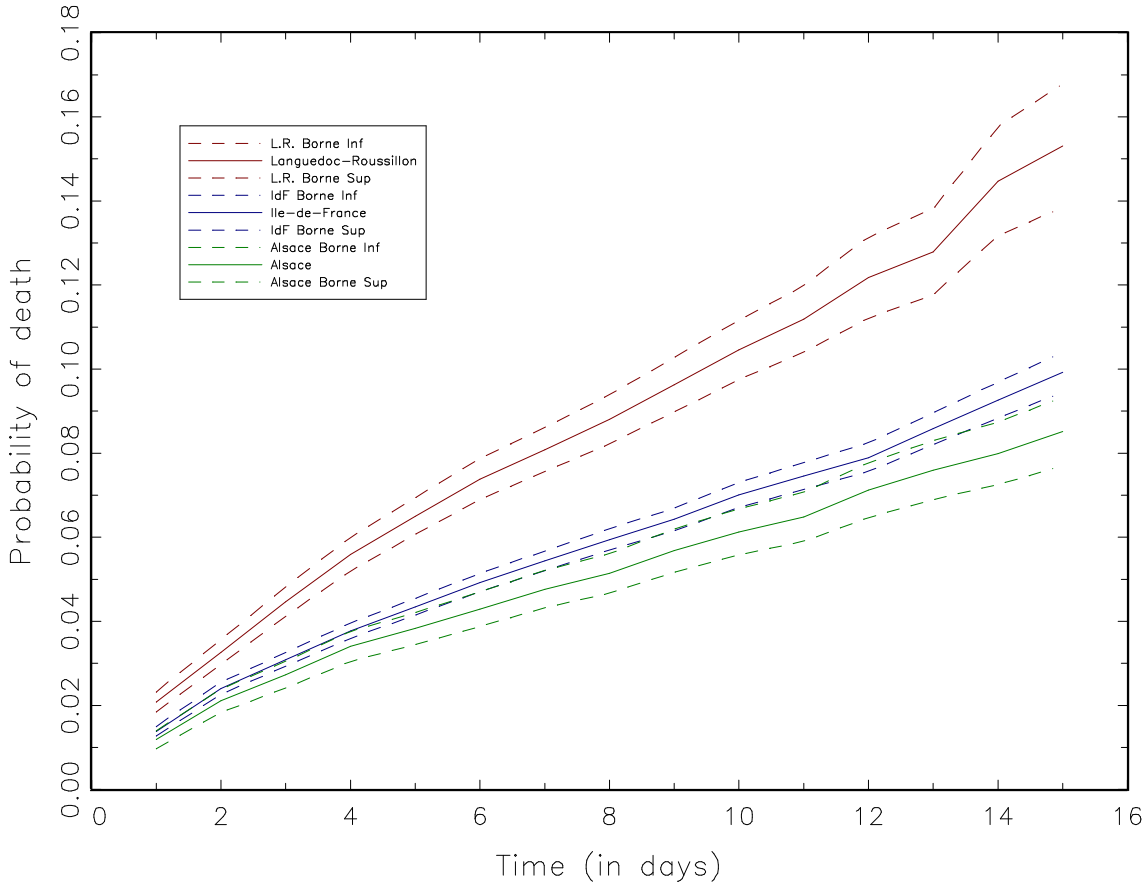
Note: we compute some regional variables from patient and hospital variables. When a variable is defined at the patient level, the corresponding regional variable is the regional average. When a variable is defined at the hospital level, the corresponding regional variable is the regional average, weighting observations by the number of AMI patients in the hospitals. The effect of a regional variable is defined as the variable times its coefficient, and the effect of a group of regional variables is defined as the sum of variables times their coefficients. We are interested in the variance of a regional variable or a group of regional variables. In the second column, we report this variance as a fraction of the variance of the average integrated hazard for the region. The higher the variance, the larger the explanatory power.

Graph 1: regional probability of death within fifteen days (in %)



Note: the probability of death within 15 days is computed for every region as follows. We first compute the survival function for each hospital using the Kaplan-Meier estimator, where all exits other than death are treated as censored. The probability of death in a hospital is then given by one minus the survival function. For a given region, the probability of death is finally defined as the average of the probabilities of death of all hospitals located in that region, weighting the observations by the number of AMI patients in the hospitals. We represent here the probability of death within 15 days. A colour is attributed to each of the five probability-of-death intervals defined in the legend. The number of regions in each interval is given in brackets. Regions are identified on the map by their administrative codes. These codes are: 11: Ile-de-France; 21: Champagne-Ardenne; 22: Picardie; 23: Haute-Normandie; 24: Centre; 25: Basse-Normandie; 26: Bourgogne; 31: Nord Pas-de-Calais; 41: Lorraine; 42: Alsace; 43: Franche-Comté; 52: Pays de la Loire; 53: Bretagne; 54: Poitou-Charentes; 72: Aquitaine; 73: Midi-Pyrénées; 74: Limousin; 82: Rhônes-Alpes; 83: Auvergne; 91: Languedoc-Roussillon; 93: Provence - Alpes Côtes d'Azur.

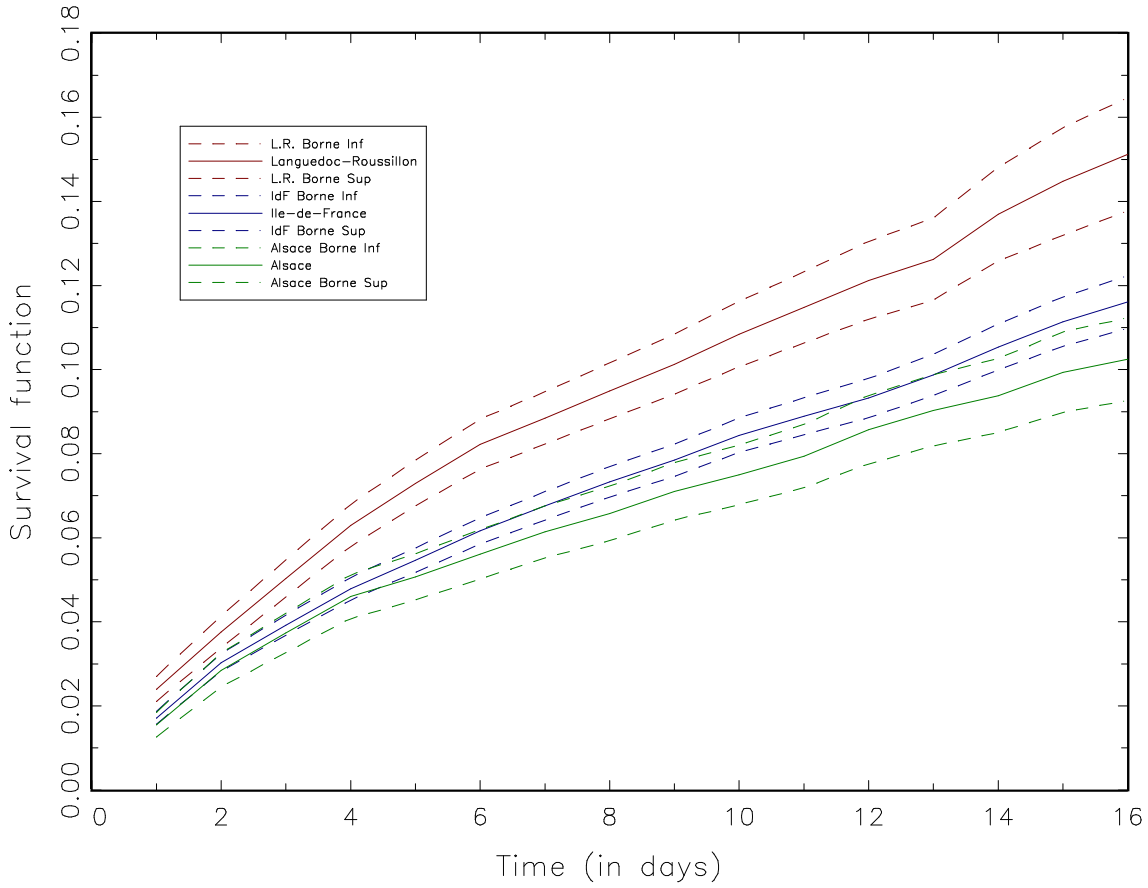
Graph 2: probability of death for extreme regions and Paris region (Kaplan-Meier)



Source: computed from the PMSI dataset (1998-2003).

Note: the probability of death is computed for every region as follows. We first compute the survival function for each hospital using the Kaplan-Meier estimator, where all exits other than death are treated as censored. The probability of death in a hospital is then given by one minus the survival function. For a given region, the probability of death is finally defined at the average of the probabilities of death of all hospitals located in that region, weighting the observations by the number of AMI patients in the hospitals. Confidence intervals for the probability of death of each region are represented by dashed lines.

Graph 3: probability of death for extreme regions and Paris region (model)



Source: computed from the PMSI dataset (1998-2003).

Note: the probability of death is computed for every region as follows. We first compute the survival function for each hospital as the exponential of minus the integrated hazard computed from the model using Breslow's estimator. The probability of death in a hospital is then given by one minus the survival function. For a given region, the probability of death is finally defined as the average of the probabilities of death of all hospitals located in that region, weighting the observations by the number of AMI patients in the hospitals. Confidence intervals for the probability of death of each region are represented by dashed lines.