Spatial disparities in hospital performance

Laurent Gobillon, Carine Milcent

Spatial disparities in hospital performance*

Laurent Gobillon† Carine Milcent‡

Keywords: spatial health disparities, economic geography, stratified duration model
JEL code: I11, C41

*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.

†INED, PSE-INRA and CREST. INED, 133 boulevard Davout, 75980 Paris Cedex. Tel: 0033(1)56062016. 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.

1 Introduction

In many countries, spatial disparities between local markets are large and raise some major policy concerns. Whereas the focus of the 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 approach to explaining the 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). Other papers are interested in the determinants of quality within a given country and exploit the spatial dimension to construct some control variables or instruments. Geweke, Gowrisankaran and Town (2003) study the effect of hospital on mortality and instrument the hospital choice with the distance between the place of residence and hospitals. A growing strand of the literature is interested in the effect of the local healthcare market structure on health outcome. Most authors try to estimate the marginal effect of local competition on health quality (see Gaynor, 2006 for a survey). However, they do not assess how spatial variations in competition can explain spatial disparities in quality.

In fact, evaluating the marginal effect of some factors on mortality and assessing how some spatial variations in these factors can explain spatial disparities in mortality are two related but different exercices. For instance, it is usually found that sex significantly affects mortality. If there was no variation in the share of females across the territory though, the differences in the local sex composition would not explain the disparities in mortality across locations. The same arguments apply when considering local determinants such as local competition indices. It may happen that local competition has a significant marginal effect on mortality but only small spatial variations. In that case, it does not explain the spatial disparities in mortality on the territory.

In this paper, we conduct a variance analysis of regional disparities in mortality by acute myocardial infarction (AMI) in France. We quantify these disparities and assess the importance of the factors which may explain them. We can distinguish three types of factors according to the literature on health. First, the spatial disparities in mortality can be explained by some 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 caused by 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, the quality of care to the elderly would be better in for-profit hospitals than in not-for-profit hospitals. Milcent

(2005) finds that in France, patients in for-profit hospitals have a lower probability of death when having 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 come from differences between local healthcare markets. In particular, the local competition measured with a Herfindahl index is often found to have a significant negative impact on mortality (Kessler and McClellan, 2000).

We estimate a model at the individual level where the propensity of patients to die during their stay in hospital is specified as a function of the three types of explanatory factors. We then average the model at the regional level and conduct a variance analysis in the spirit of the literature in economic geography (Duranton and Monastiriotis, 2002; Combes, Duranton and Gobillon, 2008; Mion and Naticchioni, 2008). Estimations are conducted on a unique matched patient-hospital dataset from some exhaustive French administrative records over the 1998-2003 period. This original dataset contains some information on the demographic characteristics of patients, their diagnoses and their treatments. It also provides some details on the hospitals where the patients are treated, like the location, the ownership status and the capacity.

More specifically, we use a very flexible econometric specification building on Ridder and Tunali (1999) and Gobillon, Magnac and Selod (2010). We first estimate a Cox duration model stratified by hospital (i.e. each hospital has a specific baseline hazard) using the stratified partial likelihood estimator (SPLE). The individual variables included in the model are the patient characteristics (demographic shifters and secondary diagnoses) and treatments (as they are patient-specific). Their effects are estimated while taking properly into account the hospital unobserved heterogeneity. Our approach also allows to recover some hospital hazard functions without specifying them parametrically. We then go further and specify the hospital hazards as the product of some hospital fixed effects and a baseline hazard. We show how to estimate the hospital fixed effects using some moment conditions. The estimated hospital fixed effects are regressed on some hospital and local variables. We finally average the model at the regional level and make a spatial variance analysis.

We show that regional disparities in mortality are quite large. In particular, the raw difference in

[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).

the propensity to die within 15 days between the extreme regions reaches 80%. After taking into account the individual variables, this difference drops to 47%. A variance analysis at the regional level shows that regional differences in innovative treatments play a major role in explaining the regional disparities in mortality. A local Herfindahl index computed from the number of patients in each hospital also plays a significant role. Results suggest that spatial differences in the local concentration of patients partly explain spatial differences in mortality.
In a first section, we present the different factors which may explain the spatial differences in healthcare quality in France and review the corresponding literature. A second section describes our dataset. We then present in a third section some descriptive statistics on the regional disparities in mortality, demand factors and supply factors. The fourth section details the econometric methodology used to identify the causes of the regional disparities in mortality. The fifth section summarizes the results of the model.

2 Heart attack in the French context

2.1 The French context

The aim of the paper is to quantify the regional disparities in the mortality of patients hospitalized in France for a heart attack and identify their key determinants. As mentioned earlier, three potential explanations of these disparities are the spatial differences in the composition of patients (case-mix), the composition of hospitals, and the local healthcare market structure (in particular the intensity of competition between hospitals). Whereas the local composition of patients can be viewed as the local demand for healthcare, the local composition of hospitals and competition are related to the local supply.
On the demand side, spatial differences in demographic shifters and secondary diagnoses possibly related to specific regional behaviours may explain regional disparities in mortality.

2.1.1 Hospitals' characteristics

On the supply side, the local composition of hospitals can affect the local propensity of AMI patients to die if hospitals differ in their efficiency to treat patients and are distributed over space according to their efficiency. We now briefly describe the French system to highlight how hospitals can differ in their efficiency.

The public sector is under a global budget system as well as part of the private sector. Private hospitals which benefit from this budget are Not-For-Profit hospitals (NFP). Every year, the government determines the global budget and chooses how to divide it between regions. The regional budget is shared between NFP and public hospitals according to the budget of the previous year and through bilateral bargaining between the regional regulator and the hospital managers. NFP and public hospitals have to grant access to hospital care to every patient and cannot make any profit. Also, public and NFP hospitals provide similar high-tech care. By contrast, the other hospitals in the private sector (namely For-Profit hospitals) are paid by fee-for-services and can select patients. The selection is usually done to maximize profit, taking into account the health status of patients. FP hospitals have no constraint on profits. *Overall, hospitals thus have different incentives to provide health care to patients depending on their status (public or private) and reimbursement rule (fee-for-service or global budget).*

Importantly, there is no segmentation of the healthcare market by insurance status interfering with the effect of the hospital status as in the US. Indeed, people with a health insurance in the US have some form of managed care insurance. Plan enrollees have to choose from a pre-approved subset of doctors and hospitals in their area. Hospitals in the subset can have a specific status. As a consequence, the effect of the hospital status is intertwined with the effect of the insurance status. In France, hospital expenditures are fully reimbursed by a unique public compulsory insurer which funds come from taxes. Patients can freely choose their hospital and there is no segmentation by insurance status. Hence, the true effect of the hospital status can be identified more easily as in the case of Taiwan (Lien, Chou and Liu, 2008).

Milcent (2005) finds for France that ownership significantly affects the mortality rate. Her results suggest that patients in FP hospitals have a lower probability of death but face a greater uncertainty on the quality of care. FP, NFP and public hospitals are distributed unevenly on the French territory, in particular for demographic and historical reasons. Hence, there are some regional disparities in the local composition of hospitals which can yield some regional disparities in mortality of AMI patients treated in hospitals.

Because of the reimbursement rules, hospitals do not have the same incentives to treat patients with innovative procedures. Indeed, FP hospitals are financed via a fee-for-service system. Innovative supplies involving expensive devices that can be used only once such as angioplasty or stent[2] for heart attack are reimbursed ex-post in addition to the fee-for-service payment. By contrast,

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

the reimbursement of public and NFP hospitals does not depend on the number of procedures which are performed. Therefore, FP hospitals have more incentives to perform innovative procedures than public and NFP hospitals. Milcent (2005) finds for France that innovative treatments decrease the mortality of AMI patients. Spatial differences in the use of innovative treatments (which are related to the spatial sorting of FP hospitals) can cause regional differences in mortality of AMI patients.

The efficiency of hospitals may also be affected by the intensity of hospital activities because of some learning by doing. As a consequence, we will investigate the role played by the spatial disparities in the occupancy rate and the proportion of patients treated for an AMI, in explaining the spatial disparities in mortality. Finally, larger hospitals can bear larger fixed costs related to equipment as these costs can be shared between more patients. Consequently, we will also evaluate whether spatial disparities in the hospital size proxied by the number of beds (in total and in surgery) are related to spatial differences in mortality.

2.1.2 Competition between hospitals

There is a growing body of US literature on the effect of local competition between hospitals on their efficiency (see Gaynor, 2006, for a survey). Whereas some papers investigate situations where prices are set by hospitals, most studies focus on cases where prices are fixed and hospitals can only choose quality of care. In France, prices are regulated in both the public and private sectors. Hence, the competition between hospitals would only be based on quality. Moreover, public and NFP hospitals which are paid under the global budget system are not allowed to make any profit. As a consequence, these hospitals do not have incentives to compete with each other even by providing a better quality of care contrary to their US counterparts. By contrast, FP hospitals can make some profit and thus have some incentives to attract patients. This can be done by providing some services of better quality than in other FP hospitals, as well as in public and NFP hospitals. The higher the number of private hospitals, the more important is competition based on quality. On the other hand, the quality of care in a hospital depends on learning-by-doing which is directly related to the number of patients treated in that hospital. When the patients in an area are scattered across many small hospitals rather than a few large ones, there is not much learning-by-doing in each hospital and the average quality of care in the area could then be lower.

When a patient chooses a hospital where to be treated, he takes into account the accomodation and catering which differ in private and public hospitals. More importantly, he is attracted by

the physicians who are the most efficient and can provide the best care. As FP hospitals want to attract patients to make profit, they will try to get the most efficient physicians. This is a specific form of competition based on quality. In fact, the best physicians have some incentives to work in FP hospitals because of the specific payment rule which differs from the one applied to public and NFP hospitals.

In the public sector, the staff (including physicians and nurses) consists of salaried civil servants. Their wages do not depend on their performance. One day a week, though, they can work outside their hospital, in particular in a FP hospital. Physicians working in NFP hospitals are also salaried but their wages are far higher than in public hospitals. In FP hospitals, some physicians are salaried and the others are self-employed. The working time of physicians as well as their wages usually depend on the number of patients. Moreover, physicians receive additional fees when performing innovative procedures. Overall, their income is far larger than in public hospitals. Consequently, physicians usually compete to get a job in the private sector and only the best of them are selected.[3] Interestingly, this competition has an effect on medical practices in public hospitals. As physicians want to get employed by private hospitals, they perform some innovative procedures to increase their reputation and skills with learning by doing. Dormont and Milcent (2006) show that hospitals under global budget perform innovative procedures at a rate close to the one of the FP hospitals.

Hospitals which ischemic service has grown large are usually those succeeding in attracting patients because of a better reputation. The best physicians have gathered there, can still improve with learning-by-doing, and perform better than in other hospitals which have become smaller. The concentration of patients in a few large hospitals rather than many small ones in an area could then be associated with a better average local quality.

In the US, the local competition among hospitals affects the patients' propensity to die (Gaynor, 2006). Was it the case in France, spatial disparities in the local competition among hospitals could partly explain the regional differences in mortality. In our study, we account for the intensity of competition between hospitals through a Herfindahl index which measures the concentration of patients in a few large hospitals rather than many small ones and is expected to have a positive effect on mortality (Kessler and McClellan, 2000; Town and Vistnes, 2001; Gowrisankaran and Town, 2003). Note however that for France, the local concentration of pa-

[3]There is a large literature on the incentives for physicians depending on the payment's rule (Hart and Holmstrom, 1987; Pauly, 1990; Blomqvist, 1991; Milgrom and Roberts, 1992; Newhouse, 1996; McGuire, 2000).

tients could also reflect some learning-by-doing or the gathering of efficient physicians in the same place. In that case, the effect of the Herfindahl index would rather be negative. Overall, the sign of its effect is theoretically ambiguous and remains an empirical issue.

Our empirical approach will allow us to assess the respective importance of the determinants in explaining the regional disparities in the mortality of AMI patients.

2.2 Treatments of heart attack

In this paper, we focus on one single disease. Indeed, evidence shows that the effect of characteristics on mortality is disease-specific (Wray et al., 1997). We select the Acute Myocardial Infarction (heart attack) for four reasons. First, it belongs to the ischemic-disease group that has been the primary cause of mortality in France, before getting second recently after cancer. Second, mortality from AMI has been widely studied in the literature to assess the quality of hospital care in the US and the UK. This literature can be used for comparison (see Goworisakaran and Town, 2003, for the US, and Propper, Burgess and Green, 2004, for the UK). Third, AMI is a well-defined pathology with only a few re-admissions due to its clinical definition. Fourth, mortality from AMI is an event frequent enough to yield some reliable statistical results.

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 (denoted as CATH hereafter), percutaneous transluminal coronary angioplasty (PTCA) and stent. A catheter is a thin flexible pipe which is installed in a vein. It may also be used for cleaning arteries in order to improve the blood flow. A bypass surgery reroute, or "bypass", is a vein or artery collected from the patient's body and set up to derive blood from coronary arteries. In some cases, the stent and the angioplasty are some alternative procedures to the bypass which yield a better quality of life after home return. An angioplasty consists in inflating a balloon in a blockage to create a channel. This procedure is costly as it induces for one stay an increase in costs which ranges from 30% to 60% (Dormont and Milcent, 2002). The stent is a spring-shaped prosthesis which is used as a complement to angioplasty. The use of stent with an angioplasty significantly improves the results. Angioplasties and stents are some innovative treatments over the 1998-2003 period. We will study how the spatial variations in treatments can explain regional differences in mortality.

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 *catheterism* refers to a

catheterism without angioplasty and stent.

2.3 Spatial features

We now propose a spatial overview of heart attack. First note that AMI patients who want to be treated in a NFP or a public hospital have to go to a hospital within their region of residence. However, some patients are sometimes transferred to a neighbouring hospital in another region. Also, a patient who gets sick in another region may be cured there. Over the 1998-2003 period, the proportion of AMI patients being treated within their region of residence is very high at 92.9%. This proportion is slightly lower for FP hospitals (91.4%) than for public hospitals (93.1%) and NFP (95.8%). These statistics support the fact that regions can be viewed as local healthcare markets for heart attack.

Depending on their residential location, patients do not face the same supply of healthcare, as the local composition of hospitals by status and mode of reimbursment varies widely across space. In 1999, the proportion of beds in public hospitals is large in the west and in Franche Comté (in the east) where it reaches 80%, whereas it is only 46% in the PACA region (the southern *French Riviera*). The proportion of NFP hospitals is the highest in some eastern regions at the German border (Alsace and Lorraine) for historical reasons. Conversely, the proportion of beds in FP hospitals is larger in the south-east (around the *French Riviera* region) where the population is older and richer.

The local proportions of patients treated for an AMI in the different types of hospitals mimic the distribution of bed capacities. For instance, Graph 1 shows that around Paris and in southern regions, the proportion of patients treated in a FP hospital is higher. These regions are often characterized by a substantial use of innovative procedures like stents, as shown in Graph 2. In fact, the rank correlation between the proportion of stents and the proportion of patients in FP hospitals is .61. When considering NFP hospitals instead of FP hospitals, the correlation is still quite high at .44.

[*Insert Graphs* 1 *and* 2]

We also computed the probability of death within 15 days (see Graph 3).[4] This probability is quite low in the Paris region, the east and south-east. It is larger in the west and south-west. There is no obvious relationship between the probability of dying and the proportions of stents or FP hospitals

[4]See below for more details on how this probability is computed.

(rank correlations: $-.09$ and .14 respectively). However, south eastern regions which have a large proportion of FP hospitals performing innovative treatments also concentrate older people who are more likely to die. Hence, it is necessary to perform an econometric analysis to disentangle the effect of age and more generally of individual attributes (demographic characteristics and secondary diagnosis) from that of innovative procedures, hospital characteristics, and local healthcare market structure.

[*Insert Graph* 3]

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 the records of all patients discharged from any French acute-care hospital over the 1998-2003 period. It is compulsory for hospitals to provide these records on a yearly basis.[5] 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 both for the public and private sectors.[6] 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.

The dataset contains some information on the demographic characteristics of patients (age and sex), as well as some very detailed information on the diagnoses and treatments. In our analysis, we can thus take into account all secondary coronary diagnoses as well as all techniques used to cure patients. One may argue that some comorbidity factors are not recorded. However, McClellan and Staiger (1999) show that much more detailed medical data on disease severity and comorbidity do not add much when taking into account the heterogeneity among patients. 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).

We only keep patients whose pathology was coded as an acute myocardial infarction in the tenth international code of disease (ICD-10-CM), i.e. the patients for whom the code was I21

[5] An exception is local hospitals for which it is not compulsory. This does not affect our study since these hospitals do not take care of AMI patients.

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

or I22. Before 35, heart attacks are often related to a heart disfunction. As a consequence, we restrain our attention to the patients more than 35 following the OMS definition, which leaves us with 421,185 stays. 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 (that are only very few), we end up with 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 some 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 reaches 97% of the patients.

The municipality code in the SAE survey also allowed us to match our dataset with some wealth variables at the municipality level coming from other sources. These variables will be used in our estimations to take into account the spatial differences in the funding of public and NFP hospitals. Indeed, local authorities sometimes take into account the local level of poverty when dispatching the budget across hospitals. Our municipality variables include 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 poor area in the municipality (poor areas being defined by a 1997 law under the label *zones urbaines sensibles*). Also, thanks to the municipality code, we could identify the urban area in which hospitals are located.[7] We computed a local index of competition between hospitals within each urban area. This index is a Herfindahl index at the urban area level using the number of patients in hospitals within each urban area.[8] In our analysis, we will also take into account the size of the healthcare market surrounding each patient's hospital as it may affect their efficiency. This size is measured by the number of beds in the urban area, the patient's hospital being excluded. When constructing urban area variables,

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

[8]The Herfindahl index for an urban area u is $H_u = \sum_{j \in u} \left(\frac{p_j}{p^u}\right)^2$ where j indices the hospitals, p_j is the number of patients in the hospital j, and $p^u = \sum_{j \in u} p_j$ 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.

we were confronted with a few hospitals in municipalities which do not belong to any areas or to several of them. We thus introduced some dummies for these two cases as controls. As we will use hospital variables which should be time-invariant in our analysis (see Section 5 and 6), all hospital and geographic variables are averaged across years.

3.2 Preliminary statistics

For each hospital, we computed a gross survival function for exit to death using the Kaplan-Meier estimator. This estimator treats other exits (home return and transfers) as censored. As we are mostly interested in hospital disparities across regions, we computed the average survival function by region.[9] Observations were weighted by the number of patients still at risk in the hospitals.[10] We selected the region with the highest survival function (Alsace), the region with the lowest survival function (Languedoc-Roussillon), and the Paris region (Ile-de-France) that is the most densely populated. Graph 4 represents the survival functions of these three regions as well as their confidence intervals (Graph A1 in appendix represents the survival functions for all the regions and Table A1 ranks the regions according to survival after 15 days). It shows that the two extreme average survival functions are significantly different.

[*Insert Graph* 4]

Table 1 reports some disparity indices between regions in the probability of dying within 1, 5, 10 and 15 days (defined as one minus the Kaplan-Meier). These indices are the max/min ratio, the Gini index and the coefficient of variation. The Gini indices and coefficients of variation are computed in two stages. First, we compute the average of a given individual variable (for instance, a death dummy) by region. Then, we compute the regional disparity indices for the resulting

[9] We could have directly computed a survival function for each region. However, we believe that the relevant unit at which the treatment of patients takes place is the hospital. Also, our approach at the hospital level parallels the model presented in Section 5.

[10] When the length of stay increases, the number of patients in a given hospital decreases. Above a given length of stay after which there is no patient at risk anymore, it is not possible to estimate the survival function. We then arbitrarily considered that the hospital survival functions remained constant after this length of stay. When we compute the average survival functions by region, this assumption does not have much effect for short/medium lengths of stays. Indeed, only small hospitals do not have any patients at risk anymore for these lengths of stays. As a consequence, we limited our analysis to lengths of stays below fifteen days to minimize the effect of our assumption.

variable (in our example, the regional proportion of deaths), weighting the observations by the number of patients in the region. Global indices like the Gini index (.07) and the coefficient of variation (.218) remain quite small and suggest that disparities are not systematic. The max/min ratio shows that regional disparities are significant. 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 transfers and home returns in the early days of AMI stays.

As mentioned earlier, the regional disparities in mortality may be explained with some regional disparities in demand factors (demographic shifters and secondary diagnosis) or in supply factors, whether they are related to hospitals (treatments and establishment characteristics) or to the local healthcare market structure. To disentangle these three types of effects, we present Gini indices which are some global measures of disparities and are not sensitive to the level of magnitude as the max/min ratio.[11] We consider in the sequel that disparities are small when the index is inferior to .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.

We first consider variables related to patients which were averaged at the regional level. There are significant disparities across regions for some demographic variables: the Gini index is moderate for females aged $35-55$ (.12) and males who are more than 85 (.11). For diagnoses, the Gini index reaches .23 for surgical French DRGs (.23), .15 for the severity index[12] and .13 for a history of vascular diseases and for stroke. Note that the Gini index is most often moderate for diagnoses related to specific behaviours before the heart attack such as obesity (.17), excessive smoking (.16), and alcohol problems (.14). Regional disparities in the use of procedures are important. The Gini index goes up to .53 for dilatations other than PTCA and .37 for the cabbage or coronary artery bypass surgery (CABG). More widespread procedures like angioplasty and stent still have a large Gini index which takes the value .28 and .21, respectively.

Overall, the Gini indices show that potential explanations of disparities in the propensity to die can be related to demographic characteristics, diagnoses and procedures. One should keep in mind

[11] alternatively, we could also comment the results obtained with the coefficients of variation which are similar.

[12] We use Deyo's adaptation of the Charlson co-morbidity index to measure the severity of co-morbidities (Deyo, 1992; Ghali, 1996). The Charlson index, which is expressed as a six-level variable, is constructed for each stay. This index is greater than 0 when a surgical procedure has been carried out on the patient. Validation exercises have shown that this index predicts well mortality in longitudinal data (Hamilton and Hamilton, 1997).

though that the explanatory power of a given variable when studying the regional disparities in mortality is closely related to its variance and its effect on death. Considering the variance, the different types of patient-specific variables still look like good candidates (even if their respective importance is different).

[*Insert Table* 1]

We then computed regional disparity indices for the hospital and geographic variables used in our regressions. Whereas hospital variables measure capacities and status (public, NFP and FP), geographic variables are mostly meant to capture the effects related to the structure of the local healthcare market (like competition). For a given variable, we constructed its regional average, weighting the observations by the number of patients in the hospitals. The resulting regional average is then used to compute disparity indices at the regional level. Results are reported in Table 2. As previously, we only comment Gini indices.

There are large disparities across regions in the size of hospitals measured by the total number of patients (.23) or the number of AMI patients (.27). Disparities are even larger for the number of beds (.49) and the number of beds in surgery (.47). These disparities point out some sorting of hospitals according to their size. Finally, disparities are smaller but still large for the hospital status and more specifically for being a FP hospital (.24). The regional disparities in hospital characteristics may thus play a role when trying to explain the regional disparities in mortality.

Concerning geographic variables, we observe some very large disparities in the number of beds in the urban area (Gini index .66) which is not surprising as there is a lot of variation in the population of regions. Disparities are also significant for the Herfindhal index computed at the urban area level (.20). Indeed, hospitals are unevenly distributed in the territory, for historical reasons, public policy and consequences of competition. This creates some regional differences in average Herfindahl Index. Regional disparities in municipality variables capturing some geographic heterogeneity in funding are at best moderate, the Gini index reaching .17 for the presence of a poor area in the municipality.

[*Insert Table* 2]

In summary, demand and supply factors are all some potential candidates to explain the regional disparities in mortality. We now present our empirical methodology to assess their respective explanatory power.

4 Econometric method

We first give a brief description of the econometric model explaining the propensity to die before turning to a more formal presentation of our approach. We build our specification around hospital units to properly take into account their heterogeneity and use a Cox duration model at the patient level stratified by hospital. Hence, each hospital has its own hazard function which captures its specific behaviour. Ridder and Tunali (1999) explain how to estimate this model using the stratified partial likelihood estimator (SPLE) and establish the theoretical properties of the estimators. Their methodology has been used in some studies related to education and unemployment, but not in health economics. Lindeboom and Kerkhofs (2000) apply their methodology to quantify the effect of school on the job sickness of teachers and Gobillon, Magnac and Selod (2010) use it to analyze the effect of location on finding a job in the Paris region.

The model contains some patient-specific explanatory variables (demographic shifters, diagnoses and treatments), as well as a specific survival function for each hospital which is left unspecified. The flexible modelling of the hospital heterogeneity allows us to recover some robust estimators of the coefficients of patient-specific explanatory variables. These coefficients are then used in the estimation of the hospital survival functions, which are in turn averaged at the regional level to study the regional disparities in mortality net of the effect of patient-specific variables.

We then link the remaining regional disparities to some local differences in hospital and geographic characteristics. For that purpose, we make the additional assumption that the hospital hazards write multiplicatively as the product of a hospital fixed effect and a baseline hazard. We show how to estimate the hospital fixed effects using moment conditions. We explain them with hospital and geographic variables and finally average the model at the regional level to perform a regional variance analysis.[13]

We now present our approach more formally. 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:

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

[13] A tempting alternative approach is to estimate all the coefficients in one stage introducing all the patient, hospital and geographic variables in a simple Cox model. However, such an approach does not take into account the hospital unobserved heterogeneity. Consequently, standard errors of the coefficients may be highly biased (see Moulton, 1990). Our approach properly takes into account the hospital unobserved heterogeneity.

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 his hospital dying after this duration. It writes:

$$P_i = \frac{\exp(X_i\beta)}{\underset{i\in\Omega_{j(i)}(t_i)}{\exp(X_i\beta)}} \tag{2}$$

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 then writes: $L = \underset{i}{\prod} P_i$. 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 writes:

$$\widehat{\Theta}_j(t) = \underset{0}{\overset{t}{}} \frac{I(N_j(s)>0)}{\underset{i\in\Omega_j(s)}{\exp\left(X_i\widehat{\beta}\right)}} dN_j(s) \tag{3}$$

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 will be averaged at the regional level to study regional disparities in mortality after any number of days. As the hospital hazards are left completely unspecified, the study of regional disparities in death using regional averages remains very general.

We then study the determinants of hospital disparities by specifying the hospital hazard rates in a multiplicative way:

$$\theta_j(t) = \alpha_j\theta(t) \tag{4}$$

where α_j is a hospital fixed effect and $\theta(t)$ is a baseline hazard common to all hospitals. We show in appendix how to estimate the parameters using empirical moments derived from (4).[14] Note

[14]In doing so, we depart from the log-linear estimation method proposed by Gobillon, Magnac and Selod (2010). Our approach is more adequate when exits are scarce as in our case. Indeed, Gobillon, Magnac and Selod split the timeline into K intervals denoted $[t_{k-1}, t_k]$. Introduce $\theta_k = \underset{t_{k-1}}{\overset{t_k}{}} \theta(t)\,dt/(t_k - t_{k-1})$ and $y_{jk} = [\Theta_j(t_k) - \Theta_j(t_{k-1})]/(t_k - t_{k-1})$. Integrating (4) 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: $\widehat{y}_{jk} = \left[\widehat{\Theta}_j(t_k) - \widehat{\Theta}_j(t_{k-1})\right]/d_k^j$

that we need an identifying restriction since α_j and $\theta(t)$ can be identified separately only up to a 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 (see appendix), we get:

$$\theta(t) = \left(\frac{1}{N^2}\sum_{j,t} N^j N_t \theta_j(t)\right)^{-1}\left(\frac{1}{N}\sum_j N^j\theta_j(t)\right) \quad (5)$$

$$\alpha_j = \left(\frac{1}{N^j}\sum_t N_{jt}\theta(t)\right)^{-1}\left(\frac{1}{N^j}\sum_t N_{jt}\theta_j(t)\right) \quad (6)$$

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 $T = 30$ for convenience). 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 (5). 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 (6). We show in appendix 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: $\alpha_j = \exp(Z_j\gamma + \eta_j)$ 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:

$$\ln\widehat{\alpha}_j = Z_j\gamma + \eta_j + \phi_j \quad (7)$$

where $\phi_j = \ln\widehat{\alpha}_j - \ln\alpha_j$ is the sampling error on the hospital fixed effect. Equation (7) can be estimated using weighted least squares where the weight is the number of patients in the hospitals. The standard errors and R-square (adjusted to take into account the sampling error),

---

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 $\psi_{jk} = \ln\widehat{y}_{jk} - \ln y_{jk}$ is the sampling error. This equation can be estimated with standard linear panel methods. The authors use weighted least square 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 it 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 scarce and a significant number of observations should be discarded when the time spent in the hospitals gets long. In practice however, the results obtained with the two approaches are quite similar.

are computed as proposed by Gobillon, Magnac and Selod (2010). Note that for a given hospital, equation (7) is well defined only when there is at least one patient dying 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 and they are discarded from our sample. We finally average equation (7) at the regional level and conduct a variance analysis for the resulting equation.

5 Results

Table 3 reports the estimation results of the first-stage equation (2). The demographic characteristics have the usual effect on the propensity to die. Females are more likely to die than males. This is consistent with care being more protective for males than for females possibly because of biological differences like the smaller target vessel size and the more important vessel tortuosity of females (Milcent et *al.*, 2007). Also, the propensity to die increases with age.

Among variables related to the diagnosis, the severity index is found to have a positive effect on the propensity to die. Intuitively, one also expects secondary diagnoses to have a positive effect as they deteriorate health. This is true empirically for renal failure, stroke, heart failure, conduction disease, alcohol. Other secondary diagnoses have a more surprising negative effect: diabetes, obesity, excessive smoking, vascular disease, peripheral arterial disease, previous coronary artery disease, and hypertension. These results may be explained by preventive health care. Indeed, these secondary diagnoses may point at patients who are monitored more carefully before and after having a heart attack (Milcent, 2005).

All treatments have the expected negative effect on the propensity to die: CABG, catheterism, PTCA, other dilatation and stent. The stent, which is the most innovative procedure, has the strongest negative effect. After taking into account these treatments, the DRG index capturing the heaviness of surgical procedures has a positive effect on the propensity to die. This can reflect the increased chances of dying because of more cumbersome and risky surgery.

[*Insert Table* 3]

From the estimated coefficients $\widehat{\beta}$, we constructed an integrated hazard for each hospital using Breslow's estimator and averaged the corresponding hospital survival functions by region (weighting by the number of patients at risk in the hospitals).[15] Regions at extremes are the same as

[15]This kind of aggregation is quite common in the labour literature. For instance, Abowd, Kramarz and Margolis

when studying the raw data: Alsace (at the German border) usually exhibits the highest survival function and Languedoc-Roussillon (in the South-East) the lowest. Graph 5 represents the survival functions (as well as their confidence intervals) for these two extremes and for Ile-de-France (The Paris region).[16] The difference between the extreme regions is smaller but still significant.

[*Insert Graph* 5]

We quantify the regional disparities computing the same disparity indices as for raw data for the probability of death within 1, 5, 10 and 15 days (defined as one minus the survival function of the model). Results reported in Table 4 show that the difference in the 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). In a variance-analysis spirit, we defined a pseudo-R^2 as one minus the ratio between the variance in the probability of dying within a given number of days computed from the model and the variance computed from raw data. At 1 and 5 days, the pseudo-R^2 is nearly 60%. Hence, patient characteristics and treatments would explain more than half of the regional disparities in early death. However, it is lower at 10 days (48%), and decreases even more to reach 40% at 15 days. These results suggest that part of the early regional disparities may be due to different timings of death events across regions. Also, there may be some specific regional behaviour for transfers and home returns which would affect the local composition of patients and hence would have an impact on the difference between the hospital survival functions obtained from the model and from the Kaplan-Meier estimators. Interestingly, the ranking of regions obtained for death within 15 days is not that different from the one obtained from the raw data (unweighted rank correlation: .70). This means that the ranking

(1999) estimate a wage equation that includes some firm fixed effects. They then compute some industry fixed effects as the averages of the estimated firm fixed effects for firms belonging to each industry (weighting the observations by the number of workers in the firms).

[16]Graph A3 in appendix represents the survival functions for all the regions and Table A2 ranks the regions according to survival after 15 days. Curves obtained with the model are not strictly comparable with those obtained from raw data with the Kaplan-Meier estimator as instantaneous hospital hazards were normalized with an *ad-hoc* rule. To get close to comparability, we multiplied instantaneous hospital hazards by a constant which was chosen in such a way that in absence of hospital heterogeneity (i.e. $\theta_j(t) = \theta(t)$ for all t), the expected integrated hazard at day 1 is equal to the expected integrated hazard obtained from the raw data (defined as minus the logarithm of the Kaplan-Meier estimator). This normalization allows to obtain an average survival function of the same magnitude as the one obtained from raw data with the Kaplan-Meier estimator.

of regions does not change much after taking into account individual variables.

[*Insert Table* 4]

We then supposed that each instantaneous hospital hazard writes multiplicatively as the product of a hospital fixed effect and a baseline hazard. The parameters of the multiplicative model are estimated using empirical moments as explained in the previous section. Graph 6 displays the baseline hazard and the confidence interval at each day. Remember that the weighted average of the instantaneous baseline hazards is normalized to zero. We obtain that the baseline hazard decreases sharply in the first two days and then more smoothly until the eighth day. It remains constant afterward. The sharp decrease just after entry in the hospital can be explained by violent deaths that are quite common in early days of heart attacks.

[*Insert Graph* 6]

We then regress the hospital fixed effects on a set of hospital and geographic variables. Results are reported in Table 5 (estimated regional dummies corresponding to the specification of Column 3 are reported in Appendix A3). When we only introduce hospital variables (Column 1), the adjusted-R^2 is quite low at .13.[17] It is larger at .23 when only geographic variables enter the specification (Column 2). Interestingly, when introducing both groups of variables (Column 3), the R^2 at .28 is below the sum of R^2 of the two separate regressions (.36), which suggests that variables are quite correlated. Also, it is higher than the R^2 of each separate regression, which suggests that each of the two groups has some explanatory power of its own.
We now comment on the sign of the estimated coefficients for the full specification (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 take into account innovative treatments (mainly angioplasty and stent). If we drop the variables related to innovative treatments from the first-stage specification, the propensity to die in public hospitals becomes higher than in FP hospitals (see Table A3 in appendix). Hence, the higher efficiency of FP hospitals would come from a wider use of innovative treatments. We also find that the propensity to die in a NFP hospital is lower than in a public or an FP hospital. The proportion of patients in the hospital treated for an AMI has a negative and significant effect. It is possible that hospitals concentrating AMI patients have specialized in heart-related

[17] The adjusted-R^2 takes into account the sampling error.

pathologies and thus have a higher efficiency. The number of beds as well as their occupation rate has no effect on mortality. The propensity to die is lower in hospitals with a higher proportion of beds in surgery (whether taking into account innovative treatments or not). In fact, hospitals with a high concentration of beds in surgery could have specialized in serious diseases and have a higher-quality staff. The propensity to die also decreases with the occupation rate of beds in surgery (significantly at 10% only). It is possible that hospitals with a higher occupation rate are more efficient and more likely to attract patients.

The Herfindahl index which measures the local concentration of patients has a significant negative effect. This result suggests that when patients in an area are distributed across a few large hospitals rather than many small ones, the mortality in that area tends to be lower. The number of beds in the urban area has a positive significant effect which turns out to be negative but not significant when innovative treatments are not taken into account. An interpretation can be that larger markets propose more innovative treatments but would also lead to some inefficiency in healing patients. These effects would compensate but after taking into account the innovative treatments, only the net inefficiency effect would remain. The municipality variables do not have much effect. The presence of a poor area in the municipality has a positive effect on mortality, but it is significant only at the 10% level.

At last, regional dummies always have a negative effect compared to the reference (Languedoc-Roussillon) and their effect is most often significant. Differences may be explained by unobserved regional factors such as the regional differences in hospital budgets and in the propensity to transfer patients when they are likely to die. Note that standard errors are quite large and two regions need to be far enough in the distribution of regional effects for the difference between their effects to be significant. The ranking of regional effects is only weakly correlated with the probability to die within 15 days obtained from raw data (unweighted rank correlation: .20) and with that obtained from the model (unweighted rank correlation: $-.11$).

[*Insert Table* 5]

We now study the variations in mortality at the regional level. Taking the logarithm of equation (1) with the multiplicative assumption (4), and computing the average for any region r gives:

$$\frac{1}{N^r}\sum_{i|j(i)\in r}\ln\lambda\left(t\,|X_i,j\left(i\right)\right)=\overline{X}^r\beta+\overline{\ln\alpha}^r+\theta\left(t\right) \tag{8}$$

where 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 an hospital for an AMI relates to their average characteristics, the average hospital effects, and the baseline hazard at t days. We qualitatively assess the relative explanatory power of right-hand side terms in (8) computing their variance and their correlation with the left-hand side term (in a way similar to Abowd, Kramarz and Margolis, 1999). In fact, the larger the variance and the correlation, the higher the explanatory power. In practice, as β and $\overline{\ln \alpha}^r$ are not observed, we use their estimators $\widehat{\beta}$ and $\widehat{\overline{\ln \alpha}}^r$ (the latter being defined as the regional weighted average of $\widehat{\ln \alpha_j}$) to compute the right-hand side terms. 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 explanatory power of $\overline{X}^r_s \widehat{\beta}$ for some sub-groups $\overline{X}^r_s$ of explanatory variables. Importantly, note that this procedure measures the explanatory power *ex ante* before any filtering process of patients through transfers or home returns. We can further assess the explanatory power of hospital and geographic variables. Taking the log of the expression of hospital fixed effects and averaging at the regional level, we get:

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

where $\overline{Z}^r$ and $\overline{\eta}^r$ are the regional averages of explanatory variables and random terms, respectively. We can assess the explanatory power of $\overline{Z}^r \gamma$ and $\overline{Z}^r_s \gamma$, for some sub-groups $\overline{Z}^r_s$ of explanatory variables, in the same way as for individual variables (replacing γ by its estimator).

We find that individual variables have a far larger power than hospital effects in explaining regional disparities in mortality (see Table 6a). Indeed, their variance is five to six times larger. Interestingly, among the individual variables, it is the innovative treatments which have the largest explanatory power. This means that regional disparities in innovative treatments are a key factor in explaining regional disparities in mortality. This has some important consequences for the regional funding of innovative equipment. Of course, the regional composition in age and sex also plays a role. Note that the sum of variances for groups of individual variables is far smaller than their sum. This comes from fairly large correlations between groups. In particular, regions where patients are aged and mostly females are also those in which more innovative treatments are performed (correlation between the demographic effects and the effect of innovative treatments: .57).

[*Insert Table 6a*]

The hospital and geographic effects have a larger variance than the demographic composition

effects, which suggests that their role in explaining regional disparities is significant. Concerning regional disparities in hospital fixed effects, the local composition by ownership status does not have a noticeable explanatory power (Table 6b). As regards geographic variables, the local size of the surrounding market (measured by the local number of beds except those in the patient's hospital) and the Herfindahl index play a significant role.[18] At last, residual local effects captured by regional dummies have a large variance. This means that some unobserved regional factors have a large effect on regional disparities in AMI death.

[*Insert Table 6b*]

6 Conclusion

In this paper, we studied the regional disparities in mortality for patients admitted in hospitals for a heart attack. This was done using a unique matched patients-hospitals dataset over the 1998-2003 period constructed from exhaustive administrative records. For patients, this dataset contains some information on demographic characteristics (sex and age), diagnoses and treatments. For hospitals, it gives some details on the location, status, rules of reimbursement and the capacity. We showed that regional disparities are fairly large. The difference in mortality rate between the extreme regions reaches 80%. We analyzed the causes of these disparities using a Cox duration model stratified by hospital. The model contains some patient-specific explanatory variables (demographic shifters, diagnoses and treatments), as well as a specific survival function for each hospital which is left unspecified. The flexible modelling of the hospital heterogeneity allows us to recover some robust estimators of the coefficients of patient-specific explanatory variables. These coefficients are then used in the estimation of the hospital survival functions which capture the differences in hospital behaviours when treating patients. Hospital survival functions are in turn averaged at the regional level to study the regional disparities in mortality net of the effect of patient-specific variables. Regional disparities are then lower but remain significant: the difference in mortality rate between the extreme regions is still 47%. Interestingly, the extent to which patients are treated with innovative procedures at the regional level plays a major role in

[18]Note that the local size of the surrounding market and the local concentration of patients have an effect that is positively correlated with hospital fixed effects. However, their correlation with the overall integrated hazard (last column in Table 6b) is negative. This is because these effects are more than compensated by regional fixed effects and the effects of innovative treatments.

the decrease of the disparities.

We then assessed to what extent the remaining regional disparities could be explained with hospitals' characteristics and competition between hospitals. This was done regressing hospital survival functions on hospital and geographic variables, and averaging the model at the regional level. We found that once treatments have been taken into account, the status of hospitals does not play much. By contrast, the local concentration of patients still plays a significant role. When patients in an area are distributed across a few large hospitals rather than many small ones, the mortality in that area tends to be lower. After hospital and geographic variables have been taken into account, some significant regional disparities still remain.

A limit of our analysis is that patients were not tracked in the data when they were transferred to another hospital. For patients who were transferred, we had to consider that the length of stay was 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. It should be possible to conduct such analyses in the future when the most recent data which track patients are made available for research.

7 Appendix: 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 (4) across time, weighting the observations by the number of patients at risk at each date. We obtain:

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

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:

$$\alpha_j = \frac{1}{N}\sum_t N_t\theta_j(t) \tag{9}$$

It could be possible to construct an estimator of hospital fixed effects 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 capture hospital specificities.

We also average equation (4) across hospitals, weighting by the number of patients at risk (summed across all dates) in each hospital. We get:

$$\frac{1}{N}\sum_{j} N^{j}\theta_{j}(t) = \frac{1}{N}\left(\sum_{j} N^{j}\alpha_{j}\right)\theta(t)$$

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 (9), we obtain: $\theta(t) = \left(\frac{1}{N^2}\sum_{j,t} N^{j}N_{t}\theta_{j}(t)\right)^{-1}\left(\frac{1}{N}\sum_{j} N^{j}\theta_{j}(t)\right)$. 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:

$$\widehat{\theta}(t) = \left(\frac{1}{N^2}\sum_{j,t} N^{j}N_{t}\widehat{\theta}_{j}(t)\right)^{-1}\left(\frac{1}{N}\sum_{j} N^{j}\widehat{\theta}_{j}(t)\right)$$

We then construct an estimator of a given hospital fixed effect α_j averaging equation (4) 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:

$$\frac{1}{N^{j}}\sum_{t} N_{jt}\theta_{j}(t) = \alpha_{j}\frac{1}{N^{j}}\sum_{t} N_{jt}\theta(t)$$

An estimator of the hospital fixed effect is then:

$$\widehat{\alpha}_{j} = \left(\frac{1}{N^{j}}\sum_{t} N_{jt}\widehat{\theta}(t)\right)^{-1}\left(\frac{1}{N^{j}}\sum_{t} N_{jt}\widehat{\theta}_{j}(t)\right) \tag{10}$$

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 noted $\widehat{V}_{\theta_J}$. 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:

$$\frac{\partial \widehat{\theta}(t)}{\partial \widehat{\theta}_k(\tau)} = \frac{NN^k}{\sum_{j,t} N^j N_t \widehat{\theta}_j(t)} 1_{\{t=\tau\}} - \frac{NN^k N_\tau}{\left[\sum_{j,t} N^j N_t \widehat{\theta}_j(t)\right]^2} \sum_j N^j \widehat{\theta}(t) \tag{11}$$

$$\frac{\partial \widehat{\alpha}_j}{\partial \widehat{\theta}_k(\tau)} = \frac{N_{k\tau}}{\sum_t N_{j,t} \widehat{\theta}(t)} 1_{\{k=j\}} - \widehat{\alpha}_j \frac{\sum_t N_{j,t} \frac{\partial \widehat{\theta}(t)}{\partial \widehat{\theta}_k(\tau)}}{\sum_t N_{j,t} \widehat{\theta}(t)} \tag{12}$$

In practice, to simplify the computations, we neglected the second term on the right-hand side of (12). This is only a slight approximation that does not have much impact on the estimated variance of $\widehat{\alpha}_j$. It amounts to neglect in (10) the variations of $\frac{1}{N^j}\sum_t N_{jt}\widehat{\theta}(t)$ with respect to the terms $\widehat{\theta}_j(t)$ compared to 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 (10).

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

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

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

[5] 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.

[6] 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.

[7] Deyo R., Cherkin D. and M. Ciol (1992), "Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases", *Journal of Clinical Epidemiology*, 45, pp. 613-619.

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

[9] Dormont B. and C. Milcent (2006), "Innovation diffusion under budget constraint", *Annales d'Economie et de Statistique*, Forthcoming.

[10] Duranton G. and V. Monastiriotis (2002),"Mind the Gaps: The Evolution of Regional Inqualities in the U.K. 1982-1987", *Journal of Regional Science*, 42(2), pp. 219-265.

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

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

[13] Ghali W., Hall R., Rosen A., Ash, A. and M. Moskowitz (1996), "Searching for an improved clinical comorbidity index for use with ICD-9-CM administrative data", *Journal of Clinical Epidemiology*, 49, pp. 273-278.

[14] Gobillon L., Magnac T. and H. Selod (2010), "The effect of location on finding a job in the Paris region", *Journal of Applied Econometrics*, forthcoming.

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

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

[17] 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.

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

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

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

[21] Johansson S.R. (1991), "The health transition: the cultural inflation of morbidity during the decline of mortality", *Health transition review*, 1, pp. 39-68.

[22] Jollis J., Peterson E., DeLong E., Mark D., Collins S., Muhlbaier L. and D. Prior (1994), "The Relation between the Volume of Coronary Angioplasty Procedures at Hospitals Treating Medicare Beneficiaries and Short-Term Mortality", *The New England Journal of Medicine*, 24(331), pp. 1625-1629.

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

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

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

[26] Lien H., Chou S. and J. Liu (2008), "Hospital ownership and performance: Evidence from stroke and

[27] cardiac treatment in Taiwan", *Journal of Health Economics*, 27, pp. 1208–1223.

[28] Lindeboom M. and M. Kerkhofs (2000), "Multistate Models for Clustered Duration Data - an Application to Workplace Effects on Individual Sickness Absenteeism", *The Review of Economics and Statistics*, 82(4), pp. 668-684.

[29] Lindeboom, M. and E. van Doorslaer (2004), "Cut-point Shift and Index Shift in Self-reported Health", *Journal of Health Economics*, 23, pp. 1083-1099.

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

[31] McClellan M. and D.O. 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.

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

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

[34] 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.

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

[36] Mion G. and P. Naticchioni, (2008), "The Spatial Sorting and Matching of Skills and Firms", *Canadian Journal of Economics*, forthcoming.

[37] 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.

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

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

[40] Pauly M.V. and M. Redish (1973), "The Not-for-Profit Hospital as a Physicians' cooperative", *American Economic Review*, 63, pp. 87-99.

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

[42] Picone G., Shin-Yi C. and F. Sloan (2002), "Are for-profit hospital conversions harmful to patients and to Medicare ?", *RAND Journal of Economics*, 33(3), pp. 507-523.

[43] 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.

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

[45] 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.

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

[47] 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.

[48] 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.

[49] 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.

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

[51] 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.

[52] 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.

[53] Wu S. (2001), "Adapting to Heart Conditions: A Test of the Hedonic Treadmill", *Journal of Health Economics*, 20, pp. 495-508.

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.

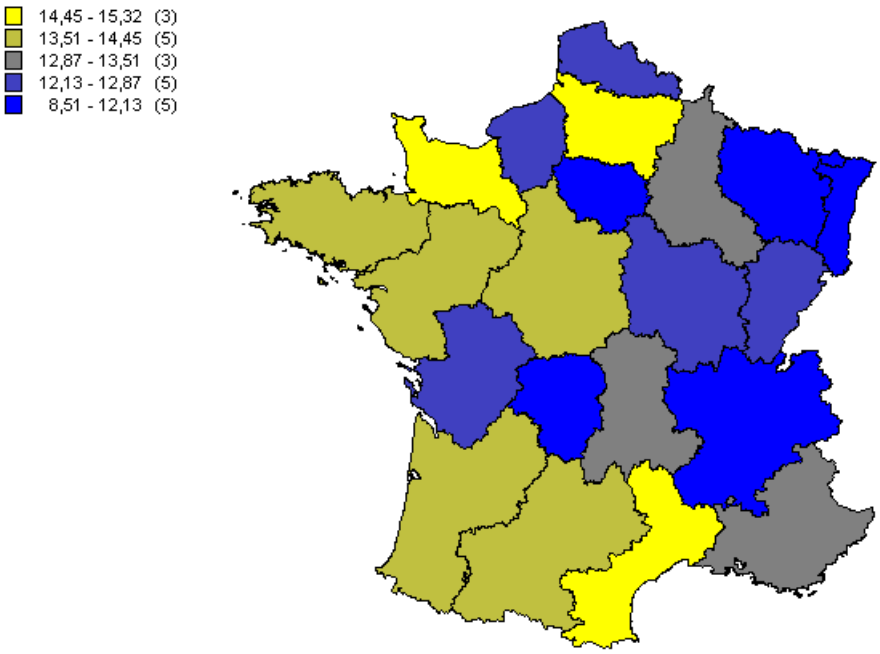
Table 9: variance analysis for the probability of death at the regional level, regression of the integrated hazard at 5 days and 15 days

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).

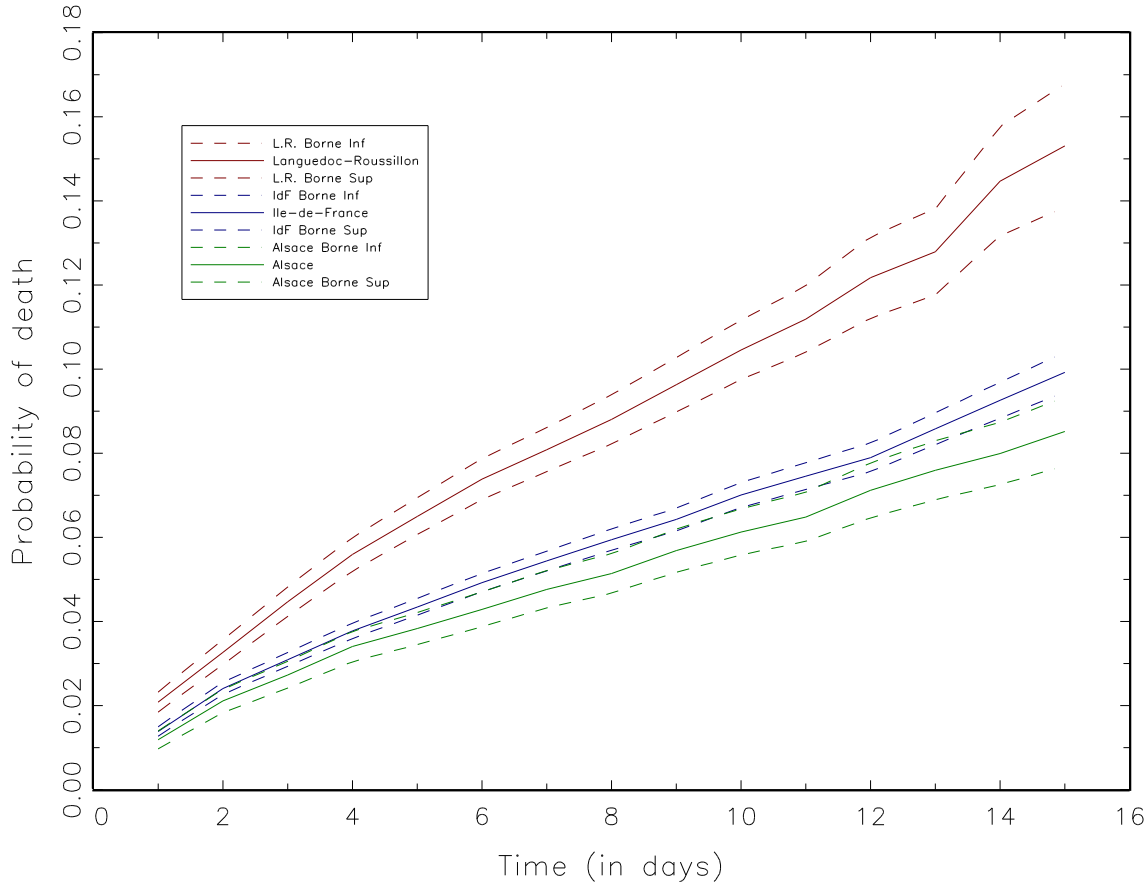
Note: we compute some regional variables from patient and hospital variables. When a variable is defined at the patient level, the corresponding regional variable is the regional average. When a variable is defined at the hospital level, the corresponding regional variable is the regional average, weighting 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.

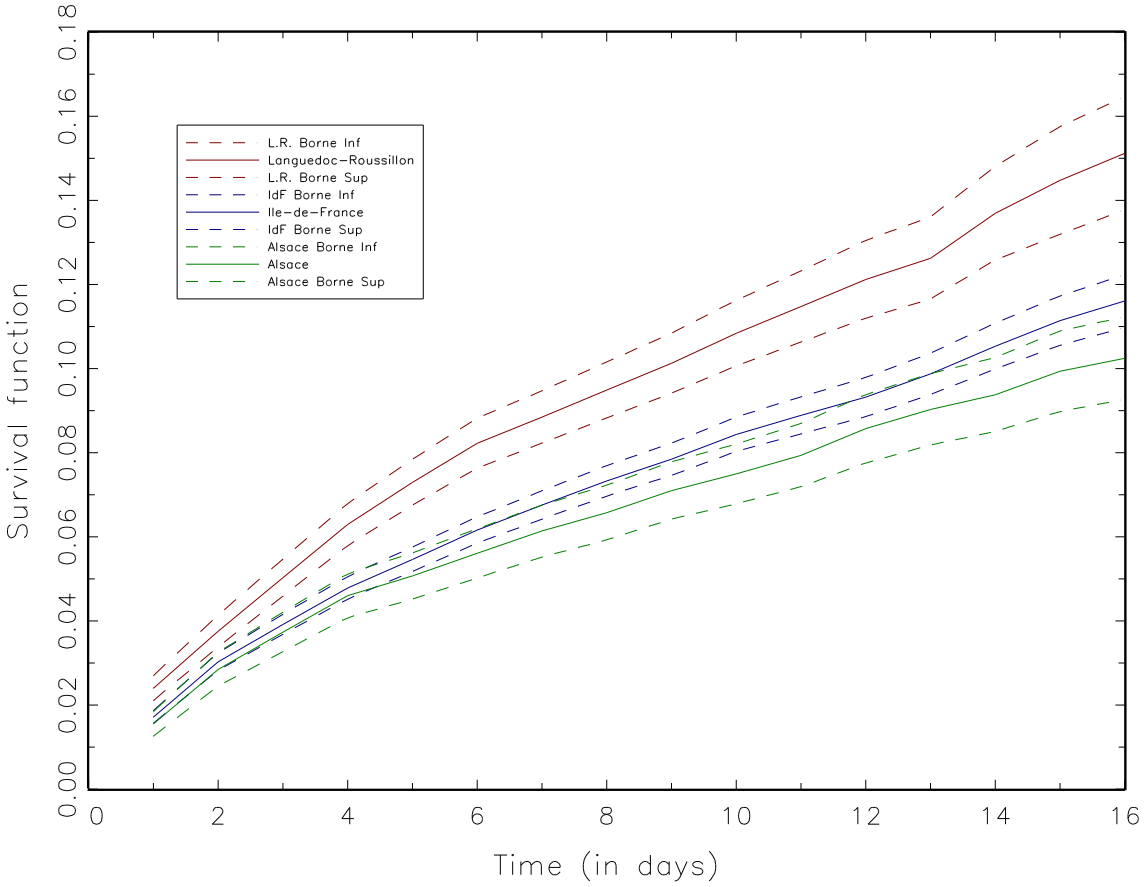
Graph 2: probability of death for extreme regions and Paris region (Kaplan-Meier)



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.

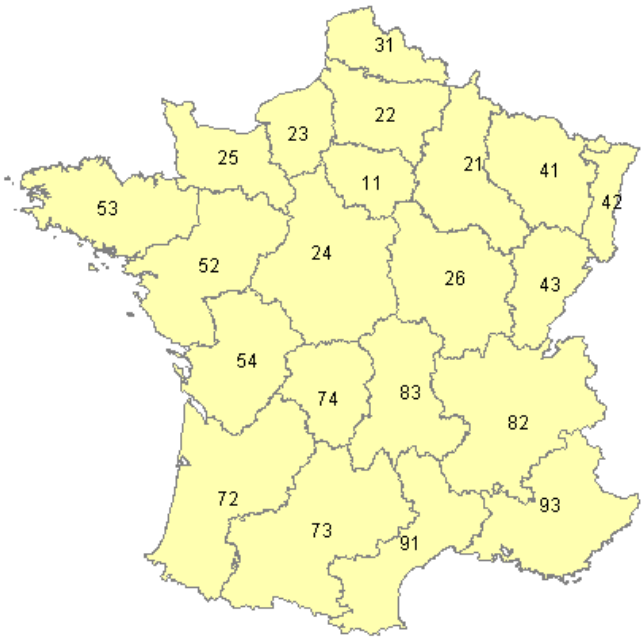
Graph 3: probability of death for extreme regions and Paris region (model)



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.

Graph A1: Map of the French Regions



Note: 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.