Network for Studies on Pensions, Aging and Retirement

Netspar THESES

Remco Mocking

Supplier Induced Demand in the Dutch Hospital Sector

MSc Thesis 2011-045

Supplier Induced Demand in the Dutch Hospital Sector

by Remco Mocking (359606)

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Econometrics and Operations Research

Faculty of Economics and Business Administration
Tilburg University

Supervisors:
Prof. dr. A.H.O. van Soest, Tilburg University
Dr. ir. R.C.M.H. Douven, CPB Netherlands Bureau for Economic Policy Analysis

August 19, 2011

Table of contents

Supplier Induced Demand in the Dutch Hospital Sector*

ABSTRACT

In this paper we empirically examine the occurrence of Supplier Induced Demand (SID) in the Dutch hospital sector. We obtained a unique panel dataset and relate the number of patients per 4-digit ZIP-code area during the years 2006-2009 to a set of demand and supply side variables including the physician density. We find that, ceteris paribus, an increase in the physician density leads to an increase in the number of patients for the major part of the sixteen treatments we analyze. The size of this effect is estimated to be substantial and dependent on the remuneration method of the physician. Depending on the type of treatment, a 1% increase in the total number of physicians is estimated to increase the number of patients by 0.1%-0.7% if these extra physicians are all paid on a fee-for-service basis.

1. Introduction

The topic of Supplier Induced Demand (SID) has been discussed extensively in the health economics literature. A clear definition of SID is given by Bickerdyke, Dolamore, Monday, and Preston (2002) who define SID as ‘the notion that doctors, in acting as agents for their patients, can use their ‘discretionary power’ to engage in demand-shifting or inducement activities such that their recommended care differs from that which an informed agent would deem appropriate’.

This definition captures the main idea that SID occurs in health care markets as a result of information asymmetries between physicians and patients (Zweifel, Breyer and Kifmann, 2009). The patient is not well enough informed to analyze his own health status or to evaluate the quality of the services supplied by the physician. In fact, the choice of medical treatment is delegated to the better-informed physician. As long as the interests of the patient and the physician are perfectly aligned, it has no consequences that demand is determined by suppliers. The concept of SID refers to these cases where interests are not perfectly aligned and the physician’s decisions on the quantity supplied are influenced by their own interests (e.g. income maintenance, full employment).

* I am grateful to CPB Netherlands Bureau for Economic Policy Analysis for giving me the opportunity to write my master thesis based on an internship. I would like to thank Rudy Douven (CPB), Arthur van Soest (Tilburg University), and Ilaria Mosca (Dutch Healthcare Authority) for their research assistance and helpful comments.

For policy makers it is important to know what role SID exactly plays in health care markets, since this enables them to prevent the possibly negative effects of SID. Negative implications of SID include the increase in health care volume and health care costs. Possible policy measures are the regulation of physician fees and the schooling of physicians.

The early paper by Evans (1974) forms the basis for the theoretical models explaining SID in health care markets. Evans assumes that physicians maximize a utility function which depends positively on income and negatively on inducement. Under these conditions, the physician is only willing to induce demand when this leads to a higher income. According to McGuire (2000), the empirical literature on SID is set predominantly within these models of physician utility maximization.

The subject of SID has been examined in a bunch of empirical studies. In one influential paper, Fuchs (1978) finds strong support for the hypothesis that US surgeons shift the demand for operations. A 10% increase in the number of surgeons increases the rate of surgery by 3% according to his estimates. In a similar study, Cromwell and Mitchell (1986) confirm these results, although the estimated inducement effect is smaller. So far, there has been very little research about SID in the Netherlands mainly due to a lack of data. Pomp (2009) finds some evidence for the occurrence of SID in the Dutch hospital sector. However, the effects on the total volume of health care are found to be only minor.

We contribute to the existing empirical literature by examining the occurrence of SID in the Netherlands using a unique panel dataset. The dataset contains per 4-digit ZIP-code area the number of patients who received a treatment in a Dutch hospital during the years 2006-2009. The analysis is restricted to sixteen different types of treatment[1]. We find strong evidence for the occurrence of SID for the major part of the treatments we analyze and the size of the inducement effects is estimated to be substantial. A 1% increase in the total number of physicians leads on average to a 0.1%-0.7% increase in the number of patients depending on the type of treatment and the remuneration method of the physician.

The set up of this paper is as follows. In the next section we briefly discuss some literature related to the empirical examination of SID in health care markets and we provide a brief description of some important characteristics of the Dutch hospital sector. We describe the datasets and present summary statistics in the third section. In the fourth section, we introduce our empirical strategy and we write down the formal model we estimate. In section

[1] The choice of treatments is in line with choice of Pomp (2009) who argues that there are a priori reasons to believe that SID might occur for the selected treatments.

five we present the results and in section six we check the robustness of our results. Finally, we conclude in section seven.

2. Theory and Dutch Setting

2.1 SID in the Netherlands

Demand inducement in the health care sector can exist as a result of information asymmetries between the physician and the patient. The patient is in general not able to evaluate the extent and the quality of the services supplied by the physician (Grytten and Sørensen, 2001). While acting as an agent of their patients, physicians have the opportunity to make decisions which are influenced by their own interests. In the literature, some conditions have been identified making the inducement of demand by physicians more likely.

Zweifel et al. (2009) mention two important conditions facilitating demand inducement. The first condition is full health care insurance coverage of patients. If patients are fully insured for the cost of medical treatment, the price of medical treatments does no longer play a role. That is, no limit is put to the consumption of medical services. The second condition deals with the risks associated with the medical treatment. Whenever it is unlikely that the medical treatment harms the patient, the physician will be more inclined to provide the treatment although the patient does not really need it. Ambiguity in the diagnosis is mentioned by Pomp (2009) as a third condition facilitating the occurrence of SID. A more trivial condition facilitating the occurrence of SID deals with the incentives of the physician. A physician will only create his own demand when an incentive is provided to do so. This idea is formalized by Evans (1974), who assumes that the physician's utility depends positively on income and negatively on the level of inducement. In other words, the physician is only willing to induce demand if this leads to a higher income.

The first condition is certainly satisfied in the Dutch case. Patients are fully insured for the costs of the medical treatments we consider. The second condition is at least partly satisfied for the treatments we selected. According to Pomp (2009), the risk of harming the patient is low for treatments such as cataract surgery and tonsillectomy. Moreover, the diagnosis is ambiguous for these two treatments, which implies that the third condition is met as well. However, we also analyze more risky treatments with a more unambiguous diagnosis (e.g. breast reduction). Consequently, we expect to find differences in the extent of inducement according to the riskiness of the medical treatment and the ambiguity regarding the diagnosis.

The fourth condition is satisfied in the Dutch case as well. In the Netherlands, there are two types of physicians. About 30% of the physicians working in the Netherlands receive a fixed wage, whereas the other 70% is paid on a fee-for-services basis. Fee-for-service physicians have a strong financial incentive to induce demand. By increasing the number of patients they treat, fee-for-service physicians are able to increase their own income. For our research it is important to notice that more than 99% of the observations in our dataset concern treatments in the so-called B-segment where financial incentives were prevalent during the period 2006-2009. This is in contrast with the A-segment, where financial incentives were introduced at the beginning of 2008. For fixed wage physicians, the financial incentive to create their own demand is not prevalent. However, they might face an indirect incentive when the hospital management stimulates them to induce demand in order to increase hospital turnover. All in all, we expect to find a stronger inducement effect for fee-for-service physicians than for fixed wage physicians.

Given that the conditions facilitating SID are all (at least partly) satisfied, it is interesting to continue with an empirical investigation of SID in the Dutch hospital sector. The next section deals with the problems related to the empirical examination of SID.

2.2 Identification

Reviewing the empirical literature, we can identify various strategies to investigate the prevalence of SID. In the spirit of Cromwell and Mitchell (1986) and Fuchs (1978) we relate the physician/population ratio to the utilization of medical services in all 4,000 4-digit ZIP-code areas in the Netherlands. A positive relationship between these two variables suggests that SID might occur in the Dutch hospital market. We include both demand factors and supply side factors in our model to explain utilization of medical services. In contrast to the study by Cromwell and Mitchell, we do not need to take price effects into account since patients are fully insured for the costs of the medical treatments we consider.

As in Grytten and Sørensen (2001) and Pomp (2009), we distinguish between two types of physicians: physicians with a fixed wage contract and fee-for-service physicians. Only the latter type of physicians has a clear financial incentive to induce demand. Compared to the study on SID in the Netherlands by Pomp (2009) our dataset is much richer. We collected data for two extra years and in contrast to the study by Pomp (2009) we obtained the actual number of physicians working in the Dutch hospitals during the years 2006-2009[2].

[2] Pomp (2009) uses the number of physicians working in Dutch hospitals in 2004 as a proxy for the number of physicians in the years 2006-2007.

The structure of the dataset allows us to use panel data techniques. The only study on SID using panel data we are aware of is the study by Delattre and Dormont (2003). They find strong evidence for the occurrence of SID in the French system for ambulatory care. The advantage of using panel data techniques is that we can take unobserved heterogeneity between ZIP-code areas into account by specifying models with random or fixed individual effects. Still, we cannot simply state that physicians induce demand when we observe that per capita utilization of medical services increases with physician density. According to Zweifel et al. (2009), there are three alternative reasons explaining this empirical observation.

The first reason is that there might be permanent excess demand for medical services. That is, the number of physicians is too low to serve all patients. When the number of physicians is higher in some area, the relative number of treatments will also be higher in that area. So when there is permanent excess demand, we empirically observe the same as we expect to observe when SID exists. Fuchs (1978) takes this point into account by referring to a report on the US surgical market. This report states that the number of operations per surgeon is far below the level that surgeons consider a full workload. We control for the possibility of demand rationing by including data on the length of waiting lists in our model.

Secondly, the outward shift of the demand curve associated with increasing physician density can be explained by a decrease in indirect costs of receiving a medical treatment. One could argue that the demand for medical treatments is linked to the indirect costs incurred by patients. When new hospitals are opened, non-financial costs such as travel time are reduced. This will shift the demand curve outward. We include data on the availability of hospitals to control for this possibility.

Finally, we need to control for the possibility of reverse causality. It is very well possible that physicians locate themselves at places where the demand for their services is expected to be high enough. We control for these location effects by including a large set of factors which are possibly related to the demand for medical services. However, it is still possible that we omitted a variable. Using Ordinary Least Squares (OLS), this will lead to biased estimates of the model parameters (omitted variable bias). According to Verbeek (2008), the advantage of the panel data we obtained is that we may (partly) resolve this problem by including ZIP-code area specific effects. Similarly, we can include fixed time effects to capture the effect of all variables that do only vary over time (and not over the individual ZIP-code areas). This explains why the use of panel data can reduce the effects of omitted variable bias.

Although the panel structure of our data reduces the effects of omitted variable bias, it is still possible that the physician density is an endogenous variable. In an attempt to deal with this problem we adopt an Instrumental Variable (IV) approach as in Fuchs (1978) and Cromwell and Mitchell (1986). We use two-stage least squares (TSLS) and generalized method of moments (GMM) to replace the actual physician/population ratio by a fitted value using a set of instruments. These instruments must be related to the physician density, but may not be directly related to the demand for medical treatments. Next, the panel structure of the data allows us to use lagged physician densities as instruments.

2.3 Control treatments

In their paper Dranove and Wehner (1994) test the power of the IV-approach by applying it to a case where induced demand surely does not exist: the demand for child births. The idea is that if they find evidence for the occurrence of SID using the same techniques as in Fuchs (1978) and Cromwell and Mitchell (1986), the results of these studies would be suspect. Dranove and Wehner (1994) indeed find evidence for the inducement of child births, which leads them to call the validity of the IV-approach into question. They give two explanations for their findings: (1) the instruments used in earlier studies are directly related to demand for medical services (instrument exogeneity is violated); (2) the failure to account for border crossing.

In order to address the first possibility we make an attempt to find instruments which are related to the supply of physicians, but not directly related to the demand for medical treatments. The fact that we obtained information on the visited hospital and the place of residence for every patient enables us to take border crossing into account. When Dranove and Wehner (1994) perform their IV-estimations at a more geographically aggregated level to take border crossing into account, they find no statistical evidence for the inducement of child births. This enforces the idea that an IV-approach might work in our case if we can find the right instruments.

In order to control for the possibility that we measure something different than the occurrence of SID we include hip fractures as a control treatment in our analysis. For hip fractures, the diagnosis is clear and an operation or treatment is required. In Wennberg (2010) it is shown that the incidence of surgical repair of the hip is closely related to the actual incidence of hip fractures in the United States. In other words, the number of physicians per capita should not be related to the per capita number of hip fractures.

3. Data Description

In this section we give a description of our data. Section 3.1 introduces our main dataset from the Dutch Healthcare Authority (NZa). This dataset contains information on the yearly number of patients in all Dutch hospitals during the years 2006-2009 for the treatments we selected. In section 3.2 we describe the dataset from Statistics Netherlands (CBS) containing the demand factors we include in our model and in section 3.3 we describe the supply side data and the instrumental variables.

3.1 Patients

We obtained DBC information system (DIS) data from the Dutch Healthcare Authority (NZa) on the number of patients in 177 hospitals, specialist hospitals, and private hospitals in the Netherlands during the years 2006-2009. In the Netherlands medical treatments are categorized using DBC's (Diagnose Behandeling Combinatie). A DBC describes the total health care package received by the patient including the diagnosis and the selected treatment. A DBC is said to be 'opened' at the moment of the first consult of the patient. The DBC is closed at the moment the patient leaves the hospital after the last medical examination. The maximum length of a DBC is one year.

We identify sixteen different treatments by means of the DBC's and a medical specialty code. In our terminology, one treatment includes in fact multiple unique DBC's corresponding to variations of the same medical treatment. As an example, Table A1 shows that the treatment of cataract consists of eight unique DBC's.

The total number of DBC's in our dataset is equal to 1,906,668. For each DBC in our dataset we know the year in which it is opened and closed, the ZIP-code of the patient (4-digit), the age and gender of the patient, and the ZIP-code of the hospital (6-digit). The dependent variable in our model is for every treatment defined as the number of patients per 1,000 inhabitants, per year, and per 4-digit ZIP-code area. We choose to analyze the data according to the year in which the DBC is opened. The reason is that the major part of the DBC's (about 75%) is opened and closed during the same year.

The data is summarized in Table A2 showing the total number of DBC's, the average age of the patients, and the share of men for the different treatments during the years 2006-2009. From Table A2 we can see that the number of patients is rising for most of the treatments during the years 2006-2008. In 2009 we observe somewhat lower numbers of patients. The reason is probably that the data for 2009 is not complete yet. For kidney stones and urethral stones there is a jump in the number of patients after 2006. It is not completely clear what causes this jump. However, we do know that from the beginning of 2007 the number of unique DBC-codes related to kidney stones and urethral stones more than doubled[3].

3.2 Demand

We merge the patient dataset with a dataset containing a wide range of variables at the 4-digit ZIP-code level that are possibly related to the demand for medical treatments. At the 4-digit level, the number of unique ZIP-code areas increases from 4,007 areas in 2006 to 4,019 areas in 2009 as a result of newly built quarters. The dataset is obtained from Statistics Netherlands (CBS) and includes, among others, demographic and socio-economic factors.

In order to make a relevant comparison between ZIP-code areas, we calculate the number of patients and the number of physicians per inhabitant. We collected data on the population size of all 4-digit ZIP-code areas in the Netherlands and we use this data to detect outliers. For very small areas, the dependent variable shows a lot of variation, whereas this variation becomes much lower when we look at larger areas. Furthermore, in small areas there are a lot of missing values and outliers for the demand variables. Therefore we decided to exclude 4-digit ZIP-code areas with less than 500 inhabitants from our analysis. This amounts to the exclusion of about 850 ZIP-code areas every year, but we lose on average only about 1% of the total number of patients. Table A3 gives an overview of the yearly number of ZIP-code areas and the average population sizes. We continue with a description of the demand variables included in our model.

For each 4-digit ZIP-code area, we can distinguish between twenty age groups ranging from 0-5 years to >95 years. The percentage of inhabitants belonging to a specific age group is captured by the variables *age0_5*,..., *age90_95*, and *age95*. The variable *men* includes the percentage of men living in each 4-digit ZIP-code area. We also have data on the relative number of foreigners living in each area. Someone is called a foreigner if at least one of his parents is born in another country than the Netherlands[4]. The group of foreigners is divided into non-western foreigners (*nonwesterns*) and western foreigners (*westerns*). The category of non-western foreigners includes foreigners originating from Turkey, Africa, Latin-America, and Asia with the exception of Japan and Indonesia. The category of western foreigners includes foreigners originating from Europe, Northern-America, Oceania, Indonesia, and Japan. The variable *mortality* refers to the mortality rate in the ZIP-code area, which is

[3] When excluding the year 2006 from the analysis we find similar results for kidney stone and urethral stones.
[4] This definition includes both the so-called first generation and the second generation of foreigners.

defined as the number of deceased per 1,000 inhabitants[5]. Since 2009 mortality rates are not available in our dataset, we approximate them by using the numbers of 2008. The variable *urbanized* is an indicator for the degree of urbanization of the ZIP-code area. The indicator depends on the number addresses per square kilometer. The variable takes values between 1 and 5, where a value of 1 denotes a very high degree of urbanization and a value of 5 denotes a very low degree of urbanization.

With respect to work and income we collected data on the working population, the distribution of income, and the number of people on welfare. The variable *lowincome* gives the percentage of people in the lowest 40% of the national income distribution, whereas the variable *highincome* gives the percentage of people in the highest 20% of the national income distribution. For 2008, no income data is available. We compute the average of 2007 and 2009 and used this as an approximation for *highincome* and *lowincome* in 2008. The number of working people as a percentage of the total population between 15 and 65 years is denoted by the variable *working*. The variable *selfemployed* gives the number of self-employed persons as a percentage of the total number of working persons. The *working* and *selfemployed* numbers are not available for 2009. Therefore we use 2008 values as a proxy for *working* and *selfemployed* in 2009. Finally, the number of people receiving social assistance per 1,000 inhabitants is captured by the variable *assistance*.

Table A3 shows for each year the mean value of all demand factors[6]. Frequency histograms showing the distributions of our demand variables are presented in Appendix B (Figure 1-Figure 8). ZIP-code areas with less than 500 inhabitants are already excluded from these figures. The histograms show that the distributions of our demand variables hardly change over time. This supports the idea of using observations in subsequent or preceding years as proxies in case of missing values.

3.3 Supply

We obtained data from Dutch Hospital Data Foundation (DHD) on the number of physicians working at the various specialties in Dutch hospitals. The number of physicians is expressed in full time equivalents, which enables us to make an appropriate comparison. The data is available for about 85% of the observations in our dataset, since physician data is not available for all hospitals. The data includes information on the type of contract of the

[5] The mortality rate could also be considered as an outcome variable in the sense that access to better health care should result in a lower mortality rate. However, for the treatments we analyze this issue does not seem to play an important role.

[6] *highincome* and *lowincome* are excluded since the averages are by definition 20% and 40% respectively.

physicians. The variable *FFS* gives the number of doctors working on a fee-for-services basis, whereas the variable *FW* denotes the number of physicians receiving a fixed wage. A substantial share of the fixed wage physicians is working in a teaching hospital. For 2006 we have no information about the number of physicians working in a teaching hospital and we use the values of 2007 as a proxy for the number of teaching hospital physicians. The total number (in fulltime equivalents) of each type of physician per specialty is presented in Table A4.

We include other variables related to the supply of hospital care in our model as well. The variable *waitinglists* gives per ZIP-code area the average number of weeks a patient has to wait before he can get a treatment in the hospital. We have waiting list data from the Dutch Healthcare Authority (NZa) for seven out of the sixteen treatments we analyze. For the other treatments, waiting list data is not available or there are no waiting lists (e.g. for hip fractures there is no waiting list since it concerns an emergency operation). Table A5 gives an overview of the waiting list data.

Furthermore, we include data on the availability of health care providers in the various ZIP-code areas. As in Fuchs (1978) and Cromwell and Mitchell (1986) we include sources of alternative care which might impact demand: the distance to the closest GP (*distGP*), the number of GP's within three kilometer of the ZIP-code area (*av3GP*), and the distance to the closest GP centre (*distGPcentre*). Next, we include the number of hospitals within twenty kilometer of the ZIP-code area (*av20hospital*) measuring the availability of hospitals. This is also in line with Cromwell and Mitchell (1986) who include the number of hospital beds per 1,000 inhabitants as an availability measure. These numbers were only available for the years 2007 and 2008. Since the numbers are relatively constant over these two years, we decided to use the values of 2007 and 2008 as proxies for the years 2006 and 2009 respectively.

Finally, we include instrumental variables in order to predict physician densities. Fuchs (1978) finds that US physicians prefer to live in non-metropolitan areas which most people consider desirable to visit. It is not clear if this will also be the case in the Netherlands. One could also argue that physicians have a preference for more urbanized areas with more (cultural) facilities. We use an indicator for house prices (*IVhouseprice*), the number of cinemas (*IVcinema*) within a ten kilometer radius, and the distance to the closest public park (*IVdist_park*) as instruments. The latter two variables are only available for 2008, so we assume that they remain constant over time. Moreover, we add a variable measuring the distance to the closest teaching hospital as an instrument (*IVdist_TH*). Cromwell and Mitchell (1986) also include this as an instrument by arguing that physicians want to reside in a professional environment. In the Dutch case, the motivation for including *IVdist_TH* as an instrument is that all physicians working in a teaching hospital receive a fixed wage. So in the neighborhood of teaching hospitals we expect a higher number of fixed wage physicians. The mean values for all years of our instrumental variables and availability factors are shown in Table A3.

4. Methodology

4.1 Demand strategy

The strategy we follow is comparable to the approach of Fuchs (1978), Cromwell and Mitchell (1986), and Pomp (2009). We are interested in the effect of the number of physicians per capita (the physician density) on the number of patients per capita (patient density). We perform our analysis on the 4-digit ZIP-code level over the years 2006-2009. In case of no SID we expect that the number of treatments in a certain area can be explained by the demand factors. Our strategy is to relate the patient density to a set of demand factors in a first step using pooled Ordinary Least Squares (OLS).

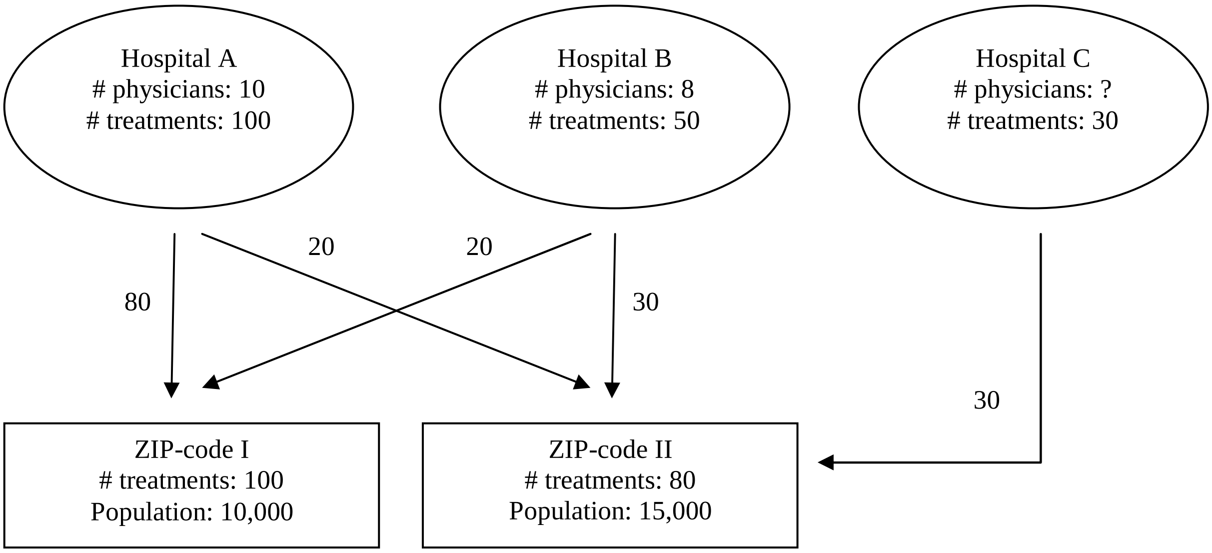
We use pooled OLS, because we are interested in the error terms. The error terms of the demand equation can help us to give a first indication for the occurrence of SID. For the moment we do not include ZIP-code specific effects, since these effects would also (partly) capture possible inducement effects. A high positive error term in a specific area might be a first indication for SID, since it means that the actual number of patients is higher than predicted by demand factors. If SID occurs, we expect to find indications for inducement in the same areas every year since there are no large fluctuations in demand factors and the location of both people and physicians hardly changes over time[7]. Next, we expect that inducement is geographically concentrated since we observe from our data that people in adjacent areas usually visit the same hospitals and thus the same physicians. To put it more formally, in case of SID: (1) the error terms of our demand equation are correlated over time; (2) the error terms are spatially correlated. It is important to notice that these two conditions are only necessary, but not sufficient conditions for the occurrence of SID.

[7] The correlations over time of the population size are about 0.99 for every year. The correlations over time of the number of physicians per hospital are between 0.87 and 0.99 for all specialties.

4.2 Supply strategy

In a second step we include supply side factors and we use panel data techniques (Random Effects and Fixed Effects) and instrumental variables in order to estimate the effect of the physician/population ration on utilization of medical care. The key problem we face is the allocation of physicians to ZIP-code areas. We have data on the number of physicians per hospital and we want to transform this data to a number of physicians per ZIP-code area. This is necessary because our dependent variable, as well as all other independent variables, is on the 4-digit ZIP-code level. The allocation problem is illustrated in Figure 1 below.

Figure 1: allocation of physicians to ZIP-code areas



In contrast to Fuchs (1978), the data allows us to exactly detect border crossing of patients. We know both the name of the hospitals visited by the patient and the place of residence of the patient. In our example the allocation of physicians to ZIP-code area I is straightforward. A total number of 100 patients receive a medical treatment in area I. The patients are divided between hospital A and hospital B: eighty patients receive a treatment in hospital A, whereas twenty patients receive a treatment in hospital B. In hospital A, 100 patients are treated. So, patients from area I use 80% of the total capacity of hospital A. Since a total number of ten physicians are working in hospital A, we allocate eight doctors of hospital A to area I. Similarly, we allocate 3.2 out of the eight physicians working in hospital B to area I. This gives us a total number of 11.2 physicians working in area I. The physician density is found by dividing the number of physicians by the population size.

For area II, the allocation of the right number of physicians is more complicated. We have no data on the number of physicians working in hospital C. In order to calculate the right physician density, we have to adjust the population size. A total number of 80 patients living in area II receive a treatment, but only for 50 patients we know the corresponding number of physicians. In order to prevent an underestimation of the physician density, we scale down the population size by a factor of 50/80.

We believe that a physician density must say something about the access people have to physicians. When we compute physician densities for every 4-digit ZIP-code area separately, there may originate differences between very small adjacent areas. For the Netherlands, it is not realistic to calculate unique physician densities for such small areas. Moreover, a problem arises when there are zero patients in a certain area. Using the allocation method described above, this will lead to a physician density of zero. This would imply that we can only find positive correlation between patient density and physician density. For that reasons we decided to calculate physician densities at the more geographically aggregated 2-digit ZIP-code level. So each 4-digit ZIP-code area within a certain 2-digit ZIP-code area is assumed to have the same physician density. In the following subsection we will write down the model we estimate in a more formal way.

We allocate physicians for each treatment separately because this is the best way to take into account that physicians working at one and the same specialty perform multiple treatments. By doing so, we take into account that physicians in one hospital might spend a higher share of their time on one specific type of treatment than physicians in another hospital. As a consequence we calculate for example six different orthopedist densities, since we analyze six different treatments performed by orthopedists.

4.3 Formal model

In this subsection we will write down the model in a more formal way. The number of patients living in 4-digit ZIP-code area i, treated in hospital h, in year t is denoted by $T_{4,iht} \in \mathbb{R}_+$ for all $i = 1, \dots, I$, for all $h = 1, \dots, H$, and for all $t = 1, \dots, T$. The same number, but then on the 2-digit ZIP-code level is denoted by $T_{2,iht}$. We do not use a subscript for the specific medical treatment, because the same model is estimated separately for all sixteen treatments included in the analysis. In our model we distinguish between two types of physicians: physicians of type FW receive a fixed wage, whereas physicians of type FFS are paid per treatment. The number of physicians of type $\theta \in \{FW, FFS\}$ working in hospital h, in year t is given by $P_{\theta ht}$ (fulltime equivalents). Z_{it} denotes the vector of demand factors for 4-digit ZIP-code area i in year t and S_{it} denotes the vector of remaining supply side variables.

Finally, $N_{4,it}$ represents the total population of 4-digit ZIP-code area i in year t and $N_{2,it}$ stands for the population size on the 2-digit ZIP-code level.

The dependent variable in our model is the patient density in 4-digit ZIP-code area i in year t and is denoted by y_{it}:

$$y_{it} = \frac{\sum_{h=1}^{H} T_{4,iht}}{N_{4,it}} \tag{1}$$

4.3.1 Demand

We start with relating our dependent variable to the vector of demand factors Z_{it}. We also include year dummies to control for year specific effects and estimate the following demand equation using pooled OLS:

$$y_{it} = \alpha + \gamma_t + Z_{it}^{'}\beta + \varepsilon_{it} \tag{2}$$

In case SID occurs, we expect that the error terms (ε_{it}) are not randomly distributed. Since the location of both people and physicians remains relatively stable over time, we expect that the error terms exhibit positive autocorrelation and positive spatial correlation in that case. If we find autocorrelation and spatial correlation in the error terms, this is a first indication for the occurrence of SID. It does not prove the existence of SID, since the correlation can also be the result of omitted demand variables.

We test for the existence of autocorrelation by relating the error terms of ZIP-code area i in successive years to each other. We estimate equation (3) using pooled OLS and formulate our first hypothesis.

$$\varepsilon_{it} = \alpha + \rho_{AC}\varepsilon_{i,t-1} + \eta_{it} \tag{3}$$

HYPOTHESIS I: *a necessary, but not sufficient, condition for the occurrence of SID is that in equation (3) the estimated coefficient* $\hat{\rho}_{AC} > 0$.

In order to test for spatial correlation we relate the error terms of ZIP-code area i to the unweighted average of the errors term in the five nearest ZIP-code areas (ε_{wit}). For every year $t = 1,\ldots,T$ we estimate equation (4) using OLS and we formulate the second hypothesis.

$$\varepsilon_{it} = \alpha + \rho_{SC}\varepsilon_{wit} + \nu_{it} \tag{4}$$

HYPOTHESIS II: *a necessary, but not sufficient, condition for the occurrence of SID is that in equation (4) the estimated coefficient* $\hat{\rho}_{SC} > 0$.

4.3.2 Supply

In this subsection we formally discuss the allocation of physicians to ZIP-code areas and the inclusion of other supply factors in our model. We have data on the number of physicians working in hospital h, in year t. However, we want to calculate physician densities for each ZIP-code area i, in year t. We will calculate these densities on the 2-digit ZIP-code level for reasons explained in section 4.2.

Before we can start with the allocation of physicians to ZIP-code areas we need to take into account that the hospital entered a non-existing patient ZIP-code for about 6% of the observations. For these patients no demand data is available as a result and therefore the observations are not included in our analysis. For that reason, we need to correct the number of physicians working in each hospital. If we do not correct the number of physicians, we would allocate physicians who treated patients with a non-existing ZIP-code to areas with an existing ZIP-code. As a result, we would overestimate the number of physicians (and thus the physician density) in the areas included in our analysis. The corrected number of physicians of type θ working in hospital h in year t is given by equation (5), where $a_i = \begin{cases} 1 \text{ if ZIP-code i exists} \\ 0 \text{ otherwise} \end{cases}$.

$$\tilde{P}_{\theta ht} = P_{\theta ht}\left(\frac{\sum_{i=1}^{I} a_i T_{4,iht}}{\sum_{i=1}^{I} T_{4,iht}}\right) \tag{5}$$

Using equation (5), we can start with the allocation of physicians to 2-digit ZIP-code areas. Following Figure 1 in section 4.2, the number of physicians allocated to 2-digit ZIP-code area i, in year t is given by equation (6).

$$P_{\theta it} = \sum_{h=1}^{H}\left(\frac{T_{2,iht}}{\sum_{i=1}^{I} T_{2,iht}}\tilde{P}_{\theta ht}\right) \tag{6}$$

From equation (6) we can calculate the corresponding physician density for area i in year t. We only need to correct the population size for the possibility that patients visited hospitals for which no physician data was available. In order to do so, we count the number of patients in area i in year t who visited hospitals for which we do know the number of physicians and divide it by the total number of patients in area i in year t. We multiply this ratio with the actual population size to obtain the corrected population size as is shown by equation (7), where $b_h = \begin{cases} 1 \text{ if we observe the number of physicians in hospital } h \\ 0 \text{ otherwise} \end{cases}$.

$$\widetilde{N}_{2,it} = N_{2,it}\left(\frac{\sum_{h=1}^{H} b_h T_{2,iht}}{\sum_{h=1}^{H} T_{2,iht}}\right) \tag{7}$$

We calculate the physician density for physicians of type θ, in ZIP-code area i, and in year t using equation (6) and equation (7). The physician density $x_{\theta it}$ is on the 4-digit ZIP-code level, but the density is the same for every 4-digit ZIP-code area within a certain 2-digit ZIP-code area.

$$x_{\theta it} = \frac{P_{\theta it}}{\widetilde{N}_{2,it}} \tag{8}$$

We add the resulting physician densities, together with the other supply side variables S_{it} (such as waiting lists, availability of hospitals), to our model of equation (2). So we try to explain the patient density from demand factors and supply factors in one equation. We do not need to take price effects into account as patients are fully insured for the costs of the medical treatments we consider. This specification is similar to Fuchs (1978) and Cromwell and Mitchell (1986), who estimate a comparable equation using instrumental variables in order to take the possible endogeneity of the physician density $x_{\theta it}$ into account. We first estimate equation (9) using pooled OLS and panel data techniques (Random Effects and Fixed Effects) allowing for a ZIP-code specific effect α_i. We finish by estimating equation (9) using two stage least squares (TSLS) and generalized method of moments (GMM) to control for the possible endogeneity of $x_{\theta it}$.

$$y_{it} = \alpha_i + \gamma_t + \delta_{FW} x_{FW,it} + \delta_{FFS} x_{FFS,it} + Z_{it}^{'}\beta + S_{it}^{'}\zeta + \varepsilon_{it} \tag{9}$$

We expect that physicians only induce demand for the services they provide in case an incentive is provided to do so. We find an indication for the occurrence of SID if the fee-for-service density is positively related to the patient density and there is no or at least a weaker relationship between the fixed wage density and the patient density. In order to compare the effect of the two physician densities, we need to calculate the relevant marginal effects. We are interested in the ceteris paribus effect on the number of patients if the number of physicians increases by 1%. Note that a 1% increase in the number of physicians (patients) is equivalent to a 1% increase of the physician density (patient density). We have to take into account that the FW-density is on average much lower than the FFS-density. That implies that a 1% increase of the FW-density would have a much lower impact on the patient density than a 1% increase of the FFS-density would have, even if both types of physicians induce demand to the same extent. Therefore, we calculate the effect on the number of patients of a 1% increase of the total number of physicians if (1) this increase is caused by an increase in the number of FW-physicians only (2) this increase is caused by an increase in the number of FFS-physicians only. The resulting estimated ceteris paribus effects are denoted by $\hat{\xi}_\theta$.

$$\hat{\xi}_\theta = \frac{\hat{\delta}_\theta \sum_\theta \bar{x}_\theta}{\bar{y}} \tag{10}$$

HYPOTHESIS III: *when SID occurs, we expect to find that* $1 \geq \hat{\xi}_{FFS} >> \hat{\xi}_{FW} \geq 0$.

Besides that, we expect to find that for our control treatment (hip fractures) the physician density is unrelated to the patient density irrespective of the remuneration method. This hypothesis controls for the possibility that we measure something different than SID.

HYPOTHESIS IV: *for hip fractures we expect to find that* $\hat{\xi}_{FFS} = \hat{\xi}_{FW} = 0$.

The panel structure of our data also allows us to use another technique to overcome the endogeneity problem of the physician density. We can use lagged values of the physician densities as instruments. This means that we can no longer use the FE-estimator since the strict exogeneity condition ($E\{x_{\theta it}\varepsilon_{is}\} = 0$ for all s, t) will be violated. The first-difference estimator is consistent under the weaker condition that $E\{(x_{\theta it} - x_{\theta i,t-1})(\varepsilon_{it} - \varepsilon_{i,t-1})\}$. Therefore, we first take first-differences to eliminate the ZIP-code specific effects and year dummies:

$$y_{it} - y_{i,t-1} = (x_{\theta it} - x_{\theta i,t-1})' \delta_\theta + (Z_{it} - Z_{i,t-1})' \beta + (S_{it} - S_{i,t-1})' \zeta + (\varepsilon_{it} - \varepsilon_{i,t-1}) \tag{11}$$

If the physician densities are indeed endogenous, then the term $(x_{\theta it} - x_{\theta i,t-1})$ is correlated with the error term $(\varepsilon_{it} - \varepsilon_{i,t-1})$. Under the assumption that $x_{\theta i,t-1}$ is not correlated with ε_{it}, we can instrument the term $(x_{\theta it} - x_{\theta i,t-1})$ by $(x_{\theta it-2} - x_{\theta i,t-3})$. We have data for four years, which is just enough to apply this method. The disadvantage of this approach is that we lose about 75% of the data. For that reason, we first estimate equation (11) using OLS and compare the results to what we find from our estimation of equation (9). Finally, we use two-stage least squares (TSLS) with the lagged first-differenced densities as instruments to estimate the effect of the number of physicians on the number of patients.

5. Results

5.1 Demand

We estimate equation (2) using pooled OLS for all sixteen treatments. As an example we present the estimation results for cataract surgery in Table A6. Since mostly older people undergo cataract surgery (see Table A2), we find a significant positive effect of the age groups *age75_80* and *age80_85* (at the 5% significance level). The coefficients of the other age groups are not reported, but they are all smaller and insignificant. The year dummies are all highly significant and the estimates are in line with the annual variation in the number of treatments as we observed in Table A2. Three other variables are significant. A higher percentage of western foreigners results, ceteris paribus, in a lower number of patients. The coefficient of *urbanized* suggests that patient density is on average higher in urban areas compared to rural areas. Also the percentage of self-employed is significant indicating that they seek less hospital care.

We are mainly interested in the error terms of the demand equation, since they can give us a first indication for the occurrence of SID. The autocorrelations and spatial correlations are provided in Table A7. We present also the Mean Absolute Error (MAE) and the average patient density ($\bar{y}$) to get an impression of the size of the error terms.

The conclusion that we draw from Table A7 is that we cannot exclude the occurrence of SID for most of the treatments we analyze. For all treatments except hip fractures in young patients, Hypothesis I and Hypothesis II are confirmed. That is, the estimated autocorrelations and spatial correlations are all positive and statistically significant at a significance level of at least 0.05, while the major part of the correlations is significant at the 0.001 level. Only for hip fractures in young patients we do not find significant positive autocorrelations and spatial correlations. From Table A7 it also follows that the relative size of the error terms is relatively larger for treatments with a low patient density. In line with our expectations, we can exclude the possibility of inducement for hip fractures in young people. We do find positive correlations for hip fractures in all patients. This is in contradiction with our expectations and this could be an indication for SID. However, it does not prove the occurrence of SID so we cannot draw a conclusion on the inducement of hip fractures yet.

5.2 Supply

We estimated equation (9) using the pooled OLS estimator, the random effects estimator, and the fixed effects estimator. We provide the full results of our estimations for cataract surgery as an example in Table A8. We find a positive and significant effect of the physician densities on the patient density in all three specifications. The effect of waiting lists in the RE and FE model is as expected: a longer waiting time leads to a higher number of treatments ceteris paribus. This effect is the same for the other treatments for which we obtained waiting list data. With respect to the other supply side variables we do not find a clear pattern. For cataract surgery, we find a significant, negative impact of the hospital availability on the patient density. The indicators for alternative suppliers of health care (*distGP*, *av3GP*, and *distGPcentre*) are not significant in the FE model. *distGP* and *distGPcentre* are significant in the OLS and RE model: we find that a higher distance to the closest GP or GP centre leads to a lower patient density holding everything else constant.

We test the RE model against the OLS model using the Breusch and Pagan Lagrange multiplier test for random effects. The null hypothesis of this test is that the variance of the random effects is equal to zero. At a significance level of 5% we reject the null hypothesis for all treatments, which implies that we reject the OLS model in favor of the RE model. Next, we conduct a generalized Hausman test in order to test the RE model against the FE model. This test is conducted in Stata using the *xtoverid* command (Schaffer and Stillman, 2010). Both the RE estimator and the FE estimator are consistent under the null hypothesis, but the RE estimator is the most efficient. Under the alternative hypothesis only the FE estimator is consistent. For all but two treatments (hip fractures in young patients (surgery) and urethral stones) we reject the null hypothesis, which implies that the FE model is favored. Therefore we will mainly focus on the results of the FE model, but we report the OLS and RE estimates as well to give an impression of the differences between the estimates of the three models.

We calculate for each treatment the ceteris paribus effects of the physician densities on the patient density using equation (10). The results are presented in Table A9. For most treatments the results of the three different specifications are quite similar. For hip fractures, our control treatment, we find no or very little evidence for SID. Almost all estimated effects for hip fractures are either statistically or economically insignificant, especially for hip fractures in young patients. That is, we find some statistically significant positive effects, but these effects are relatively small compared to the effects we find for the other treatments.

We use Wald tests to assess whether the FW-coefficients and FFS-coefficients are equal. The p-values of these tests are shown in Table A10. For hip fractures we fail to reject the null hypothesis of equal coefficients, which implies that hypothesis IV is at least partly confirmed. Moreover, for hip fractures treated by orthopedists we find a larger effect for FW-physicians than for FFS-physicians, which implies that hypothesis III is certainly not confirmed. Based on the four hypotheses formulated in section 4.3 we conclude that for hip fractures we found no evidence for the occurrence of SID. This conclusion is summarized in the last column of Table A9.

For most of the other treatments, we find strong indications for the occurrence of SID. From Table A9 and Table A10 we conclude that hypothesis III is confirmed in ten out of the fourteen cases (this holds for the FE estimator). For FFS-physicians we find estimated effects between 0.16 and 0.71, while for FW physicians the estimated effects are much smaller and vary between -0.02 and 0.32 depending on the type of treatment. These effects indicate that if the total number of physicians increases by 1%, and these extra physicians are all of the FFS-type, then the total number of patients is expected to increase by 0.16%-0.71%. We find the strongest effects for tonsils, inguinal hernias, and hernias treated by neurologists. Moreover, the estimated effects for FFS-physicians are all significant, whereas this is not the case for the estimated effects for FW-physicians. For breast reductions, malposition of the hip, malposition of the knee, and urethral stones we find no evidence for SID. A reason could be that these treatments are more risky or that the diagnosis is less ambiguous.

Although FW-physicians do not have a direct financial incentive to induce demand, we do find a positive relationship between the FW-density and the number of patients for most of the treatments. A possible reason that could explain this finding is that demand inducement takes place at the management level at the hospital. The hospital obtains money from the insurer for every treatment. For that reason, the management of the hospital has an incentive to increase the number of patients. This could result in indirect inducement incentives for FW-physicians. However, it can also be the case that our estimates are biased. It is for instance possible that the fixed effects do not capture all unobserved heterogeneity in our model because we omitted an important demand variable and that the physician density variable is endogenous. In the next section we take this possibility into account by performing an instrumental variable analysis.

5.3 IV-analysis

In this section we control for the possible endogeneity of the physician densities. Endogeneity would imply that we find a correlation between the physician densities and the patient density just because physicians locate themselves at places where they expect high demand. Although we did an attempt to control for this reverse causality by including a bunch of demand factors and ZIP-code specific effects, it is still possible that we did not completely resolve the problem. To overcome this possible problem we estimate equation (9) by using two stage least squares (TSLS) and generalized method of moments (GMM). We include *IVhouseprice*, *IVcinema*, *IVdist_park*, and *IVdist_TH* as instrumental variables in our first stage regressions.

We test for the validity of instruments using Sargan's (1958) chi-squared test (for TSLS) and Hansen's (1982) J-statistic chi-squared test (for GMM). In thirteen out of the eighteen cases we reject the null hypothesis of valid instruments. The test does not indicate which instruments are invalid. We argue that *IVhouseprice*, as an indicator for wealth, may be directly related to the demand for medical treatments, when the wealth level determines health preferences. Therefore we reduce our set of instruments to *IVcinema*, *IVdist_park*, and *IVdist_TH*. Again, we perform the overidentification test and now we find that the null hypothesis of valid instruments is rejected in seven out of the eighteen cases. We could not think of an economical reason to exclude one of the remaining three instruments, so we perform the analysis with the remaining three instruments.

Although there are a few exemptions, we find a clear pattern in our first stage regressions. For both types of physicians we find that a higher distance to the closest public park leads, ceteris paribus, to a higher physician density. This could point to higher physician densities in urban areas. Next, we find on average higher FW densities and lower FFS densities in areas close to teaching hospitals. This is explained by the fact that all physicians working in teaching hospitals receive a fixed wage. Finally, we find that the number of cinemas is on average positively related to the FW density, while the effect on the FFS density is negative. Almost all first stage coefficients are significant at the 5% level. As an example we present the first stage results for the treatment of cataract in Table A11.

The effects of the physician densities on the number of patients resulting from the second stage estimates are presented in Table A12. The estimated effects do not compare to the results we found in section 5.2. For FFS-physicians we find strong negative and significant effects in some cases (hernias treated by neurosurgeons and varicose veins treated by dermatologists), whereas we find very strong positive and significant relationships in other cases (hip fractures treated by orthopedists and tonsillectomy). Next to that, we still find positive and significant effects for FW-physicians for about 50% of the treatments.

The results of the IV-estimation do not confirm the results we found in section 5.2. Especially the fact that we find large negative effects and large positive effects cannot be explained. Next to that, we find that the standard errors of the physician density coefficients are very high compared to the corresponding standard errors in the OLS model and panel data techniques. For some treatments (e.g. hernias treated by neurosurgeons, hip fractures treated by surgeons) we reject the null hypothesis of valid instruments, which could explain that our results are not in line with what we found before. However, for other treatments the null hypothesis of valid instruments is not rejected (e.g. hip fractures treated by orthopedists) and the effects we find are still unrealistically large. Moreover, we tried to exclude either one of the remaining three instruments, but the results did not become more realistic as a result.

A possible reason for the unrealistic results of the IV-analysis is that we are not able to explain the observed physician densities well enough in the first stage regressions. The idea of including instruments such as *IVcinema* and *IVdist_park* is that we want to explain the locational choices of physicians. However, we only obtained information on the number of physicians per hospital and we allocate physicians to ZIP-code areas using the method described in section 4.2. The problem is that the allocation of physicians to ZIP-code areas is mainly determined by the distribution over the ZIP-code areas of the patients they treat. So there is no reason to believe that the physician densities as defined in our model are related to the actual locational choices of physicians. For that reason it might be a bad idea to use *IVcinema* and *IVdist_park* as instruments in our model.

Another reason why *IVcinema* and *IVdist_park* do not predict the physician densities very well could be that individual decisions of the hospital management play an important role. Consider for example the case of orthopedics. From our data we can observe that the average number of orthopedists per hospital is about four. Suppose that the management of an average hospital decides, for some reason, to hire one extra orthopedist. Such a decision would have a major impact on the physician density since the number of orthopedists at the average hospital would increase by 20% as a result. The impact of such decisions cannot be explained by the instruments we include in our model.

5.4 Lagged explanatory variables

The estimated effects of the number of physicians on patient utilization using lagged first-differenced densities as instruments are presented in Table A13. When we use OLS, the estimated effects are comparable to the effects presented in Table A9. However, we no longer find evidence for the occurrence of SID for hernias (neurosurgery), kidney stones, arthrosis (knee), and varicose veins (surgery). A possible explanation for this finding is that we have much less observations since three of the four waves of data are excluded in this analysis.

Next, we use TSLS to overcome the possible endogeneity problem. The first-differenced physician densities are instrumented by lagged first-differenced physician densities. The resulting estimated effects are also shown in Table A13. The major part of the estimated effects is not significant. Moreover, we find both very strong positive and negative estimates. For none of the treatments we find evidence for the occurrence of SID. Again, the results of the instrumental variable approach do not seem to be very realistic. This is probably caused by the fact that our instruments are not strong enough to predict physician densities in the first stage. It is not very likely that the problem is caused by the decrease in the number of observations. When we use OLS to estimate the effects, the results are much more in line with our previous estimates.

6. Robustness Checks

6.1 3-digit patient density

Since the number of patients is very low for some of the treatments we consider, we perform our analysis at a more geographically aggregated level. For example, in every year, in about 50% of the 4-digit ZIP-code areas there are no patients at all who suffer a hip fracture. This large number of zeros values of the dependent variable could possibly bias our results. Another disadvantage of the low (absolute) number of patients is that the role of coincidence is relatively high. One patient more or less causes a substantial change in the patient density, which makes it more difficult to explain patient densities from demand variables. In order to reduce the large number of zeros and to increase the (absolute) number of patients per area, we perform the same analysis at the more geographically aggregated 3-digit ZIP-code level. The total number of 3-digit ZIP-code areas is about 800 and the average population size is just above 20,000, which is substantially higher than 5,000; the average population size of the 4-digit ZIP-code areas.

The resulting effects of the physician densities on the number of patients are presented in Table A14. The estimates shown in Table A14 confirm the results of our analysis with patient densities on the 4-digit level. Again, we test the RE model against the OLS model and the RE model against the FE model. Next, we perform tests on the equality of the FW-effect and the FFS-effect. The conclusions are the same as before. Based on the summary results in the last column of Table A14 we draw exactly the same conclusions as in Table A10 and the magnitudes of the effects are very similar.

6.2 Country border areas

As a second check of our results, we consider the impact of country border areas. In the areas next to the Belgian and German border it is likely that patients cross the border to receive medical treatments. If that is the case, the observed patient densities will be lower than predicted by the demand variables. Since we fail to correct for demand properly in that case, our results could be biased. Moreover, the allocation of physicians to ZIP-code areas is based on the number of patients. This implies that we allocate a relatively low number of physicians to border areas in case of country border crossing. This, in turn, could be a source of artificial correlation between patient density and physician densities. For that reasons we exclude all 2-digit ZIP-code areas adjacent to the Belgian and German border (26 2-digit areas in total, which reduces the number of observations in our regressions by about 30%) and calculate the estimated inducement effects again.

The results are presented in Table A15. Compared to the results in Table A9, we find indications for inducement of hip fractures by surgeons and we no longer find indications for inducement of knee operations (arthrosis) and kidney stone operations when we exclude the border areas. The magnitudes of the other effects are comparable to what we find when border areas are included. For FFS-physicians we find effects between 0.14 and 0.73 and for FW-physicians we find effects between 0.02 and 0.39.

The fact that we find indications for the inducement of hip fractures is not in line with our expectations. One alternative explanations for this finding is that it could be the case that the diagnosis and related treatment of hip fractures is not completely unambiguous for elderly people. That is, when we only select the group of young people we do not find an indication for the inducement of hip fractures.

7. Conclusion

We formulated and tested a model of SID in the Dutch hospital sector. In our model, the per capita number of physicians is related to the demand for sixteen different medical treatments. We have a unique dataset at our disposal including the number of patients treated in all Dutch hospitals during the years 2006-2009. The dataset allows us to identify every patient's place of residence based on a 4-digit ZIP-code, which makes it possible to calculate the yearly number of patients per capita for all 4,000 4-digit ZIP-codes in the Netherlands.

In a first step, we use OLS to relate the number of patients per capita to a set of demand factors. The idea is to obtain a first indication for the occurrence of SID by explaining regional variations in the so-called patient density from demand factors only. If the error terms of the demand equation exhibit zero or negative autocorrelation and spatial correlation we can exclude the possibility of SID straight away. However, for all treatments, except two control treatments (hip fractures in young patients treated by surgeons and orthopedists) we found significant positive correlations.

Therefore the inclusion of supply side variables forms the second step of our approach. We obtained data on the number of physicians per hospital and the remuneration method of those physicians for the major part of the Dutch hospitals. The main problem we face is the allocation of physicians to ZIP-code areas and the calculation of the corresponding physician densities. Our dataset, however, allows us to take border crossing into account and we allocated physicians on the 2-digit ZIP-code level to make the physician density a more realistic measure of the access that people have to physicians. The idea of our estimation is that we find an indication for the occurrence of SID if a higher physician density corresponds to a higher number of patients. Moreover, we expect to find a stronger effect for fee-for-service physicians compared to fixed wage physicians since the former type of physicians has a strong financial incentive to induce demand.

The structure of the data allows us to use panel data techniques, which has the advantage that we can include ZIP-code specific effects to control for unobserved heterogeneity. We do not find indications for SID for our control treatment, hip fractures, which confirms our earlier result. However, for the majority of the other treatments we find strong indications for the occurrence of SID. We find that a 1% increase in the total number of physicians leads to an increase of 0.16%-0.71% in the number of treatments, when the extra physicians are all paid on a fee-for-service basis. This in contrast with fixed wage physicians, where we find effects between -0.02% and 0.32%. The magnitude of the effect differs per type of treatment. We find strong evidence for SID for tonsillectomy, inguinal hernias, and hernias treated by neurologists. For breast reductions, malposition of the hip, malposition of the knee, and urethral stones we find no evidence for SID.

We checked the robustness of our results by performing the same analysis at the more geographically aggregated 3-digit ZIP-code level and by excluding areas adjacent to the Belgian and German border. The robustness analysis confirms the most important results we find. Next, we took the possible endogeneity of physician densities into account by using instrumental variable techniques. The results of these estimations are not very reliable, probably because of invalid instruments that did not adequately predict the physician densities in the first stage.

For further research it is interesting to study upcoding as supplemental evidence for the occurrence of SID. That is, physicians might classify patients in more expensive DBC's than necessary in order to increase their incomes. As is shown in Table A1, 'cataract surgery' actually consists of eight different treatments. If we can rank these treatments in order of profitability for a specialist, we can investigate whether or not fee-for-service physicians are more inclined to perform more profitable treatments. Another important policy question is whether SID occurs at the hospital level or at the level of medical specialists. In order to answer this question, one must find a method to calculate the underproduction or overproduction per hospital for a number of different treatments. This number can be related to hospital dummies and the percentage of fee-for-service physicians per specialty in a hospital in order to find out which factor causes hospitals to overproduce.

Literature

Bickerdyke, I., Dolamore, R., Monday, I. and Preston, F. (2002). *Supplier-Induced Demand for Medical Services*. Productivity Commission Staff Working Paper. Canberra, November.

Cromwell, J., and Mitchell, J.B. (1986). *Physician-induced Demand for Surgery*. Journal of Health Economics, 5, 293-313.

Delattre, E. and Dormont, B. (2003). *Fixed Fees and Physician-Induced Demand: A Panel Data Study on French Physicians*. Health Economics, 12, 741-754.

Evans, R.G. (1974). *Supplier-induced Demand: Some Empirical Evidence and Implications*. In Perlman, M. (ed). The Economics of Health and Medical Care. pp 162-173. London: Macmillan.

Hansen, L.P. (1982). *Large Sample Properties of Generalized Method of Moments Estimators*. Econometrica, 50, 1029-1054.

Fuchs, V.R. (1978). *The Supply of Surgeons and the Demand for Operations*. The Journal of Human Resources, 13, 35-56.

Grytten, J. and Sørensen, R. (2001). *Type of Contract and Supplier-Induced Demand for Primary Physicians in Norway*. Journal of Health Economics, 20, 379-393.

McGuire, T.G. (2000). *Physician Agency*. In Culyer, A.J., and Newhouse, J.P. (eds). Handbook of Health Economics. pp. 461-536. Amsterdam: Elsevier.

Pomp, M. (2009). *Aanbodgeïnduceerde vraag: feit of fictie?* Onderzoeksrapport voor de ministeries van Economische Zaken en Financiën.

Sargan, J.D. (1958). *The Estimation of Economic Relationships Using Instrumental Variables*. Econometrica, 26, 393-415.

Schaffer, M.D. and Stillman, S. (2010). *xtoverid: Stata module to calculate tests of overidentifying restrictions after xtreg, xtivreg, xtivreg2, and xthtaylor*. http://ideas.repec.org/c/boc/bocode/s456779.html

Verbeek, M. (2008). *A Guide to Modern Econometrics*. Chichester: John Wiley & Sons, Ltd.

Wennberg, J.E. (2010). *Tracking medicine: a researcher's quest to understand health care*. New York: Oxford University Press.

Zweifel, P., Breyer, F. and Kifmann, M. (2009). Health Economics. Heidelberg: Springer Berlin.

Appendix A - Tables

Table A1: unique DBC's cataract surgery

DBC-code	Treatment	Description
110005540031	cataract	out-patient treatment
110005540032	cataract	one-day care surgery
110005540033	cataract	treatment in hospital and admission to hospital
110005540036	cataract	treatment in hospital, no admission to own hospital
210005540031	cataract	out-patient treatment
210005540032	cataract	one-day care surgery
210005540033	cataract	treatment in hospital and admission to hospital
210005540036	cataract	treatment in hospital, no admission to own hospital

Table A2: descriptive statistics number of treatments and patient characteristics

Treatment	Specialty	Number of Patients				Average age	% men	Number of hospitals
		2006	*2007*	*2008*	*2009*			
Hip fracture[+]	Surgery	9,428	9,648	10,141	9,493	76	30	96
Hip fracture, young patients[+]	Surgery	1,755	1,561	1,881	1,829	43	48	96
Hip fracture[+]	Orthopedics	6,364	6,643	6,825	6,279	73	28	100
Hip fracture, young patients[+]	Orthopedics	1,784	1,111	1,260	1,070	36	40	96
Cataract	Ophthalmology	134,107	144,874	160,231	145,339	70	39	117
Tonsils	Otolaryngology	61,707	62,000	57,007	53,178	10	50	103
Inguinal hernia	Surgery	29,105	30,983	30,415	27,555	48	89	109
Varicose veins	Surgery	26,551	30,721	34,491	33,912	49	25	123
Breast reduction	Plastic surgery	5,869	5,069	5,126	4,170	39	0	102
Arthrosis (knee)	Orthopedics	30,561	32,721	34,272	31,016	60	39	111
Arthrosis (pelvis/hip/upper leg)	Orthopedics	21,765	22,155	23,222	21,099	66	32	104
Malposition hip (after replacement)	Orthopedics	6,458	6,777	7,637	6,923	69	32	107
Malpostion knee (after replacement)	Orthopedics	3,162	3,704	4,247	3,967	66	33	108
Kidney stone	Urology	4,888	9,941	10,342	9,561	52	59	97
Urethral stone	Urology	3,568	6,373	6,770	6,420	51	67	98
Hernia	Neurosurgery	9,850	9,740	9,483	7,766	45	54	69
Varicose veins	Dermatology	37,366	45,526	53,510	56,051	49	15	139
Hernia	Neurology	49,682	49,180	53,130	52,001	49	51	107

[+] control treatments

Table A3: descriptive statistics ZIP-code areas

	2006	2007	2008	2009
Number of 4-digit ZIP-code areas	4,007	4,014	4,015	4,019
Number of 4-digit ZIP-code areas (population > 500)	3,162	3,167	3,171	3,167
Average population	5,105	5,105	5,114	5,145
Demand variables				
age0_5 (%)	6.1	5.9	5.7	5.5
age5_10	6.3	6.3	6.4	6.3
age10_15	6.3	6.3	6.2	6.2
age15_20	6.1	6.2	6.2	6.2
age20_25	5.5	5.5	5.5	5.6
age25_30	5.5	5.4	5.4	5.4
age30_35	6.5	6.1	5.8	5.6
age35_40	7.9	7.8	7.6	7.3
age40_45	8.1	8.1	8.0	8.0
age45_50	7.6	7.7	7.8	7.9
age50_55	7.1	7.1	7.2	7.3
age55_60	7.3	7.1	6.9	6.9
age60_65	5.4	5.9	6.4	6.6
age65_70	4.4	4.5	4.6	4.7
age70_75	3.6	3.6	3.7	3.8
age75_80	2.8	2.9	2.9	3.0
age80_85	2.0	2.0	2.0	2.0
age85_90	1.0	1.0	1.1	1.2
age90_95	0.4	0.4	0.4	0.4
age95	0.1	0.1	0.1	0.1
men (%)	49.9	49.9	49.9	49.9
mortality (per 1,000)	7.5	7.4	7.5	n/a
westerns (%)	7.7	7.8	7.8	7.9
nonwesterns (%)	7.1	7.1	7.2	7.4
assistance (per 1,000)	35.4	32.3	29.7	28.9
working (% of population between 15 and 65)	71.3	73.4	74.3	n/a
selfemployed (% of working people)	9.2	9.4	9.5	n/a
Supply variables				
distGP (km)	n/a	1.3	1.4	n/a
av3GP	n/a	7.0	6.8	n/a
distGPcentre (km)	n/a	7.3	7.3	n/a
av20hospital	n/a	4.3	4.4	n/a
Instrumental variables				
IVhouseprice (in 1,000s)	220.2	237.3	254.2	262.4
IVcinema	n/a	n/a	1.8	n/a
IVdist_park (km)	n/a	n/a	0.9	n/a
IVdist_TH (km)	47.0	47.0	47.0	46.9

Table A4: descriptive statistics physicians

Specialty	Total number of physicians											
	2006			*2007*			*2008*			*2009*		
	FFS	FW	TH	FFS	FW	TH	FFS	FW	TH	FFS	FW	TH
Ophthalmology	298	38	n/a	293	28	65	286	28	68	293	36	76
Otolaryngology	250	29	n/a	239	30	67	235	29	67	257	25	71
Surgery	560	71	n/a	546	68	250	549	71	249	595	60	238
Plastic surgery	101	7	n/a	97	8	34	105	7	37	115	9	37
Orthopedics	310	54	n/a	308	48	54	309	41	59	337	41	57
Urology	201	20	n/a	192	23	53	194	18	53	215	20	55
Neurosurgery	38	23	n/a	45	16	54	49	12	57	56	15	57
Dermatology	184	37	n/a	186	30	52	187	27	53	201	24	53
Neurology	295	101	n/a	290	90	121	279	100	120	328	87	123
Number of hospitals	84			87			84			90		

Table A5: descriptive statistics waiting lists

Treatment	Specialty	Average waiting time (in weeks)			
		2006	*2007*	*2008*	*2009*
Cataract	Ophthalmology	6.7	6.7	6.3	5.0
Tonsils	Otolaryngology	4.9	4.6	4.2	3.8
Inguinal hernia	Surgery	4.5	4.3	4.4	4.0
Varicose veins	Surgery	6.6	6.2	5.4	4.3
Breast reduction	Plastic surgery	15.9	13.1	10.5	5.6
Arthrosis (knee)	Orthopedics	9.9	8.2	8.4	6.8
Arthrosis (pelvis/hip/upper leg)	Orthopedics	8.2	7.6	7.7	6.9

Table A6: OLS results of demand equation cataract surgery (robust standard errors in parentheses)

Dependent variable: 4-digit patient density	(OLS)
age75_80	0.92*
	(0.43)
age80_85	1.00*
	(0.44)
men	-0.04
	(0.04)
mortality	-0.02
	(0.01)
westerns	-0.10***
	(0.01)
nonwesterns	0.01*
	(0.01)
urbanized	-0.14**
	(0.04)
assistance	0.00
	(0.00)
working	-0.01
	(0.01)
selfemployed	-0.04***
	(0.01)
lowincome	-0.02
	(0.01)
highincome	0.00
	(0.01)
dummy 2007	1.13***
	(0.10)
dummy 2008	1.80***
	(0.12)
dummy 2009	1.11***
	(0.13)
constant	-2.01
	(42.33)
Number of observations	12,152
R^2	0.4318

* significant at $p<0.05$; ** significant at $p<0.01$; *** significant at $p<0.001$

Table A7: autocorrelations and spatial correlations

Treatment	Specialty	Autocorrelation (ρ_{AC})	Spatial Correlation (ρ_{SC})				$\bar{y}$	(MAE)
			2006	*2007*	*2008*	*2009*		
Hip fracture^{+}	Surgery	.58***	.39***	.36***	.32***	.34***	0.5	0.4
Hip fracture, young patients^{+}	Surgery	.03	.08	.06	.10**	.08*	0.1	0.1
Hip fracture^{+}	Orthopedics	.33***	.39***	.33***	.42***	.39***	0.4	0.3
Hip fracture, young patients^{+}	Orthopedics	.07**	.21***	.00	.07	.02	0.1	0.1
Cataract	Ophthalmology	.36***	.82***	.65***	.68***	.64***	8.3	2.5
Tonsils	Otolaryngology	.30***	.78***	.55***	.63***	.56***	3.3	1.3
Inguinal hernia	Surgery	.12***	.63***	.26***	.32***	.20***	1.8	0.7
Varicose veins	Surgery	.38***	.58***	.60***	.63***	.58***	1.9	0.9
Breast reduction	Plastic surgery	.05**	.18***	.15***	.09*	.19***	0.3	0.2
Arthrosis (knee)	Orthopedics	.33***	.70***	.62***	.61***	.57***	1.9	0.9
Arthrosis (pelvis/hip/upper leg)	Orthopedics	.13***	.48***	.21***	.26***	.21***	1.3	0.6
Malposition hip (after replacement)	Orthopedics	.28***	.34***	.25***	.23***	.15***	0.4	0.3
Malpostion knee (after replacement)	Orthopedics	.27***	.18***	.27***	.24***	.20***	0.2	0.2
Kidney stone	Urology	.30***	.40***	.18***	.21***	.14***	0.5	0.5
Urethral stone	Urology	.17***	.24***	.21***	.22***	.24***	0.3	0.3
Hernia	Neurosurgery	.20***	.42***	.30***	.32***	.38***	0.5	0.4
Varicose veins	Dermatology	.70***	.86***	.90***	.90***	.91***	2.8	1.8
Hernia	Neurology	.41***	.74***	.66***	.64***	.55***	2.9	1.2
Number of observations		9073	3011	3013	3028	3023	12152	12152

$^{+}$ control treatments
* significant at $p<0.05$; ** significant at $p<0.01$; *** significant at $p<0.001$ (based on robust standard errors)

Table A8: OLS, RE, and FE results of demand and supply equation cataract surgery (robust standard errors in parentheses)

Dependent variable: 4-digit patient density	*(OLS)*	*(RE)*	*(FE)*
density_FW	26.79***	33.24***	59.34***
	(3.31)	(3.56)	(7.56)
density_FFS	142.05***	146.59***	153.89***
	(4.94)	(5.78)	(7.98)
waitinglist	-0.02	0.03*	0.09***
	(0.01)	(0.01)	(0.02)
distGP	-0.10**	-0.09*	0.29
	(0.04)	(0.04)	(0.37)
av3GP	-0.01	-0.01	-0.14
	(0.01)	(0.01)	(0.10)
distGPcentre	-0.02*	-0.02*	-0.13
	(0.01)	(0.01)	(0.19)
av20hospital	-0.04***	-0.04**	-0.79***
	(0.01)	(0.01)	(0.15)
age75_80	1.11**	1.08*	1.01
	(0.41)	(0.44)	(0.55)
age80_85	1.14**	1.03*	0.82
	(0.42)	(0.46)	(0.53)
men	-0.05	-0.03	0.06
	(0.04)	(0.06)	(0.11)
mortality	-0.02	-0.01	-0.00
	(0.01)	(0.02)	(0.02)
westerns	-0.08***	-0.08***	0.04
	(0.01)	(0.01)	(0.07)
nonwesterns	0.03***	0.03***	-0.09
	(0.01)	(0.01)	(0.08)
urbanized	-0.09	-0.08	1.14**
	(0.05)	(0.07)	(0.39)
assistance	-0.00	-0.00	0.00
	(0.00)	(0.00)	(0.01)
working	-0.00	-0.00	-0.01
	(0.01)	(0.01)	(0.04)
selfemployed	-0.02	-0.03	-0.09
	(0.01)	(0.02)	(0.08)
lowincome	-0.01	-0.01	-0.01
	(0.01)	(0.02)	(0.02)
highincome	0.00	-0.00	-0.03
	(0.01)	(0.01)	(0.03)
dummy 2007	1.11***	1.13***	1.38***
	(0.09)	(0.08)	(0.13)
dummy 2008	1.73***	1.74***	2.17***
	(0.11)	(0.11)	(0.20)
dummy 2009	0.96***	1.01***	1.56***
	(0.12)	(0.12)	(0.24)
constant	-23.81	-18.73	-18.78
	(40.77)	(43.61)	(51.27)
Number of observations	12,141	12,141	12,141
Number of groups		3,063	3,063
R^2	0.4792		
R^2 within		0.1608	0.1717
R^2 between		0.6452	0.0698
R^2 overall		0.4777	0.0644

* significant at $p<0.05$; ** significant at $p<0.01$; *** significant at $p<0.001$

Table A9: estimated effects of a 1% increase in the number of physicians if it concerns (1) FW-physicians only (2) FFS-physicians only (OLS, RE, and FE)

dependent variable: 4-digit patient density

Treatment	**Specialty**	**FW** (OLS)	**FW** (RE)	**FW** (FE)	**FFS** (OLS)	**FFS** (RE)	**FFS** (FE)	**Number of observations**	**Evidence for SID?**
Hip fracture^{+}	Surgery	0.05*	0.07*	0.07	-0.03	0.12**	0.17***	12,141	no
Hip fracture, young patients^{+}	Surgery	-0.04	-0.04	-0.10	-0.10***	-0.09***	-0.04	11,600	no
Hip fracture^{+}	Orthopedics	-0.09	0.06	0.40***	-0.08**	0.05	0.26***	12,141	no
Hip fracture, young patients^{+}	Orthopedics	0.02	0.03	0.19	-0.14**	-0.13**	0.04	11,547	no
Cataract	Ophthalmology	0.09***	0.11***	0.20***	0.49***	0.51***	0.53***	12,141	yes
Tonsils	Otolaryngology	0.19***	0.20***	0.32***	0.70***	0.71***	0.71***	12,141	yes
Inguinal hernia	Surgery	0.14***	0.14***	0.15***	0.61***	0.62***	0.67***	12,141	yes
Varicose veins	Surgery	0.05***	0.06***	0.07**	0.44***	0.36***	0.16***	12,141	yes
Breast reduction	Plastic surgery	0.06	0.06	0.06	0.11**	0.11**	0.13*	12,141	no
Arthrosis (knee)	Orthopedics	0.26***	0.27***	0.31***	0.44***	0.48***	0.54***	12,141	yes
Arthrosis (pelvis/hip/upper leg)	Orthopedics	0.22***	0.22***	0.29***	0.47***	0.48***	0.53***	12,141	yes
Malposition hip (after replacement)	Orthopedics	0.46***	0.41***	0.29***	0.10*	0.13**	0.23***	12,141	no
Malpostion knee (after replacement)	Orthopedics	0.46***	0.37***	0.20**	0.11*	0.11	0.12	12,141	no
Kidney stone	Urology	0.08	0.08	0.07	0.14***	0.18***	0.28***	12,062	yes
Urethral stone	Urology	-0.14**	-0.14**	-0.05	-0.01***	-0.01***	-0.01	12,058	no
Hernia	Neurosurgery	0.01	0.02	0.03	0.38***	0.37***	0.31***	12,062	yes
Varicose veins	Dermatology	0.22***	0.09***	-0.02	0.67***	0.47***	0.30***	12,141	yes
Hernia	Neurology	0.04**	0.09***	0.21***	0.51***	0.56***	0.64***	12,141	yes

$^{+}$ control treatments

* significant at $p<0.05$; ** significant at $p<0.01$; *** significant at $p<0.001$ (based on robust standard errors)

Table A10: p-values of the test for equality of coefficients

Treatment	**Specialty**	$H_0: \hat{\xi}_{FFS} = \hat{\xi}_{FW}$ (RE)	$H_0: \hat{\xi}_{FFS} = \hat{\xi}_{FW}$ (FE)
Hip fracture+	Surgery	0.228	0.081
Hip fracture, young patients+	Surgery	0.276	0.354
Hip fracture+	Orthopedics	0.878	0.060
Hip fracture, young patients+	Orthopedics	0.127	0.272
Cataract	Ophthalmology	0.000	0.000
Tonsils	Otolaryngology	0.000	0.000
Inguinal hernia	Surgery	0.000	0.000
Varicose veins	Surgery	0.000	0.019
Breast reduction	Plastic surgery	0.221	0.342
Arthrosis (knee)	Orthopedics	0.000	0.000
Arthrosis (pelvis/hip/upper leg)	Orthopedics	0.000	0.000
Malposition hip (after replacement)	Orthopedics	0.000	0.392
Malpostion knee (after replacement)	Orthopedics	0.001	0.290
Kidney stone	Urology	0.077	0.018
Urethral stone	Urology	0.013	0.642
Hernia	Neurosurgery	0.000	0.000
Varicose veins	Dermatology	0.000	0.000
Hernia	Neurology	0.000	0.000

+ control treatments

Table A11: first stage regressions cataract

	Dependent variable: 2-digit FW-density *(OLS)*	Dependent variable: 2-digit FFS-density *(OLS)*
IVdist_park	0.0003	-0.0010***
	(0.0002)	(0.0002)
IVcinema	0.0012***	-0.0001***
	(0.0001)	(0.0000)
IVdistance_TH	-0.0001***	0.0000***
	(0.0000)	(0.0000)
waitinglist	-0.0001**	0.0001**
	(0.0000)	(0.0000)
distGP	-0.0002*	-0.0003***
	(0.0001)	(0.0001)
av3GP	-0.0000	0.0000
	(0.0000)	(0.0000)
distGPcentre	0.0000	0.0001**
	(0.0000)	(0.0000)
av20hospital	-0.0004***	0.0003***
	(0.0000)	(0.0000)
age75_80	0.0009	-0.0018*
	(0.0009)	(0.0010)
age80_85	0.0005	-0.0013
	(0.0009)	(0.0010)
men	-0.0001	0.0001
	(0.0001)	(0.0001)
mortality	-0.0001**	-0.0000**
	(0.0000)	(0.0000)
westerns	0.0003***	-0.0001***
	(0.0000)	(0.0000)
nonwesterns	-0.0001***	-0.0000*
	(0.0000)	(0.0000)
urbanized	-0.0002	-0.0004**
	(0.0002)	(0.0001)
assistance	-0.0000	0.0000***
	(0.0000)	(0.0000)
working	-0.0002***	-0.0001**
	(0.0000)	(0.0000)
selfemployed	0.0002***	0.0000
	(0.0000)	(0.0000)
lowincome	-0.0001	-0.0001***
	(0.0000)	(0.0000)
highincome	-0.0001***	0.0000
	(0.0000)	(0.0000)
dummy 2007	-0.0003	0.0003
	(0.0003)	(0.0002)
dummy 2008	0.0005	0.0007**
	(0.0003)	(0.0002)
dummy 2009	0.0014***	0.0010***
	(0.0004)	(0.0003)
constant	-0.0330	0.1845*
	(0.0874)	(0.0716)
Number of observations	12,141	12,141
R^2	0.2049	0.0641

* significant at $p<0.05$; ** significant at $p<0.01$; *** significant at $p<0.001$

Table A12: estimated effects of a 1% increase in the number of physicians if it concerns (1) FW-physicians only (2) FFS-physicians only (TSLS and GMM)

dependent variable: 4-digit patient density

Treatment	Specialty	FW		FFS		Number of observations	p-value test overidentification
		(TSLS)	*(GMM)*	*(TSLS)*	*(GMM)*		
Hip fracture^{+}	Surgery	0.14***	0.16*	-1.36***	-1.34***	12,141	0.042
Hip fracture, young patients^{+}	Surgery	0.13	0.13	-0.46**	-0.43**	11,600	0.250
Hip fracture^{+}	Orthopedics	0.53	0.53	2.21**	2.21**	12,141	0.655
Hip fracture, young patients^{+}	Orthopedics	-0.34	-0.34	0.64*	0.63*	11,547	0.763
Cataract	Ophthalmology	-0.04	0.00	-0.03	0.11	12,141	0.000
Tonsils	Otolaryngology	0.32***	0.27***	1.28***	1.00***	12,141	0.000
Inguinal hernia	Surgery	0.15***	0.16***	0.24	0.27	12,141	0.556
Varicose veins	Surgery	-0.03	-0.03	0.12	0.10	12,141	0.073
Breast reduction	Plastic surgery	0.09	0.12	-0.53**	-0.28	12,141	0.000
Arthrosis (knee)	Orthopedics	0.34***	0.35***	0.38***	0.38***	12,141	0.129
Arthrosis (pelvis/hip/upper leg)	Orthopedics	0.30***	0.27**	0.45**	0.39**	12,141	0.110
Malposition hip (after replacement)	Orthopedics	0.01	-0.03	-0.52**	-0.56***	12,141	0.376
Malpostion knee (after replacement)	Orthopedics	-0.46*	-0.44	-0.49***	-0.46***	12,141	0.484
Kidney stone	Urology	0.62*	0.63*	0.88*	0.89*	12,062	0.540
Urethral stone	Urology	-0.38**	-0.39**	0.24	0.25*	12,058	0.224
Hernia	Neurosurgery	0.58***	0.59***	-3.47***	-3.54***	12,062	0.004
Varicose veins	Dermatology	1.02***	1.31***	-2.66*	-3.87*	12,141	0.000
Hernia	Neurology	-0.11**	-0.11**	-0.18	-0.21	12,141	0.005

$^{+}$ control treatments

* significant at $p<0.05$; ** significant at $p<0.01$; *** significant at $p<0.001$ (based on robust standard errors)

Table A13: estimated effects of a 1% increase in the number of physicians if it concerns (1) FW-physicians only (2) FFS-physicians only (OLS and TSLS)

dependent variable: 4-digit patient density

Treatment	Specialty	FW (OLS)	FW (TSLS)	FFS (OLS)	FFS (TSLS)	Number of observations (TSLS)
Hip fracture+	Surgery	0.13	-0.29	0.11*	0.23	3,041
Hip fracture, young patients+	Surgery	-0.14	-0.54*	-0.04	-0.15	2,755
Hip fracture+	Orthopedics	0.12	1.87	0.03	-0.45	3,041
Hip fracture, young patients+	Orthopedics	0.16	12.53	-0.17	-7.47	2,655
Cataract	Ophthalmology	0.21***	-0.52	0.76***	-0.48	3,042
Tonsils	Otolaryngology	0.35***	8.55	0.51***	-3.51	3,042
Inguinal hernia	Surgery	0.15**	0.67	0.44***	2.24	3,041
Varicose veins	Surgery	0.09**	0.07	0.18**	-0.16	3,041
Breast reduction	Plastic surgery	-0.01	1.86*	0.14	-0.59	3,042
Arthrosis (knee)	Orthopedics	0.33***	0.36	0.47***	0.45	3,042
Arthrosis (pelvis/hip/upper leg)	Orthopedics	0.13	0.17	0.41***	-1.61	3,042
Malposition hip (after replacement)	Orthopedics	0.31***	0.26	0.24**	-0.43	3,041
Malpostion knee (after replacement)	Orthopedics	0.09	0.42	0.16	0.20	3,041
Kidney stone	Urology	0.18	-0.20	0.07	3.69*	2,961
Urethral stone	Urology	-0.06	0.08	-0.02	3.25	2,958
Hernia	Neurosurgery	0.04	-4.71	-0.08	10.53	3,002
Varicose veins	Dermatology	-0.04	-0.07	0.24***	1.16***	3,041
Hernia	Neurology	0.09*	-0.23	0.51***	0.37	3,042

+ control treatments

* significant at $p<0.05$; ** significant at $p<0.01$; *** significant at $p<0.001$ (based on robust standard errors)

Table A14: estimated effects of a 1% increase in the number of physicians if it concerns (1) FW-physicians only (2) FFS-physicians only (OLS, RE, and FE)

dependent variable: 3-digit patient density

Treatment	Specialty	FW			FFS			Number of observations	Evidence for SID?
		(OLS)	*(RE)*	*(FE)*	*(OLS)*	*(RE)*	*(FE)*		
Hip fracture+	Surgery	0.08**	0.11**	0.09	0.02	0.14***	0.19***	3,112	no
Hip fracture, young patients+	Surgery	0.04	0.02	-0.04	0.00	0.00	-0.01	2,964	no
Hip fracture+	Orthopedics	-0.05	0.16*	0.44***	-0.04	0.13**	0.31***	3,112	no
Hip fracture, young patients+	Orthopedics	0.14	0.11	-0.09	-0.12	-0.10	-0.07	2,935	no
Cataract	Ophthalmology	0.10***	0.12***	0.17***	0.50***	0.50***	0.51***	3,112	yes
Tonsils	Otolaryngology	0.17***	0.21***	0.32***	0.69***	0.71***	0.73***	3,112	yes
Inguinal hernia	Surgery	0.15***	0.15***	0.14***	0.62***	0.63***	0.66***	3,112	yes
Varicose veins	Surgery	0.02	0.03	0.03	0.34***	0.24***	0.12***	3,112	yes
Breast reduction	Plastic surgery	0.01	0.02	0.05	0.10*	0.11*	0.18***	3,112	no
Arthrosis (knee)	Orthopedics	0.23***	0.25***	0.24**	0.40***	0.46***	0.53***	3,112	yes
Arthrosis (pelvis/hip/upper leg)	Orthopedics	0.21***	0.21***	0.22***	0.47***	0.48***	0.50***	3,112	yes
Malposition hip (after replacement)	Orthopedics	0.48***	0.44***	0.36***	0.18***	0.20***	0.23***	3,112	no
Malpostion knee (after replacement)	Orthopedics	0.39***	0.37***	0.35***	0.08	0.09	0.09	3,112	no
Kidney stone	Urology	0.02	-0.01	-0.14	0.16***	0.19***	0.24***	3,093	yes
Urethral stone	Urology	-0.05	-0.02	0.07	-0.01***	-0.01***	-0.01	3,093	no
Hernia	Neurosurgery	0.02	0.04*	0.04	0.32***	0.31***	0.28***	3,094	yes
Varicose veins	Dermatology	0.18***	0.05	-0.02	0.71***	0.42***	0.27***	3,112	yes
Hernia	Neurology	0.04	0.10***	0.18***	0.49***	0.55***	0.61***	3,112	yes

+ control treatments

* significant at p<0.05; ** significant at p<0.01; *** significant at p<0.001 (based on robust standard errors)

Table A15: estimated effects of a 1% increase in the number of physicians if it concerns (1) FW-physicians only (2) FFS-physicians only (OLS, RE, and FE)

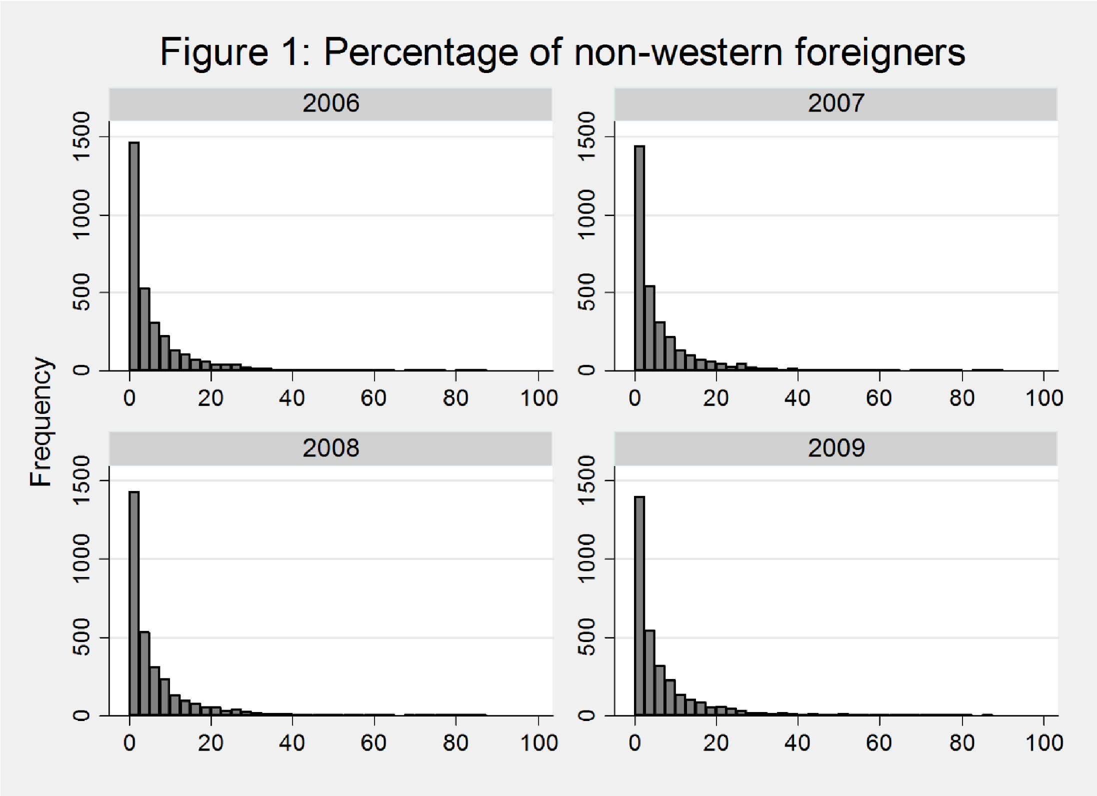
dependent variable: 4-digit patient density (excluding border areas)

Treatment	Specialty	FW			FFS			Number of observations	Evidence for SID?
		(OLS)	*(RE)*	*(FE)*	*(OLS)*	*(RE)*	*(FE)*		
Hip fracture[+]	Surgery	-0.09**	0.00	0.07	0.05	0.19***	0.26***	8,617	yes
Hip fracture, young patients[+]	Surgery	-0.11*	-0.11*	-0.10	-0.05*	-0.05*	0.00	8,445	no
Hip fracture[+]	Orthopedics	0.19**	0.27***	0.44***	0.01	0.05	0.12*	8,617	no
Hip fracture, young patients[+]	Orthopedics	0.17	0.17	0.28	-0.15**	-0.14**	-0.01	8,244	no
Cataract	Ophthalmology	0.12***	0.15***	0.29***	0.46***	0.44***	0.42***	8,617	yes
Tonsils	Otolaryngology	0.22***	0.24***	0.32***	0.64***	0.67***	0.73***	8,617	yes
Inguinal hernia	Surgery	0.13***	0.14***	0.20***	0.57***	0.60***	0.70***	8,617	yes
Varicose veins	Surgery	0.08***	0.07***	0.06*	0.41***	0.33***	0.14***	8,617	yes
Breast reduction	Plastic surgery	0.03	0.03	0.06	0.12**	0.12**	0.15*	8,617	no
Arthrosis (knee)	Orthopedics	0.59***	0.55***	0.43***	0.41***	0.45***	0.52***	8,617	no
Arthrosis (pelvis/hip/upper leg)	Orthopedics	0.27***	0.28***	0.39***	0.46***	0.48***	0.55***	8,617	yes
Malposition hip (after replacement)	Orthopedics	0.56***	0.46***	0.31***	0.17**	0.21***	0.32***	8,617	no
Malpostion knee (after replacement)	Orthopedics	0.59***	0.46***	0.25***	0.10	0.12	0.14	8,617	no
Kidney stone	Urology	0.12*	0.08*	0.08	0.05	0.11	0.18**	8,617	no
Urethral stone	Urology	-0.14*	-0.11	0.03	-0.01***	-0.01***	-0.01	8,617	no
Hernia	Neurosurgery	0.01	0.03	0.06*	0.10	0.14*	0.23**	8,617	yes
Varicose veins	Dermatology	0.26***	0.13***	0.02	0.59***	0.40***	0.25***	8,617	yes
Hernia	Neurology	0.06**	0.11***	0.20***	0.46***	0.49***	0.55***	8,617	yes

[+] control treatments

[*] significant at p<0.05; [**] significant at p<0.01; [***] significant at p<0.001 (based on robust standard errors)

Appendix B - Figures



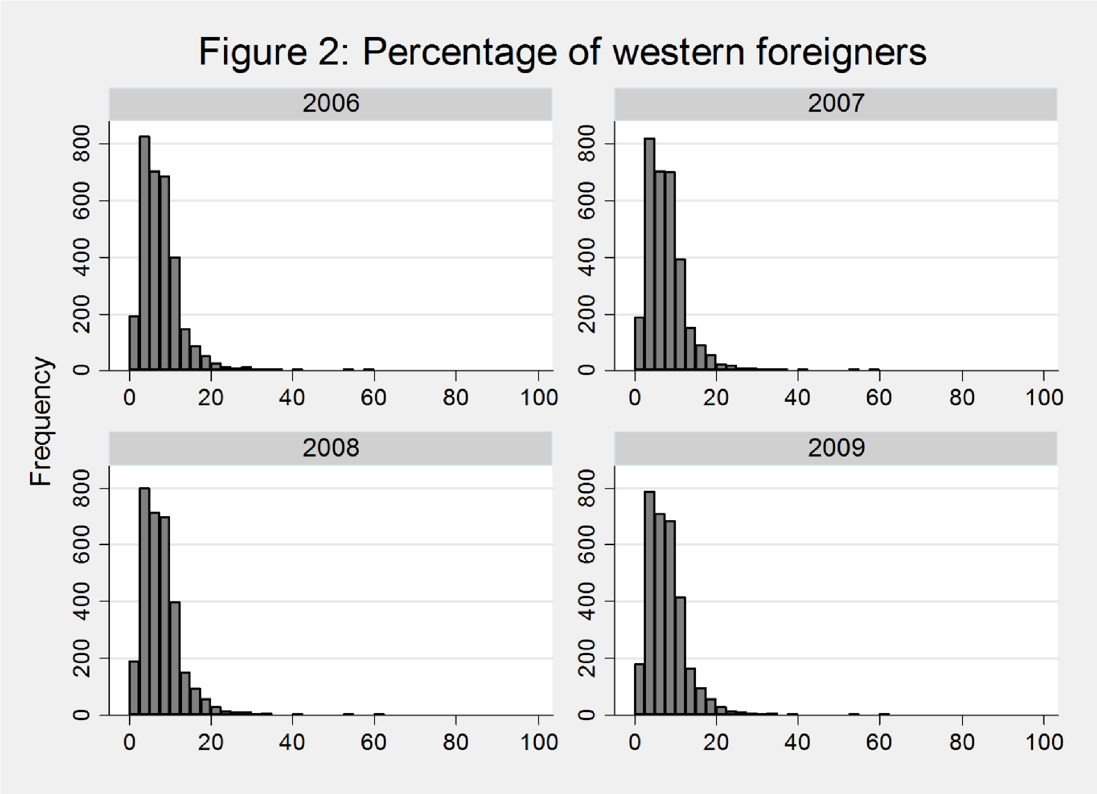


Figure 3: Mortality rate

(number of deceased per 1,000 inhabitants)

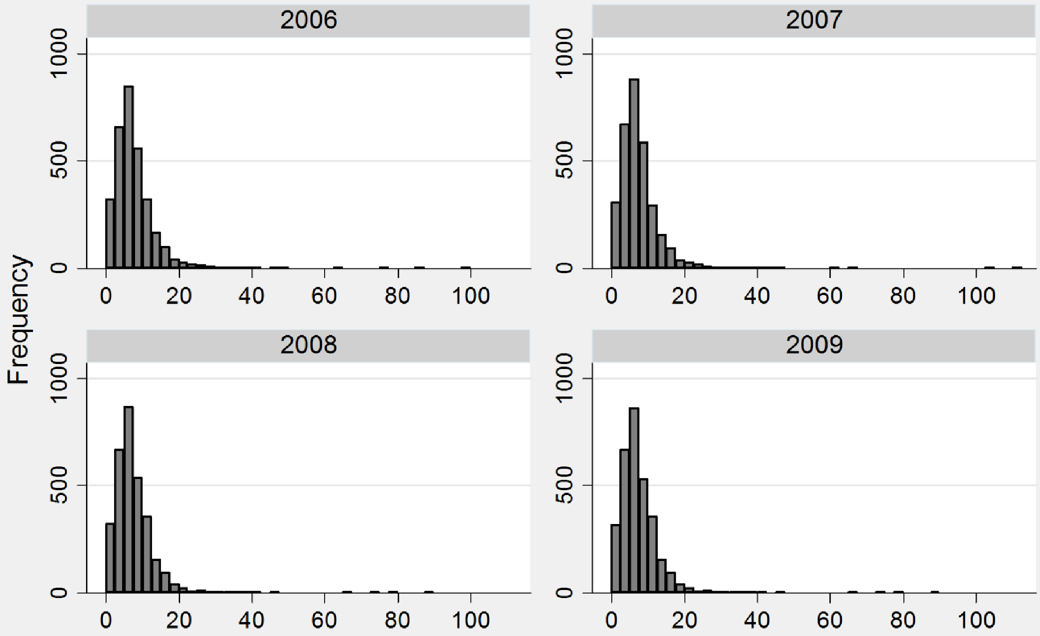


Figure 4: Low incomes

(percentage of people belonging to lowest 40% of income distribution)

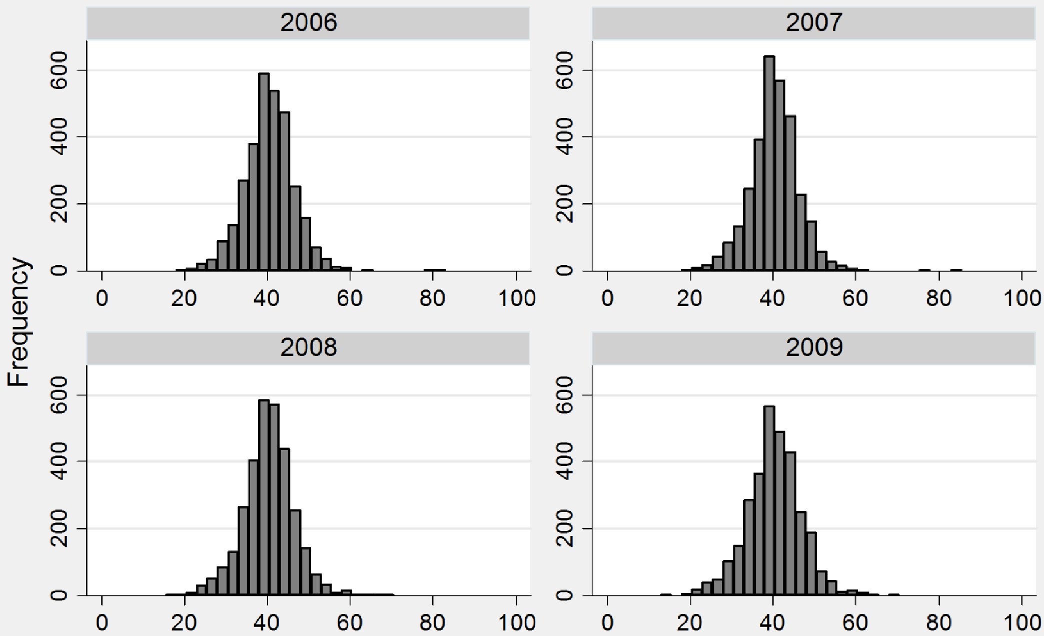


Figure 5: High incomes

(percentage of people belonging to highest 20% of income distribution)

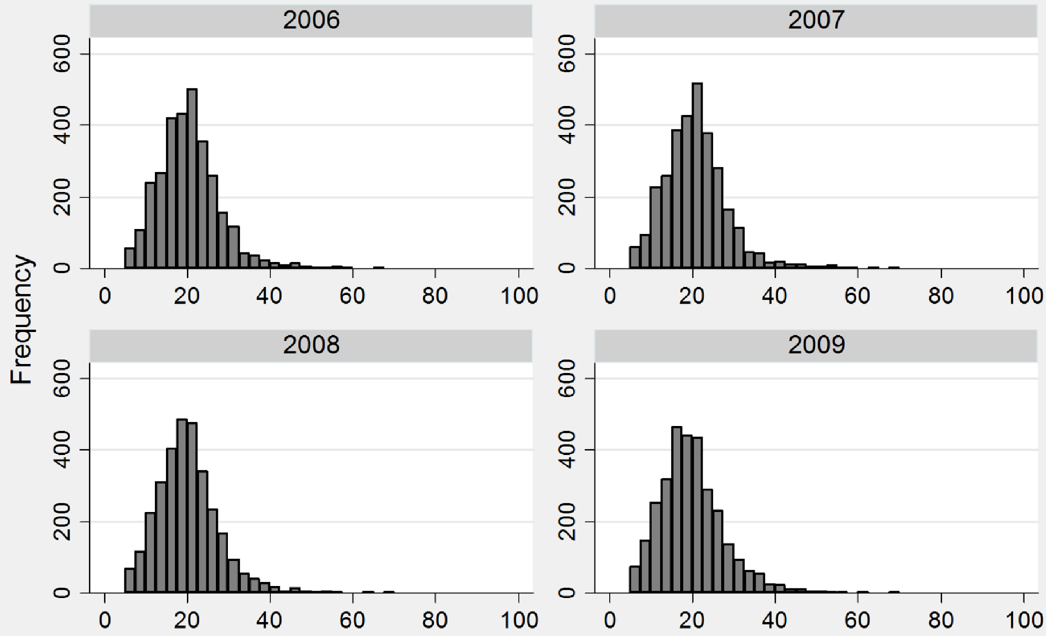


Figure 6: Percentage of working people

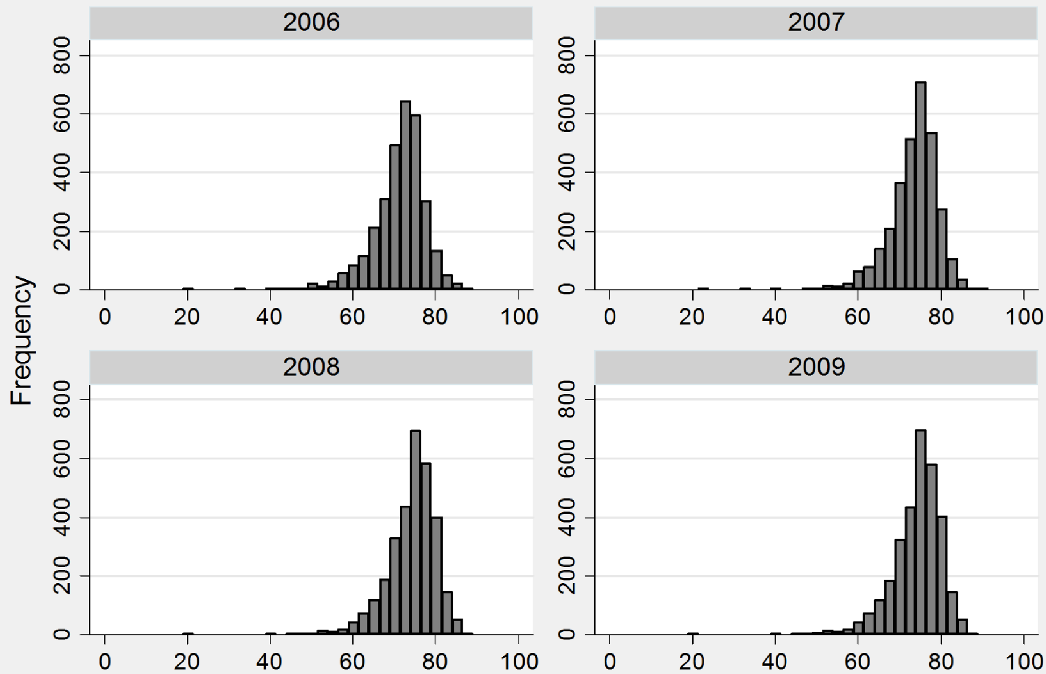


Figure 7: Percentage of self-employed people

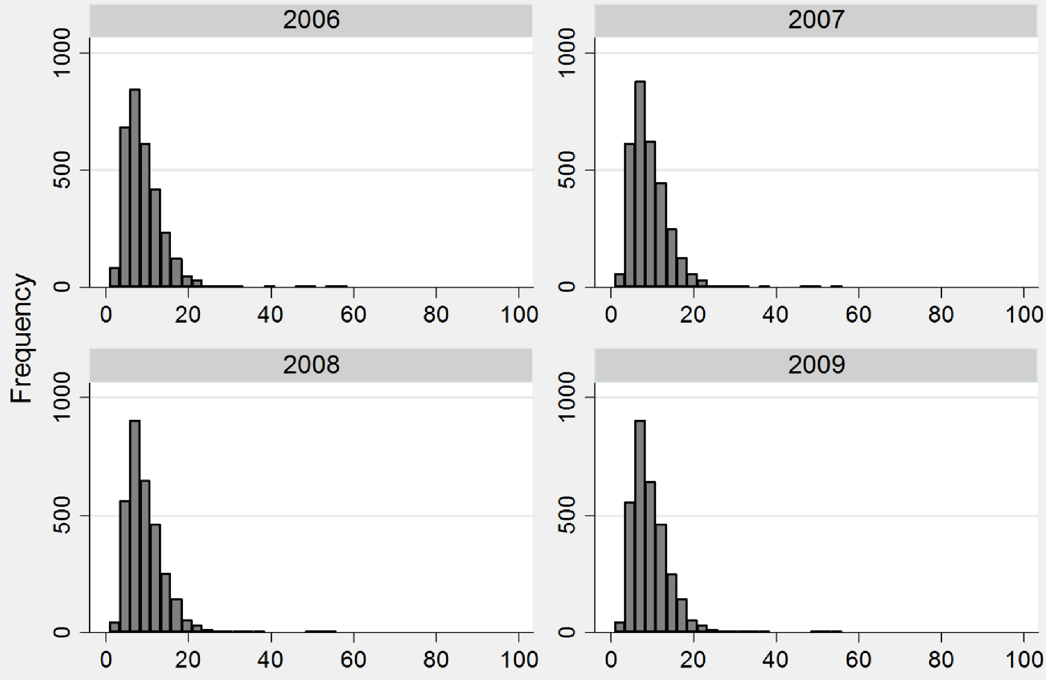


Figure 8: Number of people receiving social assistance
(per 1,000 inhabitants)