ERASMUS UNIVERSITEIT ROTTERDAM
ERASMUS SCHOOL OF ECONOMICS

Modeling the effects of a policy change on supplier induced demand for sixteen hospital treatments

Master's Thesis

Submitted to

Dr. Rudy Douven

Co-reader

Dr. Pieter van Baal

Author:

Theodoros Chatzivasileiadis

Student number: 356124

Programme: MSc Economics and Business

Specialization: Health economics

Rotterdam, 20 January 2013

Table of contents

Chapter-page

List of tables and graphs

Preface

To this point I would like to thank my supervisor Dr. Douven for the support, guidance and understanding. I would also like to thank my co-reader Dr. van Baal for the vital feedback. Last I would like to thank K. Valogianni and M. Konstantinidou for their help in shaping the grammar and syntax of this text.

Nomenclature and Abbreviations

Adm	Number of total hospital admissions (both inpatient and outpatient)
CBS	Centraal Bureau voor de Statistiek (Central Bureau of Statistics)
COPD	Chronic Obstructive Pulmonary Disease
DBC	Diagnose Dehandeling Combinatie (Diagnose Treatment Combinations)
DID	Difference-in-Difference
FE	Fixed Effects
FFS	Fee-For-Service
GLS	Generalized Least Squares
GP	General Practitioner
ISHMT	International Shortlist for Hospital Morbidity Tabulation
LSDV	Least Squares Dummy Variable
OLS	Ordinary Least Squares
PBP	Patient-Based-Payment
RE	Random Effects
SE	Standard Errors
SID	Supplier Induced Demand
TIA	Transient Ischemic Attack

Abstract

Introduction: Changes in the budgeting system of hospitals affects the admission rates. We estimated the effect of the 2001 policy change in the budgeting system of Dutch hospitals by focusing on the ambiguity of treatment. We used different statistical models to assess the effects of the policy change on 16 conditions.

Methods: The Difference-in-Difference model was used to describe the 2001 policy effect on admissions by separating the total population into two groups, where we believed ambiguity of treatments was higher, and the control group. The DID model was transformed to fit into the panel nature of our date, and to fully capture the heterogeneity between the age groups. We also conducted a robustness analysis in order to validate our assumptions.

Results: The average results of the simple DID model is 29, on the aggregated model is 35 and on the mixed effects model is 28.8. Bootstrapping had only a minimal effect on the SE.

Conclusion: On the modeling side, this model can answer some questions as to how ambiguity of treatment affects the hospital admissions of specific age groups, under the light of a policy change. Based on our results in general, the difference between the two groups identified in the population on the hospital admissions, before and after 2001, increased by 30 in 10000 hospital admissions. Based on our assumptions and the positive and significant result that the policy change of 2001 affected more the age group where ambiguity of treatment is higher.

I. Introduction

I.1 General literature review

In the international literature we can find a lot of work dedicated to the effects of a policy change on admissions both on theoretical and empirical level. Moreno-Serraa and Wagstaff, (2010) [14] concluded, based on a panel analysis of 28 counties for the period 1990-2004, that moving to fee-for-service and patient-based-payment systems affect the average admission rates and the average length of stay (FFS increased admissions but not length of stay, when PBP had no actual effect on admissions but affected negatively the length of stay). In this study we examine the effect of a policy concerning the budgeting of Dutch hospitals on the number of admission. Douven, Mocking, and Mosca, (2012) [6] analyzed the effects of changes in physicians density on regional variation. They found a significant effect of physicians density on admissions when the payment system is output based. In this study we will not focus on regional data but we will focus on a macro level relation between the number of admissions and a policy of releasing budgets. Physicians can be an important component of this relation. Last Pomp, (2009) [15], created a model to describe the magnitude of SID within the Dutch hospitals for the 150 highest, in terms of admission, conditions for the years 2006 and 2007. He concluded that in general the results are diverse and the overall effects, in terms of magnitude are small.

Supplier induced demand is a crucial aspect in Health economic literature. Starting with Arrow, (1963) [2], SID is defined as the shift in the demand curve of health resulted from information asymmetry. Information asymmetry can be a problem when patients have less knowledge of their disease than the physician. Based on that asymmetry of information, it becomes easier for the physicians to implement SID. In general SID is the situation when physicians "*knowingly induce their patients to consume other than the optimal amount of care*" Folland, Goodman, & Stano, (2010) [7]. Supplier induced demand complicates the cost containment procedure because possible gains in efficiency can be tackled by increase in productivity of more ambiguous treatments. SID is unwarranted when it increases production for treatments with no added value to the patient, something known as the "flat of the curve medicine" Getzen, (2006) [9]. At this point it is very important to define ambiguity of treatment and connect it with SID. We define ambiguity of treatment as the

grey area where it is not clear for the provider to decide if a patient should follow a treatment or not. The more ambiguous a treatment the larger the chance for SID.

Comparing the effect on the 2006 Health Insurance Act in the Netherlands (i.e a complete movement to managed competition for the health insurance) at a GP level, van Dijk, et al., (2012) [18] concluded that the 2006 Health Insurance Act had a limited effect on physician-initiated utilization of health. Focusing on the hospital admissions in the Netherlands ,van de Vijsel, Engelfriet, & Westerta, (2011) [17] concluded that the open ended budgeting system for the hospitals that was used to reduce hospital waiting times was not effective. The problem in the effectiveness of the policy change to open ended budgeting system for hospitals created due to the high increase of demand for health care. This is very important for our study because we make a link between the increased demand for health care after 2001 with SID for those treatments for which we expect that the ambiguity of treatment is high. Another analysis of the effects of a policy change in the budgeting of Dutch hospitals conducted by Folmer & Westerhout, (2002) [8]. They analyze the effects of the introduction of the lump sum for specialist's payment system that was introduced by the Dutch government in 1995 for cost containment. They concluded that the introduction of this policy had a real negative effect on welfare based on data for 1995-1999.

I.2 Institutional background of the Dutch hospitals budgeting systems

After the 90s the Dutch governments systematically tried to reform the Dutch health care system. In the last decade of the 20th century the health care system has been in a transition state from a government regulated system, toward a system of managed competition Schut & Varkevisser, (2012) [16]. During the 80s, a decade before the start of the reformation in the 90s the hospital expenses were covered by a global budgeting system. In the first half of the 80s the budget was fixed for each hospital. The budget was based solely on the operation expenses of each individual hospital. After 1985 the budget became variable (i.e was more based on specific characteristics of each hospital) and it was based on utilization of care Based on that system, a production higher than originally agreed reduced the per diem rates of the hospital (on outcome level) proportionally. This discouraged the hospital management to exceed the production above the negotiated one.

The policy change of 1985 described above did not have the desired results (i.e cost containment). "*The problem generated by the incentives given to the medical specialists directly"* [16]. In the Netherlands, medical specialists are self-employed excluding those

working in University medical centers. The medical specialists were paid on a FFS bases. In the beginning of the 90s the Dutch government introduced an annual macro budget for the total revenues of specialists. If the specialists exceeded the budget due to overproduction then all fees of specialists were reduced proportionally. With this policy the policy makers tried to remove incentives from the medical specialist for overproduction. Though this policy intensified the problem more than solving it as can be explained by the prisoners dilemma. "*... each individual specialist had an incentive to produce more to compensate a potential drop of income due to a general fee reduction as a result of increasing production by other specialists*" [16].

In 1995 there was one more attempt to reduce the incentives for overproduction. This new policy gave the medical specialists the opportunity to have a fixed budget in return for exemption from the fee reduction scheme. The result of this particular policy was the actual reduction of the incentives for overproduction. Though this new policy created another problem, it affected largely the waiting times in Dutch hospitals.

In 2001 the Dutch government changed once again the financing system of the hospitals in general. The new "Cash on the nail" policy allowed insurers to reimburse additional production than the one initial agreed. So the budgeting system moved to an activity based system, which was actually open ended. Health care provision was heavily affected by those changes. Health care provision in addition to that previously agreed with the insurers, needed to be reapproved. The additional services resulted in extra expenditure, based on recalculation of the variable costs. This way more health care could be purchased in order to reduce waiting times and to improve the given services.

This however, created a number of problems. In general we can simply say that additional health care production, was rewarded with additional resources, creating a lot of "space" for more ambiguous treatments. Moreover, the changes allowed for serious variations in the physicians' fees, depending on the total hospital production. In addition to the open ended system the new policy lifted the restrictions on the number of medical specialist in order to increase supply and further reduce waiting times in Dutch hospitals. The effects of the new policy were immediate. The admissions increase directly, decreasing the waiting times. Though the increase of admissions was unconstrained and continued for a long time.

I.3 Contribution

This paper studies the effects of a policy change affecting the payment system of hospitals and physicians on the admission rates. We aim to study the effect of various treatments. We show that fixed budgets can cause a reduction in the number treatments when treatment ambiguity is bigger, By focusing on all age groups in the Dutch population. This analysis provides a tool to examine the effects of a policy change through time and analyze the effects on admissions by condition. The idea is to concentrate on specific conditions and age groups, which are more susceptible to policy changes affecting the budgeting system of hospitals and physicians. With this analysis we are able to identify whether specialist behave differently if a treatment is more ambiguous. We quantify this effect and show the magnitude of this effect for different conditions / treatments.

I.4 Study outline

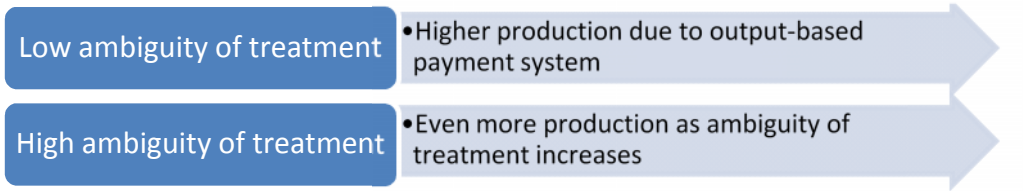
This study is organized as follows. First we identify all the background information that supports our hypotheses. Next, we explain the model and the estimation techniques that will be used. Next we present the data and through some additional graphs and descriptive statistics we will help to deeper understand the conditions described in the paper. Next we analyze the results of the regression models. Finally, we present four robustness tests that challenge the several assumptions we made in order to conduct our model.

II. Hypotheses

II.1 Main Hypotheses

Our first hypothesis is that reimbursing additional production in hospitals strongly increases hospital output for more ambiguous treatments. This hypothesis is based on the theory of SID (therefore confirmation of this hypothesis is evidence of SID). As a result, we expect to find different rate increase in hospital admission rates after 2001 for different age groups. Our second hypothesis is that ambiguity of treatment differs across age groups. Finally, we

assume that the age groups, mostly affected by the policy change, are those with the highest admission rates prior to the policy change. This is an assumption that should be discussed further. By explaining the effects of the shock in the system in 2001, we can prove that the pre-2001 budgeting system constrained admissions in treatments where ambiguity of treatments is present. We test whether the policy change increased the problem of SID for conditions in which there is more ambiguity of treatment by focusing on specific age groups. The basic idea is that in production-based payment systems for hospitals the utilization of health is higher compared to a payment system that is restricted. Another hypothesis is that the type of treatment differs for young and older patients. Based on that the main hypothesis is that the production of health is even higher if ambiguity of treatment is higher.



II.2 Supporting hypotheses

This analysis explores the effects of the policy change in 2001 at a condition level. Our focus is put on different age groups and differs from the existing literature that focuses on specific geographical areas. The reasoning behind this differentiation lies firstly on the disease side. Each condition has specific characteristics that cannot be captured on an aggregated level. On the age side, reaction can be explained by two dimensions: The “mentality” of these age groups, for example older people can be more easily manipulated concerning their health. As an example we use Cataract where the prevalence of this condition is higher amongst older people. We expect ambiguity of treatment and consequently SID to be higher because older patient can be easily convinced and rarely will seek a second opinion. The other dimension concerns the morality of the physician. Despite being immoral, "ambiguity of treatment" or encouragement of unnecessary procedures to older patients can be considered easier than younger patients under the prism of ethics and morality. Therefore, we need to analyze each disease and each age group differently. This will allow for clearing the effects of policy changes on growth of admission rates and "ambiguity of treatment". A question that we must answer is whether age is an explanatory factor for the chance of supplier induced demand.

In order to conduct this analysis we made some additional hypotheses concerning the changes in admission rates. The main idea is that the growth of admission rates is based on a number of factors:

- morbidity
- eligibility for a procedure
- technology and efficiency changes in the medical sector
- "ambiguity of treatment", under-consumption for unplanned care and among other factors that have minor effects

For this analysis we will assume that the morbidity has not been changed significantly for a specific age group between the years 2000 and 2001. This is a strong assumption, though the inclusion of morbidity will include too many dimensions to our model, therefore for simplicity and for the time being we ignore it. This assumption is of importance in our analysis. In the case were the true morbidity changed in our breaking point (i.e. year 2001) the effect we see comes from the morbidity change and not from the reaction of the physicians. Though, to our knowledge there are no significant changes in the morbidity of the population between 2000 and 2001 for the condition in question. On the other hand, one factor that affects admission rates is technology improvements and medical efficiency. Thus, the main issue is how to extract the increase of admissions that caused by technology from the increase caused by "ambiguity of treatment", so that we can test our assumptions.

In order to connect the changes of admission rates, technology and efficiency changes with "ambiguity of treatment" we make the following assumption:

“For a specific disease, which has a specific set of characteristics, any shocks in the system will affect each age group differently. On the other hand changes in technology and efficiency will have the same or similar effect to all age groups that ail from this specific condition.”

Consequently, this study follows a multidimensional analysis of the hospital admissions in the Netherlands focusing on year and age effects, by examining the effects of the 2001 policy change in a disease level. In this analysis, the effects of the policy change in 2006 have not been discussed. Additionally, we treat the whole period after 2001 as uniform because there is not enough information to generate a pattern after 2006. Therefore, we aim at proving the existence and magnitude of "ambiguity of treatment". To this end we base our

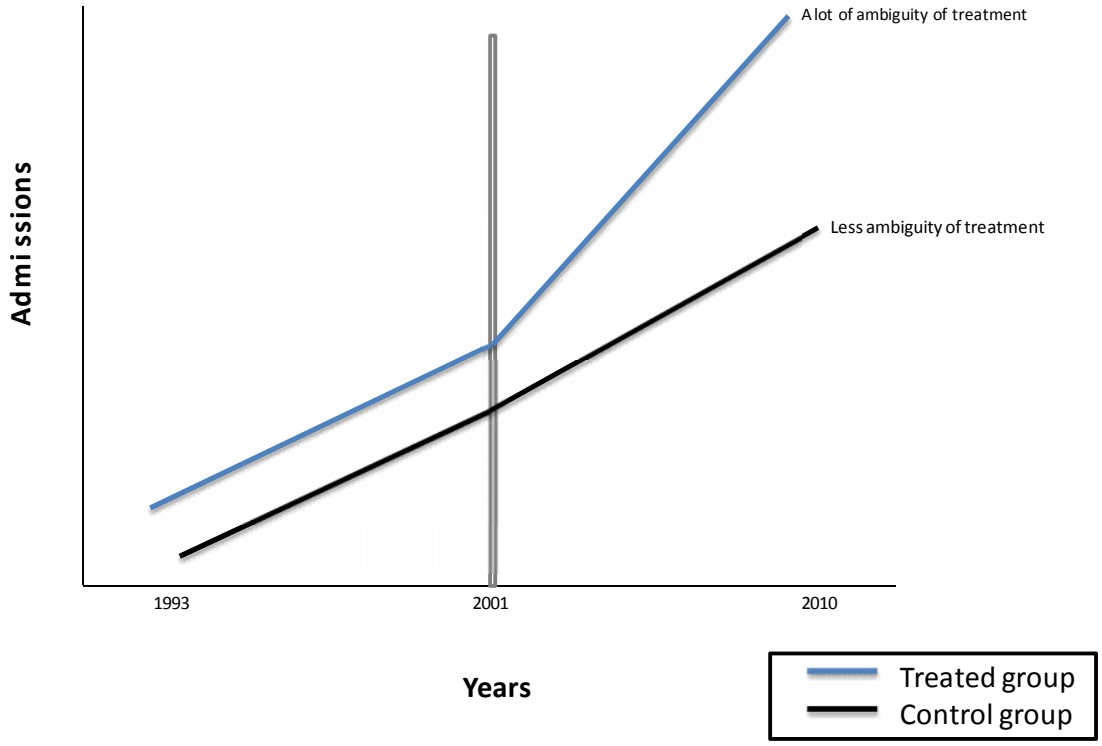
model on the effect of the policy change of 2001 on the hospital admission rates by assuming that the 2001 shock is stronger on admissions.

III. Methodology

III.1 The Difference-in-Difference regression method

As far as the methodological part of this study is concerned, we employ the Difference-in-Difference regression model. By using the Difference-in-Difference regression model, we can separate our population in two age based groups and test the effects of the 2001 policy change. Firstly, we create a simple Difference-in-Difference based on Ordinary Least Squares without any covariates (Angrist & Pischke, 2008). Then we make additional changes to the model by introducing covariates to the regression. This last model is estimated with Generalized Least Squares. As we explain below, with this approach we acquire the information possible from our dataset. With a simple Difference-in-Difference model we test for changes in the admission rates of specific age groups that resulted from the policy change. If we find evidence for an "ambiguity of treatment" effect, then we calculate this effect. If the difference in difference variable is significant and has the correct sign according to the period of interest, then "ambiguity of treatment", actually have played a role on specific age groups only (different groups according the condition). If in the years following the policy change, the slope of admission rates gets steeper for the treated group, as shown on the graph 1A below, we consider this change as being caused by "ambiguity of treatment".

2A Graphic representation of the two group / two period DID model



More analytically, in the first step we use the simple Difference-in-Difference regression model. In its simplest form, the Difference-in-Difference model consists of two groups that are observed in two different periods of time. Generally there is one group that undertakes a "treatment" or is affected by a policy called treated. Additionally, there is the "control" group. This group is not exposed to any changes. Since the same units of a group are observed in both time periods, with the simple Difference-in-Difference regression we calculate the average effect of each group, we subtract the average effects of the control group from the treated group. Employing this approach we can remove the bias, caused by preexisting differences between the groups. Moreover, this way we can compress the bias from overtime comparison of the treated group that resulted from time trends.

So for any member of any group, the model that shows the effect on admissions from the policy change is:

$$Admissions = \beta_0 + \beta_1 group + \beta_2 time + \beta_3 D + \varepsilon \quad (1.1)$$

where $Admissions$ represents the sum of outpatient and inpatient admission rates by age group at each year. The variable $group$ is a dummy that equals 1 if an age group belongs to the treated group and zero otherwise. The coefficient β_1 of the variable group will show the

difference of the average admission rates of the two groups prior to the policy change. The variable $time$ is also a dummy that equals 1 for all years after 2001. The coefficient of this variable captures time trends that affect the admission rates even if the ambiguity of treatment does not play a role. Last the variable D is the interaction between group and year as a simple product of group times year. This variable equals to 1 if an age group belongs to the treated group and it is observed after 2001. The coefficient of this interaction term is the one of interest. So the Difference-in-Difference coefficient β_3 can be calculated by:

$$\hat{\beta}_3 = (\overline{adm}_{t,a2001} - \overline{adm}_{t,b2001}) - (\overline{adm}_{c,a2001} - \overline{adm}_{c,b2001}) \quad (1.2)$$

Were $\overline{adm}$ represents the average admission rates, t and c represent the "treated" and the "control" groups respectively. Last a2001 and b2001 indicate after and before 2001 admissions respectively.

In our case the Difference-in-Difference coefficient β_3 reveals the effect of the policy change. If the coefficient is statistically or economically significant for α lower than 10% and positive then we can conclude that the treated group is affected more by the policy change in 2001 than the control group. Based on the assumptions described in the assumptions section (II.2) this is an indication that ambiguity of treatment in the treated group plays a role.

The next question of the presented analysis is whether we can quantify the effect of "ambiguity of treatment" in the population and measure it. Based on the total assumptions, we assume that the growth in admissions is caused by technology, efficiency improvements, ambiguity of treatment and other factors that we motioned above. For this analysis we assume that all other factors except "ambiguity of treatment" have the same effect on all age groups.

Difference-in-Difference regression model [1] is an analog to the population average method. The interaction term D in the equation (1.1) shows the difference of the mean admission rates for each group in the two periods as in equation (1.2). With some calculations we get the same result by the *commutative property*:

$$\hat{\beta}_3 = (\overline{adm}_{t,a2001} - \overline{adm}_{c,a2001}) - (\overline{adm}_{t,b2001} - \overline{adm}_{c,b2001}) \quad (1.3)$$

The idea is that the increase of admissions in the control group is caused by technology and efficiency improvements, increase of morbidity and other factors without the effect of "ambiguity of treatment". Thus, if we extract the effect of the control group before and after 2001 from the effects of the treated group, we get the pure effect of ambiguity of treatment in the treated group. This is based on the assumption we made above that an increase in technology and efficiency will affect all age groups the same way. But the effect of the ambiguity of treatment differs. So the equation (1.3) gives us the magnitude of ambiguity of treatment in the treated group.

III.2 The random effects estimator for Difference-in-Difference models

This Difference-in-Difference regression model with two periods and two groups is generally problematic [13]. Its validity is based on very strong assumptions. Mainly the model assumes that admissions trends of the two groups over time would be the same in the absence of the policy change. In order to relax those assumptions we change the specification of the model. Based on the panel nature of our data we can include a full set of time and group effects. This way we transform our model from a simple Difference-in-Difference to a Fixed Effects model. This model is actually the Least square dummy variable model (LSDV) which is equivalent to the fixed effects estimator [19]. By introducing all these dummies in our model, we control for different age and year trends that are irrelevant to the policy change but may have influenced the outcome. Consequently, we allow each group to follow a different trend, compared to the trend that is captured in the Difference-in-Difference coefficient. By employing this approach, the model can control for the unobserved heterogeneity between age groups and years.

To sum up, we need to explicitly include the two sources of variation in the intensity of the policy effect, the age and the year effects. With this model, the identification of the policy change is based on changes in admission trends of the different age groups that exist only in the treated group and created exactly after the change of policy. So, with many periods and groups, the fixed effects Difference-in-Difference model can be written as:

$$Admissions_{it} = \beta_0 + \sum_{n=1}^{21} \beta_n * Age_i + \sum_{n=22}^{37} \beta_n * Year_t + \beta_{38} D_{it} + e_{it} \quad (1.4)$$

Where Age_i is a specific age indicator, that captures the age specific trends. In this model each of the 21 age groups (i.e the 5 year age intervals) in our dataset are included as

dummies together with the 17 year dummies. $Year_t$ represent the year variables included in the model to capture trends of the admissions that are variable through the years but are irrelevant to the policy change in 2001. Last D_{it} indicates the difference in difference variable. This variable equals to 1 if an age group belongs to the treated group and it is observed after 2001. The coefficient of this interaction term is the one of interest that would indicate the effect of the policy change.

If we estimate (1.4) with OLS then this may yield problems in statistical inference (Hounkannounon, 2011) [11]. One of the major problems of any Difference-in-Difference model is serial correlation. Firstly, the dependent variables are highly positively serial correlated (Marianne, Duflo, & Mullainathan, 2004b) [13]. Moreover the model is based on long time-series (17 periods) which typically leads to serial correlation. So those factors affecting serial correlation in combination could make the standard errors of the estimated coefficients and severely understate the standard deviation of the estimated coefficients [13]. Up until now the model is based on the Least Square Dummy Variable (LSDV) estimator (Baltagi, 2005) [3] estimator that is estimated by OLS and it is equivalent to the fixed effects estimator. This model controls for the unobserved heterogeneity that might exist in the data. Next we discuss the inclusion of random effects in out model.

In order to control for these kind of problems we replicate the model as showed by Hansen, (2007) [10] and Marianne, Duflo, & Mullainathan, 2004a [12] with some adjustments. The equation at an aggregated level is:

$$Admissions_{git} = year_t + age_g + \beta_2 D_{gt} + \beta_3 X_{gt} + \varepsilon_{gt} + u_{git} \quad (1.5)$$

Were “I” represents a specific age cluster, ”g” indexes the specific group of age clusters (i.e the high and low admission groups) and last t represents time. In this model a full set of year effects is included, $Year_t$, alongside with dummies that represents the two age groups. As before the variable D_{gt} represents the policy variable. X_{gt} contains the individual age specific covariates. Even though in our model, this variable has been omitted due to lack of existing data. Lastly, the variable e_{gt} accounts for the unobserved random effects at a groups-time level and u_{git} is the individual age cluster error. The coefficient of interest is β_2 that will give us the difference-in-difference of the two groups. The model 1.5 is a multilevel model. It consists of two different equations that can be separated. So we can rewrite the equation 1.5 as:

$$Admissions_{git} = \xi_{gt} + \beta_3 X_{gt} + u_{git} \quad (1.6a),$$
$$\xi_{gt} = year_t + age_g + \beta_2 D_{gt} + \varepsilon_{gt} \quad (1.6b)$$

This model at an individual age cluster level allows both intercepts and slopes to differ across all groups and years. This way we moved from a simple LSDV (Baltagi, 2005) [3] model to a Random effects model that allows apart from the random effects age and year specific effects. (Hansen, 2007) [10] showed that the OLS estimator applied to 1.6b is inefficient when ε_{it} are serially correlated and heteroskedastic. The problem is that in the Fixed effects estimator, we use the residuals of the error term to calculate the variance matrix Ω_g that results in severe bias when T is relatively small. One way to control this problem is the estimation by Generalized least squares (GLS) [19].

With this model we can control for random effects, if they exist in our model. A simple Breusch-Pagan Lagrangian multiplier test for Random effects (Breusch & Pagan, 1980) [4] will be used. This way we can decide if random effects exist in our model. If so, then the LSDV model is not appropriate and the random effects model needs to be estimated. The parallel tends assumption is the main disadvantage of the DID model. By introducing random effects in out model, we relax the assumption of parallel trends and we can get a more correct estimate of the effect caused by the policy change.

Finally, as proposed in the relevant literature (Hounkannounon, 2011) [11] based on the BDM models as above, we will apply the bootstrap resampling method to 1.6 in order to correct the neglected heterogeneity and temporal correlation that lead to wrongful hypothesis testing.

Consequently we estimate the Difference-in-Difference model by using equations 1.1 and 1.6.

IV. Data

IV.1 Dataset description

Our data sample consists of information connected to hospital admissions for the years 1993-2010, from CBS (Centraal Bureau voor de Statistiek, 2012) [5]. Information derived from Pomp, (2009) [15] has been used for one other variable.

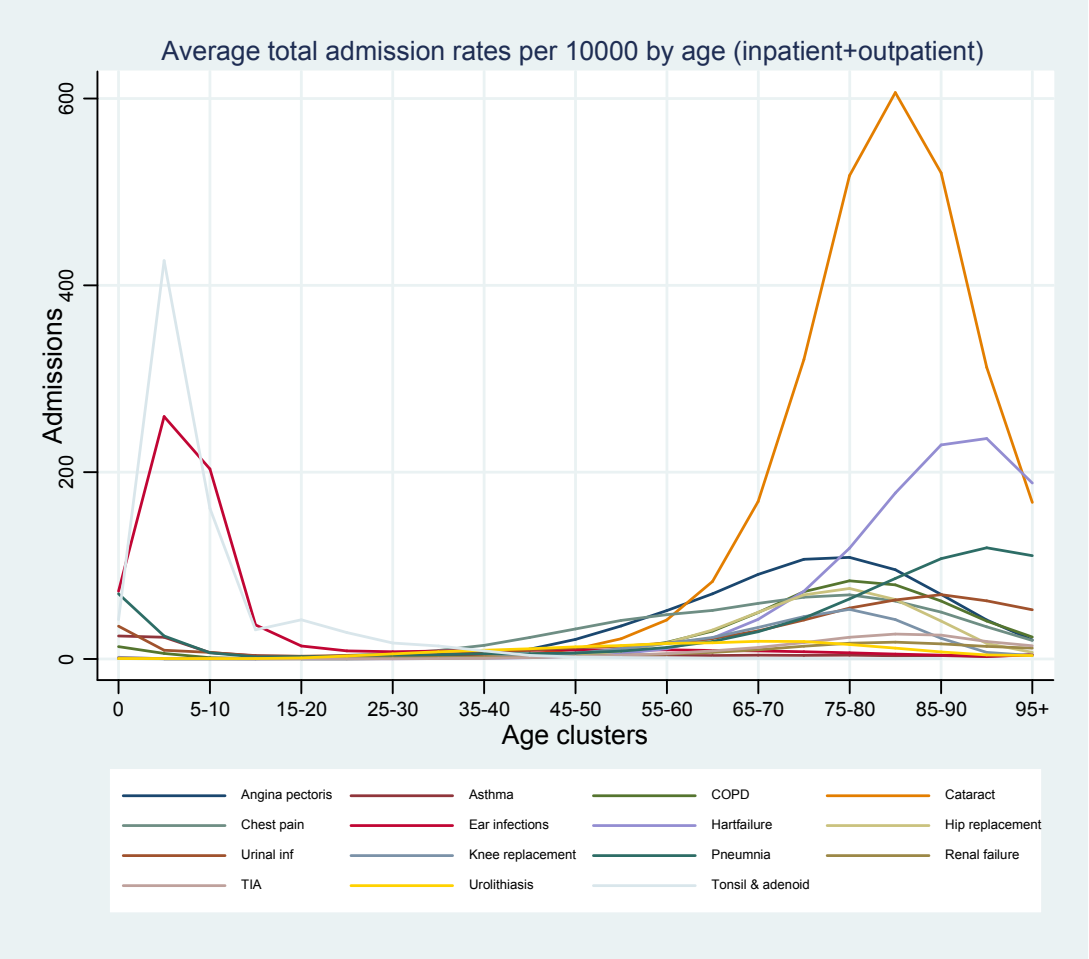
From the CBS data we use information of 16 conditions after "converting" them to the International Shortlist for Hospital Morbidity Tabulation (ISHMT) system[1]. As a result, we will assume that all conditions in a ISHMT code have the same chance for SID as described by Pomp, (2009) [15]. The variable "Chance of SID" takes three values, 0 for small chance, 1 for middle chance and 2 for high. The 16 conditions we examine are; Chest pain, Hart failure, Infections of the urinal system, Gonarthrosis (Knee replacement), Tonsil and Adenoid, Ear infections, Cataract, Cholelithiasis, TIA, Asthma, Coxarthrosis (Hip replacement), Angina-pectoris, Urolithiasis, Renal failure, COPD and Pneumonia. We chose 8 medical and 8 surgical procedures or conditions. Then according to Pomp, (2009) [15], we tried to find equal number of conditions for each "Chance of SID"; Five for high, five for middle and 6 for small chance of SID.

Given code	Condition	Chance of SID
0	Angina pectoris	0
1	Asthma	2
2	COPD	0
3	Cataract	2
4	Chest pain	2
5	Cholelithiasis	1
6	Ear infections	2
7	Hartfailure	0
8	Hip replacement	1
9	Infections of the urinal	1
10	Knee replacement	1
11	Pneumnia	0
12	Renal failure	0
13	TIA	0
14	Tonsil and adenoid	2
15	Urolithiasis	2
Sourse CBS and Pomp(2009)		

The basic outcome for our study is the total admission rate by condition, as described above, for the years 1993-2010 by 21 age cluster which contain 5 age-year intervals, except for group 0 that contains information on the newborns only. The total admission rates are per 10000 inhabitants in order to control for the population structural changes through the years. The Graph 1.1, shows the average admission rates of all years by age group for each condition. This helps us understand which conditions are more concentrated among young and which among older age groups. This clarification alongside with table 1.2, is used in separating the age groups into the two categories we discussed before, the Control and the Treated group.

[1] The ISHMT system is broader than the DBCs.

3. Graph 1.1 Average admission rates by age cluster of all years.



As expected, Ear infections and Tonsil and Adenoid prevail in rather young ages 0-3. The rest of the conditions prevail in older ages, mostly between the ages 65-90. It is important to mention that Cataract increases steeply after 60 years of age until 85 where it decreases. This condition has the highest admission rates we observe in our dataset followed by heart failure and Angina pectoris. The smallest number of admissions occurs for Asthma Renal failure and TIA.

IV.2 Treated and control group separation process

In order to run the basic model we need to form a treated and control group. We have two periods, pre and post 2001. What is very crucial is how to create the two age groups. To this end, we separate the inpatient from the outpatient admissions. We define the two groups based on the inpatient admission rates only. We believe that inpatient admissions can clearly capture the patients' "need" for health care and not the total admission rates. We

classify the groups based on the admission rates of 2000. This year is the one before the policy change and has less missing observations than all other previous years.

Then we separate three or four age clusters with the highest admission rates (15 or 20 years). Those age groups represent the treated groups. It is important to say that we choose the age categories in order to be consecutive and not separated. All other groups represent the control group. Different conditions will have different age groups as treated and control. The reason for that is simple. As it can be seen in graph 1.1 above, different conditions have different distributions among age groups. One major problem is the missing values. In most conditions very young ages have missing values which may mean zero admissions but also that we have no information. One solution is to omit those ages that have missing values for more than half of the years of question.

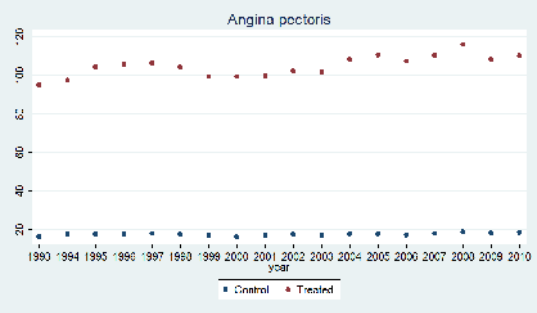
Since there is no clear scientific way to make the separation, we used two methods in order to classify the age clusters into groups. Table 1.1 depicts this procedure. First we created a table with the admission rates for every age cluster by ISHMT for the year 2000. We used that particular year as a reference year. Then, by empirical observation, we separated the age groups by identifying the age cluster with the highest admission rate. As can be seen in the table 1.1 below, in every ISHMT the neighboring age clusters also have very high admissions compared to the rest of the sample. So this way we chose the three or four age clusters with the highest admissions. In order to validate our method of separation we used a second method. We used K-means as a clustering method to separate the ages into two groups based on admission rates. The results of this method are used in the Robustness check (VI) part of this study and the K-means tables are included in the Annex.

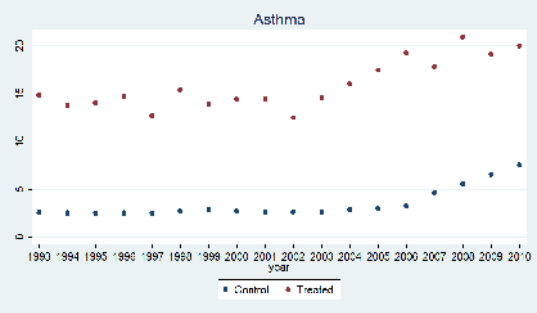
The procedure of separating the age categories into two groups is essential for our analysis. The basic requirement of the DID model is that there are two groups that are observed in two different periods. So in order to use this methodology within the context of this study, the separation of the age clusters into two groups was a conditio sine qua non.

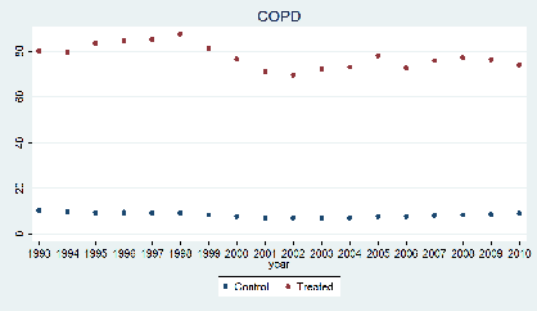
We now go a bit further by analyzing the two constructed groups for each condition/treatment separately, by focusing on the changes in total admissions by year. First, we create two new variables: total admissions as a share of the total population for each group. So we have two observations of total admissions for each year for each condition. By plotting these variables by condition we get a clear view of the changes that

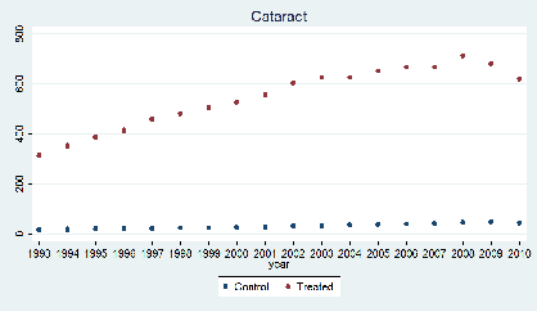
occurred in each group through time. Results are shown on graphs 1.2 through to 1.17 below.

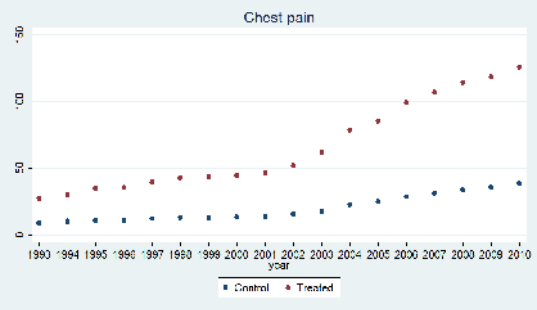
4. Graphs 1.2 through 1.17 Admission rates corrected for population changes by group.

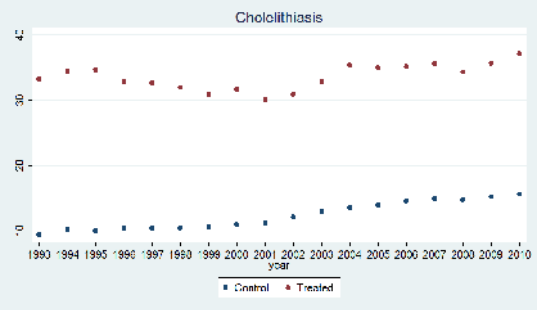


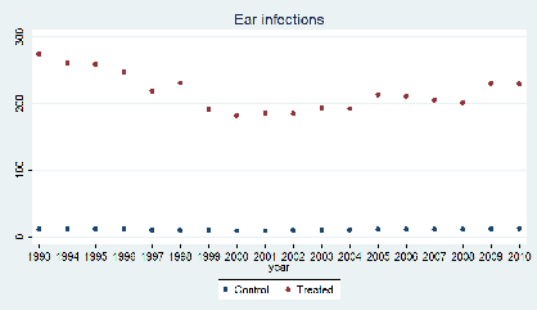


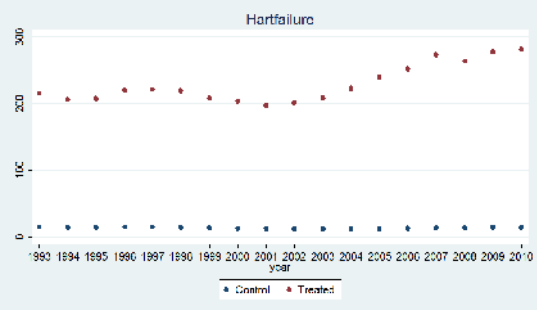


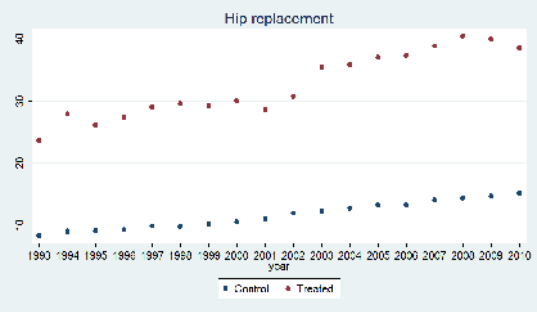


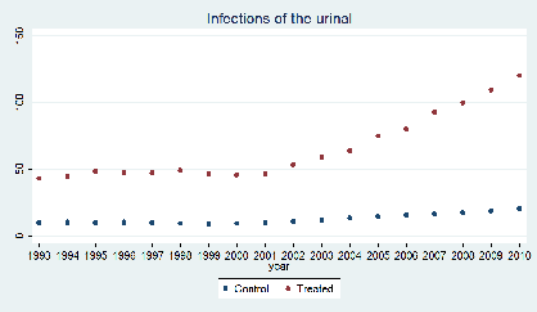


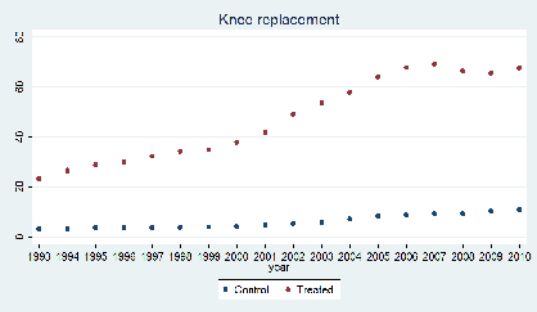


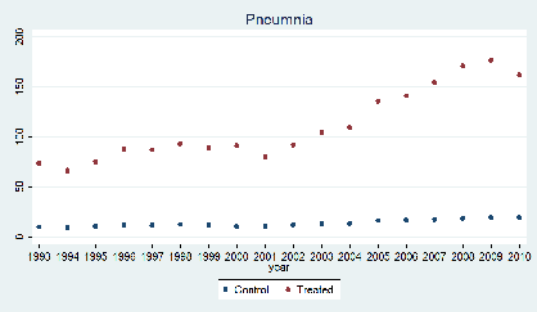


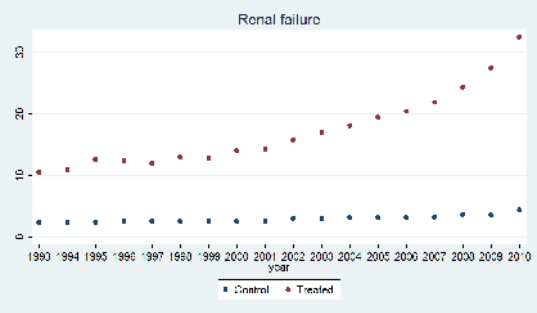


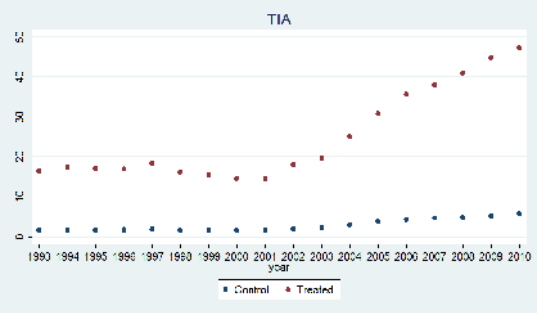

Hip replacement
year
Control
Treated

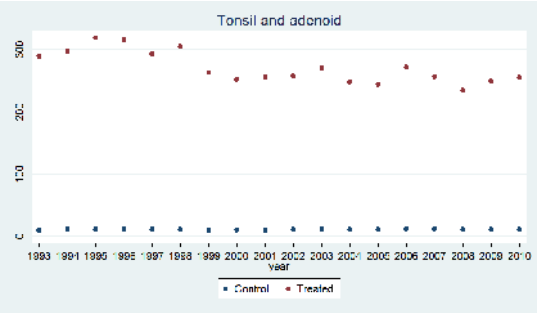

Infections of the urinal
year
Control
Treated

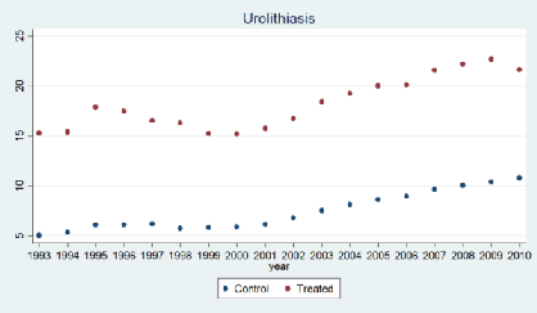

Knee replacement
year
Control
Treated


Pneumnia
year
Control
Treated


Renal failure
year
Control
Treated


TIA
year
Control
Treated


Tonsil and adenoid
year
Control
Treated


Urolithiasis
year
Control
Treated

5. Table 1.1 Total admission rates by age cluster for each condition in 2000.

Age	Angina Pectoris	Asthma	COPD	Cataract	Chest pain	Cholelithiasis	Ear Infections	Hart failure
0		18.9	10.6		0.2	0.2	14.3	1.7
1		21.2	3.7	0.1			8.5	0.1
2		5.4	0.6	0.2	0.1	0.1	9.6	
3		2.6	0.2	0.1	0.1	0.2	11.7	0.1
4		1.9	0.3	0.2	0.9	1.2	7.3	0.2
5	0.1	2.2	0.4	0.2	2	4.2	5.4	0.2
6	0.3	2	0.4	0.3	3	8.9	4.7	0.2
7	1	2.4	0.7	0.2	5.1	11.8	5	0.3
8	3.6	2.6	1.4	0.3	8.3	12.2	5.6	0.6
9	8.3	2.5	2.7	0.9	12.3	13.6	6.5	1.5
10	17.4	2.7	5	1.7	18.8	15.3	5.8	2.9
11	29	2.6	8.1	2.4	24	18	6.1	5
12	43.1	2.5	16.2	4.5	28.8	22	6.7	10.1
13	57.8	2.6	26.3	8.9	31.3	23.2	5.5	20.3
14	78.1	2.6	48.9	20.6	35.7	26.1	5	37.8
15	95.8	3	69.6	45.1	38.6	30.5	4	66.9
16	92.2	3	78.8	81.4	39.4	31.7	2.9	109.6
17	82.3	2.5	75.9	116.9	31	30.5	2.3	157
18	56	2	61	114.2	23.2	29.8	1.6	195.2
19	30	2.5	36.9	81.4	14.7	22.1	1.1	196.6
20	10.3		15.9	65.1	8.7	17.5		146.1

Age	Hip replacement	Infections of the urinal system	Knee replacement	Pneumonia	Renal failure	TIA	Tonsils and Adenoids	Urolithiasis
0		19.2		54	0.7		2.4	0.3
1		3.4		18.8	0.3		19.4	0.4
2		1.9		4.2	0.3		9.7	0.3
3		1.1		1.5	0.4		13.6	0.3
4	0.1	1.7		2	0.5		39.1	0.5
5	0.1	2.4	0.1	2	0.8	0.1	26.7	1.5
6	0.3	2.4	0.2	2.3	0.9	0.2	14.6	2.7
7	0.4	2.7	0.3	3.4	1	0.3	10.6	3.8
8	0.8	3.7	0.8	4.3	1.3	0.4	7.2	4.4
9	1.6	4.9	1.6	4.1	1.9	0.7	3.1	4
10	3.3	6.4	2.8	5.4	2.1	1.5	1.4	5.1
11	7.3	8.6	4.6	6	2.9	1.8	1.2	5.8
12	16.1	10.6	8.4	9.6	4.1	3.4	1.2	6.5
13	26.9	13.7	13.2	15.3	5.7	4.1	0.7	7.4
14	50.8	19.1	20.5	24	8.1	6	0.4	8.5
15	61.4	25.2	37.4	37.3	10	9.7	0.2	9.1
16	65.8	33.8	38.1	51.7	13.9	12	0.2	8.3
17	56.9	42.1	31.1	72.4	13.7	14.6		5.6
18	36.8	43.5	14.9	89.6	12.6	14.9		4.3
19	13.5	38.6	7.4	96.2	12.6	11.6		1.6
20		27.8		77	11.9	12.7		
			Treated	Control				

IV.3 Basic descriptive statistics

After analyzing the most important aspects of our dataset that are essential for this study, we make a short discretion of all the variables that will be used in our models.

6. Table 1.2 Basic descriptive statistics.

Variable	Obs	Mean	Std. Dev.	Min	Max
age	6048	10	6.055801	0	20
total	5548	32.59975	74.2021	0	754.8
outpatient	4698	15.14817	66.40644	0	741.3
inpatient	5475	20.00113	35.84628	0	328.4
b_segm	6048	0.246528	0.431025	0	1
SID_chance	6048	1	0.866097	0	2
year	6048	2001.5	5.188556	1993	2010

As mentioned above, the variable "age" is categorical and contains 21 age clusters (0-20) of 5 years of age each. One exception is the group 0 that contains information for the newborns only. The variable "total" represents the total admission rate per 10000 inhabitants in the Dutch population. In the total sample (all years, all conditions, and all ages), the average admissions are 32.6 and the maximum are 754.8 per 10000 inhabitants. This variable is the total of the inpatient and outpatient admission rates. As one can see in 5. Table 1.2 there are in total 500 missing observations from this variable.

Based on the table 1.2 above we observe the outpatient and inpatient admission rates per 10000 are represented by the variables "outpatient" and "inpatient" respectively. The "outpatient" and "inpatient" have 1350 and 573 missing observations respectively. We can see that the "outpatient" has a smaller mean than the "inpatient" but the maximum value is higher among the "outpatient" admissions. The variable "b-segment" is a binary variable that equals to one if a condition belongs to the B-segment in that particular year and zero otherwise. The variable "SID-chance" is the qualitative variable we extracted from Pomp's paper. This categorical variable equals to 0 if the chance of SID is small, 1 if the chance of SID is middle and 2 if the chance of SID is high. This variable is time invariant and it takes one value for each condition in the dataset. As described in Pomp's paper, those results for the 150 DBC's have been calculated based on information for the years 2006 and 2007. In our analysis we assume that the chance of SID has not changed significantly through the years. So we expand those results for all years in questions. Last the variable "year" represents the year we observe the particular admission rates. We have yearly information from 1993 until 2010 without any missing years, for all conditions and age groups.

We need to include one potential issue of this particular dataset in our model. The variable total does not follow the normal distribution. All known transformations cannot correct this problem. By looking at the detailed summary statistics (Annex XI.1) of every condition, we see that the variable total has very high variance, which in most cases is higher than 1000. Excluding the very small and very high values reduces this variation, and that problem can be solved in most cases when the ages included in the model are only between 60 and 85. So for the course of this study and the purpose of simplicity (without loss of generality), we ignore this finding, by understanding the problems it might cause in our regressions.

V. Estimation results

Table 1.3 shows the coefficients of the Difference-in-Difference regressions. We are only interested in the interaction term (D_{it}) of those regressions. Here we compare the results of models 1.1 and 1.6. In the table below we have included the qualitative variable of Pomp(2002) for comparison of results.

7. Table 1.3 Estimation output of the different models by ISHMT.

ISHMT	Simple DID aggregated	Simple DID all variation	Random Effects GLS estimation	Random Effects GLS estimation with bootstrapped variance estimation	Chance of SID Pomp(2009)
Angina pectoris	6.5***	2.9	2.3	2.3	0
Asthma	1.6	1.0	1.0	1.0	2
COPD	-5.5***	-3.2	-3.2	-3.2	0
Cataract	189***	168.5***	168.9***	168.9***	2
Chest pain	39.4***	35.4***	35.6***	35.6***	2
Cholelithiasis	-1.6*	-1.9**	-1.3	-1.3*	1
Ear infections	-21.5	-12.5	-12.5	-12.5	2
Heart failure	36.8***	36.1**	36.2**	36.2**	0
Hip replacement	5.3***	13.8***	13.2***	13.2***	1
Infections of the urinal	31.5***	25.3***	25.3***	25.3***	1
Knee replacement	25.5***	24.4***	24.1***	24.1***	1
Pneumonia	50.6***	50.3***	50.3***	50.3***	0
Renal failure	8.4***	8***	8.3***	8.3***	0
TIA	14.7***	13.1***	12.8***	12.8***	0
Tonsil and adenoid	-34.2***	-26.7***	-27.2***	-27.2***	2
Urolithiasis	1.0	1.9***	1.9***	1.9***	2
legend: * p<.1; ** p<.05; *** p<.01					

In the table 1.3 the coefficients represent the difference the two groups before and after 2001. The stars index the statistically significance at α equals to 1%, 5% and 10%. If the coefficient of an ISHMT is insignificant then there is no effect of SID that caused by the policy change. Based on the table 1.3 we analyze the effects of the policy change in 2001 by ISHMT. Comparing the coefficients we can see that the differences between the models are

rather small (almost identical excluding the aggregated simple DID model). Thus, the simple DID model would generate trustworthy results even though it ignores age and year specific effects or random trends that could not be captured by those two variables.

When the coefficient is positive and significant we classify two categories. The first one includes coefficients higher than 35, where the effect of SID was very intense within the treated group after the policy change. The other group consists of ISHMTs with coefficients lower than 35 were the intensity of SID after 2001 that was caused by the policy change was milder. Last when the coefficients are negative but statistically significant we conclude that in these ISHMTs the difference of the two groups after 2001 decreased. This indicates that the SID targeted the group with lower admission or that there are other reasons that can explain this reaction like targeted technologies for those groups, interventions etc.

First we analyze the ISHMTs with insignificant coefficients. Starting with Angina pectoris, we see that only in the simple DID model on the aggregated data as described above the interaction term is statistically significant. Based on the modeling part of this analysis we base our results on all other regression models. This also applies to all conditions / procedures in question. The results from the aggregated DID model are used only for comparison of the modeling methods. So based on the basic models, where the interaction term is not statistically significant but positive, we reach to the following conclusion: The difference between the two groups (i.e. treated and control) before and after 2001 increased by 2.9 per 10000 admissions in the simple DID model and by 2.3 per 10000 admissions in the other two methods. This difference though, is not statistically significant. This insignificance indicates that the policy change in 2001 had a benign effect on the admissions of the treated group compared to the control group. So we conclude that the change in the budgeting system of Dutch hospitals did not affect supplier induced demand for Angina pectoris. Or there is no ambiguity of treatment in Angina pectoris among the ase groups we have selected. On a more technical note, it is expected that the coefficients from the last two methods are identical. The estimation method of the coefficients is the same and the only difference is the calculation of the SEs.

Exactly the same results as in Angina pectoris we get for Asthma and COPD. Due to the chronic nature of those conditions, information asymmetry is finite. Meaning that patients know the true nature of their condition and the specifics around it. So after the lift of the budgeting restrictions in 2001, we did not find any evidence for SID that can create a problematic increase of the admission rates. Though it is important to mention that, as

Angina pectoris, the effect of SID on Cataract (besides not significant) was positive, although in COPD is negative indicating that the difference of the two groups became smaller after 2001 but due to the insignificance we conclude that the difference of the two groups before and after 2001 remain the same. From all the conditions mentioned before for Ear infections, we conclude that the lift of the budgeting restrictions did not increased the SID effect. We base this conclusion on the insignificant coefficient.

An explanation might be the pressure from the policy makers to the physicians to decrease the admissions even from 1993 in the treated group. Last in this category of ISHMTs we focus on Cholelithiasis. The coefficients are negative indicating that the difference of the two groups before and after 2001 decreased but only statistically significant for α=10%. So it is also here safe to conclude that the change of hospitals' and physicians' payments system did not affected the admissions of the two groups in indicating that effect of SID is limited or not existing in Cholelithiasis.

Secondly, we analyze the ISHMTs with highly positive and statistically significant coefficients. In this category we find high in magnitude and positive coefficients. The highest positive coefficient is observed in Cataract. In the same category but with smaller coefficients, we find Chest pain, Heart failure and Pneumonia. The coefficients of this category show the difference of the two groups before and after 2001. For cataract the difference increased by 169 per 10000 admissions between the two periods (pre and post 2001). All other differences can be seen in table 1.3.

Based on our assumptions this difference may produce evidence for supplier induced demand that was generated by the lift of the budgeting restrictions in the payment system of hospitals in 2001. So for Cataract, Chest pain, Heart failure and Pneumonia the effects of SID increased by the policy change and the age groups that were targeted for increase of demand were those with the higher admission rates before the policy change. In all those conditions, based on table 1.2, the group mostly affected, is populated by ages above 70. This is a proof of the assumption that first age is an important factor in SID and that the intensity of SID seems to be higher within older age groups.

The third group we identified through our analysis is the one with positive and significant but the coefficients' magnitude is relatively smaller than the once in the second group described above. In this category we find Hip and Knee replacement, Infections of the Urinal system, Renal failure, TIA and Urolithiasis. Within those ISHMTs the change of the budgeting

system in health care affected the treated group more than the control. The intensity of SID can be seen through the coefficient. In this category, the effects of the policy change are smaller than the previous group, indicating that there is "space" for SID, but it is limited. As in the high intensity of SID group, in this one the treated group consists of patients older than 75.

Last we need to analyze infections of Tonsils and Adenoids as a separate condition, being the only ISHMT that gave a negative and statistically significant coefficient. This indicates that the difference of the two groups was reduced by 27 admissions per 1000 inhabitance. This reflects the practitioners' evidence that even before 2000 there was a general pressure to the physicians to decrease admissions and procedures for infections of Tonsils and Adenoids. Perhaps, the physicians lowered the admissions for this particular ISHMT or, just slightly shifted the target group of "inducement" combined with a reduction of procedures.

VI. Robustness analysis

VI.1 General analysis

After our analysis we conducted a series of robustness checks in order to test the validity of our results. In the first test we excluded all variations before 1995 (Table A.1).In this particular year, fee-for-service was replaced by a fixed budgeting system for the Dutch hospitals. Next (Table A.2), we tried to solve the very high variance problem by omitting all observations below 10. Only Asthma remains unchanged since almost all observations are below 10. Last and most important, we intended to prove the validity of our most strong assumption, how the treated and control group is to be separated. So, in the last check (Table A.3) we rearranged the two groups. We calculated the total average of all years by age groups. For each condition we had one observation per age group. Then we took the mean of those observations. If an age group was higher than the mean, it transferred to the treated group. Otherwise it remained to the control group. In general the coefficients showed no significant changes. Only when the treated group was made broader, there

were changes worth mentioning. In this particular case we see that the coefficients start to differentiate from the simple difference in difference model and those using random effects. This indicates the importance of appropriate selection criteria for the choice of the two groups.

VI.2 Excluding all observation before 1995

8. Table A.1 Robustness check by excluding all information before 1995.

ISHMT	Simple DID aggregated	Simple DID aggregated excluding years lower than 1995	Simple DID all variation	Simple DID all variation excluding years lower than 1995	Random Effects GLS estimation	Random Effects GLS estimation excluding years lower than 1995	Random Effects GLS estimation with bootstrapped variance estimation	Random Effects GLS estimation with bootstrapped variance estimation excluding years lower than 1995	Chance of SID Pomp(2009)
Angina pectoris	6.5***	5.2**	3.0	2.4	2.3	1.6	2.3	1.6	0
Asthma	1.6	1.7	1.0	1.0	1.0	1.0	1.0	1.0	2
COPD	-5.5***	-5.9**	-3.2	-3.8	-3.2	-3.8	-3.2	-3.8	0
Cataract	189***	144.5***	168.5***	126.3***	168.9***	126.8***	168.9***	126.8***	2
Chest pain	39.4***	36.6***	35.6***	33.1***	35.6***	33.0***	35.6***	33.0***	2
Cholelithiasis	-1.6*	-0.5	-1.9**	-1.0	-1.3	-0.5	-1.3*	-0.5	1
Ear infections	-21.5	-3.6	-12.5	1.3	-12.5	1.1	-12.5	1.1	2
Heart failure	36.9***	35.9***	36.1**	33.9**	36.2**	34.1**	36.2**	34.1**	0
Hip replacement	5.3***	4.7***	13.8***	11.9***	13.2***	11.3***	13.2***	11.3***	1
Infections of the urinal	31.5***	30.7***	25.3***	24.8***	25.3***	24.8***	25.3***	24.8***	1
Knee replacement	25.5***	22.7***	24.4***	22.0***	24.1***	21.6***	24.1***	21.6***	1
Pneumonia	50.7***	45.8***	50.3***	46.2***	50.3***	46.2***	50.3***	46.2***	0
Renal failure	8.5***	7.9***	8.1***	7.6***	8.3***	7.9***	8.3***	7.9***	0
TIA	14.7***	15.0***	13.2***	13.5***	12.8***	13.2***	12.8***	13.2***	0
Tonsil and adenoid	-34.2***	-26.8**	-26.7***	-21.6**	-27.2***	-22.2***	-27.2***	-22.2***	2
Urolithiasis	1.0	1.2	1.9***	2.1***	1.9***	2.1***	1.9***	2.1***	2
legend: * p<.1; ** p<.05; *** p<.01									

By comparing the results before and after the exclusion of the years before 1995, we see that our results are in general consistent. The sign and significance remain the same, though the coefficients are slightly different as is to be expected (e.g. for Angina pectoris the coefficients goes from 3 to 2.4 and for COPD from -3.2 to -3.8). Those changes are expected and acceptable. So the starting year of the analysis (i.e before and after 1995) has a minor effect on the total results for all estimating methods.

VI.3 Excluding all observations below 10

One of the major problems of all admission datasets based on age groups is that some age have admission rates very close to zero. It is important to see the effect of excluding all admissions below 10 per 10000. This way we the variance is decreased which is very high as described in the data section. It is heartwarming that we see no changes in the sign of the

coefficients and we have a change in significance only for Urolithiasis and Cholelithiasis. Something that can be explained easily by the epidemiology of the two conditions, which have common characteristics. In the end based on the above findings it is safe to conclude that excluding the very small admissions will not affect the significance of our results but it will slightly change the magnitude as in the previous robustness case.

9. Table A.2 Robustness check by omitting all observations below 10.

ISHMT	Simple DID aggregated	Simple DID all variation	Simple DID all variation excluding obs with lower than 10 admission rate	Random Effects GLS estimation	Random Effects GLS estimation excluding obs with lower than 10 admission rate	Random Effects GLS estimation with bootstrapped variance estimation	Random Effects GLS estimation with bootstrapped variance estimation excluding obs with lower than 10 admission rate	Chance of SID Pompl(2009)
Angina pectoris	6.5***	3.0	1.5	2.3	1.3	2.3	1.3	0
Asthma	1.6	1.0	-0.2	1.0	1.1	1.0	1.1	2
COPD	-5.5***	-3.2	-4.3	-3.2	-2.9	-3.2	-2.9	0
Cataract	189***	168.5***	160.5***	168.9***	137.1***	168.9***	137.1***	2
Chest pain	39.4***	35.5***	31.8***	35.6***	22.9***	35.6***	22.9***	2
Cholelithiasis	-1.6*	-1.9**	-1.1	-1.3	-1.9**	-1.3*	-1.9*	1
Ear infections	-21.5	-12.5	-11.3	-12.5	-11.9	-12.5	-11.9	2
Heart failure	36.9***	36.1**	41.2**	36.2**	45.2***	36.2**	45.2***	0
Hip replacement	5.3***	13.8***	13.3***	13.2***	9.7***	13.2***	9.7***	1
Infections of the urinal	31.5***	25.4***	24.7***	25.3***	18.1***	25.3***	18.1***	1
Knee replacement	25.5***	24.5***	20.1***	24.1***	12.4***	24.1***	12.4***	1
Pneumonia	50.7***	50.4***	52.5***	50.3***	44.9***	50.3***	44.9***	0
Renal failure	8.5***	8.1***	6.9***	8.3***	5.8***	8.3***	5.8***	0
TIA	14.6***	13.2***	10.7***	12.8***	9.6***	12.8***	9.6***	0
Tonsil and adenoid	-34.2***	-26.7***	-27.7**	-27.2***	-27.9***	-27.2***	-27.9***	2
Urolithiasis	1.0	1.9***	1.2	1.9***	-0.3	1.9***	-0.3	2

legend: * p<.1; ** p<.05; *** p<.01 (Asthma has been excluded from this analysis since the admissions are all generally lower than 10)

VI.4 Rearranging the treated and control group

In the last robustness check we rearranged the age groups in the treated and control groups. The results in this case need special attention because they are different based on the estimation method that has been used. Starting with the simple DID model (equation 1.1) we find big differences both on significant as in magnitude and sign. Thus the simple OLS is very sensitive to changes in the age clusters. If that was our only estimation method we should be very cautious. Of course a clustering method could solve this problem. By using clustering by K-means. It is very important to mention that for the expansion of the treated group we also used the K-mean clustering method for the total admissions and then compared the results to all other methods from the table A3. The results are in the appendix.

10. Table A.3 Robustness check by rearranging the treated and control group.

ISHMT	Simple DID aggregated	Simple DID all variation	Simple DID all variation with groups based on the average (lower and higher)	Random Effects GLS estimation	Random Effects GLS estimation with groups based on the average (lower and higher)	Random Effects GLS estimation with bootstrapped variance estimation	Random Effects GLS estimation with bootstrapped variance estimation groups based on the average (lower and higher)	Chance of SID Pompl(2009)
Angina pectoris	6.5***	3.0	28.3***	2.3	2.7	2.3	2.7	0
Asthma	1.6	1.0	15.3***	1.0	1.3	1.0	1.3	2
COPD	-5.5***	-3.2	20.9**	-3.2	-3.0	-3.2	-3.0	0
Cataract	189***	168.5***	346.1***	168.9***	177.3***	168.9***	177.3***	2
Chest pain	39.4***	35.5***	38.5***	35.6***	36.1***	35.6***	36.1***	2
Cholelithiasis	-1.6*	-1.9**	7.8***	-1.3	-1.2	-1.3*	-1.2	1
Ear infections	-21.5	-12.5	54.9	-12.5	-12.2	-12.5	-12.2	2
Heart failure	36.9***	36.1**	91.9***	36.2**	38.7**	36.2**	38.7**	0
Hip replacement	5.3***	13.8***	26.7***	13.2***	13.4***	13.2***	13.4***	1
Infections of the urinal	31.5***	25.4***	56.1***	25.3***	26.2***	25.3***	26.2***	1
Knee replacement	25.5***	24.5***	33.2***	24.1***	24.9***	24.1***	24.9***	1
Pneumonia	50.7***	50.4***	112.8***	50.3***	52.0***	50.3***	52.0***	0
Renal failure	8.5***	8.1***	9.8***	8.3***	8.5***	8.3***	8.5***	0
TIA	14.6***	13.2***	16.8***	12.8***	13.5***	12.8***	13.5***	0
Tonsil and adenoid	-34.2***	-26.7***	94.9	-27.2***	-27.2***	-27.2***	-27.2*	2
Urolithiasis	1.0	1.9***	5.7***	1.9***	2.0***	1.9***	2.0*	2
legend: * p<.1; ** p<.05; *** p<.01								

We go now to the comparison based on the random effects estimator. What is clear from the table A3 above, small changes in the age groups that populate the treated and the control group have no effect on the significance and sign of the coefficients. What we observe is a small change in the magnitude. As a paradigm for Angina pectoris the coefficient changed from 2.3 to 2.7. For Asthma from 1 to 1.3, for Cataract from 168.9 to 177,3 and for Chest pain from 35.5 to 36.1. Thus we can conclude that the random effect method of estimation for the effect of SID is not sensitive to small changes in the age clusters that populate the treated and control group.

This particular finding is very important. One of the strongest assumptions of this study is the way that the treated and control group is separated based on age. With this sensitivity analysis we were able to relax the initially believed strong need for a perfect separation measurement.

VII. Concluding remarks and limitations

VII.1 Concluding remarks

Within the context of this study we found that half of the results are in line with (Pomp, 2009) as shown in table 1.4 below. In order to facilitate the direct comparison of our results with (Pomp, 2009) we transformed our coefficients into a three level scale. It has been proved that in 10 out of 16 cases the policy change in 2001 had a significant effect in the admission rates of the two groups, proving the existence of ambiguity of treatment in those cases. The reason is that since there is a reward in overproduction after 2001, physicians encourage increased demand. However, this demand is more evident mainly in the age groups with the higher admission rates, i.e. mostly elderly over 75 years of age. Also we concluded that since there is a significant effect, age is an important factor of this analysis.

The fixed budgeting system constrained physicians from overproduction of health care for unnecessary treatment. After this constrain was lifted, physicians increased admissions rapidly, and subsequently the costs that had to be covered. By looking at this analysis, the policy maker can focus on certain ambiguities and try to contain the costs by understanding the effects the previous policies had on the admission rates. This effect was significant in a large number of the conditions/procedures in question. In our analysis we were able to ascertain the effects in admissions caused by technology and efficiency improvement and clearly see the effects of ambiguity in specific treatments. By separating the age groups we were able to define the target of those ambiguous treatments. More specifically, the higher effect of the SID produced by the policy change, can be found in Cataract, Chest pain, Heart failure and Pneumonia. A significant but lower effect, can be found in Hip and Knee replacement, Infections of the Urinal system, Renal failure, TIA and Urolithiasis. Last by comparing different econometric models we concluded that even in the most simple for the Difference-in-Difference regression model gave the similar results as the more sophisticated models that corrected for the general problems of this method.

VII.2 Limitations

It is important to mention the limitations of this model. Firstly, the model will give us the effect of ambiguity of treatment (proof of existence) under the assumption that only the treated group is affected significantly by the policy change. If both groups are equally affected, the model will prove no influence whatsoever since it compares the two groups. As a result we would expect a very small statistically significant D coefficient. In this case, we need to interpret the result with the graph of average admission rates by year, for the two age groups. Another limitation concerns the outcome used to test the effect of SID within ambiguity of treatment. As a basic outcome we used the total admission rates i.e. inpatient plus outpatient treatments. Thus, we ignore the substitution effect that may be present between inpatient and outpatient admission rates. In future, a separate analysis of those two components should be conducted or the model should be integrated in order to capture the different trends that are generated through them.

The next limitation concerns the Random-Effects model only. This model is valid in the case that the results of the policy change have an immediate effect on the admissions rate. In case of delayed results of the policy, the model needs to include lags of the dependent variable as independent. So we need to transform our model to a dynamic fixed effects model that requires very strong assumptions. In general the Difference-in-Difference modes are biased when there are other factors that affect the difference in admissions between the treated and the control group. One other limitation is the separation of the age clusters into groups. This procedure was not based to any theoretical background and is based to very strong assumptions. So in order to check the validity of those technical assumptions we will do a robustness check by changing the age groups and other aspects of the model and we proved that in the random effects model the need for a theoretical based measurement of separation between the two groups is not needed.

The last limitation concerns the inclusion of the major policy change with the 2006 Health Insurance Act. In the future we need to extend our model so it can account for the 2006 policy change and create a Difference in Difference in Difference model that can show the clear effect of all policy changes that took place after 1993.

11. Table 1.4 Comparison of results with the qualitative variable from Pomp (2009).

ISHMT	Final result	Chance of SID Pomp(2009)
Angina pectoris	Small or no SID effect	0
Asthma	Small or no SID effect	2
COPD	Small or no SID effect	0
Cataract	High SID effect	2
Chest pain	High SID effect	2
Cholelithiasis	Small or no SID effect	1
Ear infections	Small or no SID effect	2
Heart failure	High SID effect	0
Hip replacement	Lower but existing SID effect	1
Infections of the urinal	Lower but existing SID effect	1
Knee replacement	Lower but existing SID effect	1
Pneumonia	High SID effect	0
Renal failure	Lower but existing SID effect	0
TIA	Lower but existing SID effect	0
Tonsil and adenoid	Arbitrary results	2
Urolithiasis	Lower but existing SID effect	2

VIII. References

1. Angrist, J. D., & Pischke, J.-S. (2008). Parallel Worlds: Fixed Effects, Differences-in-differences, and Panel data. In *Mostly Harmless Econometrics: An Empiricist's Companion* (pp. 166-186).

2. Arrow, K. (1963, Dec.). Uncertainty and the Welfare Economics of Medical Care. *Am. Econ. Rev.*, pp. 53, 941-73. 2.

3. Baltagi, B. H. (2005). The One-way Error Component Regression Model. In B. H. Baltagi, *Econometric Analysis of Panel Data.* Wiley & Sons.

4. Breusch, T. S., & Pagan, A. R. (1980). The Lagrange multiplier test and its applications to model specification in econometrics. *Review of Economics Studies*, pp. 239-253.

5. *Centraal Bureau voor de Statistiek*. (2012). Retrieved October 2012, from http://statline.cbs.nl/statweb/

6. Douven, R., Mocking, R., & Mosca, I. (2012). The Effect of Physician Fees and Density Differences on Regional Variation in Hospital Treatment. *NZa Research Paper, 2012 - 01*.

7. Folland, S., Goodman, A., & Stano, M. (2010). *The economics of Health and Health care.* Pearson.

8. Folmer, K., & Westerhout, E. (2002). Financing Medical Specialist Services in the Netherlands: Welfare Implications of Imperfect Agency. *CPB Discussion Paper, No. 6*.

9. Getzen, T. (2006). *Health Economics and Financing.* Wiley.

10. Hansen, C. B. (2007). Generalized Least Squares Inference in Panel and Multilevel Models with Serial Correlation and Fixed effects. *Journal of Econometrics,*, pp. 140, 670-694.

11. Hounkannounon, B. (2011). Bootstrapping differences-in-differences estimates.

12. Marianne, B., Duflo, E., & Mullainathan, S. (2004a). Are Emily and Greg More Employable than Lakisha and Jamal? A Field Experiment on Labor Market Discrimination. *The American Economic Review*, pp. 94, 991-1013.

13. Marianne, B., Duflo, E., & Mullainathan, S. (2004b). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics,*, pp. 119, 249-275.

14. Moreno-Serraa, R., & Wagstaff, A. (2010). System-wide impacts of hospital payment reforms: Evidence from Central and Eastern Europe and Central Asia. *Journal of Health Economics*, pp. 585–602.

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

16. Schut, F. T., & Varkevisser, M. (2012). The Netherlands. In *Waiting Time Policies in the Health SectorWhat Works?* OECD.

17. van de Vijsel, A. R., Engelfriet, P. M., & Westerta, G. P. (2011). Rendering hospital budgets volume based and open ended to reduce waiting lists: Does it work? *Health Policy*, pp. 60-70.

18. van Dijk, C. E., van den Berg, B., Verheij, R. A., Spreeuwenberg, P., Groenewegen, P. P., & de Bakker, D. H. (2012). Moral hazard and supplier-induced demand: empirical evidence in general practice. *Health Economic*, pp. 1099-1050.

19. Verbeek, M. (2004). Heteroskedasticity and Autocorrelation. In M. Verbeek, *A Guide to modern econometrics* (pp. 79-119). Wiley & Sons.

IX. Annex

IX.1 By condition detailed descriptive statistics

12. Descriptive statistics by ISHMT.

Angina pectoris		
total		
Percentiles Smallest		
1% .1 0		
5% .1 0		
10% .3 .1 Obs	301	
25% 4.2 .1 Sum of Wgt.	301	
50% 34.7 Mean	43.46744	
Largest Std. Dev.	39.3159	
75% 81 117.5		
90% 104.1 117.8 Variance	1545.74	
95% 109.1 119.3 Skewness	.4370418	
99% 117.5 124.6 Kurtosis	1.738493	
Variable Mean Std. Dev. Min	Max	Observations
total overall 43.46744 39.3159 0	124.6	N = 301
between 40.06458 .0923077	108.6722	n = 17
within 5.968307 28.82855	65.68411	T-bar = 17.7059
Asthma		
total		
Percentiles Smallest		
1% 1.5 1.3		
5% 2 1.4		
10% 2.2 1.4 Obs	363	
25% 2.5 1.5 Sum of Wgt.	363	
50% 3.2 Mean	5.666391	

Largest Std. Dev.	6.411805	
75% 5.1 28.7		
90% 11.9 30.1 Variance	41.11124	
95% 22.7 30.6 Skewness	2.530626	
99% 28.7 31 Kurtosis	8.315562	
Variable Mean Std. Dev. Min	Max	Observations
total overall 5.666391 6.411805 1.3	31	N = 363
between 6.151795 2.561111	24.64444	n = 21
within 1.874276 -.5780535	13.06084	T-bar = 17.2857
COPD		
total		
Percentiles Smallest		
1% .2 .1		
5% .3 .2		
10% .4 .2 Obs	378	
25% 1.2 .2 Sum of Wgt.	378	
50% 9.75 Mean	23.96032	
Largest Std. Dev.	28.38783	
75% 42.1 88.1		
90% 75.6 88.3 Variance	805.869	
95% 82.8 88.6 Skewness	1.012276	
99% 88.1 92.9 Kurtosis	2.580152	
Variable Mean Std. Dev. Min	Max	Observations
total overall 23.96032 28.38783 .1	92.9	N = 378
between 28.72818 .4111111	83.74444	n = 21
within 4.21618 10.16032	39.81032	T = 18
Cataract		
total		
Percentiles Smallest		
1% .2 .2		
5% .3 .2		
10% .4 .2 Obs	377	
25% .6 .2 Sum of Wgt.	377	

50% 11.3 Mean	132.861	
Largest Std. Dev.	202.3474	
75% 205.1 716.6		
90% 490.8 719.2 Variance	40944.47	
95% 599.5 735.5 Skewness	1.501027	
99% 716.6 754.8 Kurtosis	4.049765	
Variable Mean Std. Dev. Min	Max	Observations
total overall 132.861 202.3474 .2	754.8	N = 377
between 200.463 .2888889	606.3333	n = 21
within 50.1551 -116.9723	334.5055	T = 17.9524
Chest pain		
total		
Percentiles Smallest		
1% .1 .1		
5% .1 .1		
10% .2 .1 Obs	370	
25% 2.8 .1 Sum of Wgt.	370	
50% 19.75 Mean	28.78216	
Largest Std. Dev.	30.43647	
75% 42.6 122		
90% 75.7 124.3 Variance	926.3788	
95% 96.1 126.4 Skewness	1.249759	
99% 122 133.3 Kurtosis	3.983692	
Variable Mean Std. Dev. Min	Max	Observations
total overall 28.78216 30.43647 .1	133.3	N = 370
between 24.90088 .1411765	68.57778	n = 21
within 18.40191 -12.39562	93.50439	T-bar = 17.619
Cholelithiasis		
total		
Percentiles Smallest		
1% .1 .1		
5% .1 .1		
10% .3 .1 Obs	350	

25% 8.7 .1 Sum of Wgt.	350	
50% 19.75 Mean	18.46714	
Largest Std. Dev.	11.74351	
75% 28.7 38.3		
90% 33.5 38.5 Variance	137.9101	
95% 35.1 38.9 Skewness	-.1823286	
99% 38.3 40.5 Kurtosis	1.816778	
Variable Mean Std. Dev. Min	Max	Observations
total overall 18.46714 11.74351 .1	40.5	N = 350
between 12.31419 .1	34.81111	n = 21
within 2.512605 9.872698	26.64492	T-bar = 16.6667
Ear infections		
total		
Percentiles Smallest		
1% 1.8 1.2		
5% 2.7 1.5		
10% 4.2 1.7 Obs	362	
25% 7.3 1.8 Sum of Wgt.	362	
50% 9 Mean	34.93564	
Largest Std. Dev.	68.48568	
75% 12.8 296.1		
90% 109 299.2 Variance	4690.289	
95% 226.4 307.8 Skewness	2.645556	
99% 296.1 327.8 Kurtosis	8.744632	
Variable Mean Std. Dev. Min	Max	Observations
total overall 34.93564 68.48568 1.2	327.8	N = 362
between 68.10114 2.929412	259.5	n = 21
within 10.30388 -5.264362	103.2356	T-bar = 17.2381
Hartfailure		
total		
Percentiles Smallest		
1% .1 0		
5% .1 0		

10% .1 .1 Obs	351	
25% .4 .1 Sum of Wgt.	351	
50% 5.3 Mean	56.99259	
Largest Std. Dev.	84.46505	
75% 105.7 287.5		
90% 204.7 303.4 Variance	7134.344	
95% 230.7 310.3 Skewness	1.324862	
99% 287.5 323.6 Kurtosis	3.382887	
Variable Mean Std. Dev. Min	Max	Observations
total overall 56.99259 84.46505 0	323.6	N = 351
between 83.09958 .0666667	235.9667	n = 21
within 16.23095 -.0296278	151.3704	T-bar = 16.7143
Hip replacement		
total		
Percentiles Smallest		
1% .1 .1		
5% .1 .1		
10% .2 .1 Obs	301	
25% .8 .1 Sum of Wgt.	301	
50% 9.6 Mean	22.96545	
Largest Std. Dev.	26.6624	
75% 43.7 89.4		
90% 67.5 92 Variance	710.8837	
95% 76.3 94.1 Skewness	.9572549	
99% 89.4 94.8 Kurtosis	2.588437	
Variable Mean Std. Dev. Min	Max	Observations
total overall 22.96545 26.6624 .1	94.8	N = 301
between 26.24297 .1	75.37778	n = 19
within 5.037296 3.587671	42.38767	T-bar = 15.8421
Infections of the urinal system		
total		
Percentiles Smallest		
1% 2.4 2.1		

5% 2.8 2.2		
10% 3.4 2.2 Obs	378	
25% 5.7 2.4 Sum of Wgt.	378	
50% 13.6 Mean	24.93651	
Largest Std. Dev.	25.60432	
75% 39.5 115.4		
90% 56 116.9 Variance	655.5813	
95% 83.4 123.1 Skewness	1.558577	
99% 115.4 126.3 Kurtosis	5.204369	
Variable Mean Std. Dev. Min	Max	Observations
total overall 24.93651 25.60432 2.1	126.3	N = 378
between 23.01058 2.833333	68.87778	n = 21
within 12.24625 -6.563492	85.83095	T = 18
Knee replacement		
total		
Percentiles Smallest		
1% .1 .1		
5% .1 .1		
10% .3 .1 Obs	294	
25% 1.2 .1 Sum of Wgt.	294	
50% 7.85 Mean	16.40646	
Largest Std. Dev.	19.495	
75% 25.9 73.1		
90% 47.9 75.1 Variance	380.055	
95% 62.6 79.4 Skewness	1.375685	
99% 75.1 80.4 Kurtosis	4.047454	
Variable Mean Std. Dev. Min	Max	Observations
total overall 16.40646 19.495 .1	80.4	N = 294
between 17.38886 .1	53.16667	n = 19
within 8.873687 -9.460205	43.6398	T-bar = 15.4737
Pneumnia		
total		
Percentiles Smallest		

1% 1.7 1.5		
5% 2.2 1.5		
10% 2.4 1.6 Obs	378	
25% 4.4 1.7 Sum of Wgt.	378	
50% 12.2 Mean	34.89656	
Largest Std. Dev.	43.01761	
75% 59.7 185.8		
90% 92.9 187.4 Variance	1850.515	
95% 128.2 190.2 Skewness	1.595363	
99% 185.8 192.6 Kurtosis	5.090174	
Variable Mean Std. Dev. Min	Max	Observations
total overall 34.89656 43.01761 1.5	192.6	N = 378
between 40.57831 2.35	119.0944	n = 21
within 16.6783 -14.92011	114.4799	T = 18
Renal failure		
total		
Percentiles Smallest		
1% .3 .2		
5% .4 .2		
10% .5 .2 Obs	373	
25% 1 .3 Sum of Wgt.	373	
50% 2.7 Mean	6.083914	
Largest Std. Dev.	6.765147	
75% 10.3 28.7		
90% 15.1 29.1 Variance	45.76721	
95% 20.1 29.6 Skewness	1.493887	
99% 28.7 38.5 Kurtosis	5.139575	
Variable Mean Std. Dev. Min	Max	Observations
total overall 6.083914 6.765147 .2	38.5	N = 373
between 6.268579 .4722222	18.04444	n = 21
within 2.851877 -1.36053	26.53947	T-bar = 17.7619
TIA		
total		

Percentiles Smallest		
1% .1 .1		
5% .1 .1		
10% .2 .1 Obs	299	
25% .8 .1 Sum of Wgt.	299	
50% 5.7 Mean	9.764883	
Largest Std. Dev.	11.15898	
75% 15.6 45.5		
90% 25.1 48 Variance	124.5228	
95% 36.3 50.3 Skewness	1.481636	
99% 48 51.7 Kurtosis	4.958006	
Variable Mean Std. Dev. Min	Max	Observations
total overall 9.764883 11.15898 .1	51.7	N = 299
between 9.602185 .1	26.71667	n = 19
within 6.004549 -.9184507	35.88155	T-bar = 15.7368
Tonsil and adenoid		
total		
Percentiles Smallest		
1% .2 .1		
5% .3 .1		
10% .3 .2 Obs	312	
25% 1.2 .2 Sum of Wgt.	312	
50% 7.55 Mean	45.10353	
Largest Std. Dev.	102.0834	
75% 32.85 462.6		
90% 147.3 471.3 Variance	10421.02	
95% 385.8 481 Skewness	3.139189	
99% 462.6 493.1 Kurtosis	11.96276	
Variable Mean Std. Dev. Min	Max	Observations
total overall 45.10353 102.0834 .1	493.1	N = 312
between 100.42 .2333333	426.4167	n = 19
within 9.601085 -1.813141	111.7869	T-bar = 16.4211
Urolithiasis		
total		

Percentiles Smallest		
1% .3 .2		
5% .3 .3		
10% .4 .3 Obs	361	
25% 2.1 .3 Sum of Wgt.	361	
50% 8.3 Mean	8.861219	
Largest Std. Dev.	6.827232	
75% 14.6 22.9		
90% 18.5 23.8 Variance	46.6111	
95% 20.6 23.8 Skewness	.3122924	
99% 22.9 25.1 Kurtosis	1.911358	
Variable Mean Std. Dev. Min	Max	Observations
total overall 8.861219 6.827232 .2	25.1	N = 361
between 6.582762 .35	18.93889	n = 21
within 2.230939 4.261218	15.32789	T-bar = 17.1905

IX.2 New treated and control group based on separation from the average

13. New treated and control group based on average admissions

Age	Angina Pectoris	Asthma	COPD	Cataract	Chest pain	Cholelithiasis	Ear Infections	Hart failure
0		18.9	10.6		0.2	0.2	14.3	1.7
1		21.2	3.7	0.1			8.5	0.1
2		5.4	0.6	0.2	0.1	0.1	9.6	
3		2.6	0.2	0.1	0.1	0.2	11.7	0.1
4		1.9	0.3	0.2	0.9	1.2	7.3	0.2
5	0.1	2.2	0.4	0.2	2	4.2	5.4	0.2
6	0.3	2	0.4	0.3	3	8.9	4.7	0.2
7	1	2.4	0.7	0.2	5.1	11.8	5	0.3
8	3.6	2.6	1.4	0.3	8.3	12.2	5.6	0.6
9	8.3	2.5	2.7	0.9	12.3	13.6	6.5	1.5
10	17.4	2.7	5	1.7	18.8	15.3	5.8	2.9
11	29	2.6	8.1	2.4	24	18	6.1	5
12	43.1	2.5	16.2	4.5	28.8	22	6.7	10.1
13	57.8	2.6	26.3	8.9	31.3	23.2	5.5	20.3
14	78.1	2.6	48.9	20.6	35.7	26.1	5	37.8
15	95.8	3	69.6	45.1	38.6	30.5	4	66.9
16	92.2	3	78.8	81.4	39.4	31.7	2.9	109.6
17	82.3	2.5	75.9	116.9	31	30.5	2.3	157
18	56	2	61	114.2	23.2	29.8	1.6	195.2
19	30	2.5	36.9	81.4	14.7	22.1	1.1	196.6
20	10.3		15.9	65.1	8.7	17.5		146.1

Age	Hip replacement	Urinal inf	Knee replacement	Pneumonia	Renal failure	TIA	Tonsils and Adenoids	Urolithiasis
0		19.2		54	0.7		2.4	0.3
1		3.4		18.8	0.3		19.4	0.4
2		1.9		4.2	0.3		9.7	0.3
3		1.1		1.5	0.4		13.6	0.3
4	0.1	1.7		2	0.5		39.1	0.5
5	0.1	2.4	0.1	2	0.8	0.1	26.7	1.5
6	0.3	2.4	0.2	2.3	0.9	0.2	14.6	2.7
7	0.4	2.7	0.3	3.4	1	0.3	10.6	3.8
8	0.8	3.7	0.8	4.3	1.3	0.4	7.2	4.4
9	1.6	4.9	1.6	4.1	1.9	0.7	3.1	4
10	3.3	6.4	2.8	5.4	2.1	1.5	1.4	5.1
11	7.3	8.6	4.6	6	2.9	1.8	1.2	5.8
12	16.1	10.6	8.4	9.6	4.1	3.4	1.2	6.5
13	26.9	13.7	13.2	15.3	5.7	4.1	0.7	7.4
14	50.8	19.1	20.5	24	8.1	6	0.4	8.5
15	61.4	25.2	37.4	37.3	10	9.7	0.2	9.1
16	65.8	33.8	38.1	51.7	13.9	12	0.2	8.3
17	56.9	42.1	31.1	72.4	13.7	14.6		5.6
18	36.8	43.5	14.9	89.6	12.6	14.9		4.3
19	13.5	38.6	7.4	96.2	12.6	11.6		1.6
20		27.8		77	11.9	12.7		